Emotion in Faces and Voices: Recognition and Production by Young and Older Adults

Simone Simonetti

B. Psych (Hons)

A Thesis Submitted for the Degree of

Doctor of Philosophy

MARCS Institute for Brain, Behaviour and Development

Western Sydney University, Australia

April, 2019

Abstract

Older adults are less accurate than younger adults at emotion recognition. Given that deficits in emotion recognition have been associated with interpersonal conflict and, in turn, a reduced quality of life, understanding how older adults process emotion is vital. Thus, the research strategy of the current thesis was to investigate specific questions concerning how older adults process emotion information, e.g., how well older adults process a variety of emotional expression types and whether having multiple sources of expressive information will improve emotion recognition; how older adults extract information from emotional expressions; where problems might occur during the emotion recognition process; and whether poor emotion recognition is related to poor emotion production. My approach was different to the majority of other studies in that I tested multi-modal, dynamic spoken expressions in addition to the unimodal and static emotional expressions that are typically used for assessing emotion recognition.

The first experiment assessed emotion recognition for auditory-visual (AV), visual-only (VO), and auditory-only (AO) speech stimuli (Chapter 2). Specifically, this study investigated whether older adults would show improved recognition performance for AV relative to unimodal expressions (i.e., VO, AO); even when the emotion expressed via one modality was not well recognised. Younger adults showed near optimal AV integration whereas older adults did not. Indeed, older adults' AV recognition performance was poorer than VO recognition when AO expressions were not well recognised. These findings indicate that for older adults, adding an additional source of emotion information does not always result in improved emotion recognition performance. Ultimately, these findings suggest how older adults may perform during emotion recognition in daily life as multimodal expressions of emotion are common during daily life.

The second series of experiments focused on VO expressions of emotion and investigated how older adults extract information from such expressions. An eye-tracking technique was used to investigate age differences for gaze behaviour during facial emotion recognition (Chapter 3). Specifically, the study explored the relationships between gaze behaviour and emotion recognition as well as gaze behaviour and age-related hearing loss. Although there were age differences in gaze behaviour (e.g., younger adults looked longer at the eye than the mouth region of faces; whereas older adults showed similar eye and mouth looking), emotion recognition performance did not appear to be explained by gaze behaviour. Further, hearing loss did not appear to account for older adults' gaze behaviour. It was suggested that older adults adopt an exploratory-gaze behaviour during emotion recognition as they may have general problems realising the most appropriate gaze strategy. This suggestion was followed up by testing older adults' gaze behaviour for linguistic expressions (Chapter 4). There were no age differences in gaze behaviour during linguistic expression recognition, suggesting that older adults are capable of realising an appropriate gaze strategy and that the observed age-related changes to gaze behaviour occur only when performing an emotion recognition task.

The final two experiments investigated potential problems that older adults may encounter during the emotion recognition process. A masked priming procedure was used to evaluate the efficiency of early emotion processing in older and younger adults (Chapter 5). Happy faces were tested as they are well recognised by older adults and so initial processing should be intact. Angry faces were also tested as older adults have difficulty recognising this emotion and, as such, it was anticipated that initial processing would be compromised for such expressions. For all presented emotions (i.e., angry, happy) younger adults showed efficient early emotion processing skills whereas older adults showed this for happy faces only. These findings suggested that for angry faces, older adults' emotion processing system may be slow and may produce ambiguous emotion information. Problems with the processing system for angry faces may account for why older adults show better recognition for some emotion types (e.g., happy) relative to others (e.g., angry). The final experiment investigated the production of emotion to further understand older adults' ability to recognise emotion (Chapter 6). This experiment examined the relationship between emotion recognition and emotion production in younger and older adults. Findings indicated that if older adults are poor at recognising emotion then they are also likely to be poor at producing emotion. Taken together, the experiments in this thesis provide insight into the differences in the way older and younger adults process emotional expressions. Such insights can be used to develop reliable and ecologically valid tools to assess the emotion recognition ability of older adults.

Acknowledgments

First, I would like to thank my supervisors, Jeesun Kim and Chris Davis. Thank you for guiding and supporting me and for being patient with me as I journeyed through this process. I would not have been able to achieve this goal without you.

Sincerest thanks to all of the older adults that participated in the experiments presented in this thesis. It has been a privilege to work with each one of you and I have enjoyed every minute. Thank you for being so excited about my research.

Thank you to my office mates, Saya Kawase, Sonya Karisma, Julie Beadle, and "next-door" office mates, April Ching, Daniel Hochstrasser, Leo Chong, for the laughs, support, and general all-round good vibes. Your friendships and your willingness to let me rant about various things is very much appreciated. I have enjoyed getting to know each one of you and I look forward to nourishing these relationships in the future. Thank you to the MARCS Institute and the tech team for providing the space and (technological) resources that allowed me to run the experiments presented in this thesis.

I am grateful for my family and friends (Mandy Visser, Natasha Carnovale, Hannah Rees) for always checking in on me, offering to help, and providing me with little gift bags of delicious food.

Finally, thank you to Andy for always making me laugh, allowing me to lecture him, and for guaranteeing a bag of Twisties would greet me after a hard day of work.

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

…………………………………………

Simone Simonetti

List of published works

Simonetti, S., Kim, J., & Davis, C. (2015). Auditory, visual, and auditory-visual spoken emotion recognition in young and old adults. In Proceedings of the 18^{th} ICPhS, Glasgow, Scotland, UK.

Simonetti, S., Kim, J., & Davis, C. (2015). Cross-modality matching of linguistic and emotional prosody. In Proceedings of the 16^{th} Interspeech, Dresden, Germany.

Simonetti, S., Kim, J., & Davis, C. (2015). Cross-modality matching of linguistic prosody in older and younger adults. In Proceedings of the 1^{st} Joint Conference on Facial Analysis Animation and Audio-Visual Speech Processing, Vienna, Austria.

Simonetti, S., Kim, J., & Davis, C. (2016). Identifying visual prosody: Where do people look?. In proceedings of the 8^{th} Speech Prosody, Boston, MA, USA.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Older adults have difficulties in emotion recognition (Gonçalves et al., 2018; Ruffman, Henry, Livingstone, & Phillips, 2008). Problems in emotion recognition may play a role in inappropriate social behaviours and interpersonal conflict (Feldman, Philippot, & Custrini, 1991; Shimokawa et al., 2001); given the central place that emotion recognition has in social interaction, my thesis investigated how the process of emotion recognition is affected by aging.

In order to investigate the process of emotion recognition in older adults, the current thesis adopted a research strategy that focused on four well-defined issues: 1) how well older adults can recognise emotion in a variety of expressions (e.g., unimodal, auditory-visual, static, dynamic, clear, ambiguous) and whether emotion recognition performance can be improved by adding another source of emotion information; 2) how older adults extract emotion information from facial expressions of emotion; 3) what problems older adults might encounter during the emotion recognition process; and 4) whether poor emotion recognition is associated with poor emotion production.

To place these issues in the context of previous research, the following sections provide brief reviews on: (1.1) the influence of age on facial and vocal emotion recognition; (1.2) how older adults' recognition performance may change when presented with expressions that contain facial and vocal information; (1.3) gaze behaviour during facial expression recognition and what this may tell us about how older adults extract emotion information; (1.4) the problems that older adults may have at an early stage of emotion processing and; (1.5) possible links between emotion recognition and emotion production problems in old age. An overview of the structure of the thesis is presented at the end of Chapter (1.6).

1.1 The influence of age on emotion recognition

In general, the ability to recognise emotion declines across the adult lifespan. Older adults have particular difficulty recognising emotion from angry, sad, and fearful faces (Gonçalves et al., 2018; Ruffman et al., 2008). This finding is largely consistent across studies (e.g., Murphy & Isaacowitz, 2010; Wong, Cronin-Golomb, & Neargarder, 2005). However, older adults' emotion recognition is not uniformly worse, as there is a trend for older adults to outperform younger ones in the recognition of disgust. For happy and surprise recognition, some studies report that older adults perform similarly to younger adults (e.g., Wong et al., 2005; Ruffman, Halberstadt, & Murray, 2009); whereas others report age differences across all types of emotion (e.g., Lambrecht, Kreifelts, & Wildgruber, 2012).

The ability to recognise emotion from voices also appears to decline in old age (e.g., Lima, Alves, Scott, & Castro, 2014; Ryan, Murray, & Ruffman, 2009). According to the meta-analysis by Ruffman et al. (2008), older adults were worse than younger adults in recognising angry, sad, and happy voices, but not fear, surprise, and disgust. It is important to note that this meta-analysis of vocal expression recognition was based on the results of only five experiments. This likely reflected the limited number of studies that had investigated aging and the recognition of vocal emotions at the time of the meta-analysis. Since then, more studies have emerged and these indicate that age differences exist in the recognition of all types of emotion from voices (Dupuis, 2011; Lambrecht et al., 2012).

To date, emotion recognition in older adults has mostly been assessed using unimodal emotion expressions (typically photographs of faces). Testing with these stimuli has provided important information about how older adults process emotion (Ruffman, 2011), however because static images present a single moment (often the peak of an emotional expression) these may not tap the full extent of older adults' emotion recognition problems in interacting/communicating with others face to face. In face to face communication, emotion information is dynamic, and is most often conveyed via visual and auditory channels, and frequently also involves spoken expression. In keeping with my interest in investigating older adults' emotion recognition problems in the context of communication, I used spoken expressions as well as static stimuli in my studies. The next section will briefly review the literature on older adults' emotion recognition performance when presented with auditory-visual expressions.

1.2 Recognition of auditory-visual expressions of emotion

Only a relatively few studies have examined how the integration of face and voice information influences emotion recognition (e.g., Chaby, Luherne-du Boullay, Chetouani, & Plaza, 2015; Hunter, Phillips, & MacPherson, 2010; Lambrecht et al., 2012; Wieck & Kunzmann, 2017). In general, such studies have shown that older adults' emotion recognition performance is better from auditory and visual emotion information than from either auditory or visual information alone (and that the AV benefit is similar to that found with younger adults).

It is worth noting that many of these studies have employed AV stimuli that consisted of static pictures and short non-verbal affective vocalizations or if they used actual videos of emotion expression, only tested two emotion expressions (Wieck & Kunzmann, 2017) or tested single word utterances (Lambrecht et al., 2012). One can presume that the selection of these types of stimulus materials was to ensure that the auditory and visual emotion information presented was very clear. However, although having high quality, unambiguous stimuli is a typical prerequisite for experimental studies, using such materials may miss potentially interesting effects that may occur with less clear-cut stimuli. Moreover, it seems likely that the spoken emotion expressions that occur in daily communication convey ambiguous emotion information either in one modality or both. Given this, testing older adults with spoken expressions in which expressed emotion is ambiguous is useful for investigating older adults' emotion recognition problems in the context of typical communication. To my knowledge, no studies have specifically investigated older adults' ability to combine such variable AV emotion information.

The first study (Chapter 2) aimed to investigate whether older adults will show an AV benefit relative to unimodal recognition when information in one modality is unclear. That is, it addressed the research question of whether older adults' emotion recognition performance will always be improved by the provision of additional emotion information. Past research indicates that younger adults can show improved recognition performance for AV as compared to unimodal stimuli (i.e., an AV benefit) when the information from both modalities was not well recognised; or when the information from one modality was well-recognised and the information from the other modality was not (Kim & Davis, 2012). Older adults may not show this AV benefit if they have problems prioritising relevant information over irrelevant information (Andrés, Parmentier, & Escera, 2006). Such a problem will not impact the AV benefit when emotion information presented in both modalities is clear cut (i.e., the expressions in both modalities are well recognisable); however, an AV benefit may be reduced if one source of emotion information is ambiguous (i.e., when the expression in at least one modality is not well recognisable).

1.3 Gaze behaviour during facial expression recognition

A general interest in the research area is to understand what might cause and/or exacerbate older adults' recognition problems. One proposed answer to this question is hearing loss. That is, hearing loss may play a role in the recognition of emotion in faces, specifically by influencing the way that older adults gaze at emotion faces. Chaby, Hupont, Avril, Luherne-du Boullay, and Chetouani (2017) proposed that age-related hearing loss may cause problems in speech perception and, as a result, older adults may gaze towards the mouth region (Chaby et al., 2017). By gazing at the mouth region, older adults may obtain speech-relevant cues that can improve speech perception (Davis & Kim, 2006).

Studies that have measured the gaze behaviour of older and younger adults provide evidence that there are differences in the way older and younger adults gaze at emotion faces during emotion recognition (Chaby et al., 2017; Murphy & Isaacowitz, 2010; Sullivan, Ruffman, & Hutton, 2007; Wong et al., 2005). For instance, relative to younger adults, older adults typically look longer at the mouth region of faces when recognising emotion (e.g., Sullivan et al., 2007).

Some researchers have argued that a tendency to gaze at the mouth region may impede the recognition of anger, fear, and sad, but facilitate happy and disgust recognition (e.g., Wong et al., 2005). This is because the eyes convey essential emotion information for anger, sad, and fear recognition, whereas the mouth conveys information for disgust, and happy recognition (Calder, Young, Keane, & Dean, 2000). By gazing at the mouth, older adults may be able to use the relevant information to accurately recognise disgust and happiness; but may have poorer registration of relevant information for the accurate recognition of anger, sadness, and fear. Correlation statistics have been used to investigate the relationship between gaze behaviour and emotion recognition performance. Some studies provide evidence that older adults' gaze behaviour is significantly associated with their recognition performance (Wong et al., 2005) and others do not (Sullivan et al., 2007; Circelli, Clark, & Cronin-Golomb, 2013). Inconsistent findings across studies may result from methodological differences (as discussed in detail in Chapter 3).

The study presented in Chapter 3 tested older and younger adults' gaze behaviour during emotion recognition. The aim of the study was to investigate the cause (i.e., hearing loss) and consequence (i.e., reduced emotion recognition) of older adults' mouth gazing behaviour. Further, as most studies that investigate aging, emotion recognition, and gaze behaviour have used static expressions of emotion (i.e., images of actors expressing emotion), this study tested gaze behaviour and emotion recognition for dynamic speech expressions of emotion (i.e., videos of people uttering sentences and simultaneously conveying an expression). Older adults may be more likely to gaze at the mouth region of dynamic expressions relative to static expressions as dynamic faces convey visual speech information.

Chapter 4 presents an additional eye-tracking experiment that investigated older adults' gaze behaviour and recognition performance for non-emotion (i.e., prosodic) expressions. The rationale for this was that by investigating gaze behaviour for the expression of linguistic prosody, I could determine whether the pattern of older adults' gaze behaviour is similar for all types of facial expressions or is specific to emotion faces only.

Investigations of the gaze behaviour of older adults can shed light on how they pick up information from emotional faces, however, these cannot address the issue of how quickly and efficiently emotion information is processed. This issue will be addressed in the next two sections.

1.4 The early stage of emotion processing

The following section briefly explores why older adults may have problems at the early stages of processing. There is substantial evidence to indicate that aging is accompanied by cortical atrophy. For instance, magnetic resonance imaging scans have shown that older adults have volume loss in the frontal, particularly the prefrontal, cortex and the orbitofrontal cortex relative to other cortical regions such as the amygdala and insula, (Grieve, Clark, Williams, Peduto, & Gordon, 2005; Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010; although see Fjell et al., 2013). Further, older adult brains show reductions in the amount of neurotransmitters such as dopamine and noradrenaline (Li & Rieckmann, 2014). Given the morphological and functional changes in the cortical systems of older adults that process emotion information, it seems plausible that they will have problems in the very early stages of emotion information processing.

In apparent contrast to the suggestion that older adults will have problems in the early stages of emotion information processing, researchers suggest that older adults are able to rapidly process emotion information to the same degree as younger adults (e.g., Mather & Knight, 2006; Hahn, Carlson, Singer, & Gronlund, 2006; Ruffman, Ng, & Jenkin, 2009). However, these studies employ visual search tasks that may assess simple feature extraction or detection rather than early emotion processing. Chapter 5 used a masked priming procedure to investigate the extent of early emotion processing in older adults, independent of any problems that may occur at the later processing stages (e.g., decision stages, or response preparation).

1.5 Aging and the production of emotion

The final experimental chapter (Chapter 6) cast a broader net than most studies of older adult emotion recognition in that it investigated emotion production. The rationale for this investigation comes from recent theories of embodied cognition that propose that the process of emotion recognition may draw on emotion production processes (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). Given that emotion production can be compromised in older age (Borod et al., 2004; Hess, Adams Jr, Simard, Stevenson, & Kleck, 2012; Riediger, Voelkle, Ebner, & Lindenberger, 2011), and that emotion perception and production are linked (e.g., Deriso et al., 2012), then it could be that older adults who have problems in expressing emotions will also have problems in recognising emotions. To my knowledge, few if any studies have investigated the relationship between emotion recognition and emotion production in older adults.

1.6 Research outline (in a nutshell)

The thesis investigated how older adult's process emotion information by conducting five experimental studies. The first study (Chapter 2) investigated whether older adults' emotion recognition performance was better for auditory-visual relative to unimodal emotion expressions. In this experiment, clear and less clear emotion expressions were tested to examine how this affected the integration of AV emotion information by young and older adults. The two studies presented in Chapters 3 and 4 used eye-tracking to investigate how older adults extract visual information during expression recognition. Chapter 3 examined the relationship between gaze behaviour and emotion recognition performance and between age-related hearing loss and gaze behaviour. By way of contrast, the experiment presented in Chapter 4 investigated the gaze behaviour of older adults when recognising expressions of linguistic prosody. The study in Chapter 5 investigated the extent to which older adults can carry out basic perceptual processing of face emotion by using a masked priming technique that indexes processing relatively independent of response strategies and decision components. The final experiment (Chapter 6) aimed to further our understanding of older adults' emotion processing by investigating the link between emotion perception and production.

1.7 References

Andrés, P., Parmentier, F. B. R., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia, 44*(12), 2564-2568.

Borod, J. C., Yecker, S. A., Brickman, A. M., Moreno, C. R., Sliwinski, M., Foldi, N. S., . . . Welkowitz, J. (2004). Changes in posed facial expression of emotion across the adult life span. *Experimental Aging Research, 30*(4), 305-331.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527-551.

Chaby, L., Hupont, I., Avril, M., Luherne-du Boullay, V., & Chetouani, M. (2017). Gaze behavior consistency among older and younger adults when looking at emotional faces. *Frontiers in Psychology, 8*, 548.

Chaby, L., Luherne-du Boullay, V., Chetouani, M., & Plaza, M. (2015). Compensating for age limits through emotional crossmodal integration. *Frontiers in Psychology, 6*, 691.

Circelli, K. S., Clark, U. S., & Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in aging. *Aging, Neuropsychology, and Cognition, 20*(2), 148-173.

Davis, C., & Kim, J. (2006). Audio-visual speech perception off the top of the head. *Cognition, 100*(3), B21-B31.

Deriso, D. M., Susskind, J., Tanaka, J., Winkielman, P., Herrington, J., Schultz, R., & Bartlett, M. (2012). *Exploring the facial expression perception-production link using real-time automated facial expression recognition.* Paper presented at the European Conference on Computer Vision.

Dupuis, K. L. (2011). *Emotion in speech: Recognition by younger and older adults and effects on intelligibility (Doctoral dissertation).* Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/31738/1/Dupuis_Katherine_L_201111_PhD_thesis.pdf.

Feldman, R. S., Philippot, P., & Custrini, R. J. (1991). Social competence and nonverbal behavior. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 329-350). New York: Cambridge University Press.

Fjell, A. M., Westlye, L. T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., . . . Walhovd, K. B. (2013). Critical ages in the life course of the adult brain: nonlinear subcortical aging. *Neurobiology of Aging, 34*(10), 2239-2247.

Gonçalves, A. R., Fernandes, C., Pasion, R., Ferreira-Santos, F., Barbosa, F., & Marques-Teixeira, J. (2018). Effects of age on the identification of emotions in facial expressions: a meta-analysis. *PeerJ, 6,* e5278.

Grieve, S. M., Clark, C. R., Williams, L. M., Peduto, A. J., & Gordon, E. (2005). Preservation of limbic and paralimbic structures in aging. *Human Brain Mapping, 25*(4), 391-401.

Hahn, S., Carlson, C., Singer, S., & Gronlund, S. D. (2006). Aging and visual search: Automatic and controlled attentional bias to threat faces. *Acta Psychologica, 123*(3), 312-336.

Hess, U., Adams Jr, 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*(6), 1377-1380.

Hunter, E. M., Phillips, L. H., & MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychology and Aging, 25*(4), 779-787.

Kim, J., & Davis, C. (2012). Perceiving emotion from a talker: How face and voice work together. *Visual Cognition, 20*(8), 902-921.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*(3), 529-539.

Li, S.-C., & Rieckmann, A. (2014). Neuromodulation and aging: implications of aging neuronal gain control on cognition. *Current Opinion in Neurobiology, 29*, 148-158.

Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: how age shapes emotion processing in nonverbal vocalizations. *Emotion, 14*(5), 145-160.

Mather, M., & Knight, M. R. (2006). Angry faces get noticed quickly: Threat detection is not impaired among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(1), P54-P57.

Murphy, N. A., & Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognising emotional expressions: An in-depth look at gaze measures and covariates. *Cognition and Emotion, 24*(3), 436-452.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184-211.

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: regional and individual differences. *NeuroImage, 51*(2), 501-511.

Riediger, M., Voelkle, M. C., Ebner, N. C., & Lindenberger, U. (2011). Beyond “happy, angry, or sad?”: Age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognition and Emotion, 25*(6), 968-982.

Ruffman, T. (2011). Ecological validity and age-related change in emotion recognition. *Journal of Nonverbal Behavior, 35*(4), 297.

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions in older adults. *The Journals of Gerontology: Series B, 64*(6), 696-703.

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*(4), 863-881.

Ruffman, T., Ng, M., & Jenkin, T. (2009). Older adults respond quickly to angry faces despite labeling difficulty. *Journals of Gerontology: Series B, 64*(2), 171-179.

Ryan, M., Murray, J., & Ruffman, T. (2009). Aging and the perception of emotion: Processing vocal expressions alone and with faces. *Experimental Aging Research, 36*(1), 1-22.

Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y., & Kohno, M. (2001). Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. *Brain and Cognition, 47*(3), 423-433.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology: Series B, 62*(1), 53-60.

Wieck, C., & Kunzmann, U. (2017). Age differences in emotion recognition: A question of modality? *Psychology and Aging, 32*(5), 401-411.

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology, 19*(6), 739-749.

Chapter 2. Emotion Recognition in Young and Older Adults: Older Adults Do Not Show an Auditory-Visual Benefit

2.1 Abstract

The ability to recognise emotion from faces or voices declines with advancing age. We examined whether this decline could be offset by the presentation of Auditory-Visual (AV) expressions and whether older adults would benefit from AV presentation even when emotion presented in one modality was poorly recognised. We selected stimuli from a clear and a less clear presenter from a previous study such that unimodal emotion recognition was uniformly high, low, or mixed (i.e., high for one modality but low for the other). Young (experiment 1, n = 30) and older (experiment 2, n = 30) participants were tested on these stimuli in a forced-choice emotion recognition task. Stimuli were AV recordings of anger, happiness, sadness, surprise, and disgust; and presented in auditory-only (AO), visual-only (VO), or AV format. Similar to previous findings, younger adults showed an AV benefit even when unimodal recognition was poor. No AV benefit was found for older adults; indeed, AV was worse than VO recognition when AO recognition was poor. When compared, younger adults outperformed older adults except for clear VO expressions. We suggest that older adults' non-optimal AV integration occurred because the expression in one modality (typically the auditory one) was not well recognised. That is, low AO recognition rates may have impacted how effectively AV information was used during emotion recognition.

2.2 Introduction

It is well established that aging is associated with declines in the ability to recognise emotion (Gonçalves et al., 2018; Ruffman, Henry, Livingstone, & Phillips, 2008). Most research on aging and emotion recognition has used unimodal emotional expressions (i.e., faces or voices) as stimuli. As such, this research may not indicate what older adults' emotion recognition might be like during everyday life. This is because our everyday experience of decoding emotion from other people generally involves multiple emotion cues conveyed by both the visual and auditory channels. The current study used auditory-only (AO), visual-only (VO), and auditory-visual (AV) expressive speech stimuli and focused on a specific issue about how visual and auditory information are combined during emotion recognition in older adults.

Past research has shown that younger adults can achieve relatively high levels of emotion recognition with AV stimuli even though unimodal recognition is poor (Kim & Davis, 2012). Furthermore, emotion recognition performance for younger adults was not reduced when information from a well-recognised modality was paired with that from a poorly recognised modality (Kim & Davis, 2012). The current study tested whether older adults will show this ability. That is, the aim of the study was to determine the extent to which older adults would show improved recognition performance for AV as compared to unimodal stimuli even though the degree to which emotion recognition is supported by a modality may be marginal. For example, what will AV recognition performance be like for an emotion that is well recognised in one modality but poorly recognised in the other? Determining whether older adults can effectively combine emotion related information from different modalities is important for gauging how well they can cope with the vagaries of real-life stimuli.

Only a relatively few studies have examined AV emotion recognition in older adults (e.g., see Chaby, Luherne-du Boullay, Chetouani, & Plaza, 2015; Hunter, Phillips, & MacPherson, 2010; Lambrecht, Kreifelts, & Wildgruber, 2012; Wieck & Kunzmann, 2017). In general, these studies reported that older adults benefitted from multisensory information to the same extent as younger ones. However, several features of these studies make it difficult to draw a general conclusion about how older adults combine multimodal emotion information from unimodal sources that vary in how well they can be recognised. Unlike Kim and Davis (2012) who used emotion stimuli from five different people whose emotion productions varied in how well each was recognised in the AO, VO, and AV modalities, these studies did not specifically use AV stimuli that varied in unimodal recognition rates, and so could not explicitly test how older adults combined such information. Further, Chaby et al. (2015) and Hunter et al. (2010) used static photographs of emotional expressions and non-verbal affective vocalisations rather than spoken expressions. In Wieck and Kunzmann (2017), the recognition of only two emotions was examined. In Lambrecht et al. (2012), the results were presented in terms of linear regression fits across age for each modality rather than the AV benefit as a function of unimodal recognition scores.

The current study aimed to test what older (and younger) adults' AV emotion recognition would be like when the recognition of stimuli in the component modalities varies. It was predicted that, like that of the younger adults of the Kim and Davis (2012) study, the younger adults of the present study would use the information in AV stimuli to achieve recognition rates that were always better than those from the best unimodal stimuli. In contrast, older adults may not be able to achieve such a result when the information from either one or both modalities does not clearly specify which emotion was presented. This is because older adults seem to have problems in properly weighting the relevance of information for optimal task performance (e.g., see Dywan, Segalowitz, & Webster, 1998;

Andrés, Parmentier, & Escera, 2006). A problem in appropriately weighting information will not be evident when all information presented is task-relevant and reliable; however, if some information is task-irrelevant or unreliable then performance will suffer (e.g., Barrett & Newell, 2015; May, 1999). For instance, Ruffman, Halberstadt, and Murray (2009) investigated older and younger adults' ability to recognise emotion in faces and voices; and to match emotion faces (and bodies) to emotion voices. The authors reported that for some vocal expressions of emotion (i.e., angry, sad, happy), younger adults outperformed older adults (i.e., for older adults, auditory information for angry, sad, and happy emotion types was unreliable); and that for these same emotions, older adults were less able than younger adults to match faces to voices (see Ryan, Murray, & Ruffman, 2009, for similar findings). The findings of these studies may indicate that older adults' AV integration will suffer when the information from one (or both) modalities does not clearly convey a single emotion type.

Kim and Davis (2012) found that younger adults were able to use the joint presentation of auditory and visual emotion stimuli to reduce unimodal ambiguity (as apparent in response confusions). That is, if a particular wrong emotion (confusion) was selected only in one modality, it would not be selected with AV presentation. If older adults were unable to use AV presentation to eliminate confusions, then their AV performance would never be superior to the best unimodal score. Indeed, since the potential to confuse emotions is likely to be different for the visual and auditory modalities, AV presentation may increase the number of potential confusions that could be made. Thus, the failure to eliminate specific unimodal confusions might result in the relatively good performance for one modality being reduced when information from a poorly recognised modality is also presented.

In order to test whether older adults can effectively combine auditory and visual emotion information (especially when unimodal recognition rates vary) we used stimuli from the study by Kim and Davis (2012). Based on the average correct emotion recognition scores of 55 young participants, we selected stimuli from two presenters such that unimodal emotion recognition was either uniformly high, low, or mixed (Figure 2.1 shows the relevant data from Kim & Davis, 2012). One person (i.e., the clear presenter) had mostly clearly recognisable emotional expressions (including angry and happy expressions), the only exception was the AO expression of disgust. The expressions of the other person (i.e., the unclear presenter) were harder to recognise (although once again there were some exceptions, e.g., VO expressions of disgust, sadness, and happiness; AO expressions of sadness).

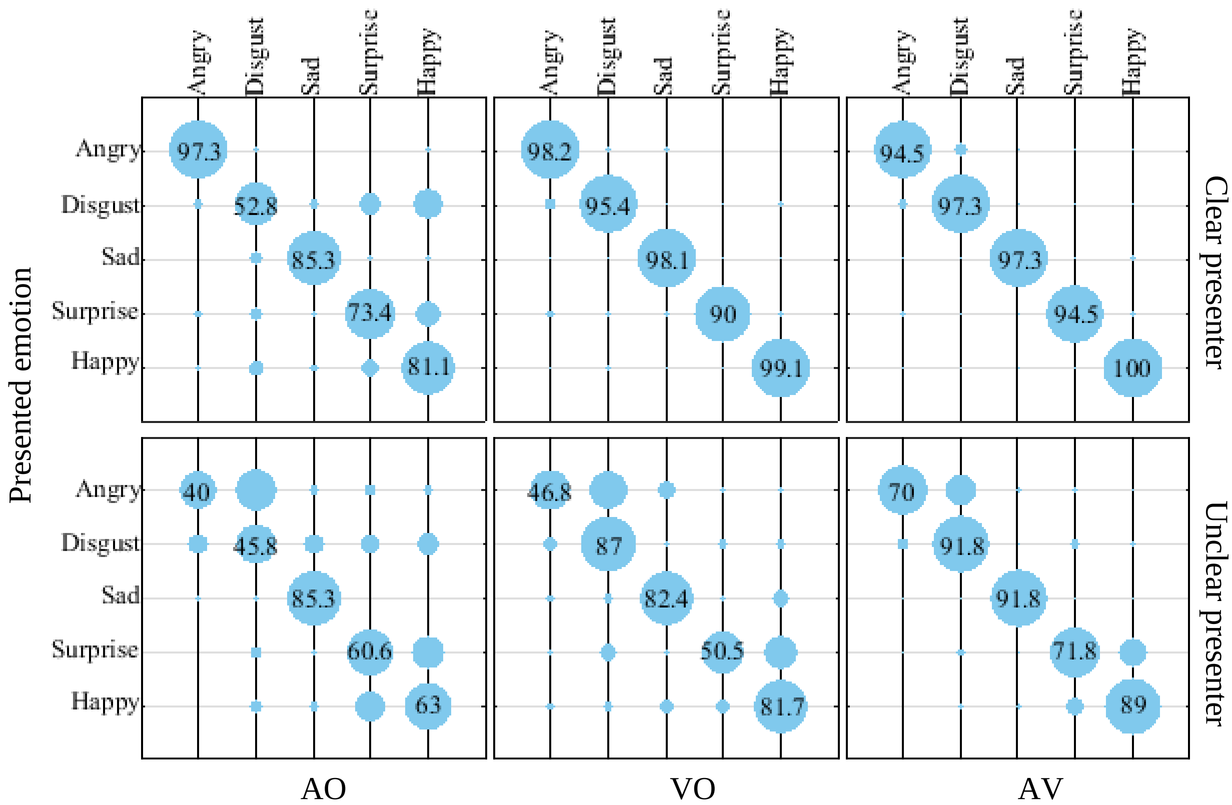

Figure 2.1. Auditory-only (AO) (left), visual-only (VO) (middle), and auditory-visual (AV) (right) response matrices for the younger adults for the two presenters selected to differ on whether AV recognition was high, low, or mixed (data from Kim & Davis, 2012). The rows show the matrices for the clear presenter (top) and the unclear presenter (bottom). The presented emotions are shown on the vertical and the responses to these are shown at the top, horizontally. Response magnitude is represented by the size of the circles; the numbers presented on the diagonal have been added as an indication of scale and indicate the percentage of times that a response was selected for a presented emotion.

As mentioned, unimodal emotion recognition performance was either uniform (i.e., accuracy scores were similar across the AO and VO modalities) or mixed (i.e., accuracy scores were different across the AO and VO modalities, e.g., accuracy was high for the VO modality and low for the AO modality) across the two presenters. For instance, anger and sad recognition scores were uniform across the auditory and visual modalities for both the clear (uniformly high) and unclear (uniformly high for sad and low for anger) presenters. In contrast, disgust and happy expressions were better recognised in the visual modality than in the auditory modality for both presenters. Surprise showed this pattern for the clear presenter only. That is, for the unclear presenter, surprise was best recognised from the auditory than the visual modality. However, as mentioned above, regardless of the clarity of the presenter's unimodal expressions, AV scores of young participants were superior to the best unimodal score (except when scores were at ceiling). That is, AV scores were superior even when one modality attracted a high score and the other did not.

In sum, the current study investigated whether older adults will benefit from AV presentation (as did younger ones, Kim & Davis, 2012), particularly when information in one modality was unclear. We measured AV scores relative to unimodal scores and constructed confusion matrices to ascertain the extent to which unimodal confusions are eliminated with multimodal presentation.

2.3 Method

2.3.1 Participants

Thirty younger adults (M_{age} = 20, SD = 2.4, 10 males) from Western Sydney University participated in this study for course credit. Thirty older adults (M_{age} = 72, SD = 6.1, 17 males) participated for monetary reimbursement. Sample sizes were determined based on Hunter et al. (2010). All participants had learnt English at age 7 or younger.

Participants were given the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) as the presence of dementia has been associated with poor emotion processing (Rosen et al., 2006). Younger (*M* = 29) and older (*M* = 28) adults scored within the normal range (above 23), indicating no presence of dementia.

Hearing and visual acuity were assessed as poor acuity may interfere with the processing of acoustic and visual signals important for emotion identification. Hearing acuity was assessed for both ears using pure tone audiometry (Diagnostic Audiometer, AD229e) for 0.5, 1, 2, and 4 kHz. Overall, older adults' hearing level was poorer than younger adults across all frequencies. Younger adults had an averaged hearing level of 12dB across all frequencies (range = 10dB-15dB) whereas older adults averaged 32dB (range = 28dB-40dB).

The Freiburg Visual Acuity Test (FrACT, Bach, 2007) was used for both eyes. Most participants showed normal visual acuity; except one younger and three older adults that showed slightly worse than normal vision (a score of 1) with scores over 0.75.

2.3.2 Materials

The stimuli selected from Kim and Davis (2012) consisted of audio and video recordings of 2 male native Australian English presenters (in their early twenties) uttering 8 Semantically Unpredictable Sentences (Benoît, Grice, & Hazan, 1996). In uttering each sentence, the presenters were instructed to express anger, sadness, disgust, surprise, happiness, or neutral as if they were communicating this emotion to an observer. The total 32 (2 x 8 x 6) recordings were edited to produce AO, VO, and AV stimuli. A female presenter's expressions were also selected for 12 practice trials.

2.3.3 Procedure

Participants were tested individually in a quiet room. They were told they would 'see', 'hear', or 'see and hear' a person conveying various emotions across trials. They were instructed to select one of 5 options on the screen that they thought best described the expressed emotion.

For each trial, a fixation point was presented (50ms), followed by an experimental item (approx. 6 secs). Then, 5 boxes labelled “Angry”, “Sad”, “Disgust”, “Surprise”, and “Happy” were presented for a response. Participants were presented with 240 trials: three blocks (AO, VO, AV) of 40 trials (8 sentences, 5 emotions) for each of 2 presenters. Prior to each presenter block (i.e., AO, VO, AV), two neutral expressions of that presenter were displayed. Participants were told to use these expressions as a baseline to compare with subsequently presented emotions. Trials were blocked by presenter to reduce the difficulty of the task. That is, by blocking the presenter trials (as per Kim & Davis, 2012) participants were given the opportunity to focus on a single presenters’ expressions and to compare emotion expressions within that presenter block (to ultimately assist emotion recognition). Practice items were presented before experimental ones.

The order of the three blocks was counterbalanced across participants. Half of the participants received the clear presenter trials first for all three blocks. The presentation order of stimuli within each presenter block was randomised. For stimulus display and response collection, we used DMDX (Forster & Forster, 2003). After the experiment, the MMSE, FrACT, and hearing test were administered, and participants completed a questionnaire detailing age, gender, and languages spoken.

2.4 Results

Overall identification scores and response confusions for the younger and older adults are presented in Figure 2.2. For younger adults, AV scores were almost always higher than the highest unimodal score (as per Kim & Davis, 2012). Conversely, older adults did not show improved AV scores relative to unimodal ones; indeed, for most cases, the AV scores for older adults were numerically lower than the VO ones. This result suggests that compared to younger adults, older adults were less adept at combining auditory and visual information.

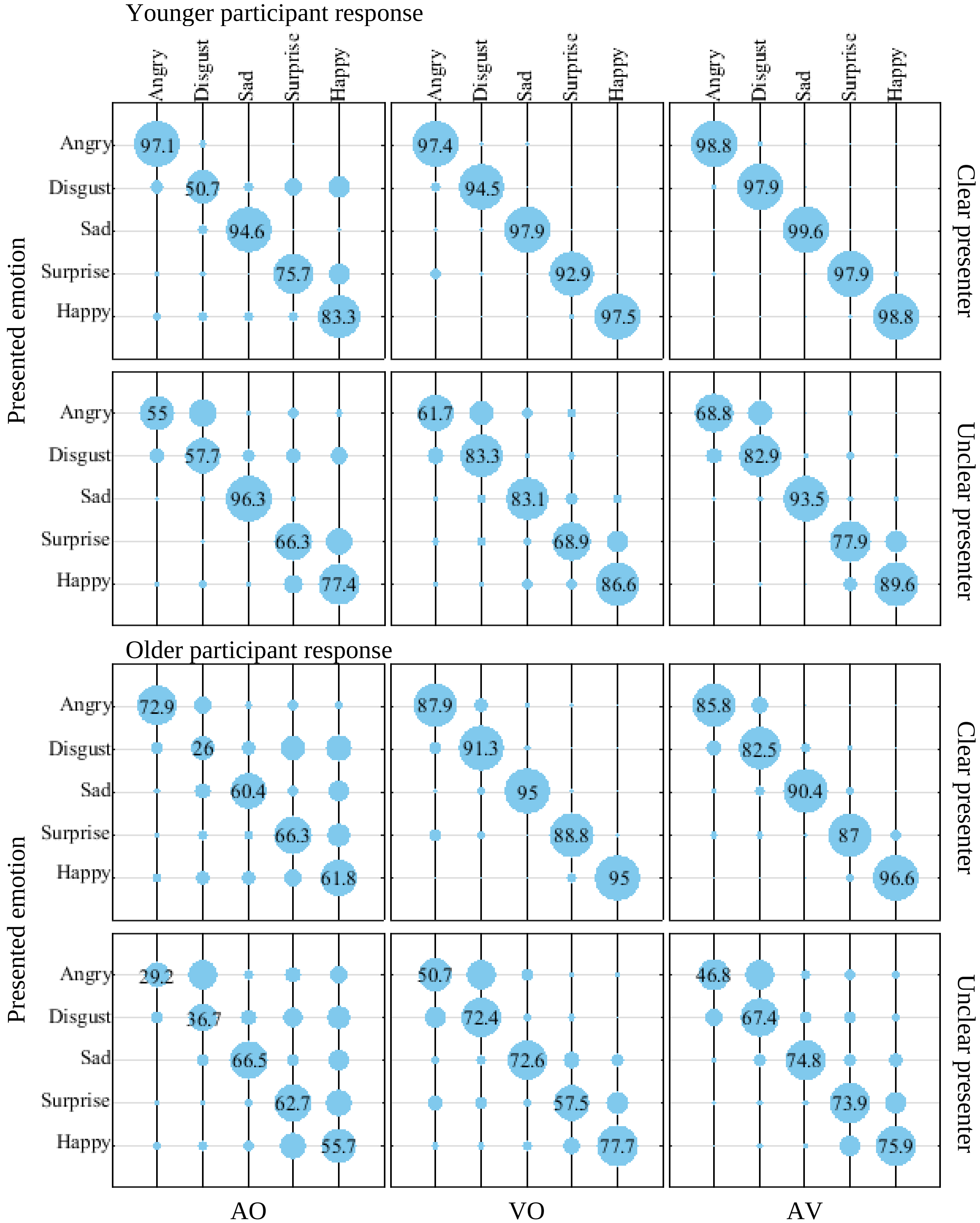

Figure 2.2. AO (left), VO (middle), and AV (right) response matrices for the younger (top) and older (bottom) adult groups. The rows show the matrices for the clear presenter (top) and the unclear presenter (bottom). The presented emotions are shown on the vertical and the responses to these are shown at the top, horizontally. Response magnitude is represented by the size of the circles; the numbers presented on the diagonal have been added as an indication of scale and indicate the percentage of times that a response was selected for a presented emotion.

2.4.1 Age-related difference in emotion recognition

Percent correct scores were analysed with a repeated measures ANOVA with modality (AO, VO, AV), presenter (clear, unclear), and emotion (angry, sad, disgust, surprise, happy) as within-participant factors, and age as a between-participants factor.

There was a significant main effect of age with younger adults (84%) outperforming older ones (70%), $F(1,58) = 42.76$, $p < .001$, $\eta p2 = .42$. The AV (84%) and VO (83%) presentation modalities attracted the highest percent correct scores, followed by the AO (65%) condition, $F(2,116) = 245.10$, $p < .001$, $\eta p2 = .81$. Participants identified the clear (85%) presenters' expressions better than the unclear (69%) presenter, $F(1,58) = 402.61$, $p < .001$, $\eta p2 = .87$. The main effect of emotion was also significant, $F(4,232) = 37.56$, $p < .001$, $\eta p2 = .39$. There was also a significant interaction between presentation modality, presenter, and age group, $F(2,116) = 5.98$, $p < .01$, $\eta p2 = .09$. This interaction was analysed further to examine the impact of expression clarity and age on the recognition of unimodal and bimodal expressions. To do this, separate repeated measures ANOVAs were conducted for the presenters with modality and emotion as within subject factors, and age as the between-subjects factor. Identification scores for the clear and unclear presenter are displayed in Figure 2.3.

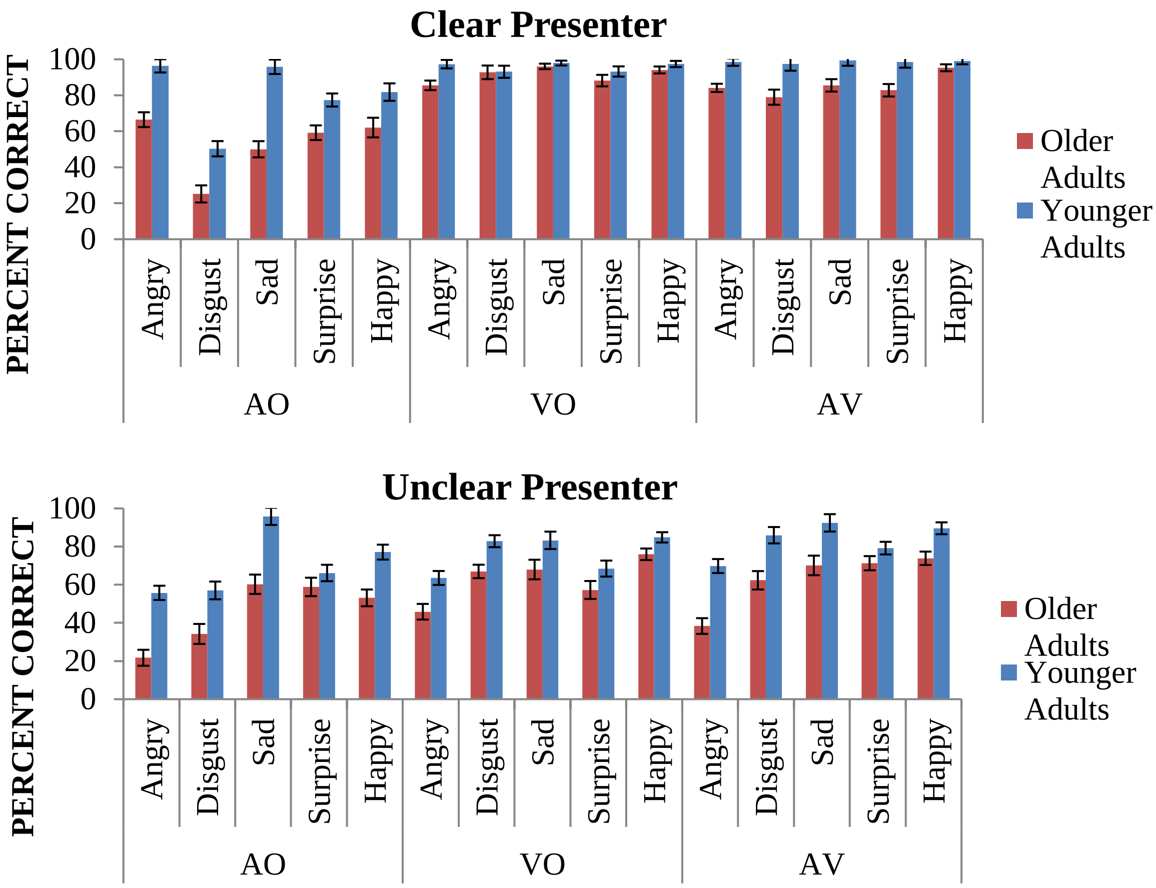

Figure 2.3. Mean percent correct scores for the clear (upper) and unclear (lower) presenters across presentation modality and emotion. Error bars represent standard error.

For the clear presenter, the ANOVA revealed that younger adults (92%) outperformed older adults (79%), $F(1,58) = 34.06$, $p < .001$, ηp2 = .37. The VO (94%) and AV (94%) conditions attracted the highest percent correct responses followed by the AO (69%) condition, $F(2,116) = 336.84$, $p < .001$, ηp2 = .85. The main effect of emotion, $F(4,232) = 37.79$, $p < .001$, ηp2 = .40, was also significant. There was a significant interaction between modality, emotion, and age group, $F(8,464) = 3.00$, $p < .01$, ηp2 = .05, and this was analysed further using a Bonferroni adjusted alpha. Post hoc analyses revealed that for disgusted expressions, older adults showed worse performance for AV relative to VO expressions ($p < .05$). For the remaining expression types (i.e., angry, sad, surprise, happy) older adults showed no performance differences when these expressions were conveyed visually or audio-visually ($p > .32$ for all comparisons). In contrast, older adults recognised AV expressions better than AO expressions across all emotion types ($p < .001$). For younger adults, there were no differences across any emotions conveyed via VO and AV modalities ($p > .27$). This was also the case for AO expressions of anger and sadness where there were no performance differences between AO and AV renditions of these emotions ($p > .41$). Only expressions of happy, surprise, and disgust were less well recognised when conveyed through the AO modality than when conveyed via the AV modality ($p < .01$). These analyses also revealed that younger adults outperformed older adults for most AO ($p < .01$) and AV emotional expressions ($p < .02$) except AO expressions of surprise ($p = .07$) and AV expressions of happy ($p = .27$) where older and younger groups did not show a difference. There were no differences between older and younger percent correct scores for most of the emotions in the VO presentation modality ($p > .16$); except anger where younger adults outperformed older ones ($p < .01$).

For the unclear presenter, the younger group (77%) performed better than the older group (62%), $F(1,58) = 40.37$, $p < .001$, ηp2 = .41. Participants showed better performance in the AV (75%) and VO (71%) modalities than the AO (60%) modality, $F(2,116) = 74.48$, $p < .001$, ηp2 = .56. The main effect of emotion was significant, $F(4,232) = 53.73$, $p < .001$, ηp2 = .48, and there was a significant interaction between modality and age group, $F(2,116) = 7.59$, $p < .01$, ηp2 = .12. This was analysed further (with a Bonferroni adjusted alpha) as the interaction between modality, emotion, and age group was not significant, $F(8,464) = 1.77$, $p = .08$, ηp2 = .03. These post hoc analyses revealed that younger adults showed better recognition of emotions when presented with AV than VO expressions, followed by AO expressions ($p < .02$). Older adults showed worse recognition in the AO compared to the AV condition ($p < .001$); but showed no significant differences when recognising emotions in the VO and AV conditions ($p > .99$). These analyses also indicated that younger adults outperformed older adults across all modality types ($p < .001$ for all comparisons).

In sum, our results showed that emotion information from multiple sensory streams improved younger adults' emotion recognition performance relative to unimodal emotion information. In contrast, older adults did not show this pattern of results and, in some instances (e.g., clear expressions of disgust), older adults showed increased recognition problems when integrating audio and visual emotion information than when identifying emotion from unimodal expressions. Indeed, older adults showed reduced AV relative to VO recognition when ambiguous AO emotion information was added to clear VO emotion information (e.g., disgust recognition in the clear presenter condition). That is, older adults' emotion recognition performance (for the clear presenter) declined for AV relative to VO emotional expressions.

2.4.2 Confusion matrices

To investigate whether younger adults were able to use auditory and visual information to reduce unimodal confusion, scores from the confusion matrices (Figure 2.2) were entered into separate repeated measures ANOVA for each emotion type with participant response (i.e., angry, sad, disgust, surprise, or happy), presenter, and modality (AO, VO, AV) as within participant factors. The interaction between presentation modality, presenter, and emotion was significant for angry, $F(8,232) = 2.98$, $p < .01$, $\eta p2 = .09$, sad, $F(8,232) = 7.08$, $p < .001$, $\eta p2 = .20$, disgust, $F(8,232) = 6.92$, $p < .001$, $\eta p2 = .19$, surprise, $F(8,232) = 5.00$, $p < .001$, $\eta p2 = .15$, and happy, $F(8,232) = 3.59$, $p < .01$, $\eta p2 = .11$, emotion types. Simple effects analyses reveal that for the younger adults, the AV condition attracted fewer response confusions than either unimodal one. This reduction appears to be due to two factors; first, the pattern of confusions was different for the auditory and visual presentations, and second, when a confusion occurred in only one presentation mode, it did not occur in the AV condition. For instance, for the clear presenter, disgust expressions were confused with happiness and surprise more so for auditory relative to visual ($p < .001$) or AV ($p < .001$)

presentation. When confusions occurred in both unimodal conditions, they remained with AV presentation (at rates reduced or rates similar to the unimodal condition with the lowest rates of confusions). For instance, expressions of disgust from the unclear presenter were confused for angry in both unimodal conditions, and hence, in the AV one ($p > .99$).

Older adults' scores from the confusion matrices (Figure 2.2) were also entered into separate repeated measures ANOVA's for each emotion type. The interaction between presentation modality, presenter, and emotion was not significant for angry expressions, $F(8,232) = 1.86$, $p = .07$, $\eta p2 = .06$, but was significant for sad, $F(8,232) = 7.60$, $p < .001$, $\eta p2 = .21$, disgust, $F(8,232) = 12.24$, $p < .001$, $\eta p2 = .30$, surprise, $F(8,232) = 7.18$, $p < .001$, $\eta p2 = .20$, and happy, $F(8,232) = 4.18$, $p < .001$, $\eta p2 = .13$, emotion types. Simple effects analyses revealed that for the older adults, the AV and VO modalities attracted fewer confusions than the AO modality. However, unlike younger adults, AV presentation did not reduce confusions compared to VO presentation. Part of the reason for this is that there was a greater spread of confusions in the AO condition (e.g., see angry, sad and surprise), so these overlapped with those of the VO condition. This meant that there were fewer cases where a confusion only occurred in one modality, and so, fewer cases where AV information would rule out a confusion. Furthermore, even when the unimodal confusions did not overlap, there were cases where confusions were not eliminated with AV presentation. For instance, for the unclear presenter, disgust expressions were confused for surprised expressions in the AO condition, but not in the VO one; despite this, the AV presentation of disgust was sometimes confused with surprise ($p < .02$) across all comparisons) (see also disgust/happy confusions with the unclear presenter). In sum, the confusion matrices suggest that older adults did not utilise the information from both the AO and VO presentations to reduce confusions as efficiently as younger adults.

2.4.3 Optimal integration

To evaluate the performance that might be expected from the "optimal" integration of the auditory and visual sources of information, we estimated a Bayesian likelihood ratio for AV presentation based on the recognition performance for unimodal presentation. In this method, ambiguous information (as indicated by different levels of performance for auditory and visual recognition) is given less weight in the auditory and visual combination. Here, we follow Massaro and Cohen (2000) by estimating performance for joint auditory and visual presentation by multiplying the recognition performance results for the unimodal responses and taking into account the combined AV support for the alternatives. That is, we used the response patterns for the confusion matrices (Figure 2.2) to calculate the probability for auditory and visual presentation given the unimodal responses. For example, the likelihood of recognising "Angry" given the separate auditory and visual recognition rates equals the results for (Aud_Angry) x those for (Vis_Angry) divided by this same product plus the AV support for each of the alternative responses. The results are shown in Table 2.1. For the clear presenter, the younger adult's actual AV performance level was similar to that estimated by the optimal integration calculation (i.e., on average, actual performance was about one percent worse than the optimal). For the unclear presenter, the actual performance was about 9 percent worse than that estimated by the optimal integration calculation.

For older adults, for the clear presenter, the actual scores were on average about 10% worse than what would be expected under an optimal Bayesian combination of the auditory and visual sources (compare this to the younger adults whose actual AV recognition scores were only about 1% worse). For the unclear presenter, the actual AV recognition score for the older adults was about 14% worse than the estimated optimal score (recall that for younger adult's, actual performance was about 9% worse). These results suggest that older adults did not integrate AV information as well as younger adults, and certainly not in an optimal fashion. Indeed, it appeared that older adults did not effectively use the evidence for the lack of candidates in the auditory modality to eliminate candidates in the visual one. The reason for this may be that the information provided by the auditory modality appeared to be much poorer for the older compared to the younger adults. That is, for the auditory presentation, older adult's correct recognition scores were on average 22% worse than the younger ones; whereas for visual presentation, the difference was only 8%. The next section aimed to investigate whether hearing loss could account for older adults' AO recognition problems.

Table 2.1. The optimal AV score based on the unimodal scores and the actual AV recognition scores for the younger (top) and older (bottom) adults. Optimal scores were calculated from the scores of the confusion matrices for the auditory and visual presentations separately.

			Angry	Disgust	Sad	Surprise	Happy
Younger Adults	Clear	Optimal AV	100.0	99.0	100.0	99.7	99.9
		Actual AV	98.8	97.9	99.6	97.9	98.8
	Unclear	Optimal AV	76.2	96.5	99.8	86.4	98.5
		Actual AV	68.8	82.9	93.5	77.9	89.6
Older Adults	Clear	Optimal AV	97.7	97.3	99.2	99.2	98.7
		Actual AV	85.8	82.5	90.4	87.0	96.6
	Unclear	Optimal AV	47.7	91.5	94.9	83.4	90.7
		Actual AV	46.8	67.4	74.8	73.9	75.9

2.4.4 Hearing acuity and emotion recognition

Older adults' hearing level scores were averaged across the right and left ears and are presented in Figure 2.4 along with younger adults' hearing level scores. As can be seen, older adults' hearing was poorer than younger adults across all (particularly higher) frequencies.

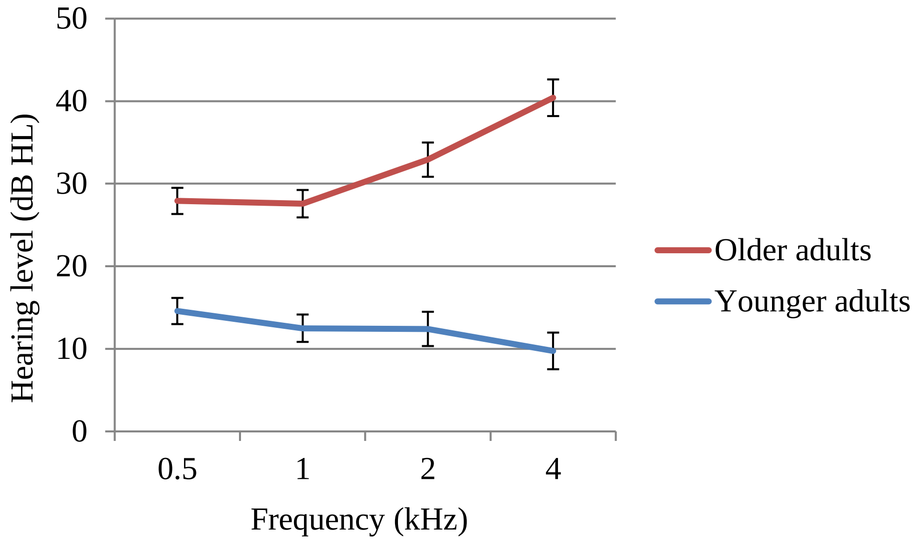

Figure 2.4. Hearing level (dB) for each age group across 0.5, 1, 2, and 4 kilohertz. Error bars indicate standard error.

To explore whether older adults' hearing ability was associated with emotion recognition performance, a multiple regression analysis was performed between the recognition scores for the AO condition and hearing level for the four hearing frequencies. The results indicate that the hearing levels were not significantly associated with performance on the AO emotion recognition task, $R = .5$, $R^2 = .3$, adjusted $R^2 = .1$, $F(4,29) = 2.1$, $p = .11$. As per Orbelo, Grim, Talbott, and Ross (2005), we separated the older adult participant group into two groups based on their hearing ability and compared them on the AO emotion recognition task. One group, consisting of 12 participants, had hearing levels above 30dB for 500 or 1000Hz or above 40dB for 2000Hz. The remaining 18 participants had hearing levels equal to or below the aforementioned thresholds. In line with Orbelo et al. (2005) these two hearing groups did not differ in their performance for the AO emotional expressions, $F(1,29) = 1.99$, $p = .17$, $\eta_p^2 = .07$, providing further evidence to suggest that hearing levels were not associated with AO emotion recognition.

2.5 Discussion

In the current study, we investigated whether older adults would show improved emotion recognition performance for AV as compared to VO or AO presentations. We were particularly interested in whether, like younger adults, older adults would show improved recognition performance for AV presentations even for stimuli where the expression in one modality was poorly recognised. Whether or not older adults are able to combine auditory and visual emotion information to better recognise an indistinct stimulus is an important issue for gauging how well they can cope with the vagaries of real-life stimuli.

Our findings showed that younger adults had better emotion recognition performance for AV relative to the unimodal expressions even when the recognition of an emotion in one modality was poor (replicating the results of Kim & Davis, 2012). This result can largely be explained if it is assumed that younger adults were able to rule out potential response candidates if they were not supported in both modalities. This would reduce the number of potential candidates and so result in an AV recognition benefit.

Older adults did not show an AV benefit. That is, for the majority of expressions, AV presentation did not improve recognition performance relative to VO performance. Indeed, for disgust (in the ‘clear’ condition), recognition performance declined when auditory information was added to visual information (i.e., AV performance was worse than VO performance). The calculation of the performance that might be expected under optimal AV integration indicated that compared to the younger adults, the older adults did not integrate auditory and visual information effectively. This finding is in line with previous studies that show that older adults have difficulty matching auditory and visual information (e.g., Ruffman et al., 2009; Ryan et al., 2009); but stands in contrast to those of previous studies in which older adults were able to benefit from AV emotion presentation, and did so to the same degree, or even more so than younger adults (e.g., Hunter et al., 2010; Lambrecht et al., 2012). These studies, however, used AV stimuli that had high VO and AO emotion recognition rates. The current results suggest that older adults may have problems integrating and recognising emotion from AV presentations when the expression in one modality (typically the auditory one) is not well recognised. That is, relative to younger adults, older adults appeared to show particular problems recognising emotion from AO expressions of emotion. It is interesting to note that older adult's AO emotion recognition scores were not related to their hearing levels (see above); this suggests that the problems in perceiving auditory emotion may be more central than peripheral.

In conclusion, we found that, unlike younger adults, older adults did not have better emotion recognition for AV as compared to unimodal presentation. Indeed, for some emotions, relatively good visual emotion recognition performance was degraded when the (less clear) auditory emotion information was added. These findings have implications for how older adults may perform in emotion recognition in everyday life. That is, in daily life, the presentation of multimodal emotion is common; moreover, expressions are often fleeting and may not be prototypical. The current results suggest that older adults will have difficulty in recognising such emotions, difficulties that may not be apparent from standard unimodal (often visual) assessment (e.g., The Diagnostic Analysis of Nonverbal Accuracy, DANVA, Nowicki & Duke, 1994; Facial Expressions of Emotion: Stimuli and Tests, FEEST, Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). Indeed, in relying on tests that use only unimodal expressions, we might be missing the problems that older adults encounter during every day, real-life situations.

2.6 References

Andrés, P., Parmentier, F. B. R., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia, 44*(12), 2564-2568.

Bach, M. (2007). The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology, 245*(7), 965-971.

Barrett, M. M., & Newell, F. N. (2015). Task-specific, age related effects in the cross-modal identification and localisation of objects. *Multisensory Research, 28*(1-2), 111-151.

Benoît, C., Grice, M., & Hazan, V. (1996). The SUS test: A method for the assessment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences. *Speech Communication, 18*(4), 381-392.

Chaby, L., Luherne-du Boullay, V., Chetouani, M., & Plaza, M. (2015). Compensating for age limits through emotional crossmodal integration. *Frontiers in Psychology, 6*, 691.

Dywan, J., Segalowitz, S. J., & Webster, L. (1998). Source monitoring: ERP evidence for greater reactivity to nontarget information in older adults. *Brain and Cognition*, *36*(3), 390-430.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*(1), 116-124.

Gonçalves, A. R., Fernandes, C., Pasion, R., Ferreira-Santos, F., Barbosa, F., & Marques-Teixeira, J. (2018). Effects of age on the identification of emotions in facial expressions: a meta-analysis. *PeerJ, 6*, e5278.

Hunter, E. M., Phillips, L. H., & MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychology and Aging, 25*(4), 779-787.

Kim, J., & Davis, C. (2012). Perceiving emotion from a talker: How face and voice work together. *Visual Cognition, 20*(8), 902-921.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*(3), 529-539.

Massaro, D. W., & Cohen, M. M. (2000). Tests of auditory–visual integration efficiency within the framework of the fuzzy logical model of perception. *The Journal of the Acoustical Society of America, 108*(2), 784-789.

May, C. P. (1999). Synchrony effects in cognition: The costs and a benefit. *Psychonomic Bulletin and Review, 6*(1), 142-147.

Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal Behavior, 18*(1), 9-35.

Orbelo, D. M., Grim, M. A., Talbott, R. E., & 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. *Journal of Geriatric Psychiatry and Neurology, 18*(1), 25-32.

Rosen, H. J., Wilson, M. R., Schauer, G. F., Allison, S., Gorno-Tempini, M.-L., Pace-Savitsky, C., . . . Miller, B. L. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia, 44*(3), 365-373.

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions in older adults. *The Journals of Gerontology: Series B, 64*(6), 696-703.

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*(4), 863-881.

Ryan, M., Murray, J., & Ruffman, T. (2009). Aging and the perception of emotion: Processing vocal expressions alone and with faces. *Experimental Aging Research, 36*(1), 1-22.

Wieck, C., & Kunzmann, U. (2017). Age differences in emotion recognition: A question of modality? *Psychology and Aging, 32*(5), 401-411.

Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion: Stimuli and Tests (FEEST).* Bury St. Edmunds, England: Thames Valley Test Company.

Chapter 3. Age-related Changes to Gaze Behaviour During Emotion Recognition

3.1 Abstract

When asked to decide what emotion a face is showing older adults look at the mouth more than do younger adults. The current study investigated a possible cause for this behaviour (i.e., hearing loss) and a possible consequence of it (i.e., reduced emotion recognition performance). Older (n = 17) and younger (n = 18) adults' eye movements were recorded as they completed an emotion recognition task for angry, disgusted, sad, fearful, surprised, happy, and neutral expressions. Stimuli consisted of static and dynamic faces. Overall, older adults gazed longer and made more fixations at the mouth region than younger adults; this occurred regardless of whether the faces were static or dynamic. Further, older and younger adults looked at the eye region of faces for a similar amount of time. Whilst older and younger adults made a similar number of fixations to the eye region of most emotional faces, in some instances (e.g., fear, sad) younger adults made more fixations to the eye region than older adults. Finally, younger adults showed better classification accuracy than older adults for most emotional expression types; however, this decline in older adults' emotion recognition was not associated to gaze behaviour patterns. Findings also indicated that older adults' gaze behaviour was not the result of hearing loss. It is suggested that younger adults adopt a more efficient focused-gaze strategy towards the eye region when recognising emotion in faces whereas older adults adopt an exploratory-gaze strategy as they may have difficulty picking up emotion information from the face regions.

3.2 Introduction

Aging is associated with changes in eye-movement behaviour with differences between young and older adults occurring at many different levels. In terms of how eye movements are performed, older adults show increased saccadic latencies and decreased saccadic accuracy and velocity (Dowiasch, Marx, Einhäuser, & Bremmer, 2015; Klein, Fischer, Hartnegg, Heiss, & Roth, 2000). In terms of what is fixated, it has been reported that the gaze pattern of older adults is different from that of younger adults when judging what emotion is expressed by a face (e.g., Sullivan, Ruffman, & Hutton, 2007). This is interesting since older adults are known to have problems in emotion recognition and under the gaze direction assumption (see Lansing & McConkie, 1999) eye gaze patterns are linked to how information is processed by a person in order to perform a task. That is, if people actively direct their gaze (the region of highest visual acuity) around a visual scene and fixate areas that provide task-relevant information then differences in the way that older adults inspect emotional faces could be used to understand differences in how younger and older adults process emotion information. In this regard, the current study examined how older and younger adults inspect emotional faces and how these gaze patterns may be related to emotion recognition performance.

Previous studies have reported that for both younger and older adults, gaze behaviour appears to be driven by the type of emotion that is displayed. For instance, younger adults tend to look longer at the upper than the lower half of angry and sad expressions; but look longer at the lower face region for disgusted expressions, and show no face region differences for fear, joy, and neutral (Chaby, Hupont, Avril, Luherne-du Boullay, & Chetouani, 2017). Older adults show a similar pattern, however, compared to younger ones, spend more time looking at the mouth and less time looking at the eye region (e.g., Chaby et al., 2017; Sullivan et al., 2007; Murphy & Isaacowitz, 2010). That is, the overall preference (i.e., greater frequency and longer fixation durations) for the eye region compared to the mouth region shown by young adults is weaker in older adults (Sullivan et al., 2007). Indeed, this relative preference that older versus younger adults show for the mouth compared to the eye region has been reported across a variety of emotion types (e.g., happy, surprise, fear, sad, disgust, anger; Sullivan et al., 2007; Chaby et al., 2017). So, for example, Sullivan et al. (2007) reported that, relative to younger adults, older adults looked longer at the mouth region of faces by approximately 700ms, with this difference maintained across all emotion types. In addition to age differences, Sullivan, Campbell, Hutton, and Ruffman (2017) suggest that there may also be gender differences in the way that older (but not younger) adults scan faces. Specifically, they found that young men and women spent equal amounts of time looking at the eyes and the mouth region of emotion faces; whereas older men gazed at the mouth region more than older women.

Emotion recognition studies that have presented restricted views of a face expressing emotion (e.g., just the top or bottom of the face) suggest that emotion information is conveyed at different spatial locations (Kim & Davis, 2012). An account of gaze patterns based on the premise that gaze fixation reflects differences in the spatial locations of critical information (see Schurgin et al., 2014) can help to explain why looking patterns are driven by the type of emotion viewed, but it is not clear why older adults show a tendency to disproportionately monitor the mouth region and what the effect of such a tendency may have on performance. To investigate these issues, the current study aimed to investigate two specific proposals, the first concerning why this behaviour occurs, the second concerning the effect this behaviour may have.

In terms of the causes underlying older adults' mouth gazing behaviour; it is unlikely to result from some general aging effect on eye-movement themselves. This is because we would expect age-related deterioration in eye-movement to be limited to motion parameters, such as eye-movement speed, and not to influence the overall gaze pattern. Further, a general age-related deterioration would be expected to result in age differences in gaze behaviour across all types of images, however, this does not seem to be the case. For example, for non-face stimuli such as landscapes (e.g., images of forests, cities etc.) gaze behaviour does not differ between the age groups (Circelli, Clark, & Cronin-Golomb, 2013).

As an alternative to a low-level account, Chaby et al. (2017) proposed that the observed gaze pattern differences are the result of an information processing strategy adopted by older adults to cope with age-related hearing loss. That is, they suggest that older adults alter their gaze behaviour as a flow-on effect of age-related hearing loss. Under this account, because hearing loss degrades the auditory speech signal, older adults (who often are affected by presbycusis) compensate by gazing at the mouth region. This strategy (adopted when viewing faces) enables older adults to obtain speech-relevant visual cues and information which then assist in speech perception (Davis & Kim, 2006). More specifically, Chaby et al. (2017) propose that older adults adopt a "focused-gaze strategy" towards the mouth region in contrast to an "exploratory-gaze strategy" typically employed by younger adults which involves constant exploration of both upper and lower face regions. Importantly, this compensatory strategy used by older adults should make the mouth a salient feature, regardless of whether the face is static or moving (Chaby et al., 2017).

Consistent with the above proposal, research with younger adults provides some evidence that a degraded speech signal could act to bias gaze behaviour when inspecting emotional faces. For example, Buchan, Paré, and Munhall (2007) presented dynamic speech stimuli to younger adults in an emotion recognition task. Overall, younger adults showed more eye than nose or mouth gazing when recognising emotion in faces. However, when the intelligibility of the auditory speech signal was reduced (using a simultaneously presented multi-talker noise masker), younger adults shifted their gaze away from the eyes and towards lower parts of the face (i.e., nose, mouth). Such a gaze shift makes sense as a compensatory strategy to improve speech perception, i.e., viewing the lower half of the face would act to increase the fidelity of speech-relevant cues attained from the mouth area. Likewise, with hearing loss, older adults may have developed a long-term strategy of gazing towards the mouth region of talking faces for speech comprehension.

The effect that such a gaze behaviour strategy may have on how face information is processed will depend upon what information is conveyed by the lower and upper face regions as a function of the given task. For emotion recognition, it has been argued that an enduring tendency to gaze at the mouth will impede performance for particular emotions (Wong, Cronin-Golomb, & Neargarder, 2005). This argument is based on the assumption that accurate recognition performance is achieved by gazing at face regions that convey emotion-relevant information. Given that the eye region contains essential cues for anger, sad, and fear recognition (Calder, Young, Keane, & Dean, 2000) and that older adults look towards the mouth region of emotional faces, older adults may not sufficiently perceive emotion information in the eye region and, as a result, may not successfully recognise these emotion types. In line with this argument, it is well reported that aging is associated with declines in the ability to recognise emotion, particularly for anger, fear, and sad relative to other emotions (e.g., happy, surprise, disgust).

Research has used correlational statistics to investigate the relationship between gaze behaviour and recognition performance. Whilst significant correlations between older adults' gaze behaviour and recognition performance could indicate that recognition performance is impeded by gaze behaviour, they could also indicate that gaze behaviour is informed by recognition performance. That is, older adults may engage in more mouth gazing because of problems identifying the presented emotion. For instance, using a forced-choice emotion recognition task and eye-tracking, Wong et al. (2005, experiment 3) found that in younger and older adults, better recognition performance for anger, fear, and sad was correlated with more gazing towards the upper half of faces. Further, they found that for older adults, poorer recognition of these emotion types was associated with more lower face looking (also reported in Sullivan et al., 2007). In contrast, Murphy and Isaacowitz (2010) found a positive correlation between mouth gazing and recognition performance in the younger adult group only, i.e., there were no significant correlations between gaze variables and accuracy for the older adult group.

Several features of these studies make it difficult to draw a conclusion about whether gaze behaviour is related to recognition performance. For example, Wong et al. (2005) measured gaze behaviour for emotional faces in a separate and subsequent study to their emotion recognition task. During the eye-tracking portion of the study, participants were not instructed to recognise the displayed emotion. It is possible that gaze patterns when passively viewing an emotional face may be different to gaze patterns when actively seeking an emotion in an emotional face. Indeed, research indicates that gaze patterns typically vary depending on the type of information to be ascertained (Itier, Villate, & Ryan, 2007). Hence, any associations or correlations between these separate tasks should be treated with caution. In the correlational analyses of Sullivan et al. (2007) looking times were grouped together for emotional expressions best recognised from the eyes (i.e., anger, sadness, and fear) and for expressions best recognised from the mouth region (i.e., happiness, disgust). Here, it is unclear whether significant correlations are specific to particular emotion types. Finally, in the study by Murphy and Isaacowitz (2010), sample sizes differed between the younger and older adult groups.

The current study followed up the above two issues, i.e., whether the age-related changes to gaze behaviour in emotion recognition might be caused by hearing loss (Chaby et al., 2017), and that the consequence of these changes is reduced emotion recognition performance (Wong et al., 2005); by determining whether such changes will be associated with hearing loss; and the extent that emotion recognition performance was affected.

We investigated older and younger adults' recognition of emotional expressions while tracking gaze patterns. Unlike previous eye-tracking studies, we also measured hearing ability in both age groups to investigate whether hearing ability is associated with gaze behaviour. Further, to date, most studies that investigate aging, emotion recognition, and gaze behaviour have used static expressions of emotion (i.e., images of actors expressing emotion). One study by Grainger, Henry, Phillips, Vanman, and Allen (2017) investigated whether older adults gaze patterns are consistent across static and dynamic (videos of people expressing an emotion) expressions of emotion. They found that younger adults showed changes to gaze behaviour for dynamic relative to static expressions whereas older adults did not.

Although it is accepted that age-related changes to gaze behaviour occur for static expressions of emotion, to our knowledge, no studies have assessed dynamic speech expressions of emotion (i.e., videos of people uttering sentences and simultaneously conveying an expression) despite that talkers' expressive faces are the ones that are often encountered in day to day communication. So, we used static and also dynamic faces to ascertain whether older adults' observed gaze behaviours present across different types (i.e., static, dynamic) of emotional faces. Further, given that the dynamic expressions in the current study convey visual speech information, such expressions may be better able to reveal whether older adults' mouth gazing behaviour is due to hearing loss.

If the above suggestions for the cause (hearing loss) and consequence (decline in emotion recognition accuracy) of older adults' observed gaze behaviour are valid, then it is likely that hearing loss will drive gaze behaviour in older adults and that gaze behaviour will determine emotion recognition accuracy. That is, across emotional expression types, older adults with hearing loss would show greater mouth gazing, relative to older adults whose hearing is relatively normal, not to mention younger adults. Further, greater mouth gazing would lead to lower recognition performance, particularly for those emotions where important information is conveyed by the eye region, e.g., anger, fear, and sad.

3.3 Method

3.3.1 Participants

Seventeen older adults (8 females, M_{age} = 70, range = 60-79) and 18 younger adults (15 females, M_{age} = 21, range = 18-33) participated in the experiment. Younger adults were undergraduate Psychology students from Western Sydney University and received course credit for their participation. Older adults were recruited from the community and received monetary reimbursement for their participation. The relatively small sample size of this experiment is due to the difficulty encountered during the data collection stage of some older adults. Seven additional older adults were recruited but did not complete the experiment as the eye-tracker could not successfully track their eye movement. Previous studies show that eye-tracking is more difficult with older adults possibly because of such things as droopy eyelids or yellowed corneas (Murphy & Isaacowitz, 2010; Isaacowitz, Wadlinger, Goren, & Wilson, 2006).

All participants reported to have learnt English at a young age (i.e., less than 7 years old) and had normal visual acuity except one older adult who had slightly worse than normal vision. Participants were screened for dementia using the Mini-Cog™ test (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000) as the presence of dementia has been associated with poor emotion processing and may act as a confounding variable (Rosen et al., 2006). All older (M = 4) and younger (M = 4) adults scored negative for cognitive impairment (where a score of at least 3 out of 5 points represents normal cognition).

3.3.2 Stimuli

Stimuli consisted of static picture and dynamic visual-only recordings of emotional expressions by two females (M_{age} = 21.5). They were native talkers of Australian English and produced varying emotional expressions (i.e., happy, surprise, fear, disgust, anger, sad), as well as a broad focused (or "neutral") expression, while uttering sentences one by one.

The dynamic recordings consisted of IEEE sentences (Rothauser, 1969) read aloud and expressing the requisite emotion. In order that the content of the sentence did not influence the emotion rendition, a large number of sentences were rated by five coders for emotional content on a Likert scale (using a 5-point scale from -2 = very negative, 0 = neutral, and +2 = very positive). Thirteen sentences that received a score of 0 from each coder were selected for recording. Eleven of these sentences were to be used throughout the experiment as experimental items (10 sentences) or practice items (1 sentence). The remaining two sentences were to be used as a presenter specific calibration for participants throughout the experiment.

The video recordings of the two presenters were captured individually in a sound attenuated booth. The presenter was seated in front of a monitor that displayed each sentence one at a time. The video camera (Sony NCCAM HXR-NX30p) was situated directly above the monitor and captured video at 1920 x 1080 full HD resolution at 50 frames per second.

The recording session was blocked by the type of expression. Each block contained eleven sentences, resulting in 77 recorded sentences (7 expression types x 11 sentences) for each of the two presenters. Prior to each block, presenters were instructed which expression was required to be produced. For the neutral expressions, the presenter was directed to read the sentence silently and then to look into the camera lens as they uttered the sentence out loud. In conjunction with these recordings, an additional two sentences with a neutral expression were also recorded (for presenter specific calibration).

For the emotional expressions, the Velten Mood Induction Procedure was used to induce a cheerful (positive) mood for surprise and happy and a depressive (negative) mood for angry, disgust, sad, and fear (Velten, 1968). That is, presenters were given a relevant positive (e.g., “I feel cheerful and lively”) or negative (e.g., “I’m discouraged and unhappy about myself”) list of sixty self-referent mood statements. They were asked to read each statement carefully and to try to feel the mood suggested by each statement. Then, presenters were asked to recall a memory where they felt the relevant emotion in an attempt to further induce the emotional state. Once the emotion induction procedure was completed, recording of the block of the relevant emotional expression commenced.

To produce the dynamic visual-only stimuli, all video recordings were stripped of audio, cropped to include just the head area (i.e., width = 600 pixels, height = 850 pixels), and segmented into individual sentences using a custom MATLAB script. The static visual stimuli were created using the dynamic stimuli, such that a single frame of the peak expression (i.e., the frame that most clearly conveyed the intended expression) was selected.

3.3.3 Experimental design

The emotion recognition task contained 280 trials consisting of 7 expression types (i.e., angry, disgust, sad, fear, surprise, happy), 2 presenters, 10 sentences, and 2 format types (static, dynamic). Trials were blocked by presenter, resulting in two blocks of 140 trials. Within each presenter block, participants viewed the expressions in one format (e.g., static) before completing the same task in the other format (e.g., dynamic). The presentation order of the presenters and format types were counterbalanced across participants. All trials of the same format type (i.e., static, dynamic) within each presenter block were randomised using the SR Research Experiment Builder software.

3.3.4 Procedure

Prior to the experiment, participants completed the Mini-Cog™ test (Borson et al., 2000). Hearing acuity was measured for the right and left ears using pure tone audiometry (Diagnostic Audiometer, AD229e) for 4 different frequencies (0.25, 0.5, 1, 2, and 4 kHz). Finally, visual acuity in both eyes was tested using the Freiburg Visual Acuity Test (Bach, 2007).

Participants were tested individually in a quiet room. They were informed that the eye movements of the right eye would be recorded during the experiment using an eye tracker (SR EyeLink® 1000) that was mounted on a tower with a chin and head rest. They were instructed to rest their heads on the head and chin rest to limit head movements that might interfere with the recording of eye movements during the experiment.

Participants were told that they would see still images or silent videos of a person conveying an emotional expression and that their task was to identify the expression by using the mouse to click one of several labels presented on the screen (i.e., “Angry”, “Disgust”, “Sad”, “Fear”, “Surprise”, “Happy”, “Neutral”). Emotion labels were presented in alphabetical order in two columns, with the “Neutral” label presented in between the two columns in the centre of the screen. Once participants indicated that they understood the instructions, the experiment began.

At the beginning of the experiment, right eye positions were calibrated using the SR EyeLink® 1000. For each new presenter and format type (i.e., static, dynamic), participants were shown two neutral expressions of that presenter and format type to act as a presenter specific calibration. Participants were told to use these neutral expressions as a baseline to compare with subsequently presented expressions. Following this, seven practice items were then presented to give participants the opportunity to be exposed to and identify each of the expression types from that format type and presenter. This was then followed by the experimental items.

To initiate the presentation of an experimental item, participants were required to gaze at a fixation box for one second. The fixation box was presented in the same location as the centre of the subsequently presented face. This was to ensure that all initial eye fixations were the same across all items and participants. Once the trial was triggered, participants were presented with a silent video (for approximately 3 to 4 seconds, depending on the length of the video) or an image (for 4 seconds). During this time, eye gaze movements were recorded. Seven emotion labels were then presented for a response. Participants were given a break after every fourteen trials. Eye calibration was performed after each break. At the conclusion of the study, participants were debriefed as to the purpose of the study.

3.3.5 Data analysis

Using the SR Research DataViewer software, two areas of interest were created for each experimental item: one area included the eyes and the eyebrows only, and the other the mouth only (see Figure 3.1). For dynamic items, similar dynamic areas of interest were used and these were inspected to ensure that they included the appropriate face areas for the length of each video. The number of fixations and the duration of fixations to each of the areas of interest was extracted using the DataViewer software and were analysed further. To identify a fixation, the Eyelink system measures the velocity of eye movements. If the velocity of the eye movement goes above a certain threshold, a saccade is marked, otherwise, a fixation is marked. Only fixations made within the areas of interest were analysed.

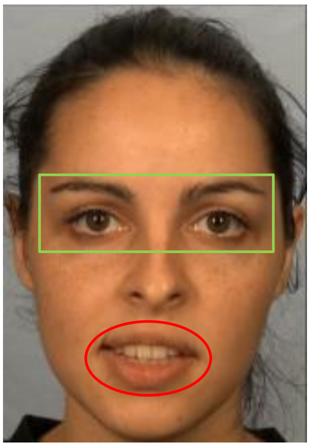

Figure 3.1. Interest areas for the eye (green box) and mouth (red oval) face areas.

3.4 Results

Accuracy scores will be presented first, followed by the participants' looking patterns (i.e., number of fixations and duration of fixations). The final analyses explore whether participants' gaze behaviour is associated with recognition performance or hearing ability.

3.4.1 Accuracy scores

Percent correct scores were collapsed across sentences and analysed in a repeated measures ANOVA containing expression type (angry, disgust, sad, fear, surprise, happy, neutral), format type (static, dynamic) and presenter as within participants factors; and age group as the between participants factor. Identification scores for each of the age groups, emotional expressions, and format types are presented in Figure 3.2.

Overall, younger adults (78%) outperformed older adults (64%), $F(1,33) = 30.95$, $p < .001$, $\eta p2 = .48$. Dynamic faces (73%) attracted higher percent correct scores than static faces (69%), $F(1,33) = 8.34$, $p < .01$, $\eta p2 = .20$. There was also a significant main effect of expression type, $F(6,198) = 55.09$, $p < .001$, $\eta p2 = .63$. Based on error rates, one presenter (79%) conveyed emotion more clearly relative to the other (63%), $F(1,33) = 240.21$, $p < .001$, $\eta p2 = .88$. The interaction between expression type, format type, and age group was significant, $F(6,198) = 2.54$, $p < .05$, $\eta p2 = .07$. To analyse this interaction further, a simple effects analysis with a Bonferroni adjusted alpha was used.

The simple effects analysis indicated that younger adults outperformed older adults across most expression types and format types ($p < .04$ for all comparisons); except for the sad static expressions, the surprised dynamic expressions, and the neutral static expressions where there were no significant differences across the age groups ($p > .07$).

Older adults showed better recognition for angry and happy emotions with dynamic compared to static format ($p < .01$). The reverse pattern was true for disgust and surprise ($p < .01$). Older adults' recognition scores appeared similar across format types for fearful, neutral, and sad expressions ($p > .16$).

Younger adults recognised anger, sad, fear, and neutral more accurately when these were presented dynamically as compared to statically ($p < .05$). For the surprised expressions, younger adults' accuracy was higher with the static compared to the dynamic format ($p < .01$). Younger adults' recognition scores were near ceiling for the disgusted and happy expressions regardless of format type ($p > .09$).

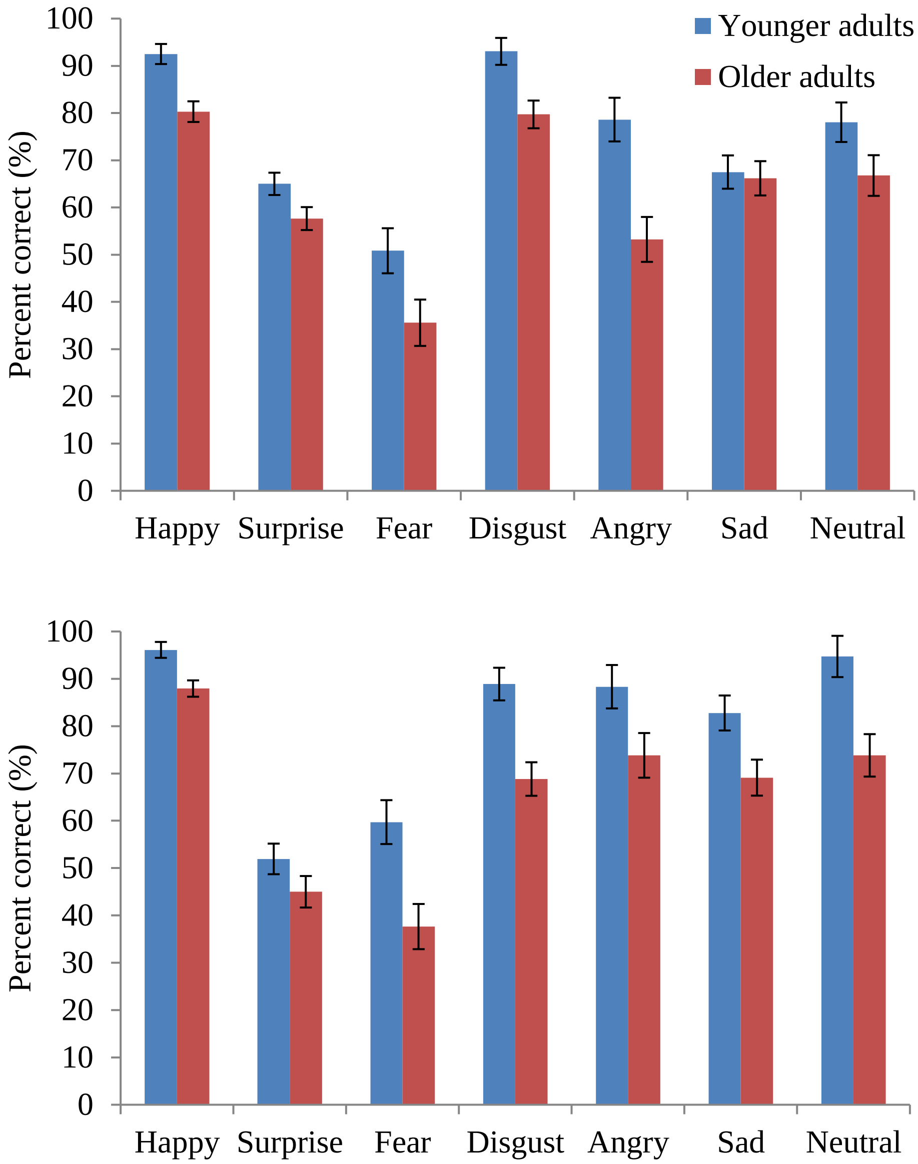

Figure 3.2. Younger and older adults' percent correct scores for each of the static (top) and dynamic (bottom) emotional expressions.

3.4.2 Looking patterns

The number and duration of the eye fixations to each of the interest areas (eye region, mouth region) were averaged across sentences. These scores were analysed through a repeated measures ANOVA with expression type, format type, presenter, and face region (eye region, mouth region) as within participants factors and age as the between participants factor.

The number of fixations for the static and dynamic expressions are presented in Figures 3.3 and 3.4, respectively. The ANOVA for the overall number of fixations (to the eye and/or mouth regions) revealed that, on average, older adults' (3.1) number of overall fixations was not significantly different to that of younger adults (3.5), $F(1,33) = 1.10$, $p = .30$, ηp2 = .03. For both groups, there was a higher number of fixations to static (3.9) relative to dynamic (2.8) faces, $F(1,33) = 113.59$, $p < .001$, ηp2 = .78, and to the eye (*4*.9) compared to the mouth (1.8) region, $F(1,33) = 28.36$, $p < .001$, ηp2 = .46. There were significant main effects of expression, $F(6,198) = 10.68$, $p < .001$, ηp2 = .24, and presenter where one presenter (i.e., the presenter that conveyed emotion more clearly than the other) attracted fewer fixations (3.2) than the other (3.4), $F(1,33) = 4.95$, $p < .05$, ηp2 = .13.

The interaction between expression type, face region, and age group was significant, $F(6,198) = 2.52$, $p < .05$, ηp2 = .07, and was analysed further with a simple effects analysis and a Bonferroni adjusted alpha. This analysis revealed that older and younger adults appeared to show a similar number of fixations for the eye regions of angry, disgust, surprise, happy, and neutral faces ($p > .05$ for all comparisons). Younger adults had a higher number of fixations than older adults for the remaining emotion types (i.e., fear, sad, $p < .05$). For the mouth region, older adults made more fixations than younger adults across all types of emotional expressions ($p < .05$). In addition, younger adults had a higher number of fixations to the eye than the mouth region across all emotion types ($p < .01$). Older adults also showed this pattern for angry, disgust, sad, and surprise expressions ($p < .05$) but showed no significant differences for fear, happy, or neutral emotional faces ($p > .05$).

To investigate overall gender differences in the number of fixations in the older adult group (as suggested by Sullivan, Campbell, Hutton, & Ruffman, 2017), a repeated measures ANOVA was conducted with face region as a within participants factor and gender as a between participants factor. Overall, men (2.8) and women (3.5) made a similar number of fixations, $F(1,15) = 1.09$, $p = .31$, ηp2 = .07. The interaction between age group and face region was not significant, $F(1,15) = 1.85$, $p = .19$, ηp2 = .11.

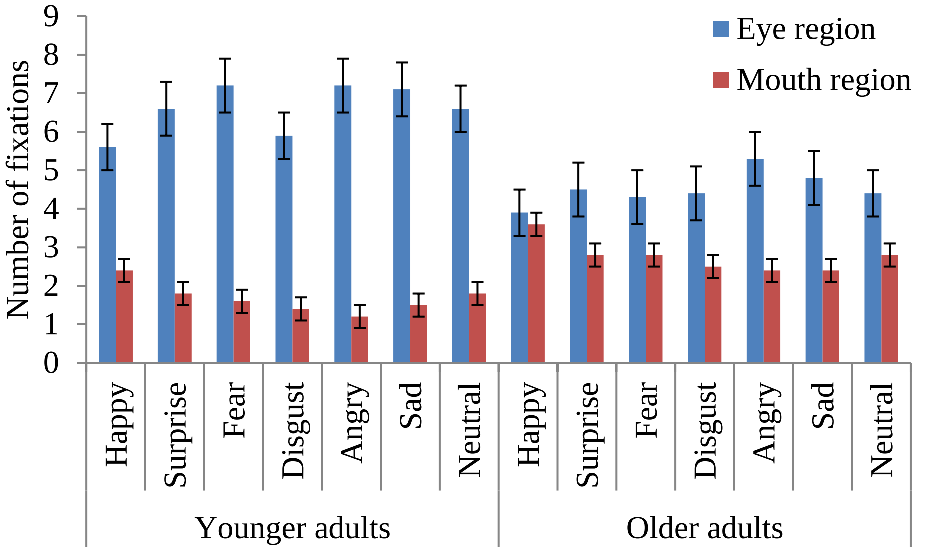

Figure 3.3. Younger (left) and older (right) adults' mean number of fixations for each face region across the static emotional expressions. The error bars represent standard error.

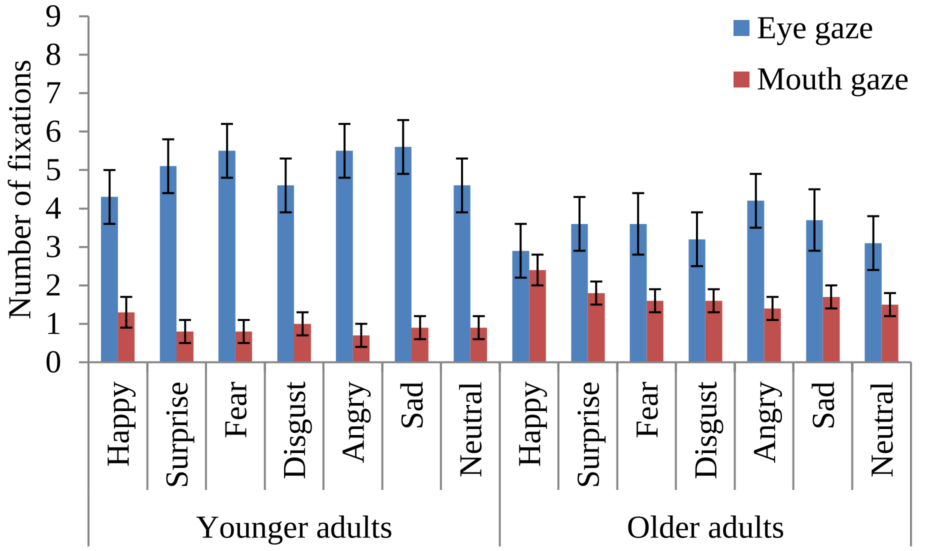

Figure 3.4. Younger (left) and older (right) adults' mean number of fixations for each face region across the dynamic emotional expressions. The error bars represent standard error.

Overall duration scores for the static and dynamic faces are presented in Figures 3.5 and 3.6, respectively. For the duration scores, younger (1035ms) and older (1012ms) adults appeared to spend a similar amount of time fixating on the eye and mouth regions of the facial expressions, $F(1,33) = .05$, $p = .83$, ηp2 = .00. Overall, participants gazed longer at static (1072ms) than dynamic (974ms) expressions, $F(1,33) = 6.28$, $p < .02$, ηp2 = .16, and longer at eye (1417ms) than mouth regions (629ms), *F*(1,33) = 15.31, $p < .001$, ηp2 = .32. The main effect of expression type was significant, *F*(6,198) = 9.04, $p < .001$, ηp2 = .22, whereas the main effect of presenter was not significant, *F*(1,33) = 3.56, $p = .07$, ηp2 = .10.

The interaction between face region and age group was significant, *F*(1,33) = 5.71, $p < .05$, ηp2 = .15. This interaction was analysed further using a simple effects analysis with a Bonferroni adjusted alpha. On average, younger adults tended to look at the eye region longer than the older adults but this difference was not significant ($p = .06$). On the other hand, older adults looked significantly longer at the mouth region of faces than younger adults ($p < .05$). Further, younger adults spent more time overall gazing at the eye region than the mouth region of faces ($p < .001$), whereas older adults did not show a significant difference ($p = .30$).

Duration scores were entered into a repeated measures ANOVA to investigate gender differences in the older adult group (as suggested by Sullivan, Campbell, Hutton, & Ruffman, 2017). Face region was entered as a within participants factor and gender as a between participants factor. Overall, men (1005ms) and women (1019ms) made a similar number of fixations, *F*(1,15) = .008, $p = .93$, ηp2 = .001. The interaction between age group and face region was not significant, *F*(1,15) = 1.48, $p = .24$, ηp2 = .09.

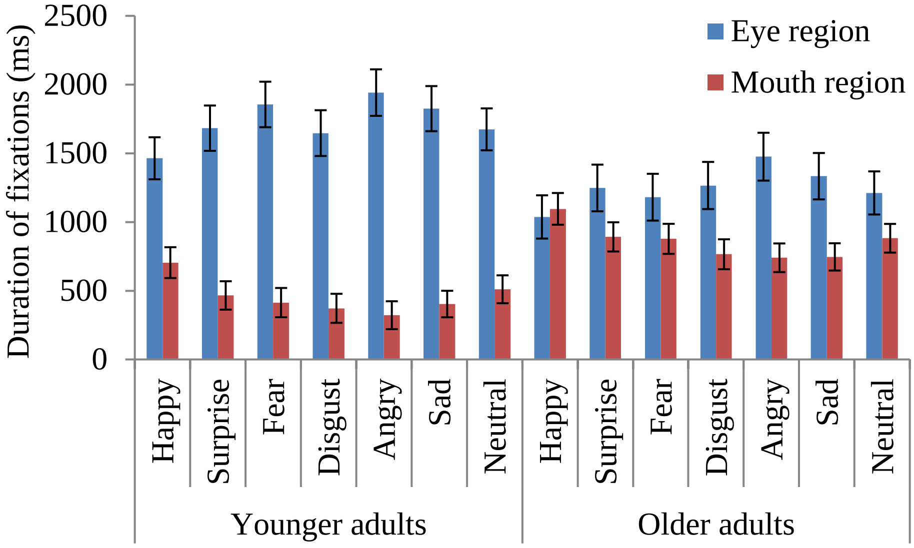

Figure 3.5. Younger (left) and older (right) adults' mean duration of fixations (in milliseconds) for each face region across the static emotional expressions. The error bars represent standard error.

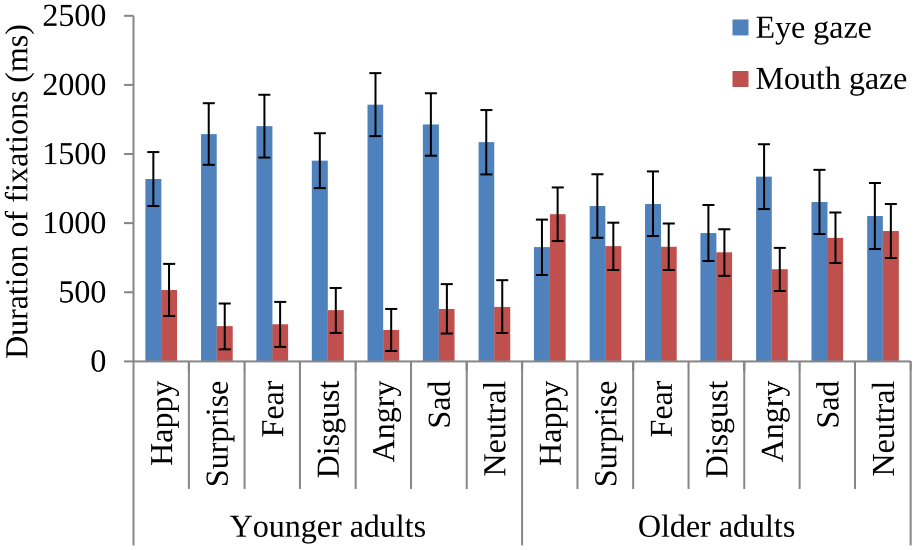

Figure 3.6. Younger (left) and older (right) adults' mean duration of fixations (in milliseconds) for each face region across the dynamic emotional expressions. The error bars represent standard error.

Difference scores between younger and older adults for the number and duration of fixations to each of the face regions were calculated and are presented in Table 3.1. These difference scores were calculated to better illustrate the overall differences between younger and older adult gaze patterns; and to identify the emotion and format types that attracted the largest age differences. As can be seen, the difference scores were similar across most of the emotion types for static faces. Age differences were the largest for the number and duration of fixations made to the eye region of fear faces whereas the smallest age differences were shown for the gaze scores to the eye region of disgust faces. Interestingly, difference scores were more consistent for dynamic faces suggesting that the pattern of gazing was more similar between the younger and older adults for the dynamic relative to the static expressions.

Table 3.1. Mean differences between the younger and older adults for the number of fixations and the duration of fixations (milliseconds) to the eye and mouth regions of static and dynamic emotional expressions. A positive number indicates that the younger group scored higher than the older group.

	Static				Dynamic			
	Number of fixations		Duration of fixations		Number of fixations		Duration of fixations	
	Eye region	Mouth region	Eye region	Mouth region	Eye region	Mouth region	Eye region	Mouth region
Angry	1.9	-1.2	466.4	-418	1.3	-.7	521.9	-438
Disgust	1.5	-1.1	381.3	-393.4	1.4	-.6	523.4	-418.5
Sad	2.3	-.9	491.6	-342.9	1.9	-.8	559.6	-514.1
Fear	2.9	-1.2	674.8	-463.7	1.9	-.8	561.7	-560.8
Surprise	2.1	-1	434.8	-426	1.5	-1	521.1	-580.1
Happy	1.7	-1.2	427	-391.1	1.4	-1.1	494.4	-545.8
Neutral	2.2	-1	462.9	-371.2	1.5	-.6	533.6	-547.2

To investigate the participants' bias to look at the eye region, gaze scores to the mouth region were subtracted from gaze scores to the eye region. Difference scores are presented in Table 3.2 and indicate that younger adults showed a bias to gaze at the eye region than the mouth region of faces; older adults also showed this bias but to a lesser extent. Despite this age difference, the pattern of looking was similar across the age groups. For instance, for both age groups, the difference between eye and mouth looking was larger for angry expressions than for happy expressions.

Table 3.2. Mean differences between the eye and mouth region gaze scores (i.e., number of fixations, duration of fixations) of static and dynamic emotional expressions for the younger and older adults. A positive number indicates that the eye region was gazed at more than the mouth region.

	Static				Dynamic			
	Number of fixations		Duration of fixations		Number of fixations		Duration of fixations	
	Younger adults	Older adults	Younger adults	Older adults	Younger adults	Older adults	Younger adults	Older adults
Angry	5.9	2.9	1619.8	735.4	4.8	2.8	1630.0	670.1
Disgust	4.6	1.9	1274.9	500.3	3.6	1.6	1082.7	140.8
Sad	5.6	2.5	1333.9	587.2	4.6	2.0	1333.9	260.3
Fear	5.6	1.5	1441.7	303.1	4.8	2.1	1432.7	310.2
Surprise	4.8	1.7	1391.9	356.1	4.3	1.9	1391.9	290.6
Happy	3.2	0.4	759.3	-58.7	2.9	0.5	801.5	-238.7
Neutral	4.8	1.6	1189.7	329.6	3.6	1.6	1189.7	108.9

3.4.3 Correlational analyses

To examine whether eye and mouth looking correlated with emotion recognition performance, we conducted simple bivariate correlations between gaze scores (i.e., number of fixations and gaze duration) for the eye and mouth regions and percent correct scores for all the emotion types (i.e., angry, disgust, sad, fear, surprise, happy). As shown in Table 3.3, no significant correlations were found between percent correct scores and the mean number of fixations for younger or older adults. This finding was replicated for the correlation between the percent correct scores and the duration of fixation scores for both the younger and older age groups (see Table 3.3).

Table 3.3. Pearson correlations for the percent correct scores and the gaze scores for each emotion type and for each age group.

		Number of fixations				Duration of fixations			
		Younger adults		Older adults		Younger adults		Older adults	
		Eye region	Mouth region	Eye region	Mouth region	Eye region	Mouth region	Eye region	Mouth region
Percent correct scores	Angry	.25	-.38	.24	.01	.33	-.45	.10	.03
	Disgust	-.03	-.39	-.08	-.04	.07	-.46	-.08	.14
	Sad	.25	.05	.31	.11	.27	.07	.17	-.01
	Fear	.21	.02	.32	-.11	.15	-.03	.28	-.11
	Surprise	.24	-.11	-.07	.02	.28	-.20	-.15	-.02
	Happy	-.23	-.02	-.01	.21	-.17	.04	.03	.01

3.4.4 Hearing acuity

Overall, older adults hearing was poorer than younger adults particularly at the higher frequencies (as is typical of presbycusis, see Table 3.4).

Table 3.4. Average hearing level (dB) in the left (top) and right (bottom) ears across 0.25, 0.5, 1, 2, and 4 kilohertz for the younger and older adult group. Average hearing level and range of hearing levels are also presented. Standard deviations are shown in parentheses.

		Frequency (kHz)					Average	Range
		0.25	0.5	1	2	4	(.25-4)	(.25-4)
Left Ear	Younger adults	18.89 (3.66)	16.11 (4.39)	14.72 (6.75)	10.56 (5.66)	7.78 (4.92)	13.61 (5.07)	0-35
	Older adults	19.71 (8.00)	22.35 (8.68)	23.24 (9.00)	27.94 (13.47)	41.18 (21.40)	26.88 (12.11)	5-75
Right Ear	Younger adults	22.78 (5.48)	18.89 (6.08)	14.72 (3.63)	12.22 (3.92)	7.78 (5.48)	15.28 (4.92)	0-35
	Older adults	21.47 (6.32)	22.94 (6.39)	25.88 (9.72)	27.65 (13.36)	33.24 (19.92)	26.24 (11.14)	5-75

To investigate whether hearing loss was associated with emotion recognition performance in the older adult group, simple bivariate correlations were conducted between hearing thresholds (in the better ear) and percent correct scores for each emotion type. Table

3.5 presents the Pearson correlations and indicates that for the older adult group, no significant correlations were found between hearing level and percent correct scores.

Table 3.5. Pearson correlations for the hearing thresholds and the percent correct scores across each emotion type for the older adult group.

		Frequency (kHz)				
		0.25	0.5	1	2	4
Percent Correct Scores	Angry	-.32	-.28	-.10	-.09	-.02
	Disgust	.001	-.03	-.11	-.11	.20
	Sad	-.23	.02	.003	-.17	-.37
	Fear	-.07	-.08	-.14	-.14	-.17
	Surprise	.34	.18	-.003	-.23	-.003
	Happy	-.14	.04	.06	-.22	-.16

Separate bivariate correlations were also conducted between hearing level and the number of fixations and hearing level and the duration of fixations (see Table 3.6). There were no significant correlations between hearing level and duration or number of fixations.

Table 3.6. Pearson correlations for the hearing thresholds and the duration of fixations (left) and the number of fixations (right) across each emotion type for the older adult group.

		Duration of fixations					Number of fixations				
		Frequency (kHz)					Frequency (kHz)				
		0.25	0.5	1	2	4	0.25	0.5	1	2	4
Angry	Eye region	-.28	.21	.19	.24	-.10	-.30	.13	.26	.31	-.12
	Mouth region	.13	-.22	-.23	-.15	.05	.20	-.21	-.21	-.16	-.06
Disgust	Eye region	-.35	.17	.19	.23	-.09	-.37	.07	.23	.29	-.11
	Mouth region	.16	-.20	-.15	-.08	.09	.22	-.19	-.15	-.12	-.06
Sad	Eye region	-.29	.18	.21	.26	-.07	-.30	.09	.26	.32	-.07
	Mouth region	.16	-.20	-.23	-.17	-.02	.22	-.16	-.20	-.18	-.09
Fear	Eye region	-.30	.22	.24	.25	-.13	-.32	.11	.26	.30	-.12
	Mouth region	.20	-.18	-.22	-.17	.01	.23	-.16	-.20	-.19	-.05
Surprise	Eye region	-.35	.16	.19	.24	-.14	-.38	.06	.22	.29	-.15
	Mouth region	.18	-.16	-.21	-.20	.04	.21	-.15	-.22	-.28	-.13
Happy	Eye region	-.45	.06	.13	.18	-.15	-.44	.01	.19	.26	-.15
	Mouth region	.19	-.13	-.22	-.20	-.01	.21	-.12	-.19	-.21	-.14

Table 3.7 presents the hearing thresholds for each participant in the older adult group along with the number and duration of fixations, and percent correct scores for comparison purposes. In this table, the data of older adults has been sorted in ascending order according to their average hearing levels. The finding that hearing ability is not associated with emotion recognition performance or eye gaze measures (i.e., number and duration of fixations) is reflected in Table 3.7. That is, performance and gaze behaviour scores do not appear to ascend in line with average hearing level scores.

Table 3.7. Hearing level (dB) in the better ear across 0.25, 0.5, 1, 2, and 4 kilohertz. Each row represents a single older adult and has been ordered according to average hearing threshold. For comparison purposes average scores for percent correct and number and duration of fixations (in milliseconds) for each face region are also presented.

Frequency (kHz)						Number of fixations		Duration of fixations		Emotion recognition scores (%)
0.25	0.5	1	2	4	Average (0.25-4)	Eye region	Mouth region	Eye region	Mouth region	
20	15	15	15	10	15	4.8	1.4	1176.6	325.0	61.8
15	15	20	15	5	14	2.4	3.1	640.0	1497.8	61.4
20	15	20	20	15	18	.8	5.6	236.2	1919.2	64.3
15	15	20	25	20	19	9.0	1.9	1895.5	636.4	63.2
5	15	25	30	20	19	9.5	.7	2649.5	151.5	67.9
15	20	15	20	30	20	4.8	1.7	1771.6	774.7	65.7
20	20	25	10	25	20	1.4	2.5	445.7	763.5	69.6
25	20	20	10	30	21	.2	5.0	39.6	2271.0	72.1
20	15	20	25	30	22	1.5	1.0	484.0	313.7	69.6
25	30	25	10	30	24	3.5	.4	1435.4	115.2	50.0
20	25	15	30	40	26	5.7	.6	2284.8	144.4	59.3
20	30	35	20	30	27	8.2	1.9	2487.0	392.2	83.6
15	15	20	30	60	28	.5	4.2	141.9	2408.2	55.7
25	25	25	25	55	31	.6	2.1	210.4	811.7	65.0
20	20	35	45	65	37	9.1	.2	2019.1	54.2	73.6
40	35	40	40	35	38	3.5	3.3	994.3	1255.1	55.4
15	25	40	55	70	41	2.7	2.3	892.8	755.0	49.3

3.5 Discussion

Our findings confirmed past research that showed that for static expressions of emotion, older adults make more fixations to and look longer at the mouth region of faces compared to younger adults (Sullivan et al., 2007; Chaby et al., 2017). We also found that younger adults looked more at the eye than the mouth region of emotional faces whereas this preference was weaker for the older adult group. These findings are in contrast to those reported in Chaby et al. (2017) that younger adults looked more at the upper face half for some emotions (angry and sad), at the lower face half for disgust, and showed no differences for the remaining emotion types (fear, joy, neutral); and that older adults showed a preference for the lower relative to the upper face region.

One methodological difference that could account for these different findings relates to the size of the areas of interest. That is, relative to the areas of interest in the current study, the areas of interest were larger in the study by Chaby et al. (2017) as they included the whole upper (i.e., from the top of forehead to middle of nose) and lower (i.e., from middle of nose to the bottom of the chin) face. Given that the mouth region contains important emotion cues for disgust and joy recognition (as well as cues to fear recognition, see Ekman, Friesen, & Hager, 2002) the younger adults in Chaby et al. (2017) may have looked at the tip of the nose (included in the lower face region) to simultaneously perceive emotion information in the mouth and eye region. Longer gazing at the nose may explain why younger adults showed similar (fear, joy) or more (disgust) looking at the lower half of faces and why such results were not found for the current study.

Despite having found age differences in the number and duration of fixations for the mouth region, the pattern of mouth looking across the different static emotion types was remarkably similar for both age groups. For instance, even though fixation times differed across the emotion types, for each emotion, the older adults consistently made about one additional fixation to the mouth region of faces than did younger adults. Further, for each emotion, the older adults looked at the mouth region of emotional faces for approximately 400ms longer than younger adults did. This finding is in line with the results of Sullivan et al. (2007) except that they found an age difference of 700ms (possibly due to differences in experimental setup).

In addition to presenting static expressions of emotion, the current study also investigated gaze behaviour for dynamic expressions of spoken emotion. For both age groups, gaze behaviour for the dynamic expressions was similar to that found for static expressions. However, there was one difference, both age groups made more fixations to and gazed longer at the mouth and eye regions of static relative to dynamic expressions (consistent with the results of Blais, Fiset, Roy, Saumure Régimbald, & Gosselin, 2017). It is possible that for dynamic expressions, the participants' gaze was drawn away from the eye and mouth regions of faces because emotion-relevant information can rapidly appear and disappear in the eye and mouth region. In order to simultaneously perceive movements in both face regions, the participants in this study may have shifted their gaze to a more central location, such as the nose. In contrast, static expressions may produce more gazing at the eye and mouth face regions as emotion information in these regions does not change and can be revisited. Taken together, these findings suggest that gaze behaviour can be influenced by format type (i.e., static, dynamic). That is, both older and younger adults may adapt their gaze strategies to perceive the most amount of emotion information from the presented expression.

It is important to note that gender was not balanced in the younger adult group (15 women, 3 men). This gender imbalance may not be problematic given that previous research indicates that for young adults, visual scanning does not differ between men and women (Sullivan et al., 2017). Similarly, we did not find gender differences in gaze behaviour for the older adult group. The finding that there were no gender differences in the older adults' group should be taken with caution as the sample size in the current study is relatively small. This may explain the difference in findings between the current study and Sullivan, Campbell, Hutton, & Ruffman, 2017, that reported that older men gazed more at the mouth region than older women.

With regards to the issue concerning the cause of older adults' observed gaze behaviour, the current findings suggest hearing loss is unlikely to explain older adults' tendency to look at the mouth when judging the emotion of a face. This is because gaze behaviour or recognition accuracy did not appear to be significantly associated with hearing ability. However, this finding should be treated with caution as some of the group sizes for this analysis were small and different from one another. A post-hoc power analysis revealed that a sample size of approximately 29 (slightly larger than the sample size of the current study) is required to obtain statistical power at the recommended .80 level.

With regards to the issue that the consequence of older adults' gaze behaviour is reduced emotion recognition performance, we suggest that the decline in older adults' emotion recognition is not simply due to their gaze behaviour patterns. One reason is that we did not find significant associations between gaze behaviour and emotion recognition performance (as mentioned previously, these findings should be treated with caution because of the small sample size). That is, although we found that overall, emotion recognition performance was worse for the older than the younger adult group and that older adults showed more mouth gazing than younger adults, the scores on these variables were not significantly associated. This finding is in line with some studies (Murphy & Isaacowitz, 2010; Sullivan et al., 2007), but in contrast to others, e.g., Wong et al. (2005) who reported significant associations between gaze and recognition performance for a number of emotions. One difference between the current study and the study by Wong et al. (2005) is that the current study measured gaze behaviour during an emotion recognition task whereas Wong et al. (2005) measured gaze behaviour for emotional faces in a separate and subsequent study to their emotion recognition task. As outlined in the introduction, any association or correlation between separate tasks should be treated with caution. Instead, measuring gaze behaviour during emotion recognition may be a more appropriate way to assess the relationship between these two variables.

There are a number of ways to interpret the finding that older adults' gaze behaviour was not related to their recognition performance. One possibility is that a connection between gaze behaviour and recognition performance was obscured because the face stimuli were presented for longer than was necessary to recognise emotion. Specifically, we presented emotional faces for a relatively long duration (approximately 3 to 4 seconds) such that participants may have recognised the presented emotion well before the time limit was reached. Having extra time to gaze at the emotional expression may have inflated the age differences in gaze behaviour as older and younger adults may have continued to gaze at the expressions in different ways. This additional looking post recognition may have obscured an association between emotion recognition performance and the measures of gaze behaviour (i.e., number of fixations, duration of fixations).

An alternative possibility is that the relationship between gaze behaviour and recognition performance is influenced by how well task-relevant information can be picked up from a region. That is, when emotion information is not clear from the eye region, the observer may adopt a strategy of exploring both face regions for task-relevant emotion information. If it is assumed that older adults have more difficulty in picking up emotion relevant information, then they may adopt an exploratory-gaze strategy whereas younger adults may adopt a focused-gaze strategy during emotion recognition. This suggestion is the opposite of that made by Chaby et al. (2017) but appears to be consistent with the findings of the current study (that it was the younger adults that showed a distinct bias in where they gazed).

The idea that older adults adopt an exploratory-gaze strategy when faced with uncertainty about the expressed emotion can be investigated by testing the gaze behaviour of older adults for other types of facial expressions that may require more eye than mouth gazing for accurate recognition. If older adults show more balanced gaze behaviour than younger adults for such expressions, then this would provide support for the idea that older adults adopt an exploratory-gaze behaviour as they may not realise that a biased looking strategy is more appropriate for emotion recognition. If older adults engage in a more effective gaze strategy (i.e., gaze towards the eye region) as younger adults do, then mouth gazing would seem to be more specifically related to categorising emotional faces. This issue warrants further investigation.[1]

[1] An investigation into older adults' gaze behaviour for non-emotional expressions was conducted in a study presented as Chapter 4.

3.6 Appendix A

Digit Span (i.e., Digit Span Forward, Backward, and Sequencing), Trail Making Tests, and the Centre for Epidemiological Studies-Depression Scale (WAIS-IV; Borson et al., 2000; Radloff, 1977) were administered to younger and older adults to check for possible covariates.

The results of the cognitive tests are summarised in Table 3.8. Separate independent *t* tests were conducted on the scores for each of the measures. The younger group were faster than the older group at completing the Trail Makings Tests indicating that the processing speed of younger adults was faster than that of older adults. There were no age differences on the depression scale or on the task measuring working memory (Digit Span).

To test for possible covariates, correlations between recognition scores for each of the emotion types and scores on the Trail Making Tests were examined. For younger adults, recognition scores were negatively correlated with scores on the Trail Making Test Part B for angry faces only, $r(16) = -.49$, $p < .05$. There were no such significant correlations for the remaining emotion types ($p > .07$). Older adults' recognition scores were not correlated with scores on either of the Trail Making Tests ($p > .12$).

Table 3.8. Younger and older adults' mean scores and standard deviations on the Cognitive measures.

	Younger adults		Older adults			
Task	*M*	*SD*	*M*	*SD*	*df*	*t*
Trail Making Test Part A (seconds)	15.78	3.08	27.59	8.41	33	-5.58*
Trail Making Test Part B (seconds)	37.72	4.96	59.47	22.25	33	-4.05*
Digit Span Raw Score	25.61	4.65	25.00	4.26	33	.41
CES-D	12.06	7.53	10.59	7.78	33	.57

Note: CES-D = Centre for Epidemiological Studies-Depression Scale
*$p < .001$

3.7 References

Bach, M. (2007). The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology, 245*(7), 965-971.

Blais, C., Fiset, D., Roy, C., Saumure Régimbald, C., & Gosselin, F. (2017). Eye fixation patterns for categorizing static and dynamic facial expressions. *Emotion, 17*(7), 1107-1119.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: a cognitive‘vital signs’ measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15*(11), 1021-1027.

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2*(1), 1-13.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527-551.

Chaby, L., Hupont, I., Avril, M., Luherne-du Boullay, V., & Chetouani, M. (2017). Gaze behavior consistency among older and younger adults when looking at emotional faces. *Frontiers in Psychology, 8*, 548.

Circelli, K. S., Clark, U. S., & Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in aging. *Aging, Neuropsychology, and Cognition, 20*(2), 148-173.

Davis, C., & Kim, J. (2006). Audio-visual speech perception off the top of the head. *Cognition, 100*(3), B21-B31.

Dowiasch, S., Marx, S., Einhäuser, W., & Bremmer, F. (2015). Effects of aging on eye movements in the real world. *Frontiers in Human Neuroscience, 9*, 46.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system. *Network Research Information, Salt Lake City, UT, USA*.

Grainger, S. A., Henry, J. D., Phillips, L. H., Vanman, E. J., & Allen, R. (2017). Age deficits in facial affect recognition: The influence of dynamic cues. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 72*(4), 622-632.

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Selective preference in visual fixation away from negative images in old age? An eye-tracking study. *Psychology and Aging, 21*(1), 40.

Itier, R. J., Villate, C., & Ryan, J. D. (2007). Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia, 45*(5), 1019-1028.

Kim, J., & Davis, C. (2012). Perceiving emotion from a talker: How face and voice work together. *Visual Cognition, 20*(8), 902-921.

Klein, C., Fischer, B., Hartnegg, K., Heiss, W., & Roth, M. (2000). Optomotor and neuropsychological performance in old age. *Experimental Brain Research, 135*(2), 141-154.

Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, and Hearing Research, 42*(3), 526-539.

Murphy, N. A., & Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognising emotional expressions: An in-depth look at gaze measures and covariates. *Cognition and Emotion, 24*(3), 436-452.

Radloff, L. S. (1977). The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Rosen, H. J., Wilson, M. R., Schauer, G. F., Allison, S., Gorno-Tempini, M.-L., Pace-Savitsky, C., . . . Miller, B. L. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia, 44*(3), 365-373.

Rothauser, E. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17*, 225-246.

Schurgin, M., Nelson, J., Iida, S., Ohira, H., Chiao, J., & Franconeri, S. (2014). Eye movements during emotion recognition in faces. *Journal of vision, 14*(13), 14-14.

Sullivan, S., Campbell, A., Hutton, S. B., & Ruffman, T. (2017). What's good for the goose is not good for the gander: Age and gender differences in scanning emotion faces. *The Journals of Gerontology: Series B, 72*(3), 441-447.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology: Series B, 62*(1), 53-60.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6*(4), 473-482.

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology, 19*(6), 739-749.

Chapter 4. Gaze Behaviour Consistency Among Older And Younger Adults When Recognising Visual Prosody

4.1 Abstract

During emotion recognition, younger adults gaze at the eye region more often than the mouth region presumably because the eye region conveys more task-relevant information. Given the importance of the eye region for emotion recognition, it is unclear why older adults look at the eye and mouth regions a similar number of times during emotion recognition (See Chapter 3). One proposal is that in general, older adults are not able to evaluate the degree of information content in each face region and so, adopt a general strategy of exploring both face regions equally. The current study tested this proposal by measuring gaze behaviour during the recognition of non-emotional expressions (i.e., a prosody recognition task) where the eye region provides more task-relevant information than the mouth region. As such, the appropriate gaze behaviour for such expressions would be to gaze at the eye region. Older (n = 17) and younger (n = 18) adults completed a forced-choice recognition task with questioning, narrowly focused, and broad focused expressions whilst their eye movements (i.e., number of fixations, duration of fixations) were recorded. There were no age differences for gaze behaviour; both age groups showed more eye than mouth gazing across all expression types. However, younger adults outperformed older adults on the recognition task. These findings indicate that older adults are capable of adopting an appropriate gaze strategy when recognising facial expressions but that adopting this strategy does not in itself lead to equivalent performance levels as young adults.

4.2 Introduction

Compared to younger adults, older adults display a different pattern of eye movements when they look at faces to determine what emotion is being expressed (see Chapter 3). That is, older adults tend to look at the eye and mouth region a similar number of times whereas younger adults look at the eyes more often than the mouth. This difference between young and older adults is typically described in the literature as older adults showing a bias for looking the mouth region (e.g., Chaby, Hupont, Avril, Luherne-du Boullay, & Chetouani, 2017). However, in Chapter 3 of this thesis, another way of thinking about this difference was proposed. Rather than view the older adult pattern as a bias, it was suggested that it is the younger adults that show the bias (i.e., more uneven looking patterns between the two regions). It was argued that younger adults show this pattern because this behaviour is appropriate for the emotion recognition task i.e., a bias to gaze at the eye region makes sense as this region conveys cues relevant to the classification of more emotions.

It is unclear why older adults do not show a bias to gaze at the eye region during emotion recognition. It could be due to changes in the way that older adults evaluate emotion information. For instance, older adults may not be able to pick up emotion information from this region and so, may search the mouth region for emotion information. Or it could be that, in general, older adults are unable to exploit regional differences in information content and instead adopt a general strategy that roughly equal looks/time should be given to all information containing regions.

To investigate whether the latter is the case, the current experiment investigated the gaze patterns of older and younger adults on a non-emotion processing task that involved using information from the eye and the mouth regions. We specifically selected the expression of linguistic prosody namely, narrowly focused (i.e., a word in a phrase is stressed) and echoic questioning (i.e., a phrase is uttered as a question). These expression types were presented to older and younger adults in a forced-choice recognition task. Importantly, in this recognition task, the eye region provides more task-relevant information than the mouth region (Lansing & McConkie, 1999). That is, although task-relevant information can be gleaned from the lower face (particularly for narrowly focused expressions), overall, task-relevant information is most likely conveyed via the eye rather than the mouth region. As such an appropriate gaze behaviour would be to look more at the eyes than the mouth region (as the younger adults showed in Lansing & McConkie, 1999).

The key issue of the present study is whether older adults would show this gaze pattern. If older adults do show more eye than mouth gazing, then this may suggest that age-related changes to gaze behaviour may be specifically related to categorising emotional faces. If older adults do not show this gaze behaviour, and instead, show a more equal eye and mouth gazing pattern, it would suggest that older adults are unable to determine the face region with the most relevant information and so, may adopt a general gaze strategy.

4.3 Method

4.3.1 Participants

The participants recruited for this study were the same as those recruited for our eye-tracking study with emotional faces (Chapter 3).[2] Eighteen young undergraduate students (3 males, M_{age} = 21, range = 18-33) of Western Sydney University participated in the experiment for course credit. Seventeen older adults (8 females, M_{age} = 70, range = 60-79) from the community participated in this study and received monetary reimbursement. All reported to have learnt English at a young age (i.e., less than 7 years old) and had visual normal acuity except one older adult who had slightly worse than normal vision.

[2] The order of these experiments was counterbalanced across participants so that half of participants completed the emotion recognition experiment first and the other half completed the linguistic experiment first.

Participants were screened for dementia using the Mini-Cog™ test (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000) as the presence of dementia has been associated with poor expression recognition and may act as a confounding variable (Rosen et al., 2006). All older (M = 4) and younger (M = 4) adults scored negative for cognitive impairment (where a score of at least 3 out of 5 points represents normal cognition).

4.3.2 Stimuli

Stimuli consisted of static picture and dynamic visual-only recordings of linguistic expressions by two females (M_{age} = 21.5). These presenters were native talkers of Australian English and produced varying linguistic expressions (i.e., narrow focus, question, and broad focus) while uttering sentences one by one.

The dynamic recordings consisted of thirteen IEEE sentences (Rothauser, 1969) read aloud and conveying the required expression. Eleven of these sentences were used throughout the experiment as experimental items (10 sentences) or practice items (1 sentence). The remaining two sentences were used as a presenter specific calibration for participants throughout the experiment.

The video recordings of the two presenters were captured individually in a sound attenuated booth. The presenter was seated in front of a monitor that displayed each sentence one at a time. The video camera (Sony NCCAM HXR-NX30p) was situated directly above the monitor and captured video at 1920 x 1080 full HD resolution at 50 frames per second.

The recording session was blocked by the type of expression. Each block contained eleven sentences, resulting in 33 recorded sentences (3 expression types x 11 sentences) for each of the two presenters. Prior to each block, presenters were instructed which expression was required to be produced. For the broad focus or “neutral” expressions, the presenter was directed to read the sentence silently and then to look into the camera lens as they uttered the sentence out loud. In conjunction with these recordings, an additional two sentences with a neutral expression were also recorded (for presenter specific calibration).[3]

To produce naturalistic versions of the narrowly focused and questioning expressions, presenters were required to interact with an interlocutor (as in Cvejic, Kim, & Davis, 2010). Within each sentence, a word (e.g., “HOLE” in examples 1 and 2) received a narrow focus (example 1) or questioning (example 2) intonation.

For a sentence with a narrow (corrective) focus, presenters were required to correct an error made by the interlocutor (example 1).

Example 1

(a) Interlocutor: “A round **POLE**[error] was drilled through the thin board?”

(b) Presenter: “A round **HOLE**[focused] was drilled through the thin board.”

For an echoic question, presenters were required to question a narrowly focused statement produced by the interlocutor (example 2).

Example 2

(a) Interlocutor: “A round **HOLE**[focused] was drilled through the thin board.”

(b) Presenter: “A round **HOLE**[question] was drilled through the thin board?”

To produce the dynamic visual-only stimuli, all video recordings were stripped of audio, cropped to include just the head area (i.e., width = 600 pixels, height = 850 pixels), and segmented into individual sentences using a custom MATLAB script. The static visual stimuli were created using the dynamic stimuli. That is, a single rater selected one frame from each dynamic emotional expression that they judged to be of the peak expression (i.e., the frame that most clearly conveyed the intended expression). Figure 4.1 provides examples of the static stimuli for each expression type.

[3] All neutral expressions described in the current experiment were also used in the experiment presented in Chapter 3.

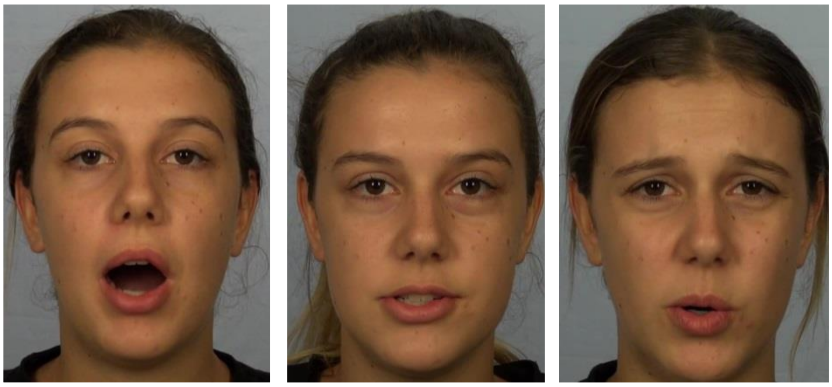

Figure 4.1. Examples of the narrowly focused (left), neutral (middle), and echoic questioning (right) expressions in the static format.

4.3.3 Experimental design

Altogether, there were 120 trials consisting of 3 linguistic expressions (i.e., focus, question, neutral), 2 format types (static, dynamic), 2 presenters, and 10 sentences. Trials were blocked by presenter, resulting in two blocks of 60 trials. Within each block, participants viewed the expressions in one format (e.g., static) before viewing the expressions in the other format (e.g., dynamic). The presentation order of the presenter blocks as well as the format types was counterbalanced across participants. All trials of the same format (i.e., static, dynamic) within each presenter block were randomised and presented using the SR Research Experiment Builder software.

4.3.4 Procedure

Participants were tested individually in a quiet room. They were informed that the eye movements of the right eye would be recorded during the experiment using an eye tracker (SR EyeLink® 1000) that was mounted on a tower with a chin and head rest. They were instructed to rest their heads on the head and chin rest to limit head movements that might interfere with the recording of eye movements during the experiment.

Participants were told that they would see still images or silent videos of a person conveying a linguistic expression and that their task was to identify the expression by using the mouse to click one of a three labels presented on the screen (i.e., “Focus”, “Neutral”, “Question”). Labels were presented in alphabetical order in a single row across the screen, with the “Neutral” label presented in the centre of the screen. The experimenter explained to the participant that each label referred to a particular type of expression. For instance, participants were told to click the “Focus” label when they thought the presenter was placing emphasis (or stress) on a word in an utterance. Similarly, they were told to click the “Question” label if they thought the presenter was asking a question. Once participants indicated that they understood the instructions, the experiment began.

At the beginning of the experiment, right eye positions were calibrated using the SR EyeLink® 1000. For each new presenter and format type (i.e., static, dynamic), participants were shown two neutral expressions of that presenter and format type to act as a presenter specific calibration. Participants were told to use these neutral expressions as a baseline to compare with subsequently presented expressions. Following this, three practice items were presented to give participants the opportunity to be exposed to and identify each of the expression types from that format type and presenter. This was then followed by the experimental items.

To initiate the presentation of an experimental item, participants were required to gaze at a fixation box for one second. The fixation box was presented in the same location as the centre of the subsequently presented face. This was to ensure that all initial eye fixations were the same across all items and participants. Once the trial was triggered, participants were presented with a silent video or an image for approximately 3 to 4 seconds (depending on the length of the video). During this time, eye gaze movements were recorded. Three labels were then presented for a response. Participants were given a break after every fifteen trials. Eye calibration was performed after each break. At the conclusion of the study, participants were debriefed as to the purpose of the study.

4.3.5 Data analysis

The data analysis used in the current experiment was identical to the analysis used in the previous eye-tracking experiment (Chapter 3). That is, using the SR Research DataViewer software, two areas of interest were created for each experimental item: one area included the eyes and the eyebrows only, and the other the mouth only (see Figure 4.2). For dynamic items, similar dynamic areas of interest were used and these were inspected to ensure that they included the appropriate face areas for the length of each video. The number of fixations and the duration of fixations to each of the areas of interest was extracted using the DataViewer software and were analysed further.

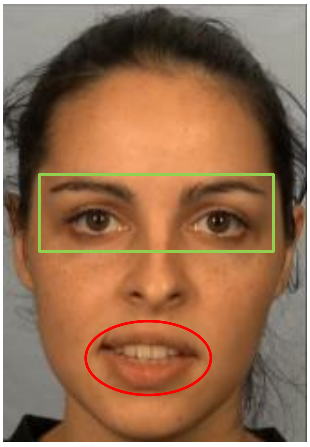

Figure 4.2. Interest areas for the eye (green box) and mouth (red oval) face areas.

4.4 Results

4.4.1 Accuracy scores

Overall identification scores for the linguistic task are presented in Figure 4.3. The ANOVA revealed that overall, younger adults (90%) outperformed older ones (77%), $F(1,33) = 6.11$, $p < .02$, $\eta p2 = .16$, the dynamic format (86%) attracted higher percent correct scores than the static format (81%), $F(1,33) = 5.87$, $p < .05$, $\eta p2 = .15$, and the neutral (89%) were better recognised than the focused (82%) or questioning (80%) expressions, $F(2,66) = 5.81$, $p < .01$, $\eta p2 = .15$. Linguistic expressions from one presenter (86%) were better recognised than the expressions of the other presenter (81%), $F(1,33) = 11.72$, $p < .01$, $\eta p2 = .26$. There were no significant two- or three-way interactions between format type and age group, $F(1,33) = .78$, $p = .38$, ηp2 = .02, expression type and age group, $F(2,66) = .27$, $p = .76$, ηp2 = .008, presenter and age group, $F(1,33) = 1.66$, $p = .21$, ηp2 = .05, and format type, expression type, and age group, $F(2,66) = 1.47$, $p = .24$, ηp2 = .04, format type, presenter, and age group, $F(1,33) = .50$, $p = .48$, ηp2 = .02, and expression type, presenter, and age group, $F(2,66) = .10$, $p = .91$, ηp2 = .003. There was also no significant interaction between age group, format type, expression type, and presenter, $F(2,66) = .96$, $p = .39$, ηp2 = .03.

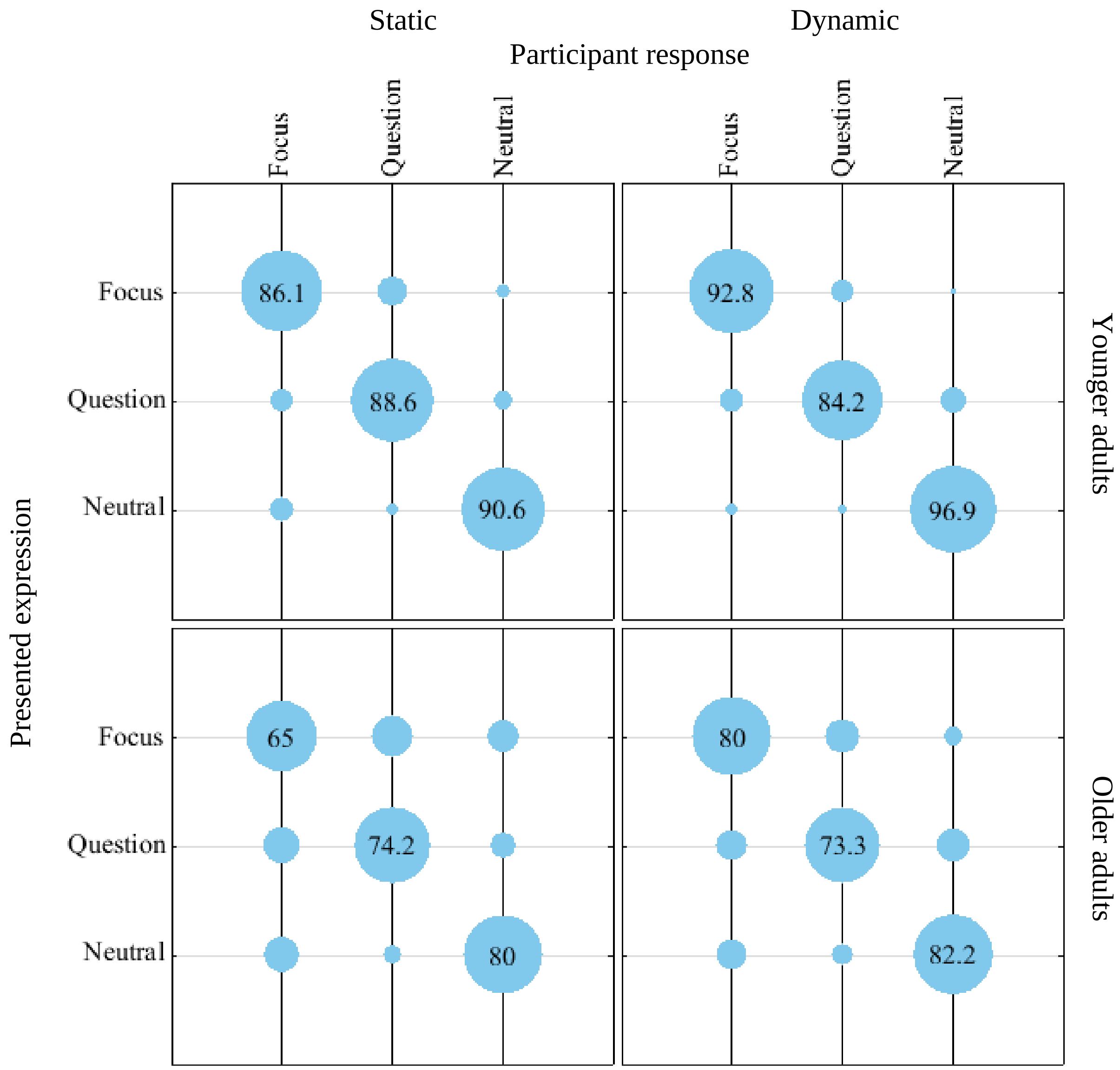

Figure 4.3. Static (left) and dynamic (right) response matrices for the younger (top) and older (bottom) adult groups. The presented linguistic expressions are shown on the vertical and the responses to these are shown at the top, horizontally. Response magnitude is represented by the size of the circles; the numbers presented on the diagonal have been added as an indication of scale and indicate the percentage of times that a response was selected for a presented linguistic expression.

4.4.2 Looking patterns

The overall number of fixations are presented in Figure 4.4. The number of fixations made to the faces was similar across the younger (2.9) and older (2.5) groups, $F(1,33) = 1.00$, $p = .32$, ηp2 = .03. Static faces (3.4) attracted a higher number of fixations than dynamic ones (2.0), $F(1,33) = 105.23$, $p < .001$, ηp2 = .76. Overall, the eye region (3.7) of faces showed more fixations than the mouth region (1.7), $F(1,33) = 17.38$, $p < .001$, ηp2 = .35. The main effect of expression was not significant, $F(2,66) = 1.73$, $p = .19$, ηp2 = .05. The presenter that attracted higher percent correct scores also attracted a higher number of fixations (2.9) than the other presenter (2.6), $F(1,33) = 14.41$, $p < .01$, ηp2 = .30. There were no significant interactions between format type and age group, $F(1,33) = .12$, $p = .74$, ηp2 = .004, face region and age group, $F(1,33) = 2.91$, $p = .10$, ηp2 = .08, expression type and age group, $F(2,66) = 2.81$, $p = .07$, ηp2 = .08, presenter and age group, $F(1,33) = .000192$, $p = .99$, ηp2 = .000006, format type, face region, and age group, $F(1,33) = 1.92$, $p = .18$, ηp2 = .06, format type, expression type, and age group, $F(2,66) = .63$, $p = .54$, ηp2 = .02, format type, presenter, and age group, $F(1,33) = .51$, $p = .48$, ηp2 = .02, face region, expression type, and age group, $F(2,66) = 1.49$, $p = .23$, ηp2 = .04, face region, presenter, and age group, $F(1,33) = 1.69$, $p = .20$, ηp2 = .05, and expression type, presenter, and age group, $F(2,66) = .66$, $p = .52$, ηp2 = .02. Finally, the interaction between age group, format type, face region, expression type, and presenter was also not significant, $F(2,66) = 1.26$, $p = .29$, ηp2 = .04.

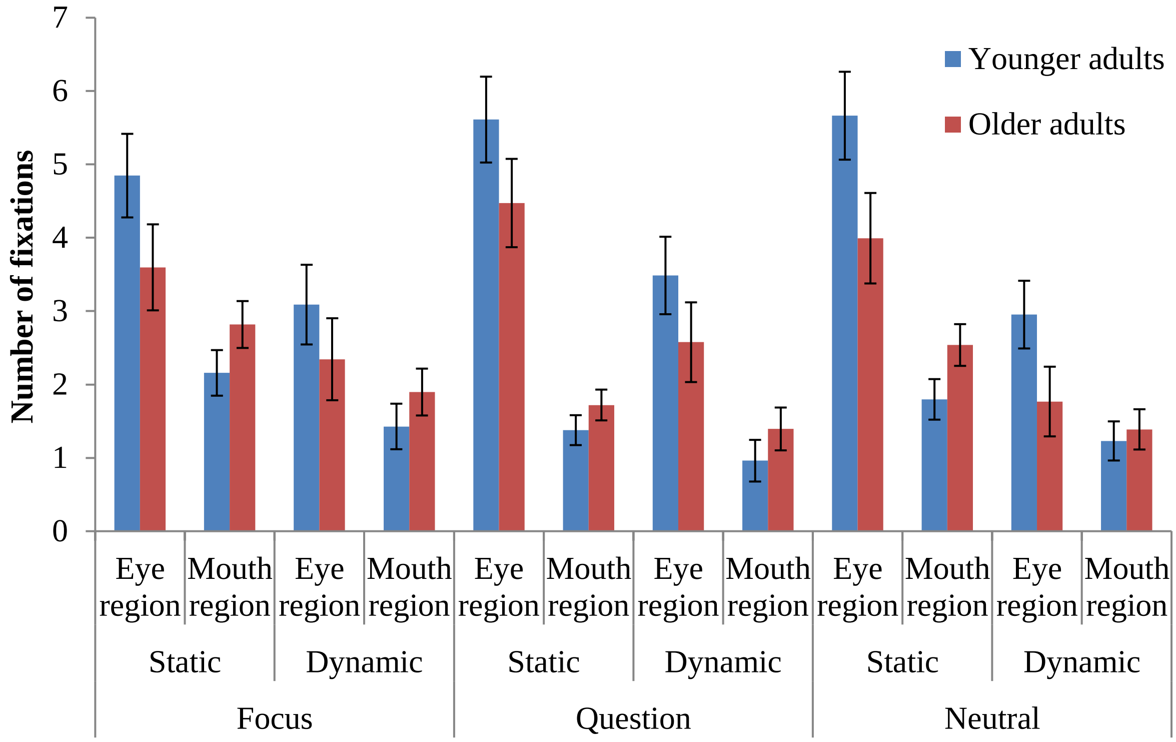

Figure 4.4. Older and younger adults' mean number of fixations for each face region across the static and dynamic linguistic expressions. The error bars represent standard error.

The overall duration of fixation scores are presented in Figure 4.5. For the duration scores expressions, younger (963ms) and older (866ms) adults showed similar amounts of gazing at the eye and mouth regions of the linguistic expressions, $F(1,33) = .98$, $p = .33$, ηp2 = .03. Static faces (989ms) attracted longer gaze times than dynamic faces (840ms), $F(1,33) = 9.33$, $p < .01$, ηp2 = .22. Participants showed greater amounts of gazing to the eye (1139ms) than the mouth (690ms) regions, $F(1,33) = 5.60$, $p < .05$, ηp2 = .15. Further, gaze duration appeared to be consistent across the focused (919ms), questioning (890ms), and neutral (934ms) expression types, $F(2,66) = 2.04$, $p = .14$, ηp2 = .06. Participants spent more time looking at the facial expressions of the clearest presenter (963ms) relative to the less clear presenter (866ms), $F(1,33) = 9.64$, $p < .01$, ηp2 = .23. There were no significant two- or three-way interactions between format type and age group, $F(1,33) = .06$, $p = .82$, ηp2 = .002, face region and age group, $F(1,33) = 2.95$, $p = .10$, ηp2 = .08, expression type and age group,

$F(2,66) = .46$, $p = .64$, ηp2 = .01, presenter and age group, $F(1,33) = .11$, $p = .75$, ηp2 = .003, format type, face region, and age group, $F(1,33) = .000232$, $p = .99$, ηp2 = .000007, format type, expression type, and age group, $F(2,66) = .80$, $p = .45$, ηp2 = .02, format type, presenter, and age group, $F(1,33) = .63$, $p = .43$, ηp2 = .02, face region, expression type, and age group, $F(2,66) = 1.15$, $p = .32$, ηp2 = .03, face region, presenter, and age group, $F(1,33) = 1.80$, $p = .19$, ηp2 = .05, and expression type, presenter, and age group, $F(2,66) = .32$, $p = .73$, ηp2 = .009. There was no significant interaction between age group, format type, face region, expression type, and presenter, $F(2,66) = .97$, $p = .39$, ηp2 = .03.

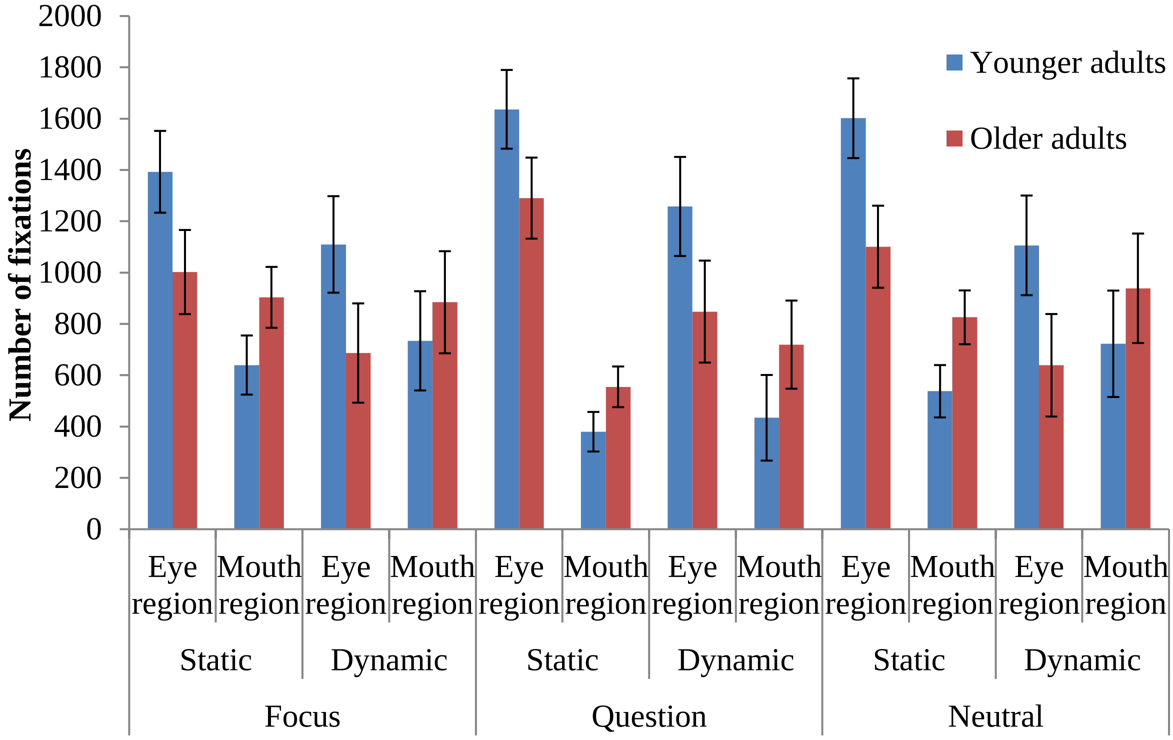

Figure 4.5. Older and younger adults' mean duration (in milliseconds) of fixations for each face region across the static and dynamic linguistic expressions. The error bars represent standard error.

4.5 Discussion

We examined younger and older adults' gaze behaviour for linguistic expressions. Overall, we found that both younger and older adults showed more eye than mouth gazing during linguistic recognition. Further, the gaze patterns for the expressions (with respect to the eye and mouth region) did not differ. This finding indicates that, like younger adults, older adults are capable of engaging in an appropriate gaze strategy for a face expression recognition task.

If older adults are capable of engaging in an appropriate gaze strategy during linguistic expression recognition, then why did older adults not show such a strategy (i.e., more eye than mouth gazing) for emotion recognition? As mentioned, one potential reason is that during emotion recognition, older adults may not be able to effectively evaluate the benefit of looking longer at the eye region (possibly because older adults cannot extract emotion information from the eye region, e.g., Sullivan, Ruffman, & Hutton, 2007) and so, look elsewhere.

An alternative proposal is that older adults may have difficulty learning the face region that conveys the most task-relevant information when the task contains many expression types where emotion information is distributed across the face regions. That is, the emotion recognition task presented six expressions where task-relevant information was distributed across the eye and mouth regions. For example, angry, sad, and fear, are best recognised from the eyes whereas happy and disgust are best recognised from the mouth region (Calder, Young, Keane, & Dean, 2000). Given this variability, older adults may have adopted a balanced looking strategy to get the best chance of picking up emotion-related information that can be used in making the classification judgement. In contrast, the linguistic recognition task presented only two expressions and the eye region provided important cues for both types of expressions. When given this simple task, older adults may have been able to discern an appropriate gaze behaviour strategy.

To further investigate the factors that modulate older adults' gaze behaviour during emotion recognition, future research could investigate gaze behaviour during the recognition of only two emotions where the eye region provides more task-relevant information than the mouth region. This setup closely mimics the setup of the linguistic recognition task and can test the impact that task constraints may have on gaze behaviour. If older adults show a biased gaze strategy (i.e., more eye than mouth gazing) with only two emotional expressions, then it would suggest that older adults can adopt an appropriate gaze strategy when this strategy is easily realised (i.e., task-relevant information is best recognised from the same face region across the expressions). If they show a more balanced gaze strategy, then this may indicate that older adults do not realise that a biased looking strategy is better than a balanced looking strategy (possible because of poor emotion recognition in the eye region).

With regards to the recognition scores, younger adults outperformed older adults across all expression types. This occurred even though older adults showed the same gaze behaviour as younger adults, suggesting that this gaze strategy does not in itself result in similar performance as younger adults. It is not clear why younger adults showed better recognition than older adults. One simple reason could be that general aging of cognitive abilities impacted older adults' ability to complete this task. For instance, older adults are less able than younger adults to ignore task-irrelevant information and to inhibit inappropriate responses (Healey, Campbell, & Hasher, 2008). Thus, deficits in these skills may have impeded their ability to accurately recognise linguistic expressions. This warrants further investigation.

It is important to note that the present study may be limited in terms of statistical power because of the small sample sizes in the older (n = 17) and younger (n = 18) adult groups. Regardless of this limitation, this study reflects a preliminary but important step towards interpreting the gaze behaviour of older adults for emotion faces.

4.6 References

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: a cognitive‘vital signs’ measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15*(11), 1021-1027.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527-551.

Chaby, L., Hupont, I., Avril, M., Luherne-du Boullay, V., & Chetouani, M. (2017). Gaze behavior consistency among older and younger adults when looking at emotional faces. *Frontiers in Psychology, 8*, 548.

Cvejic, E., Kim, J., & Davis, C. (2010). Prosody off the top of the head: Prosodic contrasts can be discriminated by head motion. *Speech Communication, 52*(6), 555-564.

Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. *Progress in Brain Research, 169,* 353-363.

Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, and Hearing Research, 42*(3), 526-539.

Rosen, H. J., Wilson, M. R., Schauer, G. F., Allison, S., Gorno-Tempini, M.-L., Pace-Savitsky, C., . . . Miller, B. L. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia, 44*(3), 365-373.

Rothauser, E. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17*, 225-246.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology: Series B, 62*(1), 53-60.

Chapter 5. The Extent of Early Emotion Processing in Young and Older Adults

5.1 Abstract

Evidence indicates that older adults are less accurate than younger adults at facial emotion recognition. The present study, using a masked priming paradigm, investigated whether older adults have problems at an early stage of emotion processing. In the experiment, a prime emotional face preceded a target emotional face with prime and target emotion repeated (happy-happy, angry-angry), different (neutral control: neutral-happy, neutral-angry), or similar (surprise-happy, disgust-angry). Participants classified target faces as ‘positive’ or ‘negative’ and response times and error rates were measured. In Experiment 1, younger (n = 30) and older (n = 30) adults were tested with primes presented for 50ms. For younger adults, priming effects were found for both happy and angry target faces, i.e., response times were significantly faster for the repeated relative to the neutral prime condition. Priming effects were not found for the similar relative to the neutral prime condition. Older adults showed the same pattern of priming effects for happy faces, but there were no priming effects for angry faces. In Experiment 2, only older adults (n = 30) were tested with primes presented for three longer durations (58ms, 67ms, and 83ms). A significant masked priming effect was found but again only for happy faces. These findings suggest that the lack of an angry priming effect for older adults may be due to the reduced capacity of the emotion processing system to output clear category information for angry faces.

5.2 Introduction

Research indicates that the ability to recognise facial emotion declines with age (Gonçalves et al., 2018; Ruffman, Henry, Livingstone, & Phillips, 2008). Further, older adults have difficulty recognising particular emotions, e.g., angry relative to other emotions. However, it is not clear exactly why and at what stage these recognition problems occur. This is because the majority of behavioural studies that have investigated emotion recognition by older adults have used explicit recognition tasks, and in such tasks overt recognition responses are the end product of a number of different processes, each of which might contribute to poor recognition performance.

That is, in order for a person to correctly perform an emotion recognition task, she/he must be able to efficiently and accurately process facial expressions; map the results of this processing to stored representations; be able to evaluate the results of this mapping in terms of a match to a single best emotion category; and hold the outcome of this processing in mind while an appropriate response option is selected (usually from multiple alternatives). Older adults' poor recognition performance could be due to problems at any (or all) of these sub-processing levels. Given the inherent causal ambiguity associated with making an explicit recognition response, what is needed to better understand the causes that underpin older adult's recognition problems, is a research paradigm that can evaluate the functioning of the sub-processes involved in an emotion recognition task.

Studies have investigated aging and early processing of emotion (e.g., Mather & Knight, 2006; Hahn, Carlson, Singer, & Gronlund, 2006; Ruffman, Ng, & Jenkin, 2009). For instance, Mather and Knight (2006) used a visual search task in which older and younger adults were required to detect whether there was a discrepant emotion expression amongst an array of faces. They reported that, like younger adults, older adults have an intact ability to rapidly detect anger. In a similar study, Ruffman et al. (2009) found that, relative to younger adults, older adults showed similar efficiency for anger detection. Based on this finding, the authors suggest that implicit emotion processing is intact in older adults. However, it is possible that the visual search paradigm does not require emotion processing or recognition. That is, a visual search task may not assess early emotion processing in older adults; but instead, may assess simple feature extraction or detection. In other words, the findings of studies like that of Ruffman et al. (2009) and Mather and Knight (2006) may suggest that simple feature detection (not implicit or early emotion processing) remains intact in the older adult groups.

The aim of current study was to investigate whether older adults have problems at the early mapping stage of emotion recognition by using an implicit measure of how well a stored representation has been triggered by the presentation of an emotion stimulus. Before giving details of this procedure, the following section will briefly outline why older adults may have problems at an early processing level.

One reason for why older adults may have problems at the early stage of emotion processing relates to age-related changes in the integrity of the cortex. That is, there are typically reductions in the amount of cortical white and grey matter in older (e.g., age 60 and older) compared to younger adults, as well as reductions in the density of neurotransmitters such as dopamine and noradrenaline (Fjell & Walhovd, 2010; Li & Rieckmann, 2014; Mather, 2016). These changes may result in a decline in the efficiency of the operation of brain networks including those responsible for emotion processing. Indeed, some researchers have argued that the disproportionate difficulty that older adults have in recognising angry faces may be related to earlier and more rapid decline in the functioning of the neural circuits that subserve anger recognition (Gonçalves et al., 2018; Ruffman et al., 2008).

Given that the evidence for a decline in cortical structure in old age is relatively clear, why is a behavioural investigation of the early stages of emotion processing necessary? The first reason is that the link between brain structure and function is complex; indeed, changes in brain morphology do not always lead to changes in function (see the concept of neural reserve, Steffener & Stern, 2012). Thus, it is important to investigate whether older adults do have behavioural deficits in the early processing of facial emotion. The second reason is that, there is evidence that older adults may also experience problems at the performance level of an emotion recognition task. That is, older adults were able to show similar recognition performance to younger adults in a simple recognition task that presented only two emotion response options (Orgeta, 2010); this raises the possibility that early emotion recognition processes may be intact and that recognition problems occur due to age-related changes in other skills (e.g., in attention, memory or executive functions). Thus, once again it is important to test the extent to which older adults do have problems in the early processing of facial emotion.

To assess the extent of early emotion processing, relatively independent of problems that may be reflected at the performance level, the current study used a masked priming procedure (see Forster & Davis, 1984 for a general approach to the masked priming paradigm). In the masked priming procedure, a target emotional face is immediately preceded by a prime emotional face that is presented in such a way that viewers are largely unaware of its presentation (see Rohr, Degner, & Wentura, 2012 for an example of the masked priming procedure with emotional faces). The emotion portrayed by the prime face is selected to either be the same (repeated) or different (incongruent) to the target emotion which participants are required to categorise as quickly and as accurately as they can.

Despite participants being unaware of the prime stimulus, the presentation of a repeated prime has been shown, in younger adults, to facilitate the response (i.e., faster response times and fewer errors) to the subsequent target stimulus compared to the target preceded by an incongruent prime (e.g., Harry, Davis, & Kim, 2012). Masked priming effects are short-lived (lasting only a second or so, Forster & Davis, 1984) and are typically interpreted in terms of a processing savings, i.e., the work done in processing prime being transferred to the processing of the target (via their memory representation). Given this, masked priming can be used to assess how efficiently older adults can access the emotion information of the prime. That is, the production of a masked priming effect can provide information about how quickly and efficiently emotion information from the prime can be triggered. The assumption is that if the initial processing of emotion information is intact, then prime processing should occur rapidly and result in a robust priming effect. Little or no masked priming would result if the initial emotion processing system has been compromised.

In the current experiment, happy and angry faces were selected as target expressions (with happy and angry primes, respectively), and faces with a neutral expression were selected as a control prime condition. Happy and angry faces were selected in order to provide a contrast between where a masked priming effect would be expected and where it would not. That is, given that older adults typically do not have problems in recognising happy emotional expressions, we presume that the initial processing of a happy face is intact and so will produce a masked priming effect. On the other hand, older adults have problems in recognising angry emotional expressions; if the initial processing of an angry face is compromised, then no masked priming effect would be produced. Finding a priming effect for happy faces will demonstrate that older adults are able to generate masked emotion priming effects, and it is only when this occurs that the priming result for angry primes and targets is able to unambiguously reveal the state of early emotion processing for this emotion.

One additional prime condition was selected to provide information about the specificity of any masked priming effect found. This condition consisted of prime emotions that were selected to have some similarity to the happy and angry emotions (henceforth the ‘similar’ prime condition). That is, surprise and disgust were selected to be prime emotions for happy and angry target faces, respectively, as these emotions are frequently confused, i.e., surprise for happy and disgust for angry (Palermo & Coltheart, 2004; Widen, Russell, & Brooks, 2004). If masked priming is specific to the same emotion being repeated in the prime and the target, then priming should be found in the repeated prime condition relative to the similar one (which may not differ from the neutral control). Finding no differences between the repeated prime and similar prime conditions, and where both are facilitated compared to the neutral condition would indicate that the masked priming is based on a somewhat cursory analysis of the prime emotion.

5.3 Experiment 1

5.3.1 Method

5.3.1.1 Participants

Thirty young university students (M_{age} = 20, range = 18-29, 4 males) and thirty older adults (M_{age} = 70, range = 61-83, 13 males) from the local community participated in this study. Younger adults received course credit for their participation whereas older adults received monetary reimbursement. All participants reported to have normal or corrected to normal vision. Participants were screened for dementia using the Mini-Cog^{TM} test (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000) as the presence of dementia has been associated with poor emotion processing and may act as a confounding variable (Rosen et al., 2006). Older (M = 5) and younger (M = 5) adults scored within the normal range on the Mini-Cog^{TM} test (where a score of at least 3 out of 5 points represents normal cognition) indicating no symptoms of dementia (Borson et al., 2000).

5.3.1.2 Stimuli

Thirty faces (15 female) were chosen from the Radboud database (Langner et al., 2010). For each face, two emotional expression variations (happy, angry) were chosen to be target images. These faces depicted the person turned slightly to the right, with their left cheek leading and their gaze directed to the front. The leading left cheek was shown for target faces as this pose has been suggested to be more emotionally expressive (Lindell & Savill, 2010). In addition, five emotional expression variations (happy, angry, surprise, disgust, neutral) were also taken from the Radboud database to be used as prime faces (Langner et al., 2010). These stimuli showed the same people, but their faces were differently posed, i.e., front-facing, compared to the target faces so as to ensure that any priming effects result from the overlap of emotion information, and not from low-level, form-based physical repetition.

Faces were selected based on inter-rater agreement scores with regard to the intended expression. That is, faces that attracted the highest inter-rater agreement scores averaged across all emotion types were selected for this experiment. On average, happy ($M_{agreement\ (\%)}$ = 99%), surprise ($M_{agreement\ (\%)}$ = 92%), angry ($M_{agreement\ (\%)}$ = 93%), and neutral ($M_{agreement\ (\%)}$ = 92%) faces showed similar inter-rater agreement scores whereas the disgusted faces had a slightly lower average ($M_{agreement\ (\%)}$ = 80%). A further four faces with three emotional expression variations (i.e., 2 female and 2 male faces x happy, angry, and neutral expression, n =12) were selected from the database to be used as practice items.

In total, there were 222 images (12 practice images, 60 target faces, 150 prime faces). Images were cropped to include the face region and hairline. Images were then converted to gray-scale and the SHINE MATLAB toolbox was used to normalise intensity levels and spatial frequencies across all stimuli (Willenbockel et al., 2010). Target faces were resized to 6.56 x 8.68 cm and prime faces were resized to 4.37 x 5.79 cm as per Harry et al. (2012). The forward and backward masks were taken from the study by Harry et al. (2012). The forward mask consisted of an assortment of scrambled facial features (eye, noses, mouths) and the backward mask consisted of a blurred and scrambled face superimposed on a checkerboard.

5.3.1.3 Procedure

Participants were tested individually in a sound attenuated booth with dim lighting. Participants were told they would see a series of images, followed by a person expressing a positive or negative emotion and that they would indicate which emotion was expressed by pressing labelled buttons on a button box. They were asked to be as fast and as accurate as possible.

A windows-based display program DMDX was used to present stimuli in the specified sequence, and durations and also to randomise the order of the trials (Forster & Forster, 2003). DMDX provides accurate presentation rates as it synchronises the display to the monitor's display refresh cycle. Participants were randomly assigned to one of two versions of the experimental list. That is, for half of participants, the right button corresponded with a "negative" response and the left button with a "positive" response.

Participants were first presented with eight practice trials followed by the experimental items. Altogether, there were 180 trials as each of the 60 target faces appeared 3 times accompanied by either a repeated prime emotion (i.e., angry-angry, happy-happy), a neutral (control) prime emotion (i.e., neural-angry, neutral-happy), or a similar emotion (i.e., disgust-angry, surprise-happy). Within each trial, participants were presented with a fixation cross (500ms), followed by a forward mask (500ms), a prime face (50ms), a backward mask (33ms), and a target face (700ms). Response time and error rates were recorded as the priming effect can be determined by comparing these scores between repeated prime and incongruent trials. Participants were given a break every 18 trials to avoid fatigue effects. At the end of the experiment, participants were debriefed regarding the purpose of the study.

5.3.2 Results

Overall response times and error rates for the younger and older adults are presented in Table 5.1. Response times and error rates were analysed separately for each age group. To establish whether there were priming effects for happy and angry faces, scores were analysed in a 2 (target emotion, happy, angry) x 2 (prime emotion, repetition, control) repeated measures ANOVA. Following this, a second repeated measures ANOVA was conducted to investigate potential priming effects in the similar prime condition relative to the control prime condition. The same factors as the previous ANOVA were entered except that prime emotion included similar and control primes only.

Table 5.1. Mean response times, ms (SE) and percent error rates (SE) for the happy and angry target faces in each of the prime emotion conditions (repeated, control, similar).

Target emotion	Prime emotion	Younger adults		Older adults	
		Response time	Error rate	Response time	Error rate
Happy	Repeated (happy)	448.7 (41)	4.7 (4)	573.2 (100)	2.9 (4)
	Control	460.2 (37)	5.2 (5)	584.6 (96)	3.2 (3)
	Similar (surprise)	462.9 (38)	6.4 (7)	582.0 (97)	3.8 (5)
Angry	Repeated (angry)	460.6 (34)	5.0 (5)	595.4 (101)	3.8 (4)
	Control	474.0 (34)	6.0 (5)	591.3 (84)	4.6 (5)
	Similar (disgust)	472.7 (34)	6.0 (6)	591.8 (91)	3.9 (4)

5.3.2.1 Response time data

Prior to analysing the response time data, lower and upper response time cutoffs were applied to each participant's response time data (lower cutoff = 150 ms; upper cutoff = 1500 ms). A winsorisation procedure was then used to curtail the influence of outlying response times. In this procedure, response times that were +/- 3 SD from each participant's mean response time were brought back to a prespecified boundary (see Dixon & Yuen, 1974). Across the 30 younger adults, winsorisation was applied on average 2.30 times and cutoffs were applied on average 0.03 times. For older adults, winsorisation was applied on average 2.50 times and cutoffs were applied on average 0.10 times.

Repeated and control primes. Overall, younger adults responded faster to happy faces (454ms) than angry faces (467ms), $F(1,29) = 7.22$, $p < .02$, $\eta p2 = .20$. Repeated prime emotions attracted faster response times (455ms) than control primes (467ms), $F(1,29) = 15.52$, $p < .001$, ηp2 = .35. There was no interaction between target emotion type and prime emotion type, $F(1,29) = .12$, $p = .74$, ηp2 = .004.

Older adults were also faster at responding to happy (579ms) than angry (593ms) target faces, $F(1,29) = 8.91$, $p < .01$, ηp2 = .24. Unlike younger adults, older adults did not show any significant response time differences for repeated (584ms) or control primes (588ms), $F(1,29) = .86$, $p = .36$, ηp2 = .03. There was a significant interaction between target emotion and prime emotion, $F(1,29) = 5.87$, $p < .05$, ηp2 = .17, that was analysed further using a simple effects analysis with a Bonferroni adjusted alpha. This analysis revealed that response times were faster for the repeated prime (happy) relative to the control prime (neutral) for happy faces, $F(1,29) = 4.61$, $p < .05$, ηp2 = .14. For angry faces, response times did not significantly differ for repeated (angry) and control primes (neutral) conditions, $F(1,29) = .72$, $p = .40$, ηp2 = .02.

Similar and control primes. For the younger group, happy faces (462ms) attracted faster response times than angry faces (473ms), $F(1,29) = 7.05$, $p < .02$, ηp2 = .20. There were no significant response time differences for the similar prime (468ms) and control prime (467ms) conditions, $F(1,29) = .05$, $p = .83$, ηp2 = .002. There was no interaction between target emotion type and prime emotion type, $F(1,29) = .62$, $p = .44$, ηp2 = .02.

Older adults' response times for happy (583ms) and angry (592ms) faces were similar, $F(1,29) = 1.92$, $p = .18$, ηp2 = .06. Response times were also comparable across the similar prime (587ms) and control prime (588ms) emotions, $F(1,29) = .14$, $p = .71$, ηp2 = .01. The interaction between target emotion and prime emotion was not significant, $F(1,29) = .24$, $p = .63$, ηp2 = .01.

5.3.2.2 Error rates

Repeated and control primes. Younger adults had low error rates for both happy (5%) and angry (6%) targets, and these error rates did not significantly differ, $F(1,29) = .46$, $p = .50$, ηp2 = .02. There were no significant error rate differences when the target was preceded by a repeated (5%) or a control (6%) prime emotion, $F(1,29) = 1.03$, $p = .32$, ηp2 = .03. There was also no significant interaction between target emotion and prime emotion, $F(1,29) = .13$, $p = .73$, ηp2 = .004.

Older adults also had low error rates for happy (3%) and angry (4%) target emotions, and these error rates were not significantly different, $F(1,29) = 3.67$, $p = .07$, ηp2 = .11. There was also no significant difference in error rates to targets as a function of prime type (i.e., targets preceded by repeated (3%) or control (4%) prime emotions), $F(1,29) = 1.12$, $p = .30$, ηp2 = .04. No significant interaction between target emotion and prime emotion was found, $F(1,29) = .17$, $p = .68$, ηp2 = .01.

Similar and control primes. For the younger group, error rates for happy (6%) and angry (6%) target emotions did not differ significantly, $F(1,29) = .04$, $p = .85$, ηp2 = .001. Error rates also did not differ significantly for similar (6%) and control (6%) prime emotions, $F(1,29) = .57$, $p = .46$, ηp2 = .02. There was also no significant interaction between target and prime emotion, $F(1,29) = .58$, $p = .45$, ηp2 = .02.

Like younger adults, older adults' error rates were not significantly different for happy (4%) and angry (4%) targets, $F(1,29) = 1.39$, $p = .25$, ηp2 = .05. Similar (4%) and control (4%) primes attracted comparable error rates, $F(1,29) = .01$, $p = .93$, ηp2 = .0002. The interaction between target emotion and prime emotion was not significant, $F(1,29) = 1.02$, $p = .32$, ηp2 = .03.

5.3.3 Discussion

This experiment used masked priming to investigate the early stage of emotion processing in younger and older adults. Angry and happy target faces were used. Given that older (and young) adults can accurately recognise happy emotional expressions, it was expected that masked priming would be found for old (and younger) adults for happy primes and targets (versus the neutral emotion control), whereas it was expected that only younger adults would show priming for angry primes and targets.

As expected, there was priming for both happy and angry for the younger adults. Moreover, this masked priming effect was specific to repeated prime and target emotions, as responses to targets preceded by the similar primes did not significantly differ from responses to targets preceded by the neutral emotion baseline. This finding suggests that in the case of younger adults, priming effects are not due to the repetition of similar facial features but instead, occur when the prime and target expressions map onto the same emotion category.

It is important to note that gender was not balanced in the younger adult group (i.e., 4 males, 26 females). In the current study, it is unlikely that gender differences would be revealed for accuracy as the task was simple and participants had low error rates (floor effects). However, previous research has shown that there are gender differences in the efficiency with which emotion is detected (e.g., men detect angry faces faster than women, Williams & Mattingley, 2006). Despite gender differences in efficiency, it is anticipated that the younger group (regardless of gender makeup) would show robust priming effects. That is, even though women may be slower at detecting angry faces, priming effects are still likely to occur.

Like younger adults, older adults showed priming for happy faces and this priming was specific to the repeated prime condition only, i.e., response times were comparable across similar and control prime conditions. In contrast to happy faces, older adults did not show priming for angry faces, i.e., response times were similar for the repeated (angry) and neutral prime conditions. Further, response times did not differ across the similar (disgust) prime and the neutral prime conditions.

One interpretation of these results is that relative to happy faces, angry faces are slower to trigger a response from the older adults' early emotion processing system. In other words, for older adults, the early processing system for angry faces may not be as efficient as that for happy faces. It seems plausible then that older adults may need to be exposed to angry prime faces for a longer duration to trigger a response picked up as a masked priming effect. A related but potentially different interpretation of these results is that the limiting factor is that emotion processing system is less able to produce clear emotion category information for angry faces. To test whether the former interpretation is the case (i.e., that older adults simply require more time to trigger a response to angry prime faces), the prime duration was increased in the following experiment.

5.4 Experiment 2

To investigate whether a lack of angry priming effects in Experiment 1 was due to a limitation in processing rapidly presented angry prime faces, Experiment 2 tested only older adults for three different longer durations (i.e., 58ms, 67ms, and 83ms). Three durations were selected as it is not clear at what duration older adults would be able to process angry prime faces. In line with the results of the previous experiment, we anticipated priming effects for happy faces across all the prime durations. For angry faces, if older adults' early emotion processing systems are relatively intact but just require more time to get going, then priming effects would be observed for the angry primes with the longest duration. Finding no priming effects, regardless of prime duration, would indicate that older adults' early emotion processing system for angry faces may have more problems than simply being sluggish.

5.4.1 Method

5.4.1.1 Participants

Thirty older adults (M_{age} = 72, range = 61-83, 14 males) participated in the current experiment for monetary reimbursement. As in the previous experiment, older adults (*M* = 5) were screened for dementia using the Mini-Cog™ test (Borson et al., 2000).

5.4.1.2 Stimuli

All face stimuli used in the previous experiment were retained for the current experiment.

5.4.1.3 Procedure

The procedure from the previous experiment remained largely the same as the first experiment except that the presentation duration of the prime stimuli was manipulated. That is, prime stimuli were presented for three different durations; 58ms, 67ms, and 83ms. This resulted in a total of 540 trials consisting of 30 faces, 2 target emotion types, 3 prime emotion types, and 3 prime durations.

5.4.2 Results

Response times and error rates were analysed in separate repeated measures ANOVA's. As in the analyses of the first experiment, two repeated measure ANOVA's were conducted. First, target emotion (happy, angry), prime emotion (repeated, control), and prime duration (58ms, 67ms, and 83ms) were entered as factors. The additional repeated measures ANOVA was conducted with the same factors except prime emotion included similar and control prime emotions. Overall response times and error rates are presented in Table 5.2.

Table 5.2. Mean response times (SE) and error rates (SE) for older adults across the happy and angry target faces for each of the prime emotion types (repeated, control, similar) and prime durations.

Target emotion	Prime emotion	Response time (ms)			Error rate (%)		
		58ms	67ms	83ms	58ms	67ms	83ms
Happy	Repeated (happy)	548.3 (77)	546.4 (85)	542.9 (87)	4.4 (5)	3.0 (4)	3.9 (5)
	Control (neutral)	563.8 (84)	566.2 (78)	568.1 (75)	3.2 (3)	3.3 (4)	5.1 (7)
	Similar (surprise)	553.1 (81)	556.3 (80)	564.3 (80)	3.8 (4)	4.0 (4)	3.8 (4)
Angry	Repeated (angry)	568.5 (100)	569.5 (94)	561.4 (101)	3.4 (5)	4.6 (5)	2.7 (3)
	Control (neutral)	575.0 (98)	568.7 (87)	569.9 (90)	4.0 (5)	4.3 (6)	5.0 (5)
	Similar (disgust)	576.1 (93)	575.8 (88)	576.4 (98)	3.4 (4)	3.8 (4)	3.6 (5)

5.4.2.1 Response time data

As per Experiment 1, lower and upper response time cutoffs were applied to each participant's response time data (lower cutoff = 150 ms; upper cutoff = 1500 ms) and a winsorisation procedure was used. Across the 30 older adults, winsorisation was applied on average 6.83 times and cutoffs were applied on average 0.03 times.

Repeated and control primes. Overall, there were no response time differences for target faces as a function of prime duration, i.e., for 58ms (564ms), 67ms (563ms), or 83ms (561ms), $F(2,58) = 1.09$, $p = .34$, ηp2 = .04. Older adults responded to happy target faces (556ms) faster than angry target faces (569ms), $F(1,29) = 7.18$, $p < .02$, ηp2 = .20. Repeated primes (556ms) were responded to faster than control primes (569ms), $F(1,29) = 18.74$, $p < .001$, ηp2 = .39. There was a significant interaction between target emotion and prime emotion, $F(1,29) = 12.70$, $p < .02$, ηp2 = .31. There was no significant interaction between prime duration, target emotion, and prime emotion, $F(2,58) = .85$, $p = .43$, ηp2 = .03.

The interaction between target emotion and prime emotion was analysed using a simple effects analysis (with a Bonferroni adjusted alpha). This analysis showed that for happy faces, older adults responded faster to the repeated prime condition (happy) than the control (neutral) prime condition, $F(1,29) = 23.12$, $p < .001$, ηp2 = .44. There was no significant masked priming effect for angry faces, i.e., response times to angry targets in the repeated (angry) prime condition were not significantly different to response times to angry targets in the control (neutral) prime condition, $F(1,29) = 2.62$, $p = .12$, ηp2 = .08.

Similar and control primes. Overall response times to target faces did not differ across the 58ms (567ms), 67ms (567ms), or 83ms (570ms) prime durations, $F(2,58) = 1.19$, $p = .31$, ηp2 = .04. Response times were faster for happy targets (562ms) relative to angry targets (574ms), $F(1,29) = 6.17$, $p < .02$, ηp2 = .18. Response times were not significantly different between the similar prime condition (567ms) and the control prime condition (569ms), $F(1,29) = .52$, $p = .48$, ηp2 = .02. The interaction between target emotion and prime emotion was significant, $F(1,29) = 7.94$, $p < .01$, ηp2 = .22. The three-way interaction between prime duration, target emotion, and prime emotion was not significant, $F(2,58) = .41$, $p = .67$, ηp2 = .01.

A simple effects analysis (with a Bonferroni adjusted alpha) was used to analyse the interaction between target emotion and prime emotion. For happy faces, response times were faster in the similar (surprise) prime condition relative to the control (neutral) prime condition, $F(1,29) = 6.05$, $p < .05$, ηp2 = .17. Response times did not significantly differ for angry target faces preceded by similar (disgust) or control (neutral) prime faces, $F(1,29) = 2.47$, $p = .13$, ηp2 = .08.

Repeated and similar primes. Given that response times were faster in the similar relative to the control condition for happy faces, an additional repeated measures ANOVA was conducted on the response times for the repeated and similar conditions for both happy and angry target faces. Overall, there were no significant response time differences across the 58ms (562ms), 67ms (562ms), or 83ms (561ms) prime durations, $F(2,58) = .04$, $p = .96$, ηp2 = .001. Responses to happy target faces (552ms) were faster than angry target faces (571ms), $F(1,29) = 16.29$, $p < .001$, ηp2 = .36. Response times were faster for the repeated (556ms) relative to the similar (567ms) prime conditions, $F(1,29) = 24.43$, $p < .001$, ηp2 = .46. The interactions between target emotion and prime emotion, $F(1,29) = .18$, $p = .68$, ηp2 = .006, and prime duration, target emotion, and prime emotion, $F(2,58) = .49$, $p = .62$, ηp2 = .02, were not significant.

5.4.2.2 Error rates

Repeated and control primes. Error rates did not significantly differ across the prime durations, 58ms (4%), 67ms (4%), or 83ms (4%), $F(2,58) = .44$, $p = .65$, ηp2 = .02. There was no significant difference between error rates for the happy (4%) and angry (4%) target emotions, $F(1,29) = .13$, $p = .72$, ηp2 = .004, or for the repeated (4%) or control (4%) prime emotions, $F(1,29) = 1.36$, $p = .25$, ηp2 = .05. Finally, there were no significant interactions between target emotion and prime emotion, $F(1,29) = .76$, $p = .39$, ηp2 = .03, or between prime duration, target emotion, and prime emotion, $F(2,58) = .81$, $p = .45$, ηp2 = .03.

Similar and control primes. Error rates were similar across the 58ms (4%), 67ms (4%), and 83ms (4%) prime durations, $F(2,58) = 1.32$, $p = .28$, ηp2 = .04. Happy targets (4%) attracted similar response times to angry targets (4%), $F(1,29) = .13$, $p = .72$, ηp2 = .01. Response times were also similar across similar prime (4%) and control prime (4%) conditions, $F(1,29) = 1.58$, $p = .22$, ηp2 = .05. There were no significant interactions between target emotion and prime emotion, $F(1,29) = 1.48$, $p = .23$, ηp2 = .05, and prime duration, target emotion, and prime emotion, $F(2,58) = .18$, $p = .84$, ηp2 = .01.

5.4.3 Discussion

The current experiment was a modified replication of Experiment 1 and manipulated the duration of the primes and only investigated the performance of older adults. More specifically, longer prime durations were introduced to investigate whether the lack of angry priming effects in the previous experiment were due to a slow processing system.

For happy faces, significant priming effects were found (as in experiment 1) and the size of this masked priming effect did not significantly vary as a function of prime duration. Further, priming was not simply due to low-level physical repetition as the repeated happy prime and target emotional faces were different poses of the same person. Taken together, these findings indicate that the emotion processing system of older adults can reliably produce priming effects from happy prime faces. Unexpectedly and in contrast to the previous experiment, priming effects were also found for happy face targets preceded by similar (surprise) face primes (relative to neutral baseline); although priming in the repeated (happy) condition was larger than that for the similar (surprise) condition. This variable form-priming effect possibly reflects individual differences in how well older adults can process emotional expressions and may be similar to the variation found for masked orthographic form-priming effects that occur in reading as a function of reading skill (Andrews & Lo, 2012).

In contrast to happy faces, angry faces did not produce significant priming effects for repeated or similar prime emotions (as measured against the neutral baseline), regardless of the prime duration. A lack of angry priming effects, particularly at the longer prime durations, indicates that in addition to being slow at triggering a response, the processing system for angry faces is compromised in other ways (e.g., the system may produce ambiguous emotion information). This finding stands in contrast to the findings of previous studies that suggest that the processing system for angry faces is intact in older adults (e.g., Mather & Knight, 2006; Hahn et al., 2006; Ruffman et al., 2009). One reason for these inconsistent findings relates to differences in the task/paradigm used to assess emotion procoessing. Unlike the current study that used a masked priming paradigm, previous studies used a visual search paradigm that may not assess early emotion processing but may, instead, assess rapid feature extraction.

5.5 General discussion

The present study investigated whether older adults have problems at the early stages of emotion processing. A masked priming procedure was used with happy and angry target faces. Overall, younger adults showed priming effects for happy and angry faces whereas older adults showed priming effects only for happy faces. To understand the implications of these findings for older adults, we must first interpret what a masked priming effect, and the lack of a priming effect, represent.

Masked priming effects, where the effect of a repeated prime is measured against a neural control prime, i.e., a prime that does not match the alternative response option, have been typically interpreted in terms of the prime triggering a saving in information processing related to the target representation (rather than priming response selection). That is, when priming is measured from a neutral control prime, masked priming is thought to reflect a savings in the time taken to process task-relevant target information (Kim & Davis, 2003). Given this interpretation, then in the current paradigm, masked priming can be viewed as indicating how quickly and clearly the emotion processing system can produce information that is relevant to recognising the category of a target emotion.

In older adults, the priming effect for happy faces indicates that the emotion processing system can rapidly produce clear task-relevant emotion information from happy faces. Further, the finding that there were no priming effects for the similar relative to the neutral control prime (experiment 1; but not experiment 2) supports the assumption that the priming effect is mediated by representations of emotion categories that link the prime and the target, i.e., both prime and targets represent a happy emotion.

In contrast to happy faces, the lack of priming effects for angry faces indicates that for older adults, the emotion processing system that produces information about whether a face is angry or not is not able to rapidly and clearly output this information. The results from experiment 2, where the processing system was given more time to process and produce this information, suggests that for angry faces something in addition to slow processing speed may be involved (i.e., the bottom-up specialised system may be compromised for processing angry faces so that it does not provide clear emotion information).

Although priming effects were found for happy but not angry faces (in terms of response time), error rates were low and similar across all conditions for both older and younger adults. A system that is sluggish and that produces vague emotion information (i.e., a weak output signal) in response to an angry face may not result in poorer recognition performance for angry faces when there are only two response options. This is because there are no response competitors particularly when the alternative emotion is a positive one (e.g., happy). However, when there are six response options as is common in recognition tasks, a weak signal may impede one’s ability to successfully perform the task. That is, having many response options increases the number of potential response competitors. In addition, having a weak signal may impair the selection of the correct response option particularly for older adults whose emotion processing system may not be able to efficiently prioritise one competitor and inhibit competition from other response candidates. This account can explain the Orgeta (2010) finding that older adults were able to show similar anger recognition performance to younger adults in a simple recognition task that presented only two emotion response options (similar to the current study).

In conclusion, older adults may have problems with the bottom-up (quality of signal) and top down (selecting a response option) processes involved in angry (but not happy) emotion recognition. Problems at both of these levels for angry but not happy faces can account for why older adults show better recognition for some emotion types (e.g., happy) and worse recognition for others (e.g., angry). Future research plans include replicating the current study to ascertain how rapidly and how well the system can yield accurate information for emotion types other than happy and angry.

5.6 References

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 152.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: a cognitive'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15*(11), 1021-1027.

Dixon, W. J., & Yuen, K. K. (1974). Trimming and winsorization: A review. *Statistische Hefte, 15*(2-3), 157-170. doi:10.1007/BF02922904

Fjell, A. M., & Walhovd, K. B. (2010). Structural brain changes in aging: courses, causes and cognitive consequences. *Reviews in the Neurosciences, 21*(3), 187-222.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680-698.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*(1), 116-124.

Gonçalves, A. R., Fernandes, C., Pasion, R., Ferreira-Santos, F., Barbosa, F., & Marques-Teixeira, J. (2018). Effects of age on the identification of emotions in facial expressions: a meta-analysis. *PeerJ, 6*, e5278.

Hahn, S., Carlson, C., Singer, S., & Gronlund, S. D. (2006). Aging and visual search: Automatic and controlled attentional bias to threat faces. *Acta Psychologica, 123*(3), 312-336.

Harry, B., Davis, C., & Kim, J. (2012). Subliminal access to abstract face representations does not rely on attention. *Consciousness and Cognition, 21*(1), 573-583.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49*(4), 484-499.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377-1388.

Li, S.-C., & Rieckmann, A. (2014). Neuromodulation and aging: implications of aging neuronal gain control on cognition. *Current Opinion in Neurobiology, 29,* 148-158.

Lindell, A. K., & Savill, N. J. (2010). Time to turn the other cheek? The influence of left and right poses on perceptions of academic specialisation. *Laterality, 15*(6), 639-650.

Mather, M. (2016). The affective neuroscience of aging. *Annual Review of Psychology, 67,* 213-238.

Mather, M., & Knight, M. R. (2006). Angry faces get noticed quickly: Threat detection is not impaired among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(1), P54-P57.

Orgeta, V. (2010). Effects of age and task difficulty on recognition of facial affect. *Journal of Gerontology: Psychological Sciences, 65B*, 323-327.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers, 36*(4), 634-638.

Rohr, M., Degner, J., & Wentura, D. (2012). Masked emotional priming beyond global valence activations. *Cognition & emotion*, *26*(2), 224-244.

Rosen, H. J., Wilson, M. R., Schauer, G. F., Allison, S., Gorno-Tempini, M.-L., Pace-Savitsky, C., . . . Miller, B. L. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia, 44*(3), 365-373.

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*(4), 863-881.

Ruffman, T., Ng, M., & Jenkin, T. (2009). Older adults respond quickly to angry faces despite labeling difficulty. *Journals of Gerontology: Series B, 64*(2), 171-179.

Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1822*(3), 467-473.

Widen, S. C., Russell, J. A., & Brooks, A. (2004). *Anger and disgust: Discrete or overlapping categories.* Paper presented at the 2004 APS Annual Convention, Boston College, Chicago, IL.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*(3), 671-684. doi:10.3758/BRM.42.3.671

Williams, M. A., & Mattingley, J. B. (2006). Do angry men get noticed? *Current Biology, 16*(11), R402-R404.

Chapter 6. Older Adults' Perception and Production of Facial Emotion are Linked

6.1 Abstract

The current study investigated whether older adults' ability to recognise emotion is related to their ability to produce it. Eight older adult participants were selected from a previous recognition study based on their average correct face emotion recognition scores: Four had relatively high scores (74% accuracy) – the High recognition group; and four had relatively low scores (61%) – the Low recognition group. These older adults and eight younger ones were audio and video recorded expressing various emotions (i.e., anger, disgust, fear, happy, sad, surprise, and neutral) while saying emotionally neutral phrases. Recordings were presented to 78 younger adults in auditory-only (AO), visual-only (VO), and auditory-visual (AV) formats in an emotion recognition task. Findings indicate that for AO, VO, and AV formats, participants showed better emotion recognition of the stimuli from the younger and the High perception older adults compared to that of the Low perception older adults. These findings indicate that emotion recognition may be related to emotion production in older adults.

6.2 Introduction

It is well established that older adults are less accurate than younger ones at recognising facial emotion (Gonçalves et al., 2018). This decline in the ability to recognise emotions raises the issue of why old age has such an effect but it also poses questions about the nature of the emotion recognition process itself. One such question concerns the possible link between emotion perception and (implicit) emotion production. Traditional theories of emotion recognition only considered perception processes; i.e., it was typically proposed that visual emotion recognition involves mapping a facial expression code to stored representations of facial emotional expressions (Bruce & Young, 1986). More recently, it has been suggested that the recognition of emotion from a face is not simply a passive process, but also involves an implicit production of the observed emotion (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). That is, under this view, part of the process of recognising a facial emotion involves implicitly simulating the observed emotion. It is this simulation of another's emotional expression that facilitates access to knowledge of the presented emotion.

In support of the view that emotion perception and production are linked, studies with younger adults provide evidence that the somatosensory and motor systems play a role in both processes. For instance, studies has shown that when the automatic activation of facial muscles is prevented (by holding a pen in the mouth) the recognition performance for certain emotional expressions is reduced (Oberman, Winkielman, & Ramachandran, 2007; Experiment 2, Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001). It has also been shown that facial emotion recognition performance can be improved by facial motor production training in younger adults (Deriso et al., 2012). Furthermore, it has been reported that using transcranial magnetic stimulation on the face region of the somatosensory cortex to disrupt sensory feedback impairs the discrimination of facial emotional expressions (Pitcher, Garrido, Walsh, & Duchaine, 2008). In addition, it has been found that lesions to the somatosensory cortices are associated with low emotion recognition performance (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). Taken together, these studies indicate that for younger adults, emotion recognition and emotion production may be linked through a common use of a variety of critical systems (e.g., somatosensory or motor systems).

If recognising an emotion from a face is in some way linked to the ability to produce an expression, then it would seem to follow that age-related declines in emotion recognition will be accompanied by problems in producing emotion. In this regard, it is interesting to note that emotions are perceived less accurately from older, relative to young faces (Borod et al., 2004; see Ruffman, Halberstadt, Murray, Jack, & Vater, 2019 for an investigation of the recognition of genuine expressions of emotion; Hess, Adams Jr, Simard, Stevenson, & Kleck, 2012; Riediger, Voelkle, Ebner, & Lindenberger, 2011; see the FACES database by Ebner, Riediger, & Lindenberger, 2010). Although this finding could be due to the influence of peripheral factors (such as less muscle tone in older faces, etc.), or to perceiver bias, it could also indicate that some older adults have problems producing emotion. If it is the case that emotion recognition involves an element of emotion production (as suggested by theories of embodied cognition, Niedenthal et al., 2005), then the ability to recognise and produce emotion should be related in older adults. The current study investigated this relationship. That is, we explored whether older adults' emotion recognition problems may be related to their emotion production problems.

Older adults provide an interesting case to test the proposition that emotion recognition is linked with emotion production. This is because the ability to recognise emotion varies across older adults. That is, some older adults show preserved emotion recognition abilities whereas others show deficits in emotion recognition (e.g., Ebner, He, & Johnson, 2011). According to the embodied cognition theory, older adults that show poorer emotion recognition should also show poorer emotion production, whereas those that show better recognition should also show better production. According to perception only theories of emotion recognition, there would be no such relationship.

For the current study, only eight older adult participants from our previous face emotion recognition study (Chapter 3) were able to participate in the current experiment. Table 6.1 displays the overall percent correct scores for the previous emotion recognition experiment. Four of the older adults had relatively high average recognition scores (74% accuracy – the High recognition group) and four had relatively low scores (61% – the Low recognition group). As can be seen in Table 6.1, recognition scores for some emotions (i.e., angry, disgust, fear) were higher in the High recognition group relative to the Low recognition group whereas scores for other emotion types (i.e., sad, surprise, happy, neutral) appeared to be similar between the older adult groups. Further, accuracy scores appeared to be similar between the younger adults and the High recognition older adult group for some emotions (i.e., angry, disgust, sad, fear, surprise) but not others (i.e., happy, neutral) whereas younger adults showed better recognition scores than the Low recognition group across all emotion types.

Table 6.1. Emotion recognition scores (SD) for the younger adults, the High recognition older adults and the Low recognition older adults. Average age and gender are also presented for each group.

		Younger adults (*N* = 18)	Older High recognition group (*N* = 4)	Older Low recognition group (*N* = 4)
Percent correct (%) scores (SD) for the emotion recognition task (young faces only)	Average	78 (15)	74 (14)	61 (20)
	Angry	83 (15)	86 (14)	57 (30)
	Disgust	91 (11)	85 (17)	69 (20)
	Sad	75 (19)	74 (11)	68 (18)
	Fear	55 (25)	56 (15)	26 (17)
	Surprise	58 (14)	52 (20)	51 (17)
	Happy	94 (7)	85 (13)	79 (15)
	Neutral	86 (14)	77 (9)	76 (22)
Average age		21 (4)	70 (7)	71 (5)
Gender		15 females	3 females	1 female

The current study tested these eight older adults, as well as an additional eight younger adults on their ability to produce clearly recognized emotion. For this, they were required to express various emotions one at a time while uttering a semantically neutral phrase; these expressions were video recorded for assessments. There is an issue about how emotion production should be assessed in the current study. Automatic emotion recognition systems (e.g., Openface, Baltrušaitis, Robinson, & Morency, 2016) are typically trained on young faces/voices and so, may not be able to adequately deal with the age-related variance that occurs in older faces and voices (e.g., reduced muscle tone). Given that younger adults are capable of accurate emotion recognition and that they have likely been exposed to older adult emotional faces and voices (unlike automatic emotion recognition systems), then, they may provide a better index of how well older (and younger) adults produce emotion. It could be argued that older adults have more experience with emotional faces than younger adults, however, evidence indicates that they are less able than younger adults to accurately recognise emotion (Gonçalves et al., 2018). So, in the current study, expressions were presented to a newly recruited group of younger adults who were instructed to identify each emotional expression.

In line with previous research, it was expected that across all conditions, the younger adult group would outperform the two older groups on their ability to produce emotion (i.e., produce emotions that would be more accurately recognised). If the ability to recognise and produce emotion is linked, then emotional expressions from the High recognition older adult group should be better recognised than emotional expressions from the Low recognition group across all modalities. If recognition and production are not related, then we should see no differences between High or Low recognition older adult groups across all modalities.

6.3 Method

6.3.1 Emotion production

6.3.1.1 Participants

Eight younger adults (M_{age} = 23, 4 females) and eight older adults (M_{age} = 72, 4 females) were recruited to record the audio and video stimuli. All received monetary reimbursement for their participation. All presenters were native speakers of Australian English. The eight older adults, selected from our previous recognition task based on their average correct face emotion recognition scores, consisted of High (M_{accuracy} = 74%; SD = 14) and Low (M_{accuracy} = 61%; SD = 20) recognition groups.

6.3.1.2 Stimuli

Audio and video recordings consisted of the sixteen participants expressing various emotions (i.e., anger, disgust, sad, fear, surprise, happy, and neutral) while saying emotionally neutral phrases. To select the sentences for recording, a large number of IEEE sentences were rated by five coders for emotional content using a 5-point Likert scale (i.e., from -2 = very negative to 0 = neutral to +2 = very positive) (Rothauser, 1969). The ten sentences that received a score of 0 from each coder were selected to be used in the recording session.

6.3.1.3 Recording equipment

A monitor was used to display each sentence one at a time. The video camera (Sony NCCAM HXR-NX30p) was positioned directly above the monitor and recorded video at 1920 x 1080 full HD resolution at 50 frames per second. The microphone (AT 4033a Transformerless Capacitor Studio Microphone) was placed approximately 15 cm away from the presenters' mouth out of the cameras view. All audio recordings were relayed through a Motu Ultralite mk3 audio interface with FireWire connection to a PC running CueMix FX digital mixer and then to Adobe Audition.

6.3.1.4 Instruction

To assist presenters to produce the emotional expressions, the Velten Mood Induction Procedure was used to induce a depressive (negative) mood for angry, disgust, sad, and fear (Velten, 1968) and a cheerful (positive) mood for surprise and happy. A general induction procedure was used (instead of an emotion-specific procedure) to allow presenters to express emotion without constraint and so that presenters varied naturally in their emotion production. Presenters were given a relevant negative (e.g., "I'm discouraged and unhappy about myself") or positive (e.g., "I feel cheerful and lively") list of sixty self-referent mood statements. They were asked to read each statement carefully and to try to feel the mood suggested by each statement. Following the completion of the Velten Procedure, presenters were asked to recall a memory where they felt the relevant emotion in an attempt to further induce the emotional state. Once the emotion induction procedure was completed, recording of the relevant emotional expression commenced.

6.3.1.5 Recording procedure

The older and younger presenters were individually recorded in a sound attenuated booth. The presenter was seated in front of the monitor that displayed each sentence one at a time. To produce the expressions, presenters were asked to first read the sentence presented on the monitor silently and then to gaze into the camera lens as they uttered the sentence out loud whilst expressing the required emotion type.

The recording sessions were blocked by type of emotional expression. Presenters recorded the angry block first, followed by the disgust, fear, sad, neutral, happy, and surprise blocks. Each block contained 10 sentences, resulting in 70 recorded sentences for each presenter (i.e., 7 expression types x 10 sentences). A further two sentences (selected from the IEEE sentence list) were recorded with a neutral expression. These sentences were to be used as a presenter specific calibration throughout the recognition experiment.

6.3.2 Emotion recognition task

6.3.2.1 Participants

Seventy-eight undergraduate students (62 females, M_{age} = 20, age range = 17-29) from Western Sydney University participated in the recognition experiment for course credit. All reported to have learnt English at an early age (i.e. less than 9 years old) and reported normal hearing and normal or corrected-to-normal vision.

6.3.2.2 Stimuli

Stimuli consisted of the audio and video recordings from the eight older adults and the eight younger adults. These recordings were segmented into each sentence using a customised MATLAB script. Audio recordings were normalised so that the intensity of all sound files was equal to 75 decibels. Video recordings were cropped so that just the head and neck area were visible. Recordings were manipulated using a customised MATLAB script to produce auditory-only (AO), visual-only (VO), and auditory-visual (AV) experimental items.

Altogether, there were 3360 experimental items consisting of 7 expression types (anger, disgust, sad, fear, surprise, happy, and neutral), 16 presenters, 3 presentation modalities (AO, VO, AV), and 10 sentences.

6.3.2.3 Procedure

In the recognition study, younger participants were tested individually in a sound-attenuated booth. They were informed that they would hear (AO), see (VO), or hear and see (AV) people conveying various emotions. They were instructed to select (using the mouse) one of the seven options (angry, disgust, sad, fear, surprise, happy, neutral) presented on the screen they thought best described the expressed emotion.

Once participants indicated they understood the instructions, they were allocated to one of five different versions of the experimental list. Within each version, two sentences from each of the 16 presenters (8 younger, 8 older) were used (i.e., version one contained sentences 1-2, version two contained sentences 3-4 etc.). Across all versions, participants were presented with three presentation conditions (AO, VO, and AV). Trials within each AO, VO, and AV presentation condition were blocked by presenter resulting in 16 blocks of 14 trials. The order of the presentation modalities were counterbalanced across participants. The presentation order of the presenter blocks and the stimuli within each presenter block was randomised using the DMDX display and response collection software (Forster & Forster, 2003).

In each AO, VO, and AV presentation condition, participants were first presented with 7 practice trials (7 expression types) produced by a young female presenter taken from Kim and Davis (2012), followed by the experimental items. Prior to the commencement of each new or different presenter, two neutral expressions of that presenter were displayed. Participants were told to use these expressions as a baseline to compare with subsequently presented emotions.

Within each trial, a fixation point was presented followed by an experimental item (approximately 3 to 6 seconds depending on the length of the utterance). Finally, 7 boxes labelled “Angry”, “Disgust”, “Fear”, “Neutral”, “Happy”, “Sad”, and “Surprise” were presented on the screen until a response was made. Emotion labels were presented in alphabetical order in two rows; except the “Neutral” label which was presented in the centre of the screen. At the conclusion of the experiment, participants completed a questionnaire detailing age, gender, and languages spoken.

6.4 Results

Percent correct scores were collapsed across sentences and analysed with a repeated measures ANOVA with modality (AO, VO, AV), group (younger adults, High recognition older adults, Low recognition older adults), and emotion (angry, disgust, sad, fear, surprise, happy, neutral) as within subject factors. Overall, participants recognised the stimuli from the younger adults (61%) and the High recognition older adults (60%) better than those of the Low recognition older adults (46%), $F(2,154) = 322.05$, $p < .001$, $\eta p2 = .81$. The AV modality (65%) attracted the highest percent correct score followed by the VO (56%) and then the AO (47%) modalities, $F(2,154) = 349.83$, $p < .001$, $\eta p2 = .82$. There was also a significant main effect of emotion, $F(6,462) = 93.02$, $p < .001$, $\eta p2 = .55$. Finally, the three-way interaction between modality, group, and emotion was significant, $F(24,1848) = 11.70$, $p < .001$, $\eta p2 = .13$, and was analysed further using a simple effects analysis with a Bonferroni adjusted alpha.

Overall identification scores for the VO presentation conditions are presented in Figure 6.1. The analysis of these data revealed that participants recognised VO emotional expressions (i.e., anger, disgust, fear, surprise, happy) from the High recognition group better than those from the Low recognition group ($p < .001$); except for the sad and neutral emotion types where scores did not significantly differ between the groups ($p > .91$). Recognition scores for disgust, sad, happy, and neutral expressions were higher for the younger adult group relative to the High recognition older adult group ($p < .001$). The reverse pattern occurred for the expressions of anger, fear, and surprise ($p < .01$). Across all emotional expressions, percent correct scores were higher for the younger adult group than for the Low recognition older adults' group ($p < .01$).

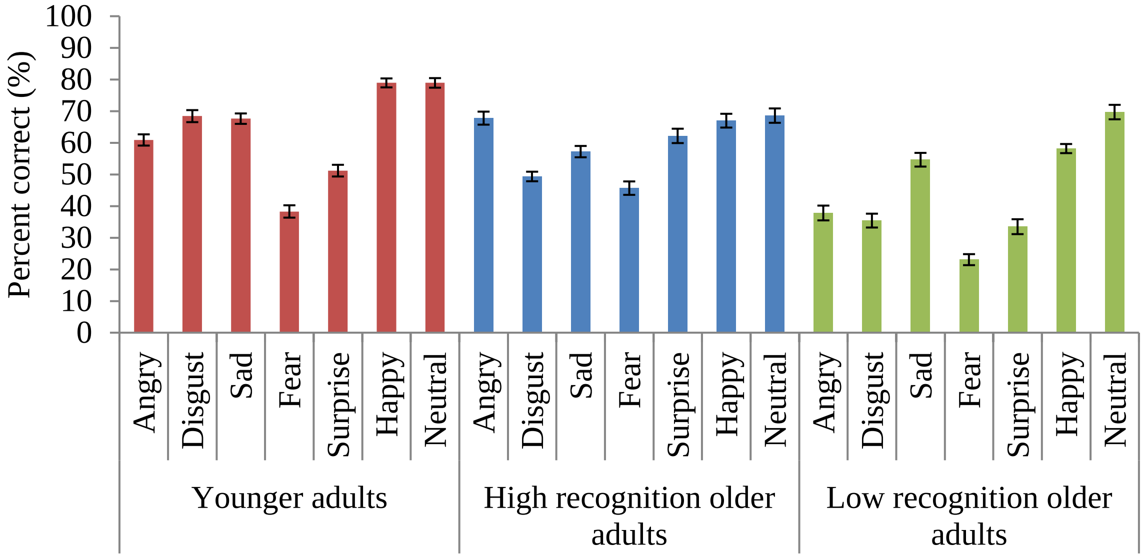

Figure 6.1. Mean percent correct scores for each VO emotion type across the younger adult (left), High recognition older adult (middle), and Low recognition older adult (right) groups. The error bars represent standard error.

AO identification scores are presented in Figure 6.2. For most AO emotion types (i.e., anger, sad, fear, surprise), the High recognition group attracted higher percent correct scores than the Low recognition group ($p < .01$ for all comparisons). There was no significant difference between these two groups for disgust, happy, or neutral AO recognition ($p > .12$). The High recognition older adults attracted higher percent correct scores relative to the younger adults for AO expressions of disgust, sad, and fear ($p < .05$). In contrast, percent correct scores for AO expressions of anger, surprise, and happiness were not significantly different between the younger adults and the High recognition older adults ($p > .08$). The neutral expression was the only AO expression where recognition scores were significantly higher for younger adults than the High recognition older adults ($p < .02$). Finally, participants recognised sad, fearful, surprised, and neutral AO expressions from the younger adults better than the Low recognition older adults ($p < .01$). Accuracy scores for the remaining emotion types (i.e., anger, disgust, happiness) did not significantly differ across the two groups ($p > .18$).

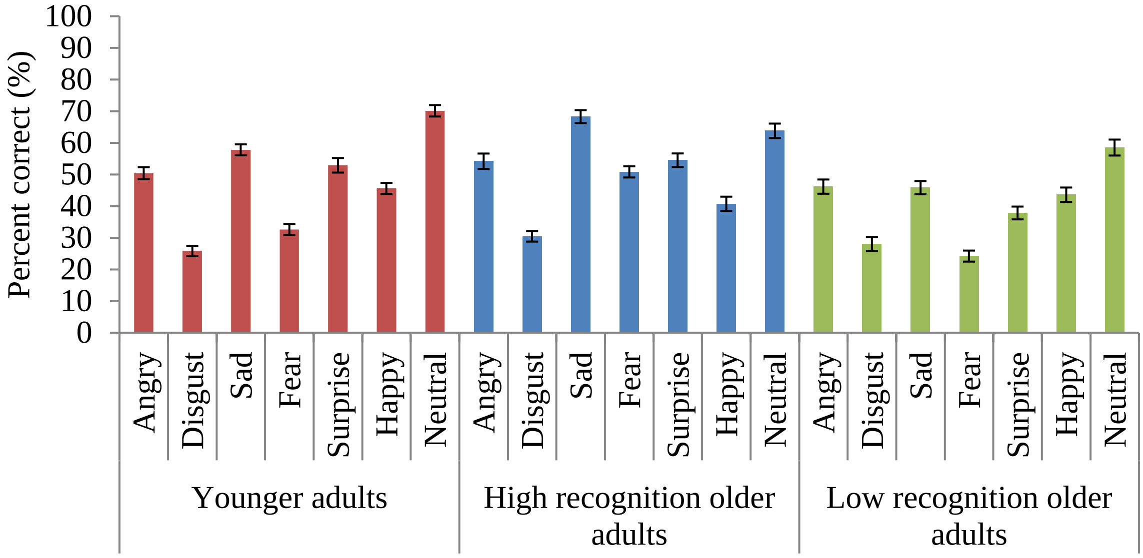

Figure 6.2. Mean percent correct scores for each AO emotion type across the younger adult (left), High recognition older adult (middle), and Low recognition older adult (right) groups. The error bars represent standard error.

Figure 6.3 presents the identification scores for the AV presentation conditions. Across most of the AV expressions (i.e., angry, disgust, sad, fear, surprise, and happy), emotion was better recognised from the High recognition older adults than the Low recognition ones ($p < .01$). Only the scores for neutral expressions did not significantly differ across the older adult groups ($p = .70$). Participants recognised disgust, sad, and neutral expressions from younger adult faces better than from High recognition older adults ($p < .001$). The opposite pattern was found for fear expressions ($p < .01$). No significant differences were found for angry, surprised, and happy expressions ($p > .15$). All expressions conveyed by younger adults attracted higher percent correct scores than expressions conveyed by the Low recognition older adults (p < .001).

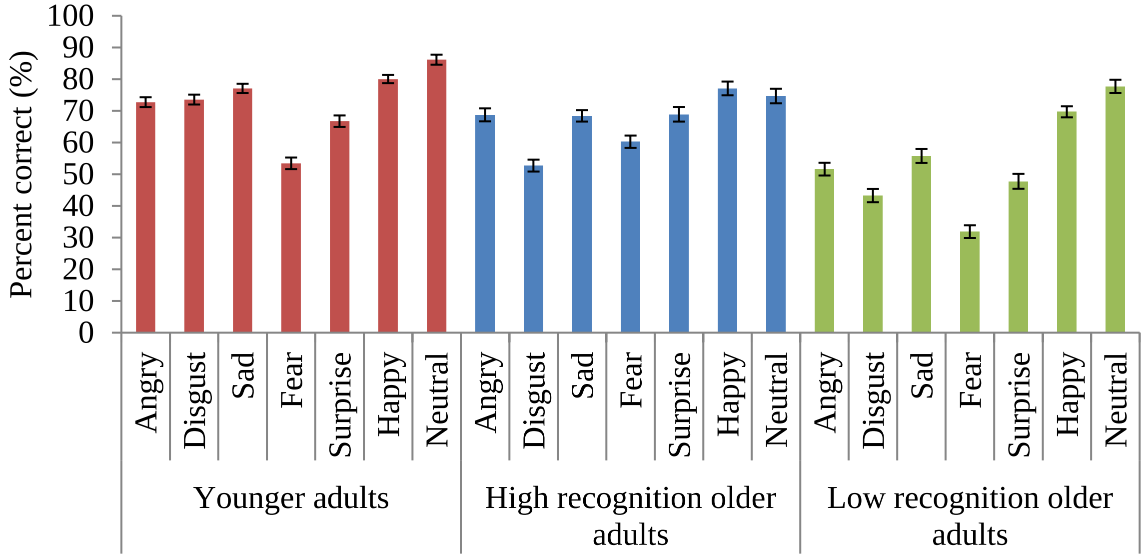

Figure 6.3. Mean percent correct scores for each AV emotion type across the younger adult (left), High recognition older adult (middle), and Low recognition older adult (right) groups. The error bars represent standard error.

Table 6.2 presents the order of older adults (from best to worst performer) based on recognition scores from the previous experiment along with their ranks for the VO production portion of this experiment. As can be seen, the ability to produce VO expressions of emotion closely matched the ability to recognise VO expressions of emotion. For instance, the best performing older adult from the previous emotion recognition task was the highest performer in the current study, and vice versa (i.e., the older adult that scored the lowest on the previous emotion recognition task attracted the lowest percent correct scores in the current study). To examine whether the ability to recognise emotion was associated with the ability to produce emotion in older adults, we conducted a simple bivariate correlation between percent correct scores for older faces from the current study and their emotion recognition scores from the previous experiment. Percent correct scores from the present study were positively correlated with older adults' emotion recognition scores from the previous experiment, $r(6) = .90$, $p < .01$.

Table 6.2. Older adults rank for the previous VO emotion recognition experiment in order of the best to the worst performer. For comparison purposes ranks have also been assigned to older adults based on their VO production scores. Each row represents a single older adult.

	Rank for the previous VO emotion recognition experiment	Ranks for the VO production scores
High recognition group	1	1
	2	4
	3	2
	4	6
Low recognition group	5	5
	6	3
	7	7
	8	8

Table 6.3 presents the VO production and VO recognition scores across each emotion for each older adult group. For the emotions of angry, disgust, and fear, the High recognition group outperformed the Low recognition group for both production and recognition. Surprise and happy recognition were somewhat similar across the older adult age groups and yet, the ability to produce these emotions differed (i.e., the High recognition group outperformed the Low recognition group). The ability to recognise and produce a sad or neutral emotion appeared to be consistent across the groups.

Table 6.3. Mean percent correct scores (SD) for each VO emotion type from the previous emotion recognition experiment and from the VO portion of the current experiment across the High recognition and Low recognition older groups.

	Emotion type	VO emotion recognition scores (%) from previous experiment	VO production scores (%)
High recognition group	Angry	86 (14)	68 (18)
	Disgust	85 (17)	49 (13)
	Sad	74 (11)	57 (16)
	Fear	56 (15)	46 (19)
	Surprise	52 (20)	62 (20)
	Happy	85 (13)	67 (19)
	Neutral	77 (9)	69 (20)
Low recognition group	Angry	57 (30)	38 (21)
	Disgust	69 (20)	35 (19)
	Sad	68 (18)	55 (19)
	Fear	26 (17)	23 (15)
	Surprise	51 (17)	33 (21)
	Happy	79 (15)	58 (13)
	Neutral	76 (22)	70 (20)

6.5 Discussion

The present study investigated whether, in older adults, the ability to recognise and produce emotion is linked. Four older adults that had relatively high recognition scores and four that had relatively low recognition scores on a previous VO emotion recognition task (as well as eight younger adults) were tested on how well their produced emotions could be recognised.

Overall, the productions of the High recognition group were better recognised than those of the Low recognition group; this was the case for VO, AV, and to some extent, the AO emotion presentations. Further, better performance on the previous recognition task was significantly related to better emotion production recognition in the current study. These findings are in line with previous findings and are consistent with the idea that emotion recognition and emotion production are linked (Adolphs et al., 2000; Deriso et al., 2012; Dimberg, Thunberg, & Elmehed, 2000; Experiment 2, Niedenthal et al., 2001; Oberman et al., 2007; Pitcher et al., 2008).

However, the results of the current study showed that the above association was not clear for all emotions. That is, as anticipated, the productions of the High recognition group were better recognised than those of the Low recognition group for VO angry, disgust, and fear. Unexpectedly, even for happy and surprise faces for which recognition performance was more or less similar across the two older adult groups, the High recognition group produced more recognisable VO productions than the Low recognition group. One way of explaining this incomplete association between perception and production across emotions, is to assume that the primary way that emotions are recognised are via the perceptual recognition system tuned to the input modality (e.g., the visual recognition system) but that the emotion production system can provide additional evidence that can assist recognition. That is, the emotion production system can support the process of emotion recognition, even in healthy recognition systems, but it does not necessarily play a role in the recognition process all the time.

So, how does this view of the role of the emotion production system in the process of emotion recognition help explain the pattern of results? Consider the case of VO expressions of surprise and happy in the current study, where both older groups showed similar recognition levels, but the High recognition groups' emotion production seemed better (indicating a problem in the Low recognition group's emotion production abilities). Here, we presume that these particular emotions can be relatively easily recognised via visual perceptual recognition processes and so both groups perform at a similar level. The recognition of the production of these emotions by the two groups differ because (by presumption) the Low recognition group's ability to produce emotions is compromised. The effect of having a compromised production system only becomes clear for those emotions where backup information from the production system is useful (i.e., emotions where the sensory analysis can be aided by the simulation of the observed emotion). For these cases, the High recognition group (who it is presumed have an intact ability to simulate/produce the to-be-recognised emotion) will do better than the Low recognition group who cannot effectively draw on this resource. What is not clear in this interpretation is a clear specification of the conditions when the knowledge triggered by this simulation is not utilised for emotion recognition. This warrants further investigation.

One reason that the production system of High and Low recognition older adults may be differentially compromised relates to the ability of the brain to resist disease or injury (i.e., cognitive reserve). That is, relative to those with compromised emotion production, older adults with preserved production abilities may have more efficient brain mechanisms due to genetic endowment, and/or have engaged in healthier lifestyle choices (e.g., aerobic exercise, not smoking, etc.) and experiences (e.g., high level of education, intellectually challenging occupations) that may have increased cognitive reserve (see Whalley, Deary, Appleton, & Starr, 2004). In contrast, older adults that do not have a predisposition for efficient neural processing (genes) or did not engage in protective behaviours may be less able to resist disease or injury and, as a result, may show declines on a variety of tasks including emotion production.

The previous eye-tracking/recognition study provides some evidence that the High recognition group may have intact cognitive capacity relative to the Low recognition group. That is, the previous emotion recognition study tested the older adult groups on two cognitive measures; the Trail Making tests that measure processing speed, task switching, and visual attention, and the Digit Span tests that measure working memory ability (see Chapter 3,

Appendix A). For the Trail Making tests, the High recognition older adults appeared to outperform the Low recognition older adults, suggesting that the Low recognition older adults have reduced cognitive functioning (see Table 6.4). In contrast, the High recognition older adults have maintained efficient functioning as their scores were similar to that of the younger adults. Indeed, across a number of conditions in the current study, the High recognition older adults' production scores appeared similar to that of younger adults. Instances where younger expressions were better recognised than that of the High recognition group may be due to peripheral factors or to perceiver bias.

Unlike the scores on the Trail Making test, scores on the Digit Span tests were similar across all age groups and were even somewhat higher for the older adult group relative to the younger adults. These scores provide evidence that the poorer performance in the Low recognition group was not due to some general performance factor (like anxiety or lack of motivation). Further, these older adults are familiar with the experimenter as they participate in studies with the experimenter on a regular basis and the experimenter was not present in the room during the recording. Instead, these older adults appear to be highly motivated as they volunteered to participate in the current study.

Table 6.4. Digit Span and Trail Making test scores for the younger adults, the High recognition older adults and the Low recognition older adults.

	Younger adults	Older High recognition group	Older Low recognition group
Digit Span Raw Score (WAIS-IV)	25.61	27.25	26.00
Trail Making Tests Part A and B (seconds)	26.75	29.00	56.25

It is important to note that the sample size was small for the older adult group and gender was not balanced between the high and low recognition groups. The small sample size and the gender imbalance of the older adult's group was unavoidable as only 8 older adult participants from the previous recognition experiment were able to return to participate in the current study. Future research could address this issue by testing the recognition and production of larger, gender-balanced older adult groups.

In conclusion, emotion recognition and emotion production may be linked but only under circumstances when the expressed emotion is difficult to recognise. This provides further evidence for the notion that simulating an observed emotion can assist with emotion recognition.

6.6 References

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience, 20*(7), 2683-2690.

Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016). *Openface: an open source facial behavior analysis toolkit.* Paper presented at the Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on.

Borod, J. C., Yecker, S. A., Brickman, A. M., Moreno, C. R., Sliwinski, M., Foldi, N. S., . . . Welkowitz, J. (2004). Changes in posed facial expression of emotion across the adult life span. *Experimental Aging Research, 30*(4), 305-331.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(3), 305-327.

Deriso, D. M., Susskind, J., Tanaka, J., Winkielman, P., Herrington, J., Schultz, R., & Bartlett, M. (2012). *Exploring the facial expression perception-production link using real-time automated facial expression recognition.* Paper presented at the European Conference on Computer Vision.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86-89.

Ebner, N. C., He, Y., & 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 and Emotion, 25*(6), 983-997.

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*(1), 351-362.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*(1), 116-124.

Gonçalves, A. R., Fernandes, C., Pasion, R., Ferreira-Santos, F., Barbosa, F., & Marques-Teixeira, J. (2018). Effects of age on the identification of emotions in facial expressions: a meta-analysis. *PeerJ, 6*, e5278.

Hess, U., Adams Jr, 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*(6), 1377-1380.

Kim, J., & Davis, C. (2012). Perceiving emotion from a talker: How face and voice work together. *Visual Cognition, 20*(8), 902-921.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184-211.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition and Emotion, 15*(6), 853-864.

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience, 2*(3-4), 167-178.

Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *Journal of Neuroscience, 28*(36), 8929-8933.

Riediger, M., Voelkle, M. C., Ebner, N. C., & Lindenberger, U. (2011). Beyond “happy, angry, or sad?”: Age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognition and Emotion, 25*(6), 968-982.

Rothauser, E. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17*, 225-246.

Ruffman, T., Halberstadt, J., Murray, J., Jack, F., & Vater, T. (2019). Empathic Accuracy: Worse Recognition by Older Adults and Less Transparency in Older Adult Expressions Compared With Young Adults. *The Journals of Gerontology: Series B*.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*(4), 473-482.

Whalley, L. J., Deary, I. J., Appleton, C. L., & Starr, J. M. (2004). Cognitive reserve and the neurobiology of cognitive aging. *Ageing Research Reviews, 3*(4), 369-382.

Chapter 7. Summary and Conclusions

7.1 Thesis overview

The thesis investigated four specific questions concerning how older adults process emotion information: 1) how well do older adults process expressions and does having multiple sources of expressive information improve recognition? 2) How do older adults extract emotion information from emotional expressions? 3) Where do problems occur during the emotion recognition process? 4) Is poor emotion recognition associated with poor emotion production? To address these research questions, I conducted a series of experiments to measure older adults' emotion perception and production using a variety of research methods including eye-tracking and masked priming techniques, as well as those that are more typically used. To test older adults' emotion recognition within the broad context of communication, and to test stimuli that naturally varied in how well emotion information was conveyed, I also used dynamic (spoken) expressions of emotion.

In the first experiment, I demonstrated that for older adults the provision of multiple sources of emotion information did not always improve recognition performance. This was the case particularly when emotion information was conveyed from one modality that was poorly recognised. The second and third experiments investigated how older adults extract expressive information from faces (i.e., emotion and visual prosodic information respectively). The results indicated that when older adults carried out an emotion recognition task, they adopted an exploratory-gaze strategy; whereas in a prosody judgement task they showed a bias to inspect the eye region. The remaining two experiments investigated the problems that older adults may have in the process of emotion recognition. The first of these experiments suggested that older adults' emotion recognition deficits may be related to problems at an early processing stage. The second experiment revealed that poor emotion recognition was associated with poor emotion production.

Below, these research findings will be briefly discussed under the headings of the initial research questions. Following this, the limits and practical applications of this research will be considered.

7.2 Older adults do not always benefit from AV emotion information

The first experiment investigated whether emotion information from multiple modalities would improve emotion recognition performance for older adults compared to information from only one. The typical finding with younger adults is that AV emotion recognition exceeds that from the best single modality. I found that, in general, older adults did not benefit from AV presentation relative to the best unimodal one. Indeed, for stimuli where AO recognition was poor, AV recognition was worse than the VO condition. These results contrast with previous ones that have reported an AV benefit for older adults; my finding, that AV recognition can suffer when the emotion information from one modality is poor, suggests a reason why. If previous studies only used very clear emotion expressions in each modality, then they may have missed the AV suppression effect that I found (e.g., Hunter, Phillips, & MacPherson, 2010; Wieck & Kunzmann, 2017; Lambrecht, Kreifelts, & Wildgruber, 2012; Chaby, Luherne-du Boullay, Chetouani, & Plaza, 2015).

The finding that in older adults AV emotion recognition can be worse than VO recognition raises the possibility that this could occur in real life face to face communication. That is, older adults' emotion recognition problems may be exacerbated in face to face communication where it is common for emotion to be simultaneously conveyed via the face and voice; and for emotion information to not always be clearly conveyed.

7.3 How do older adults extract emotion information from faces?

To investigate how older adults extract emotion information from faces, I used an eye-tracking technique to measure gaze behaviour during emotion recognition. In line with previous findings, I found that younger adults looked more at the eye than the mouth region of faces (e.g., Sullivan, Ruffman, & Hutton, 2007; although see Chaby, Hupont, Avril, Luherne-du Boullay, & Chetouani, 2017 where differences may be due to the size of the face areas of interest); whereas older adults showed similar amounts of eye and mouth gazing. That is, compared to younger adults, older adults made more fixations and spent more time gazing at the mouth region of faces (Chaby et al., 2017; Sullivan et al., 2007; Murphy & Isaacowitz, 2010). These findings suggested that older adults adopted an exploratory-gaze strategy during emotion recognition; whereas younger adults adopted a focused-gaze strategy towards the eye region of faces. The idea that older adults adopt an exploratory gaze strategy is in direct contrast to the suggestion made by Chaby et al. (2017) and was investigated by testing older adults' gaze behaviour for other types of expressions that required more eye gazing than mouth gazing for accurate recognition. That is, to determine whether older adults' exploratory-gaze strategy was specific to emotion processing, I conducted a follow-up experiment to examine gaze behaviour in a prosody recognition task. I found that older adults showed similar gaze behaviour to younger adults (i.e., more eye than mouth gazing) for the linguistic recognition task (this finding is in line with Circelli, Clark, & Cronin-Golomb, 2013, that reported that gaze behaviour did not differ for non-face stimuli such as landscapes), suggesting that the exploratory-gaze behaviour is adopted during emotion recognition only.

One reason that older adults may adopt an exploratory-gaze strategy during emotion recognition is because they may experience changes in the way they evaluate emotion information from the eye region. That is, older adults may no longer be picking up more emotion information from the eye region relative to the mouth region and so, may look elsewhere for emotion information.

It is important to note that the emotion recognition task and the linguistic recognition task differed in terms of the number of categories of facial expression tested: the linguistic expression task presented two expression categories, whereas the emotion recognition task presented six. As such, differences in older adults' gaze behaviour between the two studies may be due to the number of response options available. That is, during the linguistic recognition task older adults may learn to focus their gaze more towards the eye region because there are only two expression categories and the eye region contains important cues for recognising both. In contrast, the emotion recognition task contains six emotion expression categories that may differ in where on the face important emotion information is presented. In response to this more complex mix of where diagnostic emotion information is located, older adults may adopt a balanced looking strategy. Younger adults, on the other hand, may be better at distilling out a gaze strategy for inspecting multiple facial expressions, and so may engage in focused-gaze behaviour.

To tease apart this interaction between the content of the expression (i.e., emotion vs. linguistic) and the task (i.e., two expression categories vs. six), future research could measure older adults gaze behaviour during the recognition of only two emotions where important emotion information is presented via the eye region. Such expressions would replicate the task conditions of the linguistic experiment and test the impact of task constraints on gaze behaviour. If older adults show more eye than mouth gazing, then this would suggest that older adults are able to realise the most appropriate gaze behaviour under relatively simple conditions. If older adults show more balanced gaze behaviour (i.e., an exploratory gaze behaviour), then this may indicate that older adults are not able realise the most appropriate gaze behaviour for emotion recognition regardless of task demands.

7.4 Where do problems occur during the emotion recognition process?

To assess the extent of early processing in older adults, I used a masked priming paradigm. The results suggested that the emotion processing system of older adults is slow to process angry faces and may produce ambiguous emotion information. This finding stands in contrast to previous studies that indicate that older adults are able to rapidly detect emotion in faces to the same extent as younger adults (e.g., Mather & Knight, 2006; Ruffman, Ng, & Jenkin, 2009). I suggested that paradigms used by previous research (e.g., visual search tasks) may not accurately measure early levels of emotion processing; instead, the masked priming paradigm may be a more sensitive measure of early emotion processing. I also found that initial processing problems may not impact emotion recognition performance when the task is simple (e.g., deciding between two response options), but may impede recognition when the task is complex (e.g., deciding between six response options).

In this study, I investigated angry and happy emotion expressions only. I will follow up this in future research by testing other emotion expressions. I predict that, for instance, the processing system may be relatively intact for disgust expressions given that disgust recognition appears to be maintained in older adulthood.

7.5 Poor emotion recognition is associated with poor emotion production

The final experiment investigated whether emotion recognition was related to emotion production in older adults. I recruited eight older adults that had participated in a previous emotion recognition task; four older adults had relatively high recognition accuracy (the High recognition group) and four had relatively low accuracy (the Low recognition group). These eight older adults, as well as eight younger adults, were video and audio recorded producing emotion. Recordings were then presented to a group of younger adults in a forced-choice emotion recognition task. I found that expressions from the younger adult group and the High recognition group were more easily recognised than expressions from the Low recognition group. These results suggest that emotion recognition is related to emotion production in older adults and supports recent theories suggesting that emotion production and recognition are linked (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005).

Given that this study investigated the recognition of VO expressions only, future research could specifically investigate the relationship between the recognition and production of AO and AV expressions.

7.6 Limits

The current thesis used dynamic, multi-modal spoken expressions of emotion as these expressions differ from static expressions typically used in emotion recognition studies in that they are closer to those encountered in real-world emotion recognition scenarios. However, there are many other aspects of real-world emotion expressions that were not considered. For instance, the meaning of the sentences used in the production of the emotion expressions did not convey any emotional information or contextual information. Such information may improve older adults' emotion recognition performance, and, as such, it could be argued that expressions (such as those used in the current thesis) over-estimate older adults' problems during real-world emotion recognition. This warrants future investigations, particularly given that adding emotion information does not always improve emotion recognition performance (see the AV experiment in Chapter 2).

The studies in the current thesis were conducted using acted expressions of emotion (instead of genuine expressions of emotion). Given that eliciting genuine expressions would raise ethics issues, the emotional stimuli in the current study were produced in a controlled laboratory setting where participants were asked to imagine each emotional state. One could raise a question about the validity of using such acted expressions, perhaps arguing that such may be more prototypical than genuine expressions (Isaacowitz & Stanley, 2011) or that acted expressions may not provide as much emotion information as genuine expressions. However, the exact nature of what differs between acted and genuine expressions is unclear. That is, even though actors may not "feel" the required emotion, it is likely that they can identify what a specific emotion expression involves and produce it. Indeed, the younger adults in the studies reported in the current thesis were able to accurately recognise the emotion expressions.

7.7 Practical applications

One of my long-term aims for investigating emotion processing in older adults is to develop tools to assess their emotion recognition ability. Currently, standard assessments of emotion recognition typically present unimodal (often visual) emotion expressions of emotion to participants (e.g., The Diagnostic Analysis of Nonverbal Accuracy, DANVA, Nowicki & Duke, 1994; Facial Expressions of Emotion: Stimuli and Tests, FEEST, Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). Such assessments do not take into account the variety of expressions older adults may encounter (e.g., multi-modal expressions, clearly recognisable and ambiguous expressions) and thus may not accurately estimate older adults' emotion recognition ability and reflect everyday emotion recognition problems. Further, these assessments provide a description of older adults' emotion recognition problems but do not assess the nature of the problems older adults may have during the emotion recognition process. In the future, assessment tools could include a variety of expressions similar those tested in the current thesis. Further, assessment tools could also include research paradigms to assess a variety of processes required for emotion recognition. For instance, the masked priming paradigm could be incorporated as an assessment of initial processing in older adults.

7.8 Concluding remarks

The overall aim of the thesis was to investigate older adults' emotion recognition ability. The experimental program provided evidence that adding information does not always improve emotion recognition performance; that the way older adults may extract emotion information from faces differs to that of younger adult; and that emotion recognition problems may be due to deficits with a number of different processes during the emotion recognition task.

The research presented in this thesis is an important step towards understanding how older adults process emotion information. Such research can provide the foundation for the development of tools to improve emotion recognition in older adults.

7.9 References

Chaby, L., Hupont, I., Avril, M., Luherne-du Boullay, V., & Chetouani, M. (2017). Gaze behavior consistency among older and younger adults when looking at emotional faces. *Frontiers in Psychology, 8*, 548.

Chaby, L., Luherne-du Boullay, V., Chetouani, M., & Plaza, M. (2015). Compensating for age limits through emotional crossmodal integration. *Frontiers in Psychology, 6*, 691.

Circelli, K. S., Clark, U. S., & Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in aging. *Aging, Neuropsychology, and Cognition, 20*(2), 148-173.

Hunter, E. M., Phillips, L. H., & MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychology and Aging, 25*(4), 779-787.

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*(4), 261-278.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*(3), 529-539.

Mather, M., & Knight, M. R. (2006). Angry faces get noticed quickly: Threat detection is not impaired among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(1), P54-P57.

Murphy, N. A., & Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognising emotional expressions: An in-depth look at gaze measures and covariates. *Cognition and Emotion, 24*(3), 436-452.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184-211.

Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal Behavior, 18*(1), 9-35.

Ruffman, T., Ng, M., & Jenkin, T. (2009). Older adults respond quickly to angry faces despite labeling difficulty. *Journals of Gerontology: Series B, 64*(2), 171-179.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology: Series B, 62*(1), 53-60.

Wieck, C., & Kunzmann, U. (2017). Age differences in emotion recognition: A question of modality? *Psychology and Aging, 32*(5), 401-411.

Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion: Stimuli and Tests (FEEST).* Bury St. Edmunds, England: Thames Valley Test Company.