

PERCEPTION AND COGNITION: INTERACTIONS IN THE AGEING BRAIN

EDITED BY: Harriet A. Allen and Katherine L. Roberts
PUBLISHED IN: Frontiers in Aging Neuroscience





frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-937-2

DOI 10.3389/978-2-88919-937-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PERCEPTION AND COGNITION: INTERACTIONS IN THE AGEING BRAIN

Topic Editors:

Harriet A. Allen, University of Nottingham, UK

Katherine L. Roberts, University of Warwick, UK



The cover is an image from Scarborough beach, 1956. We are grateful to all the participants who have given up their time to take part in research.

Copyright (c) 1956 G Barron.

Permission granted to reproduce, intact and without distortion or editing of the image, for personal and educational use only.

Healthy ageing can lead to declines in both perceptual and cognitive functions. Impaired perception, such as that resulting from hearing loss or reduced visual or tactile resolution, increases demands on 'higher-level' cognitive functions to cope or compensate. It is possible, for example, to use focused attention to overcome perceptual limitations. Unfortunately, cognitive functions also decline in old age. This can mean that perceptual impairments are exacerbated by cognitive decline, and vice versa, but also means that interventions aimed at one type of decline can lead to improvements in the other. Just as improved cognition can ameliorate perceptual deficits, improving the stimulus can help offset cognitive deficits. For example, making directions and routes easy to follow can help compensate for declines in navigation abilities.

In this Topic, we bring together papers from both auditory and visual researchers that address the interaction between perception and cognition in the ageing brain. Many of the studies demonstrate that a broadening of representations or increased reliance on gist underlie perceptual and cognitive age-related declines. There is also clear evidence that impaired perception is associated with poor cognition although, encouragingly, it can also be seen that good perception is associated with better cognition. Compensatory cognitive strategies were less successful in improving perception than might be expected. We also present papers which highlight important methodological considerations that are required when studying the older brain.

Citation: Allen, H. A., Roberts, K. L., eds. (2016). Perception and Cognition: Interactions in the Ageing Brain. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-937-2

Table of Contents

06 Editorial: Perception and Cognition: Interactions in the Aging Brain

Harriet A. Allen and Katherine L. Roberts

09 Perception and Cognition in the Ageing Brain: A Brief Review of the Short- and Long-Term Links between Perceptual and Cognitive Decline

Katherine L. Roberts and Harriet A. Allen

Section: Impaired perception and impaired cognition

16 The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners

Tim Schoof and Stuart Rosen

30 Cognitive spare capacity in older adults with hearing loss

Sushmit Mishra, Stefan Stenfelt, Thomas Lunner, Jerker Rönnerberg and Mary Rudner

43 The effect of functional hearing loss and age on long- and short-term visuospatial memory: evidence from the UK biobank resource

Jerker Rönnerberg, Staffan Hygge, Gitte Keidser and Mary Rudner

56 What's on TV? Detecting age-related neurodegenerative eye disease using eye movement scanpaths

David P. Crabb, Nicholas D. Smith and Haogang Zhu

66 Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition

Christian Füllgrabe, Brian C. J. Moore and Michael A. Stone

Section: Common Decline

91 The ups and downs of global motion perception: a paradoxical advantage for smaller stimuli in the aging visual system

Claire V. Hutchinson, Tim Ledgeway and Harriet A. Allen

99 Preserved fine-tuning of face perception and memory: evidence from the own-race bias in high- and low-performing older adults

Jessica Komes, Stefan R. Schweinberger and Holger Wiese

109 Reduced audiovisual recalibration in the elderly

Yu Man Chan, Michael J. Pianta and Allison M. McKendrick

118 How does aging affect the types of error made in a visual short-term memory 'object-recall' task?

Raju P. Sapkota, Ian van der Linde and Shahina Pardhan

127 The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration

David P. McGovern, Eugenie Roudaia, John Stapleton, T. Martin McGinnity and Fiona N. Newell

Limited Benefit from Compensatory Cognition

136 *A rightward shift in the visuospatial attention vector with healthy aging*

Christopher S. Y. Benwell, Gregor Thut, Ashley Grant and Monika Harvey

147 *The drive-wise project: driving simulator training increases real driving performance in healthy older drivers*

Gianclaudio Casutt, Nathan Theill, Mike Martin, Martin Keller and Lutz Jäncke

161 *Age-related increases in false recognition: the role of perceptual and conceptual similarity*

Laura M. Pidgeon and Alexa M. Morcom

178 *Holistic face perception in young and older adults: effects of feedback and attentional demand*

Bozana Meinhardt-Injac, Malte Persike and Günter Meinhardt

Methodological considerations

191 *Spread of activation and deactivation in the brain: does age matter?*

Brian A. Gordon, Chun-Yu Tse, Gabriele Gratton and Monica Fabiani



Editorial: Perception and Cognition: Interactions in the Aging Brain

Harriet A. Allen^{1*} and Katherine L. Roberts²

¹ School of Psychology, University of Nottingham, Nottingham, UK, ² Department of Psychology, University of Warwick, Coventry, UK

Keywords: aging, perception, cognition, vision, audition

The Editorial on the Research Topic

Perception and Cognition: Interactions in the Aging Brain

Healthy aging can lead to declines in both perceptual and cognitive functions. Many of the studies in this Topic demonstrate such age-related declines, but also identify links between them. Encouragingly, these links suggest that improving perception could benefit cognition. In addition, while compensatory cognitive strategies were mainly unsuccessful in improving perception, cognitive training was effective under certain conditions.

COMMON AGE-RELATED DECLINE

Cognitive and perceptual change may be linked because they are susceptible to the same age-related factors. In the present Topic, several studies suggest an age-related widening of tuning or a decrease of inhibition between representations. Both Chan et al. and McGovern et al. find that older adults are more likely to judge sound and lights presented asynchronously as synchronous. At the same time, Sapkota et al. find that older adults report more distractor items in a memory test. Pidgeon and Morcom also find that older adults are more susceptible to intrusions from distractor items, particularly if they are conceptually or perceptually similar to the target item. Similarly, Meinhardt-Injac et al. find that the unattended half of a face has a bigger influence on older (compared to younger) adults' judgments of face identity. These have some conceptual similarity with cognitive researchers' suggestions of age-related loss of focus, increased distractibility, and increased effects of similarity. For example, increased interference from distractors could be due to increased ambiguity in the way that target items are represented in the brain. In perceptual regions this could present as widening of tuning of receptors (Hua et al., 2006; Yu et al.; Betts et al., 2007). In higher-level areas this could present as less distinct category representations, leading to a higher reliance on "gist."

The results of Hutchinson et al. and Komes et al. might be considered inconsistent with this broadening of tuning account. Hutchinson's finding of enhanced motion perception in older people could be interpreted as reflecting an age-related narrowing, rather than broadening, of tuning, while Komes' results indicate that poor memory is not associated with poor perceptual representations. In both cases, further research will help discriminate between a purely perceptual account and explanations that rely on attention and strategy changes. It therefore seems that while widening of tuning and more gist-based processing offer an explanation for a wide range of results, it is not yet a comprehensive explanation for all changes that occur with aging. It is not clear, for example, whether higher cognitive processes can offset the effects of perceptual tuning changes.

OPEN ACCESS

Edited by:

Gemma Casadesus,
Kent State University, USA

Reviewed by:

Matthieu P. Boisgontier,
University of Leuven, Belgium

*Correspondence:

Harriet A. Allen
h.a.allen@nottingham.ac.uk

Received: 16 February 2016

Accepted: 25 May 2016

Published: 08 June 2016

Citation:

Allen HA and Roberts KL (2016)
Editorial: Perception and Cognition:
Interactions in the Aging Brain.
Front. Aging Neurosci. 8:130.
doi: 10.3389/fnagi.2016.00130

FROM IMPAIRED PERCEPTION TO IMPAIRED COGNITION

Poor perception may lead to, or exacerbate, cognitive impairment (see Roberts and Allen for a review). As perceptual input gets harder to discriminate, more cognitive effort, or processes are required to decode the incoming signal. This may lead to worse performance by older adults because, effectively, the task they are performing is harder. Mishra et al. attempt to quantify this loss of cognitive resources through their Cognitive Spare Capacity Test.

When perceptual and cognitive deficits co-occur, cognitive skills may appear worse due to knock-on effects of poor perception on cognition. Thus, it might be tempting to ascribe a cognitive cause to what is, in fact, a perceptual deficit. Perceptual deficits can impact on cognition to the extent that a measure of (cognitively controlled) eye movements has potential to be used to diagnose eye disease (Crabb et al.). It is therefore important to fully account for perceptual impairment before concluding that there is a cognitive deficit, as was done in this Topic by Schoof and Rosen and Füllgrabe et al. In these studies, older adults were recruited who, unusually, had normal hearing sensitivity as measured by audiogram. Auditory temporal perception was also evaluated, as it is often impaired in older age and impacts on speech-in-noise perception. Even in these audiometrically-normal adults, the ability to understand speech in noise was predicted by a combination of temporal processing ability and cognitive performance (Füllgrabe et al.). Accounting for perception may therefore need to go beyond perceptual sensitivity to also consider suprathreshold processing (Allen et al., 2010; Füllgrabe et al.).

In the longer term, older adults with hearing impairment can show a faster rate of cognitive decline than those without (Lin et al., 2013). While this could be mediated by a separate factor, such as social isolation (Strawbridge et al., 2000), it may be that continual exposure to impoverished perceptual input leads to decrements in cognitive processes. A test of this idea is offered by Rönnberg et al., who discuss whether continued mismatch between perceptual input and long-term memory (LTM) representations could lead to less efficient LTM function. They find that even performance on a visually presented memory task is correlated with hearing loss, suggesting the impact of perceptual loss may be supramodal (Rönnberg et al.).

These studies suggest that the constant, and cumulative effort of coping with impaired perception could impact substantially on age-related cognitive decline.

COGNITIVE IMPROVEMENTS FROM TRAINING AND PERCEPTUAL INTERVENTIONS

In this Research Topic, while there was evidence for compensatory cognitive strategies, in most cases this led to different, rather than equivalent performance to younger adults. For example, reliance on gist led to increased false recognition (Pidgeon and Morcom), and reduced right-hemispheric dominance led to reduced pseudoneglect on a line

bisection task (Benwell et al.). Komes et al. found that older adults who had more accurate memory for faces had a more bilateral electroencephalographic response than those with less accurate memory, presumably indicating some sort of cognitive compensatory process, but they still performed worse than younger adults.

In contrast, optimizing the perceptual input did, in some cases, lead to unimpaired cognition (e.g., Hutchinson et al.; Schoof and Rosen). This is demonstrated here by Rönnberg et al., who found that the impact of hearing loss on cognition could be mitigated via hearing aids. Although direct comparisons are limited by the difficulty of equating the complexity of perceptual and cognitive tasks, these results certainly put some limits on the extent of possible compensation mechanisms.

Older adults did show a benefit from cognitive training, particularly when that training was optimized for their needs. Casutt et al. improved older adults' on-road driving behavior and cognitive performance through training in a driving simulator. Meinhardt-Injac et al. found that older adults could benefit from feedback on a face-processing task when cognitive demands were low, although they were unable to benefit when cognitive demands were high.

METHODOLOGICAL CONSIDERATIONS

To clarify the relative contributions of perception and cognition, it may be tempting to use perceptual tests in one modality and cognitive tests in another (e.g., Rönnberg et al.; Schoof and Rosen; Füllgrabe et al.). Sometimes this is deliberate, so that cognitive measures will not be confounded by perceptual difficulties, but sometimes it is simply due to an established cognitive test being in a particular modality. It should be noted, though, that cognitive skills are not necessarily supramodal. Cognitive impairments such as neglect can arise in one modality but not another (e.g., Sinnett et al., 2007). Furthermore, some perceptual features are more critical to one modality than another (e.g., spatial location, Roberts et al., 2006, 2009).

It is also important to consider whether outcome measures are optimized to detect age-related effects. Gordon et al. report that functional magnetic resonance imaging measures of the peak and spread of activation are separately informative about older adults' performance on a cognitive task. During performance on a Sternberg task, some age-related effects were only apparent when looking at measures of spread. Komes et al. also found differences in the distribution of (EEG) activation in older adults with good vs. poor memory.

CONCLUSIONS

The link between perception and cognition is not well understood. It is possible that perception and cognition are affected by the same, superordinate cause. How much of age-related change can be explained by gradual widening of tuning or categories is a promising and interesting route for further work. This could also help to improve

older adults' performance on cognitive tasks. For example, where older adults show a broadening of representations or increased reliance on gist, cognitive performance could be improved by providing more distinctive or spatially-separated stimuli (Pidgeon and Morcom; Sapkota et al.).

While broadening of tuning appears to have multimodal effects, other interactions between perception and cognition appear to differ between modalities. Impaired perception may

put a load on cognition, perhaps eventually leading to capacity loss. It is interesting to note, however, that these studies are predominantly from the auditory domain. It remains to be seen whether this generalizes to visual stimuli.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Allen, H. A., Hutchinson, C. V., Ledgway, T., and Gayle, P. (2010). The role of contrast sensitivity in global motion processing deficits in the elderly. *J. Vis.* 10:15. doi: 10.1167/10.10.15
- Betts, L. R., Sekuler, A. B., and Bennet, P. J. (2007). The effects of aging on orientation discrimination. *Vision Res.* 47, 1769–1780. doi: 10.1016/j.visres.2007.02.016
- Hua, T., Li, X., He, L., Zhou, Y., Wang, Y., and Leventhal, A. G. (2006). Functional degradation of visual cortical cells in old cats. *Neurobiol. Aging* 27, 155–162. doi: 10.1016/j.neurobiolaging.2004.11.012
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868
- Roberts, K. L., Summerfield, A. Q., and Hall, D. A. (2006). Presentation modality influences behavioral measures of alerting, orienting, and executive control. *J. Int. Neuropsychol. Soc.* 12, 485–492. doi: 10.1017/S1355617706060620
- Roberts, K. L., Summerfield, A. Q., and Hall, D. A. (2009). Covert auditory spatial orienting: An evaluation of the spatial relevance hypothesis. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1178–1191. doi: 10.1037/a0014249
- Sinnett, S., Montserrat, J., Rafal, R., Azanon, E., and Soto-Faraco, S. (2007). A dissociation between visual and auditory hemi-inattention: evidence from temporal order judgements. *Neuropsychologia* 45, 552–560. doi: 10.1016/j.neuropsychologia.2006.03.006
- Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., and Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: a longitudinal analysis. *Gerontologist* 40, 320–326. doi: 10.1016/j.neuropsychologia.2006.03.006

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Allen and Roberts. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Perception and Cognition in the Ageing Brain: A Brief Review of the Short- and Long-Term Links between Perceptual and Cognitive Decline

Katherine L. Roberts^{1*} and Harriet A. Allen²

¹ Department of Psychology, University of Warwick, Coventry, UK, ² School of Psychology, University of Nottingham, Nottingham, UK

Ageing is associated with declines in both perception and cognition. We review evidence for an interaction between perceptual and cognitive decline in old age. Impoverished perceptual input can increase the cognitive difficulty of tasks, while changes to cognitive strategies can compensate, to some extent, for impaired perception. While there is strong evidence from cross-sectional studies for a link between sensory acuity and cognitive performance in old age, there is not yet compelling evidence from longitudinal studies to suggest that poor perception causes cognitive decline, nor to demonstrate that correcting sensory impairment can improve cognition in the longer term. Most studies have focused on relatively simple measures of sensory (visual and auditory) acuity, but more complex measures of suprathreshold perceptual processes, such as temporal processing, can show a stronger link with cognition. The reviewed evidence underlines the importance of fully accounting for perceptual deficits when investigating cognitive decline in old age.

OPEN ACCESS

Edited by:

Gemma Casadesus,
Kent State University, USA

Reviewed by:

Daniel Ortuño-Sahagún,
Centro Universitario de Ciencias de la
Salud, Mexico

Archana Mukhopadhyay,
University of Kansas, USA

*Correspondence:

Katherine L. Roberts
k.roberts@warwick.ac.uk

Received: 17 September 2015

Accepted: 15 February 2016

Published: 01 March 2016

Citation:

Roberts KL and Allen HA (2016)
Perception and Cognition in the
Ageing Brain: A Brief Review of the
Short- and Long-Term Links between
Perceptual and Cognitive Decline.
Front. Aging Neurosci. 8:39.
doi: 10.3389/fnagi.2016.00039

Keywords: perception, cognition, ageing, audition, vision

INTRODUCTION

It is well known that ageing is associated with declines in both perception and cognition. As we age, there is increased need for perceptual aids such as glasses and hearing aids, and we start to find cognitive tasks such as paying attention and remembering more difficult. It is possible that perception and cognition decline in tandem due to common effects of ageing, but there is also evidence to suggest that declines in perception and cognition impact on each other, both in the short and long term (**Figure 1**). Here we review evidence that impoverished perceptual input can increase the cognitive difficulty of tasks, while changes to cognitive strategies and processes can compensate, to some extent, for impaired perception. We also review evidence that both visual and auditory perceptual impairments are associated with a faster rate of cognitive decline, and consider whether perceptual aids, such as hearing aids and glasses, can provide protection against cognitive decline.

AGE-RELATED DECLINES IN HEARING, VISION AND COGNITION

The ability to hear faint sounds in a quiet environment deteriorates with age. Peripheral hearing sensitivity, as measured by the audiogram, is impaired in around a third of 61–70 year olds

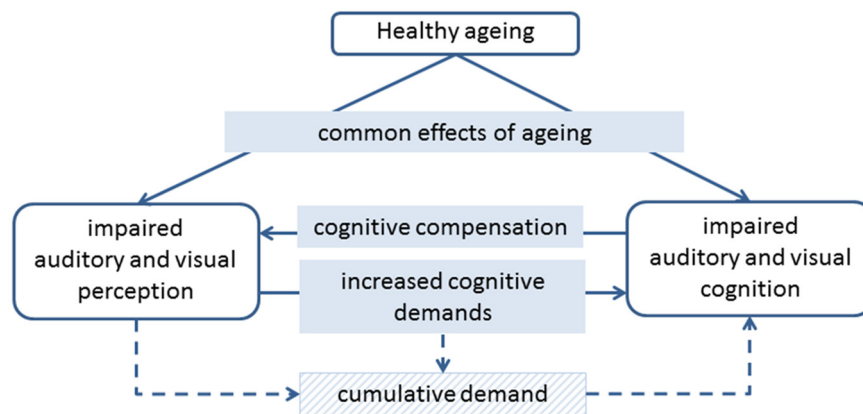


FIGURE 1 | Diagram of the potential links between healthy ageing, auditory and visual perception, and auditory and visual cognition.

and almost two thirds of over-70 year olds (Davis, 1989; Cruickshanks et al., 1998; Wilson et al., 1999). Ageing can also affect suprathreshold auditory processing. Sensitivity to temporal fine structure is impaired in older adults, even in those with normal audiometric thresholds (Grose and Mamo, 2010; Füllgrabe, 2013; Füllgrabe et al., 2015), as is sensitivity to changes in the temporal envelope (Füllgrabe et al., 2015).

Vision also declines in old age. As with hearing, many of these changes are peripheral, and other changes affect central processing. Hardening of the lens leads to presbyopia, which makes it more difficult to focus on near objects. As well as age-related eye diseases such as cataracts, glaucoma and macular degeneration, healthy ageing is linked to a thickening and yellowing of the lens (Said and Weale, 1959; Ruddock, 1965). Ageing is also associated with changes in color perception (Page and Crognale, 2005), temporal resolution (Wright and Drasdo, 1985; Kim and Mayer, 1994; Culham and Kline, 2002), visual acuity (Spear, 1993), motion perception (Snowden and Kavanagh, 2006; Hutchinson et al., 2012), and a loss of fine detail (high spatial frequency) pattern vision (Elliott, 1987; Pardhan, 2004).

Deficits in one sensory modality can sometimes be offset through other senses, e.g., using visual speechreading to support impaired hearing (Dickinson and Taylor, 2011), or increased multisensory integration (Freiherr et al., 2013). This may be limited in older age when both hearing and vision decline, particularly as declines in vision and audition appear to be linked, with higher than expected rates of dual-sensory decline (Dawes et al., 2014).

Old age also brings decline in a number of cognitive abilities, including working memory, memory, attention, and executive control (Schaie, 1996; Hedden and Gabrieli, 2004). Cross-sectional and longitudinal studies consistently find that ageing is related to poor performance on tasks involving cognitive control and switching, manipulation of information, visuospatial processing, processing speed and more (for an extensive list of studies, see Hofer et al., 2003). There are a number of proposed explanations for this decline, including generalized

slowing (e.g., Salthouse, 1996) and dedifferentiation (e.g., Wilson et al., 2012). In this brief review, we focus on the role of sensory function.

A number of cross-sectional, population-based studies have demonstrated a link between (auditory and visual) sensory impairment and poor cognition in old age (Lindenberger and Baltes, 1994; Baltes and Lindenberger, 1997; Lin et al., 2004, 2011; Tay et al., 2006; Moore et al., 2014), that does not simply reflect the visibility or audibility of the task materials (Lindenberger et al., 2001). The association between perception and cognition is also present in younger adults, but is stronger in older adults (Baltes and Lindenberger, 1997). While the evidence for a link between perception and cognition is compelling, it is not universally found. Two studies have found a link between vision and cognition but not hearing and cognition (Anstey et al., 2001; Gussekloo et al., 2005), one study failed to find a link between hearing and cognition (Gennis et al., 1991), and one found no link between visual or auditory function and cognition in a narrow-age cohort sample of 75 year olds (Hofer et al., 2003).

The majority of studies that have attempted to link sensory and cognitive decline have used only visual (or auditory) acuity as a proxy for more general perceptual decline (e.g., Anstey et al., 2001; Hofer et al., 2003). If a common factor underlies age-related changes in performance, then we would expect changes to also affect more central, suprathreshold perceptual skills. However, few studies have used more higher order, cortical measures of perception. Humes et al. (2013) used an index of temporal sensory processing and found a clear link to performance on temporal cognitive tasks. Glass (2007) found that visual contrast sensitivity was a better predictor of performance in cognitive tasks with unfamiliar, multiple or complex stimuli than simple stimuli. These results suggest that going beyond simple sensory acuity may reveal a more direct link to cognitive task performance.

Various hypotheses have been proposed to account for the link between perception and cognition (Table 1; see Wayne and Johnsrude, 2015 for a recent review of these hypotheses in the

TABLE 1 | Proposed hypotheses for the link between perceptual and cognitive decline in old age.

Hypothesis	Description
Common cause	A third, general factor underlies the decline in both perception and cognition (Lindenberger and Baltes, 1994; Baltes and Lindenberger, 1997).
Cognitive load on perception	Poor cognition results in poor performance on perceptual tests (Lindenberger and Baltes, 1994).
Information degradation	Impoverished perceptual input impacts on performance on cognitive tasks (Schneider and Pichora-Fuller, 2000).
Sensory deprivation	Over time, impoverished perceptual input leads to cognitive decline (Lindenberger and Baltes, 1994).

Note that these hypotheses are not mutually exclusive.

auditory modality). Here we consider both vision and audition as well as possible effects of compensatory cognition.

EFFECTS OF AGEING INFLUENCE BOTH PERCEPTION AND COGNITION (“COMMON CAUSE”)

Sensory systems are subject to the same molecular and cellular processes as the rest of the body and as such, any general age-related change will affect both sensory and cognitive mechanisms (e.g., Pathai et al., 2013). Cardiovascular risk factors, for example, are associated with both hearing loss (Shargorodsky et al., 2010) and cognitive decline (Knopman et al., 2001).

The link between sensory and cognitive impairment is typically found after adjusting for age, suggesting that it is not simply an effect of general ageing. In a detailed study, Humes et al. (2013) found that the link between age and cognitive function was entirely mediated by sensory function, based on a composite measure of auditory, visual and tactile perception. It therefore seems that while common effects of ageing do affect both perception and cognition, it is likely that perception and cognition directly impact on each other (Figure 1).

POOR COGNITION AFFECTS PERFORMANCE ON PERCEPTUAL TASKS (“COGNITIVE LOAD ON PERCEPTION”)

Perception and cognition are highly interrelated, such that even measures that might be considered to be entirely sensory, such as the audiogram, have been shown to be influenced by cognition (Zwislocki et al., 1958). More complex perceptual tasks are likely to be more strongly influenced by the participant's cognitive abilities. While it is important that researchers consider the impact of cognition on their perceptual measures, it seems unlikely that poor cognition drives purely perceptual decline. Schneider and Pichora-Fuller (2000) report empirical evidence of this, based on a study that showed age-related impairment for some perceptual tasks, but not others, despite cognitive demands being held constant across conditions (Pichora-Fuller

and Schneider, 1991). In contrast, there is strong evidence that cognition has a compensatory role in adapting to impaired perception (see below and Figure 1). A more likely causal link between perception and cognition suggests that impoverished perceptual input impacts on cognitive function, both directly and in the longer term.

IMPOVERISHED PERCEPTUAL INPUT DIRECTLY IMPACTS COGNITIVE RESOURCES (“INFORMATION DEGRADATION”)

When the perceptual signal is poor, either through degraded stimuli or impaired perception, additional cognitive resources are required to decipher the signal. For example, cognitive load in listening tasks, as measured by the pupil response, is higher for those with hearing loss (Zekveld et al., 2011). This leaves fewer cognitive resources available for performing the cognitive task. A number of studies have demonstrated that recall of target words is impaired if the words are degraded, either through the addition of noise or hearing impairment, even when ensuring that the words can still be identified (e.g., Rabbitt, 1968, 1991; Pichora-Fuller et al., 1995; McCoy et al., 2005; Tun et al., 2009; Ng et al., 2013). This is generally considered to be evidence that the additional effort required to decode degraded stimuli takes up cognitive resources that would otherwise be involved in encoding and rehearsal (Rabbitt, 1968, 1991). Similar effects are found in young and older adults, and it is not yet clear whether ageing confers an additional deficit, over and above that of perceptual loss. Several studies have confounded age and hearing impairment by comparing young, normally-hearing adults with older, hearing-impaired adults (e.g., Pichora-Fuller et al., 1995; Mishra et al., 2014). Verhaegen et al. (2014) found that young and older adults with matched hearing impairment showed similar recall performance on a verbal short-term memory task, which was worse than that found in young adults with normal hearing. On the other hand, the link between working memory capacity and speech-in-noise perception has been shown to be weaker for young adults than normally-hearing older adults (Füllgrabe and Rosen, 2016) who differed in their sensitivity to temporal-fine-structure information (Füllgrabe, 2013).

Recently there has been a concerted effort to describe and quantify the loss of cognitive resources when the perceptual input is degraded (Rudner and Lunner, 2014). This includes development of a “cognitive spare capacity” (CSC) test (Mishra et al., 2013), which is a measure of auditory working memory incorporating both storage and executive processing. Using this task under a variety of listening conditions, Mishra et al. (2014) found that older, hearing-impaired adults had similar CSC to young, normally-hearing adults when listening conditions were optimal. This again suggests that task difficulty in old age can relate to sensory, rather than cognitive, impairment.

It is worth noting that all of the above evidence comes from auditory studies; the impact of impoverished visual input on visual cognition has been less thoroughly explored.

Visual sensory quality has been suggested as an explanation for age-related changes in performance of the Stroop task (Ben-David and Schneider, 2009, 2010), providing a sensory explanation for what has previously been considered a change in (cognitive) inhibition.

There is also evidence that hearing impairment affects performance on visual tasks. Rönnberg et al. (2014) found that older adults with hearing loss had worse performance on visuospatial short- and long-term memory (LTM) tasks. Similarly, an earlier study found a link between hearing loss and LTM function when the LTM task required motor encoding (Rönnberg et al., 2011). These findings point to more general effects of sensory loss than simply increasing the difficulty of cognitive tasks. Rönnberg et al. (2011) hypothesize that LTM representations are used less by people with hearing loss due to a mismatch between the input signal and the stored signal, and that this disuse leads to a decline in LTM over time.

SENSORY IMPAIRMENT LEADS TO COGNITIVE DECLINE (“SENSORY DEPRIVATION”)

A handful of longitudinal studies have shown that impaired perception is associated with cognitive decline, providing support for a causal link between perceptual and cognitive decline over time (Lin et al., 2004, 2013; Ghisletta and Lindenberger, 2005). For example, women with impaired (corrected) vision at baseline had a faster rate of cognitive decline over a 4 year period than those without visual impairment (Lin et al., 2004). In addition, Lin et al. (2013) found that poor hearing at baseline was associated with a higher rate of cognitive impairment and a faster rate of cognitive decline over a 6 year period. In contrast, Anstey et al. (2001) found a link between visual, but not auditory, decline and cognition over a 2 year period, and Valentijn et al. (2005) found that although there was a link between perceptual and cognitive decline, there was no convincing evidence that baseline auditory and visual acuity predicted cognitive decline over the following 6 years. The current evidence is therefore mixed, but it is important to note that only the effects of peripheral sensory changes, such as acuity, have been considered. There appears to be a stronger link with cognition when measures of perception go beyond simple sensory acuity (e.g., Humes et al., 2013).

Several plausible mechanisms have been proposed to explain how impaired perception could lead to worsening cognition over time (e.g., Lin et al., 2013). One possibility is that poor perception leads to social isolation (Strawbridge et al., 2000), which in turn leads to cognitive decline. However, Dawes et al. (2015) found that while hearing-aid use was associated with better cognition, this was not mediated by social isolation.

A further possibility is that performance on cognitive tasks is reduced by the ongoing effort from sensory deprivation (cf. Rönnberg et al., 2011). In this case, the ongoing effort puts a strain on cognition, which eventually leads to performance breakdown. This hypothesis predicts that short-term improvements in stimulus quality (or perceptual

abilities) will have limited effectiveness, but that longer-term improvements in perception will help maintain cognition.

There is not yet convincing evidence to suggest that hearing aids or glasses offer protection against cognitive decline. Hearing-aid use is generally associated with better cognition (Lin, 2011; Rönnberg et al., 2014; Dawes et al., 2015), but it is possible that the people who seek treatment differ from those who do not, both cognitively and functionally. Two studies have shown no cognitive benefit from cataract surgery (Hall et al., 2005; Valentijn et al., 2005) or hearing aids (Valentijn et al., 2005). A randomized controlled trial of hearing-aid use showed a positive change in communication, social and emotional function, and depression, but no longer-term improvement in cognition (Mulrow et al., 1990, 1992). There is therefore no compelling evidence that correcting for peripheral sensory deficits improves cognition.

COGNITION CAN COMPENSATE FOR THE EFFECTS OF IMPOVERISHED PERCEPTUAL INPUT

While a lack-of-use explanation may account well for effects of poor perception on memory function, it is less clear how this might extend to other cognitive skills such as executive control. It has been widely established that older adults use compensatory cognitive strategies to deal with, among other things, impoverished perceptual input. It is not yet clear if perception-related cognitive decline occurs as a function of the increased cognitive demands, or despite them.

Older adults can engage compensatory cognitive processes in order to perform at a similar level to younger adults. This is reflected in different patterns of cortical activity (Reuter-Lorenz and Lustig, 2005) and is associated with better performance compared to older adults who show less compensatory activation (Cabeza et al., 2002). For example, when listening to speech, older adults showed reduced activation in auditory cortex, but increased activation in prefrontal and precuneus regions associated with working memory and attention (Wong et al., 2009). The increased activation in cognitive regions was associated with better behavioral performance, although older adults were still at a disadvantage relative to younger adults when the signal-to-noise ratio was unfavorable.

Few studies have investigated the direct effect of sensory degradation on compensatory mechanisms. There is some evidence that increasing the difficulty of encoding the perceptual task can lead to increased compensatory activity. For example, Schulte et al. (2011) manipulated perceptual and cognitive demands in a match-to-sample Stroop task. Older adults showed a different pattern of fronto-parietal and visuomotor activation, indicating recruitment of additional regions to cope with increased cognitive and perceptual difficulty. This suggests that increasing visual complexity increases the requirement for compensatory activity.

Recruitment of additional cortical regions is usually linked to maintained performance on cognitive tasks (e.g., Schulte et al., 2011). As task difficulty increases, however, performance

can collapse. In a letter memory task, older adults had similar performance to younger adults, but showed widespread additional cortical activations (in comparison to younger adults). As task difficulty increased, younger adults also showed these additional activations. At the highest task difficulty the performance of older adults declined, as did their brain activations (Schneider-Garces et al., 2010). This is consistent with a peak or ceiling for performance, which falls with age. The longer term consequences of repeatedly reaching this ceiling are not yet clear.

CONCLUSIONS AND FUTURE DIRECTIONS

There is clear evidence for a link between perception and cognition in old age, in terms of both their impact on task performance and their age-related decline. While there are clearly common and general factors acting on both sensory and cognitive decline, there also appears to be a more direct link between impaired perception and cognitive decline. Degraded input leads to a higher load on cognition, reducing resources available for cognitive processing. It has been proposed that, over time, this sensory deprivation leads to cognitive decline. At the same time, compensatory processes have been shown that allow people to ameliorate the effects of age-related perceptual decline.

There is therefore something of a paradox, that while staying cognitively active can protect against cognitive decline in old age (Hultsch et al., 1999), needing to be more cognitively active to overcome poor perceptual input is associated with cognitive decline. It should be noted that the majority of evidence for the sensory deprivation hypothesis comes from the auditory

domain, whereas evidence for compensatory processes comes predominantly from the visual domain. Future research in both domains could draw on findings from the other modality.

A further avenue for future work is to consider what is evaluated when assessing sensory abilities. Where perception has been linked to cognition, typically only sensory acuity has been assessed. Where measures have gone beyond simple sensory acuity, a clearer link to cognitive change can often be seen.

Improving the quality of the perceptual input should reduce cognitive impairment in the short term and reduce the involvement of compensatory mechanisms. In the longer term, there is little suggestion that cataract surgery and hearing aids can improve cognition. This may simply reflect limits on the number of longitudinal studies and their follow-up periods, but it may also indicate that interventions should be aimed at more central perceptual processes. An alternative approach would be to investigate whether improving cognition through (for example) brain training could have a beneficial effect on auditory or visual perception. Differentiating between stimulation-based interventions and compensatory cognitive training (Kim and Kim, 2014) could provide further insights into the mechanism that links perceptual and cognitive decline in older age.

AUTHOR CONTRIBUTIONS

KLR and HAA reviewed the literature and wrote the article.

ACKNOWLEDGMENTS

Thanks to Christian Füllgrabe for helpful discussion and comments on the manuscript.

REFERENCES

- Anstey, K. J., Luszcz, M. A., and Sanchez, L. (2001). Two-year decline in vision but not hearing is associated with memory decline in very old adults in a population-based sample. *Gerontology* 47, 289–293. doi: 10.1159/000052814
- Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12
- Ben-David, B. M., and Schneider, B. A. (2009). A sensory origin for color-word Stroop effects in aging: a meta-analysis. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 16, 505–534. doi: 10.1080/13825580902855862
- Ben-David, B. M., and Schneider, B. A. (2010). A sensory origin for color-word Stroop effects in aging: simulating age-related changes in color-vision mimics age-related changes in Stroop. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 17, 730–746. doi: 10.1080/13825585.2010.510553
- Cabeza, R., Anderson, N. D., Locantore, J. K., and McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., et al. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The epidemiology of hearing loss study. *Am. J. Epidemiol.* 148, 879–886. doi: 10.1093/oxfordjournals.aje.a009713
- Culham, J. C., and Kline, D. W. (2002). The age deficit on photopic counterphase flicker: contrast, spatial frequency and luminance effects. *Can. J. Exp. Psychol.* 56, 177–186. doi: 10.1037/h0087395
- Davis, A. C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *Int. J. Epidemiol.* 18, 911–917. doi: 10.1093/ije/18.4.911
- Dawes, P., Dickinson, C., Emsley, R., Bishop, P. N., Cruickshanks, K. J., Edmondson-Jones, M., et al. (2014). Vision impairment and dual sensory problems in middle age. *Ophthalmic Physiol. Opt.* 34, 479–488. doi: 10.1111/opo.12138
- Dawes, P., Emsley, R., Cruickshanks, K. J., Moore, D. R., Fortnum, H., Edmondson-Jones, M., et al. (2015). Hearing loss and cognition: the role of hearing aids, social isolation and depression. *PLoS One* 10:e0119616. doi: 10.1371/journal.pone.0119616
- Dickinson, C. M., and Taylor, J. (2011). The effect of simulated visual impairment on speech-reading ability. *Ophthalmic Physiol. Opt.* 31, 249–257. doi: 10.1111/j.1475-1313.2010.00810.x
- Elliott, D. B. (1987). Contrast sensitivity decline with aging: a neural or optical phenomenon. *Ophthalmic Physiol. Opt.* 7, 415–419. doi: 10.1016/0275-5408(87)90065-2
- Freiherr, J., Lundström, J. N., Habel, U., and Reetz, K. (2013). Multisensory integration mechanisms during aging. *Front. Hum. Neurosci.* 7:863. doi: 10.3389/fnhum.2013.00863
- Füllgrabe, C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 22, 313–315. doi: 10.1044/1059-0889(2013)12-0070
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347

- Füllgrabe, C., and Rosen, S. (2016). "Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing," in *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, eds P. van Dijk, D. Baskent, E. Gaudrain, E. de Kleine, A. Wagner, and C. Lanting (Heidelberg: Springer), 29–36.
- Gennis, V., Garry, P. J., Haaland, K. Y., Yeo, R. A., and Goodwin, J. S. (1991). Hearing and cognition in the elderly. New findings and a review of the literature. *Arch. Intern. Med.* 151, 2259–2264. doi: 10.1001/archinte.151.11.2259
- Ghisletta, P., and Lindenberger, U. (2005). Exploring structural dynamics within and between sensory and intellectual functioning in old and very old age: longitudinal evidence from the berlin aging study. *Intelligence* 33, 555–587. doi: 10.1016/j.intell.2005.07.002
- Glass, J. M. (2007). Visual function and cognitive aging: differential role of contrast sensitivity in verbal versus spatial tasks. *Psychol. Aging* 22, 233–238. doi: 10.1037/0882-7974.22.2.233
- Grose, J. H., and Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear Hear.* 31, 755–760. doi: 10.1097/AUD.0b013e3181e627e7
- Gussekloo, J., de Craen, A. J. M., Oude, C., van Boxtel, M. P. J., and Westendorp, R. G. J. (2005). Sensory impairment and cognitive functioning in oldest-old subjects: the Leiden 85+ Study. *Am. J. Geriatr. Psychiatry* 13, 781–786. doi: 10.1176/appi.ajgp.13.9.781
- Hall, T. A., McGwin, G., Jr., and Owsley, C. (2005). Effect of cataract surgery on cognitive function in older adults. *J. Am. Geriatr. Soc.* 53, 2140–2144. doi: 10.1111/j.1532-5415.2005.00499.x
- Hedden, T., and Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nat. Rev. Neurosci.* 5, 87–97. doi: 10.1038/nrn1323
- Hofer, S. M., Berg, S., and Era, P. (2003). Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance and cognitive functioning. *Psychol. Aging* 18, 285–305. doi: 10.1037/0882-7974.18.2.285
- Hultsch, D. F., Hertzog, C., Small, B. J., and Dixon, R. A. (1999). Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging? *Psychol. Aging* 14, 245–263. doi: 10.1037/0882-7974.14.2.245
- Humes, L. E., Busey, T. A., Craig, J., and Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten. Percept. Psychophys.* 75, 508–524. doi: 10.3758/s13414-012-0406-9
- Hutchinson, C. V., Arena, A., Allen, H. A., and Ledgeway, T. (2012). Psychophysical correlates of global motion processing in the aging visual system: a critical review. *Neurosci. Biobehav. Rev.* 36, 1266–1272. doi: 10.1016/j.neubiorev.2012.02.009
- Kim, E. Y., and Kim, K. W. (2014). A theoretical framework for cognitive and non-cognitive interventions for older adults: stimulation versus compensation. *Aging Mental Health* 18, 304–315. doi: 10.1080/13607863.2013.868404
- Kim, C. B. Y., and Mayer, M. J. (1994). Foveal flicker sensitivity in healthy aging eyes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 11, 1958–1969. doi: 10.1364/josaa.11.001958
- Knopman, D., Boland, L. L., Mosley, T., Howard, G., Liao, D., Szklo, M., et al. (2001). Cardiovascular risk factors and cognitive decline in middle-aged adults. *Neurology* 56, 42–48. doi: 10.1212/wnl.56.1.42
- Lin, F. R. (2011). Hearing loss and cognition among older adults in the United States. *J. Gerontol. A Biol. Sci. Med. Sci.* 66A, 1131–1136. doi: 10.1093/gerona/glr115
- Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., and Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore longitudinal study of ageing. *Neuropsychology* 25, 763–770. doi: 10.1037/a0024238
- Lin, F. R., Gutierrez, P. R., Stone, K. L., Yaffe, K., Ensrud, K. E., Fink, H. A., et al. (2004). Vision impairment and combined vision and hearing impairment predict cognitive and functional decline in older women. *J. Am. Geriatr. Soc.* 52, 1996–2002. doi: 10.1111/j.1532-5415.2004.52554.x
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868
- Lindenberger, U., and Baltes, P. B. (1994). Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339–355. doi: 10.1037/0882-7974.9.3.339
- Lindenberger, U., Scherer, H., and Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: not due to sensory reductions operating during cognitive assessment. *Psychol. Aging* 16, 196–205. doi: 10.1037/0882-7974.16.2.196
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A* 58A, 22–33. doi: 10.1080/02724980443000151
- Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013). Visual information can hinder working memory processing of speech. *J. Speech Lang. Hear. Res.* 56, 1120–1132. doi: 10.1044/1092-4388(2012)12-0033
- Mishra, S., Stenfelt, S., Lunner, T., Rönnberg, J., and Rudner, M. (2014). Cognitive spare capacity in older adults with hearing loss. *Front. Aging Neurosci.* 6:96. doi: 10.3389/fnagi.2014.00096
- Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., et al. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS One* 9:e107720. doi: 10.1371/journal.pone.0107720
- Mulrow, C. D., Aguilar, C., Endicott, J. E., Tuley, M. R., Velez, R., Charlip, W. S., et al. (1990). Quality-of-life changes and hearing impairment: a randomized trial. *Ann. Intern. Med.* 113, 188–194. doi: 10.7326/0003-4819-113-3-188
- Mulrow, C. D., Tuley, M. R., and Aguilar, C. (1992). Sustained benefits of hearing-aids. *J. Speech Hear. Res.* 35, 1402–1405. doi: 10.1044/jshr.3506.1402
- Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181
- Page, J. W., and Crognale, M. A. (2005). Differential aging of chromatic and achromatic visual pathways: behavior and electrophysiology. *Vision Res.* 45, 1481–1489. doi: 10.1016/j.visres.2004.09.041
- Pardhan, S. (2004). Contrast sensitivity loss with aging: sampling efficiency and equivalent noise at different spatial frequencies. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 21, 169–175. doi: 10.1364/josaa.21.000169
- Pathai, S., Shiels, P. G., Lawn, S. D., Cook, C., and Gilbert, C. (2013). The eye as a model of ageing in translational research - molecular, epigenetic and clinical aspects. *Ageing Res. Rev.* 12, 490–508. doi: 10.1016/j.arr.2012.11.002
- Pichora-Fuller, M. K., and Schneider, B. A. (1991). Masking-level differences in the elderly: a comparison of antiphasic and time-delay dichotic conditions. *J. Speech Hear. Res.* 34, 1410–1422. doi: 10.1044/jshr.3406.1410
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Rabbitt, P. M. A. (1968). Channel-capacity intelligibility and immediate memory. *Q. J. Exp. Psychol. A* 20, 241–248. doi: 10.1080/14640746808400158
- Rabbitt, P. (1991). Mild hearing-loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol. Suppl.* 476, 167–176. doi: 10.3109/00016489109127274
- Reuter-Lorenz, P. A., and Lustig, C. (2005). Brain aging: reorganizing discoveries about the aging mind. *Curr. Opin. Neurobiol.* 15, 245–251. doi: 10.1016/j.conb.2005.03.016
- Rönnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternäng, O., Wahlin, Å., et al. (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *J. Speech Lang. Hear. Res.* 54, 705–726. doi: 10.1044/1092-4388(2010)09-0088
- Rönnberg, J., Hygge, S., Keidser, G., and Rudner, M. (2014). The effect of functional hearing loss and age on long- and short-term visuospatial memory: evidence from the UK biobank resource. *Front. Aging Neurosci.* 6:326. doi: 10.3389/fnagi.2014.00326
- Ruddock, K. H. (1965). The effect of age upon colour vision II. Changes with age in light transmission of the ocular media. *Vision Res.* 5, 47–58. doi: 10.1016/0042-6989(65)90074-x
- Rudner, M., and Lunner, T. (2014). Cognitive spare capacity and speech communication: a narrative overview. *Biomed Res. Int.* 2014:869726. doi: 10.1155/2014/869726
- Said, F. S., and Weale, R. A. (1959). The variation with age of the spectral transmissivity of the living human crystalline lens. *Gerontologia* 3, 213–231. doi: 10.1159/000210900
- Salhouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295x.103.3.403

- Schaie, K. W. (1996). *Intellectual Development in Adulthood: The Seattle Longitudinal Study*. Cambridge: Cambridge University Press.
- Schneider, B. A., and Pichora-Fuller, M. K. (2000). "Implications of perceptual deterioration for cognitive aging research," in *The Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (London: Lawrence Erlbaum Associates), 155–219.
- Schneider-Garces, N. J., Gordon, B. A., Brumback-Peltz, C. R., Shin, E., Lee, Y., Sutton, B. P., et al. (2010). Span, CRUNCH and beyond: working memory capacity and the aging brain. *J. Cogn. Neurosci.* 22, 655–669. doi: 10.1162/jocn.2009.21230
- Schulte, T., Müller-Oehring, E. M., Chanraud, S., Rosenbloom, M. J., Pfefferbaum, A., and Sullivan, E. V. (2011). Age-related reorganization of functional networks for successful conflict resolution: a combined functional and structural MRI study. *Neurobiol. Aging* 32, 2075–2090. doi: 10.1016/j.neurobiolaging.2009.12.002
- Shargorodsky, J., Curhan, S. G., Eavey, R., and Curhan, G. C. (2010). A prospective study of cardiovascular risk factors and incident hearing loss in men. *Laryngoscope* 120, 1887–1891. doi: 10.1002/lary.21039
- Snowden, R. J., and Kavanagh, E. (2006). Motion perception in the ageing visual system: minimum motion, motion coherence and speed discrimination thresholds. *Perception* 35, 9–24. doi: 10.1068/p5399
- Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision Res.* 33, 2589–2609. doi: 10.1016/0042-6989(93)90218-1
- Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., and Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: a longitudinal analysis. *Gerontologist* 40, 320–326. doi: 10.1093/geront/40.3.320
- Tay, T., Wang, J. J., Kifley, A., Lindley, R., Newall, P., and Mitchell, P. (2006). Sensory and cognitive association in older persons: findings from an Australian population. *Gerontology* 52, 386–394. doi: 10.1159/000095129
- Tun, P. A., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity and the attentional costs of effortful listening. *Psychol. Aging* 24, 761–766. doi: 10.1037/a0014802
- Valentijn, S. A. M., van Boxtel, M. P., van Hooren, S. A., Bosma, H., Beckers, H. J., Ponds, R. W., et al. (2005). Change in sensory functioning predicts change in cognitive functioning: results from a 6-year follow-up in the maastricht aging study. *J. Am. Geriatr. Soc.* 53, 374–380. doi: 10.1111/j.1532-5415.2005.53152.x
- Verhaegen, C., Collette, F., and Majerus, S. (2014). The impact of aging and hearing status on verbal short-term memory. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 21, 464–482. doi: 10.1080/13825585.2013.832725
- Wayne, R. V., and Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res. Rev.* 23, 154–166. doi: 10.1016/j.arr.2015.06.002
- Wilson, R. S., Segawa, E., Hizel, L. P., Boyle, P. A., and Bennett, D. A. (2012). Terminal dedifferentiation of cognitive abilities. *Neurology* 78, 1116–1122. doi: 10.1212/WNL.0b013e31824f7ff2
- Wilson, D. H., Walsh, P. G., Sanchez, L., Davis, A. C., Taylor, A. W., Tucker, G., et al. (1999). The epidemiology of hearing impairment in an Australian adult population. *Int. J. Epidemiol.* 28, 247–252. doi: 10.1093/ije/28.2.247
- Wong, P. C. M., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., and Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* 47, 693–703. doi: 10.1016/j.neuropsychologia.2008.11.032
- Wright, C. E., and Drasdo, N. (1985). The influence of age on the spatial and temporal contrast sensitivity function. *Doc. Ophthalmol.* 59, 385–395. doi: 10.1007/bf00159172
- Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2011). Cognitive load during speech perception in noise: the influence of age, hearing loss and cognition on the pupil response. *Ear Hear.* 32, 498–510. doi: 10.1097/AUD.0b013e31820512bb
- Zwislocki, J., Maire, F., Feldman, A. S., and Rubin, H. (1958). On the effect of practice and motivation on the threshold of audibility. *J. Acoust. Soc. Am.* 30, 254–262. doi: 10.1121/1.1909559

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Roberts and Allen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners

Tim Schoof* and Stuart Rosen

Speech, Hearing and Phonetic Sciences, University College London, London, UK

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Jonathan E. Peelle, Washington University in Saint Louis, USA
Ramesh Rajan, Monash University, Australia

*Correspondence:

Tim Schoof, Speech, Hearing and Phonetic Sciences, University College London, Chandler House, 2 Wakefield Street, London, WC1N 1PF, UK
e-mail: t.schoof@ucl.ac.uk

Normal-hearing older adults often experience increased difficulties understanding speech in noise. In addition, they benefit less from amplitude fluctuations in the masker. These difficulties may be attributed to an age-related auditory temporal processing deficit. However, a decline in cognitive processing likely also plays an important role. This study examined the relative contribution of declines in both auditory and cognitive processing to the speech in noise performance in older adults. Participants included older (60–72 years) and younger (19–29 years) adults with normal hearing. Speech reception thresholds (SRTs) were measured for sentences in steady-state speech-shaped noise (SS), 10-Hz sinusoidally amplitude-modulated speech-shaped noise (AM), and two-talker babble. In addition, auditory temporal processing abilities were assessed by measuring thresholds for gap, amplitude-modulation, and frequency-modulation detection. Measures of processing speed, attention, working memory, Text Reception Threshold (a visual analog of the SRT), and reading ability were also obtained. Of primary interest was the extent to which the various measures correlate with listeners' abilities to perceive speech in noise. SRTs were significantly worse for older adults in the presence of two-talker babble but not SS and AM noise. In addition, older adults showed some cognitive processing declines (working memory and processing speed) although no declines in auditory temporal processing. However, working memory and processing speed did not correlate significantly with SRTs in babble. Despite declines in cognitive processing, normal-hearing older adults do not necessarily have problems understanding speech in noise as SRTs in SS and AM noise did not differ significantly between the two groups. Moreover, while older adults had higher SRTs in two-talker babble, this could not be explained by age-related cognitive declines in working memory or processing speed.

Keywords: aging, speech perception, auditory processing, cognition, noise

1. INTRODUCTION

Older adults often experience increased difficulties understanding speech in noisy environments (CHABA, 1988), even in the absence of hearing impairment (Dubno et al., 2002; Helfer and Freyman, 2008). One type of masker that seems particularly detrimental to older adults is competing speech (Tun and Wingfield, 1999; Helfer and Freyman, 2008; Rajan and Cainer, 2008). It has similarly been suggested that normal-hearing older adults benefit less from fluctuations in the masker compared to young adults (Takahashi and Bacon, 1992; Stuart and Phillips, 1996; Peters et al., 1998; Dubno et al., 2002, 2003; Gifford et al., 2007; Grose et al., 2009). It remains unclear, however, what is specific to aging, independent of hearing loss as defined in terms of the audiogram, that explains these difficulties in the perception of speech in noise.

One possible explanation is that speech in noise difficulties in part arise from an age-related auditory temporal processing deficit (e.g., Frisina and Frisina, 1997; Pichora-Fuller and Souza, 2003; Pichora-Fuller et al., 2007). A useful way to think about auditory temporal processing is in terms of the decomposition of

sound in the time domain into a slowly varying envelope (ENV) superimposed on a more rapidly varying temporal fine structure (TFS) (Moore, 2008). Aging has in fact been associated with declines in both ENV and TFS processing. Age-related declines in ENV processing become apparent, for instance, in terms of increased amplitude-modulation (Purcell et al., 2004; He et al., 2008) and gap detection thresholds (Snell, 1997; Schneider and Hamstra, 1999). Similarly, support for an age-related decline in TFS processing comes from a variety of psychophysical measures, such as frequency modulation (FM) detection (He et al., 2007), pitch discrimination using harmonic and inharmonic complex sounds (Vongpaisal and Pichora-Fuller, 2007; Füllgrabe, 2013), and the detection of inter-aural phase or time differences (Pichora-Fuller and Schneider, 1992; Grose and Mamo, 2010). These temporal processing deficits, and ultimately the increased difficulties understanding speech in noise, may be the result of disrupted neural sound encoding that are manifest even in the absence of any elevation in audiometric thresholds (Pichora-Fuller et al., 2007; Anderson et al., 2012; Sergeyenko et al., 2013).

While it is reasonable to assume that envelope processing is particularly important for the perception of speech in fluctuating maskers, age-related declines in temporal envelope processing may not necessarily be the cause of the decreased fluctuating masker benefit (FMB) in older adults. For one, FMB may be reduced in older adults at relatively slow modulation rates (e.g., 10 Hz; Dubno et al., 2002, 2003; Gifford et al., 2007), while age-related declines in ENV processing only become apparent at higher modulation rates (above about 200 Hz, e.g., Purcell et al., 2004; Grose et al., 2009). Instead, older adults might simply be less able to make use of the information in the dips of the fluctuating masker (Grose et al., 2009).

A perhaps more compelling theory is that an age-related decline in TFS processing partly explains difficulties understanding speech in noise. It has been argued that while ENV information may be sufficient for the perception of speech in quiet (Shannon et al., 1995), TFS may be required to successfully understand speech in the presence of interfering sound sources (Lorenzi et al., 2006; Moore, 2008, 2012). Although it has previously been suggested that TFS may be important to benefit from amplitude dips in fluctuating maskers (Schooneveldt and Moore, 1987; Lorenzi et al., 2006; Moore, 2008), the role of TFS may not necessarily be in detecting glimpses, but it may instead allow efficient auditory scene analysis and/or spatial release from masking (Bernstein and Brungart, 2011; Moore, 2012). In other words, age-related declines in TFS processing could be equally important in accounting for difficulties in steady-state noises as well as those that fluctuate. Furthermore, age-related declines in TFS processing may particularly impact speech perception in the presence of competing talkers as TFS provides pitch cues for sound source segregation (e.g., Bregman, 1990; Darwin and Carlyon, 1995).

An alternative explanation is that speech in noise difficulties in part arise from age-related declines in cognitive processing. Aging is associated with declines in several cognitive abilities that are thought to be important for the perception of speech in noise, such as working memory, attention, and processing speed (Craik and Byrd, 1982; Cohen, 1987; Kausler, 1994; Salthouse, 1996). However, older adults may in fact require more cognitive resources, putting higher demands on top-down processing to interpret the speech signal in the presence of background noise. Such demands may increase further when the input signal is further degraded as a result of auditory temporal processing declines (Rönnberg et al., 2010).

Working memory capacity, which refers to the ability to simultaneously store and process task-relevant information (Daneman and Carpenter, 1980; Baddeley, 1986), is perhaps most important for the perception of speech in noise (Rönnberg, 2003; Akeroyd, 2008). In a review of 20 studies looking at the role of cognition in speech perception in noise, Akeroyd (2008) found that working memory capacity, especially as assessed by the reading span test (Daneman and Carpenter, 1980; Rönnberg et al., 1989), was most predictive of speech perception in noise. Given that working memory capacity decreases with age (e.g., Craik and Jennings, 1992; Van der Linden et al., 1994), it is not unreasonable to assume that a decline in working memory plays an important role in the difficulties older adults experience when understanding speech in noise (Pichora-Fuller et al., 1995).

Similarly, selective attention, the ability to focus on relevant information and ignore irrelevant information, is probably equally important for successful speech understanding, especially in the presence of competing talkers which requires the suppression of meaningful competing information. Older adults may be less successful, however, at ignoring competing talkers as a result of an age-related decline in executive function, and more specifically a decline in inhibitory control (Hasher and Zacks, 1988; Hasher et al., 1999).

Underlying these age-related changes in working memory and attention may be a decline in processing speed. Salthouse (1985, 1996) argued that age-related declines in cognitive function may be the result of “cognitive slowing.” A reduction in processing speed means that relevant operations cannot be executed successfully in the time available and that the amount of simultaneously available information required for higher level processing is reduced. An age-related decline in processing speed may thus in part explain the speech in noise difficulties (Schneider et al., 2010).

Another factor thought to be important for speech perception in noise is linguistic closure, a supra-modal linguistic capacity thought to reflect the ability to fill in missing information (c.f. Zekveld et al., 2007). Linguistic closure is often assessed using the Text Reception Threshold (TRT), a visual analog of the SRT task, in which participants read sentences masked by bars of varying widths. This task was developed, more generally, to assess the extent to which inter-individual differences in speech in noise performance can be attributed to non-auditory factors. It remains unclear, however, whether the ability to read masked text decreases with age (see Besser et al., 2013, for a review).

The aim of this study was to assess why older adults, even in the absence of hearing impairment, experience increased difficulties understanding speech in noise. This study is novel in two ways. Firstly, relatively strict criteria for normal hearing were used (thresholds <25 dB HL up to 6 kHz). Secondly, while the majority of studies examining the effects of aging on speech perception in noise have used simple target stimuli, such as syllables (e.g., Stuart and Phillips, 1996; Dubno et al., 2002) or simple sentences (e.g., Peters et al., 1998; Gifford et al., 2007), this study used more complex targets (IEEE sentences; Rothauser et al., 1969).

Speech perception was assessed in the presence of different types of background noise. First, to examine whether normal-hearing older adults indeed benefit less from amplitude fluctuations in the masker, speech reception thresholds (SRTs) were measured in steady-state and amplitude-modulated noise (c.f. Takahashi and Bacon, 1992; Stuart and Phillips, 1996; Peters et al., 1998; Dubno et al., 2002, 2003; Gifford et al., 2007; Grose et al., 2009). Second, SRTs were also measured in the presence of two-talker babble since competing speech is both ecologically valid and particularly detrimental for older adults (Tun and Wingfield, 1999; Helfer and Freyman, 2008; Rajan and Cainer, 2008). In addition, various measures of auditory temporal (ENV and TFS) and cognitive processing (working memory, attention, processing speed, linguistic closure, and reading skills) were assessed to examine the relative contribution of declines in both domains on speech perception difficulties.

Individual differences in cognitive processing appear to be the most important factor explaining aided speech understanding in noise, after accounting for differences in audiometric thresholds, for hearing impaired older adults (see reviews by Humes et al., 2007; Akeroyd, 2008; Houtgast and Festen, 2008; Humes and Dubno, 2010). Therefore, age-related cognitive declines may also be expected to be the primary contributor to increased difficulties in speech perception in noise for normal-hearing older adults.

2. MATERIALS AND METHODS

2.1. PARTICIPANTS

Nineteen young (19–29 years old, mean 23.7 years, *SD* 2.9 years, 10 males) and 19 older (60–72 years old, mean 64.1 years, *SD* 3.3 years, 3 males) monolingual native English speakers participated in this study. All participants had near-normal hearing defined as (air-conducted) pure-tone thresholds of 25 dB HL or better at octave frequencies from 0.25 to 4 kHz in both ears and at 6 kHz in at least one ear (**Figure 1**). In addition, all participants over the age of 65 had normal cognitive function [scores ≥ 17 MMSE telephone version (Roccaforte et al., 1992)] and normal or corrected-to-normal vision. None of the participants reported a history of language or neurological disorders. All participants signed a consent form approved by UCL Research Ethics Committee and were paid for their participation.

2.2. SPEECH PERCEPTION IN NOISE

Speech reception thresholds (SRT) were measured for sentences in different types of background noise. The target stimuli were pre-recorded IEEE sentences (Rothauser et al., 1969) produced by

a male talker with a standard Southern British accent. Each sentence contained five keywords. The sentences were presented in steady-state speech-shaped noise (SS), speech-shaped noise sinusoidally amplitude modulated at 10 Hz (AM) with a modulation depth of 100%, and two-talker babble [see Rosen et al. (2013), for a description of the speech-shaped noise and two-talker babble]. The masker always started 600 ms prior to stimulus onset and was gated on and off across 100 ms.

To rule out possible contributions of differences in audiometric thresholds above 6 kHz, the stimuli were low-pass filtered at 6 kHz using a 4th order Butterworth filter. In addition, for six older participants with thresholds > 25 dB HL at 6 kHz in one ear the stimuli were spectrally shaped using the National Acoustics Laboratories-Revised (NAL-R) linear prescriptive formula based on their individual thresholds (Byrne and Dillon, 1986).

The participants were seated in a soundproof booth and listened to the stimuli over Sennheiser HD 25 headphones. They were asked to repeat verbatim what they heard. The experimenter scored responses using a graphical user interface (GUI) which showed the five key words. The scoring screen was not visible to the participants and no feedback was provided.

The SNR was varied adaptively following the procedure described by Plomp and Mimpen (1979). The first sentence was presented at an SNR of -10 dB. Until at least 3 out of 5 key words were correctly repeated, the SNR was increased by 6 dB on the next presentation. The initial sentence was repeated until at least 3 out of 5 keywords were repeated correctly or the SNR reached 30 dB. For each subsequent sentence the SNR increased by 2 dB when 0–2 key words were correctly repeated or decreased by the

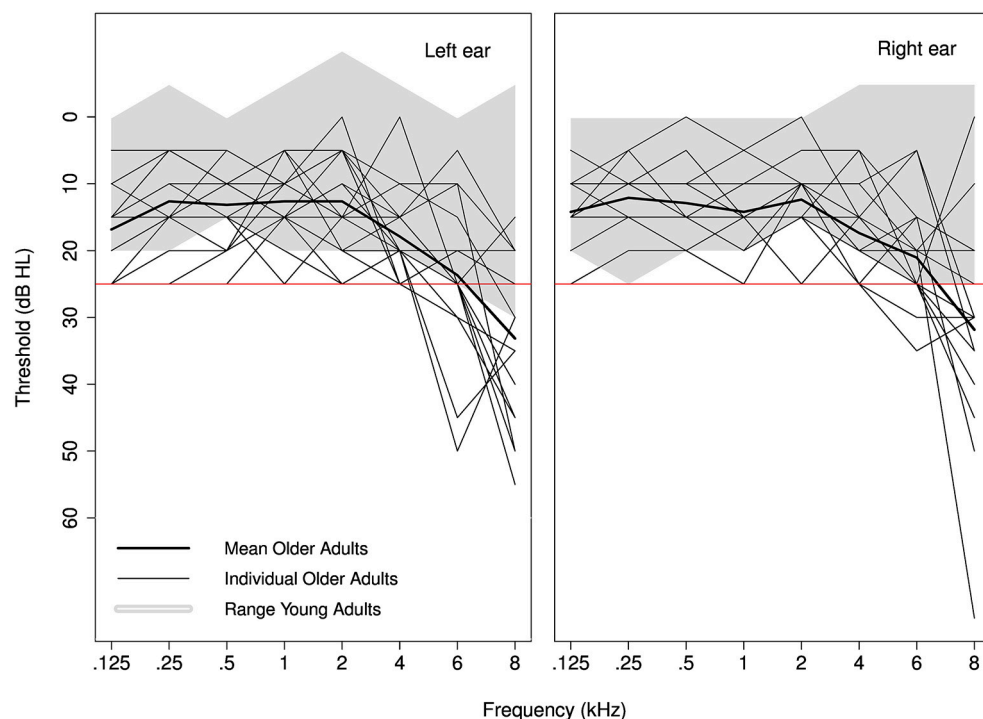


FIGURE 1 | Individual audiograms for older adults are plotted for the left and right ear separately. The shaded area represents audiometric thresholds of the younger adults. The red line indicates the inclusion criterion of 25 dB HL.

same amount for 3–5 correct repetitions. The number of trials was fixed at twenty, tracking 50% correct.

SRTs for each condition were measured twice. A measurement was repeated, with a different set of sentences, when fewer than 3 reversals were obtained or when the standard deviation across the final reversals exceeded 4 dB. Thresholds for each run were computed by taking the mean SNR (dB) across the reversals at the final step size of 2 dB.

Participants were given brief training on the different conditions to familiarize them with the different types of background noise. Practice consisted of 5 trials and started at 0 dB SNR. The order of conditions in the experiment proper was counterbalanced across participants following a Latin square design. Stimuli were presented binaurally at 70 dB SPL.

2.3. SUBJECTIVE MEASURE OF SPEECH PERCEPTION IN NOISE

Participants were asked to complete section one of the Speech, Spatial, and Qualities of Hearing Scale (SSQ; Gatehouse and Noble, 2004), which addresses listeners' abilities to understand speech in quiet as well as in the presence of different types of noise. Composite scores were calculated for each participant by averaging across all questions.

2.4. TEMPORAL PROCESSING

Participants completed three tasks that assess temporal processing: gap detection, amplitude modulation (AM) detection and frequency modulation (FM) detection. While the gap and AM detection tasks are concerned with temporal resolution in the envelope domain, the FM detection task assesses processing of TFS. The general procedure was similar for all three tasks. More details on the different tasks are provided below.

In all three tasks, a 3AFC paradigm was used and participants were asked to identify the stimulus that either contained a gap, or was modulated in amplitude or frequency. The duration of the gap or the depth of modulation was varied adaptively following the adaptive three-down, one-up procedure thus tracking 79% (Levitt, 1971).

Thresholds were obtained across two runs. A run was terminated after six reversals or after a maximum of 50 trials. Thresholds were computed by taking the mean gap duration or modulation depth across the last four reversals of each run. Thresholds reported here are the mean across the two runs.

Participants received training on five trials to familiarize themselves with the task. During this brief training they received visual feedback. During the experiment proper no feedback was provided.

Stimuli were presented binaurally over Sennheiser HD 25 headphones at 70 dB SPL. The order of the three tasks was counterbalanced across participants following a Latin square design.

2.4.1. Gap detection

Gap detection thresholds were measured using three 3-kHz-wide noises bandpass filtered between 1 and 4 kHz. A relatively wide band of noise was used as this limits the confounding effect of inherent fluctuations of the noise source on gap detection thresholds. The stimuli had a duration of 400 ms with a 10 ms rise-fall time and an inter-stimulus interval of 500 ms. The bands of noise

were generated online at the start of each trial. All three noise bursts were thus based on the same underlying 400 ms section of noise. When a temporal gap was present in the stimulus, it was centered 300 ms after stimulus onset. Gap durations were varied from 0.5 to 7 ms in 20 logarithmic steps. Gaps were created by zeroing the waveform. Since this results in spectral cues that could aid the listener in identifying the presence of a gap, the stimuli were filtered to the required bandwidth after the insertion of the gap using a 4th order Butterworth filter. It should be noted that this procedure causes some temporal smearing of the gap. However, for relatively shallow filters this should not affect gap detection thresholds too much (c.f. Eddins et al., 1992).

The initial gap duration was 7 ms and was decreased after each trial until an error was made. Subsequently, three consecutive correct responses were required to decrease the gap duration, while one incorrect response increased the gap duration. The initial step size was 3 logarithmic steps and was decreased to 2 and finally 1 logarithmic step after each reversal. To prevent the gap duration from decreasing too far below the participant's threshold during the first few runs, the step size was automatically set to 1 logarithmic step once the gap duration was ≤ 1 ms. A run was repeated when fewer than 3 reversals were obtained or when the standard deviation across the final reversals exceeded 2 ms.

2.4.2. AM detection

As in the gap detection task, AM detection thresholds were measured using three 3-kHz-wide noises bandpass filtered between 1 and 4 kHz. The temporal-modulation transfer function was determined on the basis of AM detection thresholds for five (sinusoidal) AM rates: 10, 20, 40, 80, and 160 Hz. These modulation rates are all multiples of 10 Hz, which is the modulation rate of the masker used in the speech perception in noise task. The duration of the stimuli was 500 ms, which resulted in a whole number of AM cycles in all four conditions. The stimuli had a 10 ms rise-fall time and a 500 ms inter-stimulus interval. As in the gap detection task, the bands of noise were generated online at the start of each trial, which meant that the three stimuli in each trial were composed of the same noise sample. Amplitude modulation depths varied in 25 steps of 1 dB from -8 to -32 dB for rates up to 80 Hz and from -5 to -29 dB for the 160 Hz modulation rate. Since AM of bandpassed noise produces spectral side bands, the stimuli were filtered using a 4th order Butterworth filter after modulation. It should be noted that this may have reduced the effective modulation depth, especially for higher AM rates, although the filtering used should not have much of an effect (c.f. Eddins, 1993, 1999).

On the initial trial, the modulation depth was set to -8 dB, or -5 dB for the 160 Hz modulation rate, and was decreased after each trial until the participant gave an incorrect response. Subsequently, three consecutive correct responses were required to decrease the AM depth, while one incorrect response increased the AM depth. The initial step size was 6 dB, and was decreased in four steps after each reversal to the final step size of 1 dB. To prevent the AM depth from overshooting the participant's threshold during the initial runs, the step size was automatically set to 1 dB once the AM depth reached ≤ -25 dB for modulation rates of 10 and 160 Hz, and ≤ -20 dB for modulation rates of 20, 40,

and 80 Hz. A run was repeated when fewer than 3 reversals were obtained or when the standard deviation across the final reversals exceeded 3 dB. The order of conditions was counterbalanced across participants following a Latin square design.

Since the temporal modulation transfer function (TMTF) resembles the form of a low-pass filter (Viemeister, 1979), the AM detection thresholds were fitted with an equation describing the frequency response of a low-pass Butterworth filter using a non-linear least-squares regression (Eddins, 1993):

$$y = 10\log_{10} \left(\frac{1}{1 + (\alpha f)^2} \right) + c \quad (1)$$

where y is the gain of the imputed filter (in dB) and f is the modulation rate in Hz. The inverse of α gives the -3 dB cutoff frequency (TMTF cutoff frequency) and c (the y -intercept) provides a measure of efficiency (AM efficiency). Note that a higher α (i.e., a higher cutoff frequency) and a lower c (i.e., better efficiency) indicate better performance.

2.4.3. FM detection

FM detection thresholds were determined using a 1 kHz sinusoidal carrier modulated at 2 Hz. A relatively low carrier frequency and modulation rate were used to ensure participants could only detect FM based on temporal cues (Moore and Sek, 1995, 1996). Frequency modulation depths varied logarithmically between 0.02 and 4.5 dB in 30 steps. The stimuli had a duration of 1 s, which is equal to 2 FM cycles. The interstimulus interval was set to 500 ms.

On the initial trial the modulation depth was set to 4.5 dB and was decreased after each trial until the listener made an error. Subsequently, three consecutive correct responses were required to decrease the FM depth, while one incorrect response increased the FM depth. The initial step size was three logarithmic steps, and was decreased in three steps after each reversal to the final step size of one logarithmic step. In addition, the step size was automatically set to one logarithmic step once the FM depth reached ≤ 0.57 dB to prevent the FM depth from overshooting the participant's threshold during the initial runs. A run was repeated when fewer than 3 reversals were obtained or when the standard deviation across the final reversals exceeded 2 dB.

FM detection thresholds are reported as modulation indices, which is the modulation depth divided by the modulation rate (2 Hz).

2.5. COGNITIVE SKILLS

Cognitive skills were assessed in the visual domain to ensure that auditory factors did not influence these measures.

2.5.1. Working memory

A reading span task was used to examine participants working memory capacity (Rönnberg et al., 1989). This task was designed to tax not only information storage and rehearsal (as do, for example, digit span and word span tasks) but also information processing. The reading span task developed by Rönnberg and colleagues is an extension of the task developed by Daneman and Carpenter (1980). Here, participants were asked to read sequences of 3–6 three-word sentences and judge whether the sentence was

semantically sensible or not (e.g., “The train sang a song,” or “The girl brushed her teeth”). At the end of each sequence of sentences, participants were asked to recall either the first or last word of each sentence in the correct order. The typeface of the text was Helvetica with font size 40. Words were presented in black on a gray background at 0.8 s/word. The inter-sentence interval, during which participants are required to make a semantic judgment, was 1.75 s. Participants were given one sequence of three sentences as a practice trial. During the testing phase, participants were presented with three runs of each sequence length (i.e., 3–6 sentences). The number of correctly remembered words was recorded.

2.5.2. Attention

Participants were assessed on the Visual Elevator task, a subtask of the Test of Everyday Attention (TEA; Robertson et al., 1996). It is thought to reflect an ability to switch attention, which is important for understanding speech in noise, especially in the presence of competing talkers. In essence, the participants' task was to count in a certain direction and at a given cue start counting in the opposite direction. The task consists of 10 trials. Participants were asked to determine the floor number for each item and complete the task as fast as they could. The responses for each trial and the total time required to complete all 10 trials were recorded. The total number of reversals for all correct responses were subsequently recorded. The final score was calculated by dividing the total duration required to complete the task (in seconds) by the total number of reversals for the correct responses.

2.5.3. Processing speed

To assess processing speed, participants were asked to complete the Letter Digit Substitution Test (LDST; Van der Elst et al., 2006). Participants were asked to complete the written version of the LDST. They were provided with a key in which the numbers 1–9 are each paired with a different letter. The test items, consisting of eight rows of 15 randomized letters, were printed below the key. The letters and digits were printed in font size 14. None of the participants had difficulties reading the items. The participants were asked to replace the letters by the corresponding digits as quickly as possible in sequential order. The first 10 items were practice items. After completion of the test items they were given 60 s to substitute as many items as possible. The score is the number of correctly substituted items. Note that potential age-related declines in motor performance were not controlled for.

2.5.4. Text reception threshold

The text reception threshold (TRT) is a visual analog of the speech reception threshold (SRT), especially in fluctuating noise (Zekveld et al., 2007; Besser et al., 2012). This task was developed to measure the variance in speech perception in noise abilities that are associated with supra-modal cognitive and linguistic skills. In this task sentences that are partly masked by a vertical bar pattern are presented on a computer screen.

As in the speech perception in noise task (measuring SRTs), the target stimuli were IEEE sentences (Rothauser et al., 1969). While the target stimuli were taken from the same corpus, the specific sentences used in the two tasks were different. The participants were seated approximately 50 cm from the screen. The

typeface used to present the sentences was Arial, with a font size of 28. The background color was white, the masked bar pattern was black, and the sentences were presented in red. The participants were asked to read the sentence out loud. The experimenter scored responses using a graphical user interface (GUI) which showed all the words in the sentence. The scoring screen was not visible to the participants and no feedback was provided.

The degree of masking was varied adaptively following the procedure described by Plomp and Mimpen (1979). The first sentence was presented with 16% unmasked text. Until the sentence was correctly repeated, the percentage of unmasked text was increased by 12% on the next presentation. Subsequent sentences were only presented once. When a sentence was correctly repeated, the degree of masking was increased by 6%. Conversely, the degree of masking was decreased by 6% when a sentence was not repeated correctly, thus tracking 50% correct.

TRTs were measured in response to two lists of twenty sentences each. Thresholds for each run were computed by taking the mean percentage unmasked text across sentences 5–20. The thresholds reported here are the mean across the two trials.

2.5.5. Reading skills

Given that both the text reception threshold and the reading span tasks rely heavily on reading, participants were assessed on reading ability using the Test of Word Reading Efficiency (TOWRE, Torgesen et al., 1999). Participants were asked to read out a list of 104 English words as fast as they could. Subsequently, they were asked to do the same for a list of 84 non-words. The words were presented in Arial font size 20. While the first subtask assesses participants' sight reading skills, the second subtask addresses their phonemic decoding efficiency. The TOWRE is aimed at children and normally assesses the number of words that can be correctly identified within 45 s. However, to avoid any ceiling effects in adults, participants read out all the words on the list and reading ability was assessed in terms of the time it took them to read the whole list. The score for this task was calculated by dividing the total duration required to complete the task by the number of correctly read items.

3. RESULTS

Data points that fell outside the mean ± 3 SD were considered outliers and excluded from the analyses reported below. In total, ten data points were excluded [data points from the older group were excluded for AM detection threshold at 160 Hz (one), TMTF cut-off (one), AM efficiency (one), TEA (two), TRT (one); data points from the young group were excluded for AM detection threshold at 160 Hz (one), TEA (two), and non-words TOWRE (one)].

Descriptive statistics for all measures as well as confidence intervals for the group differences are summarized in **Table 1**.

3.1. AUDIOMETRIC THRESHOLDS

While both groups had near-normal hearing, defined as pure-tone thresholds ≤ 25 dB HL up to 4 kHz in both ears and at 6 kHz in at least one ear, their thresholds were significantly different. Independent *t*-tests indicated that pure-tone averages (PTA) across 0.5–4 kHz (all ≤ 25 dB HL) were significantly higher (i.e., worse) by 7.1 dB for the older age group

Table 1 | Descriptive statistics.

Dependent variable	Mean (SD) young	Mean (SD) older	CI
SRT SS	−3.7 (1.5)	−3.8 (1.2)	(−1.0, 0.8)
SRT AM	−6.7 (2.4)	−6.4 (1.2)	(−1.0, 1.5)
SRT babble	−1.6 (2.0)	−0.3 (0.8)	(0.4, 2.4)
SSQ	7.5 (1.0)	7.1 (1.2)	(−1.2, 0.3)
AM cutoff	110 (50)	111 (62)	(−36, 38)
AM efficiency	−19.4 (2.0)	−19.5 (1.0)	(−1.2, 0.9)
FM detection	1.0 (0.4)	1.1 (0.3)	(−0.2, 0.3)
Gap detection	4.3 (0.8)	4.6 (1.2)	(−0.4, 0.9)
Working memory	32 (5.6)	24 (5.0)	(−11.6, −4.6)
Attention	4.2 (1.9)	5.0 (1.7)	(−0.4, 2.0)
Processing speed	39 (6.7)	34 (6.6)	(−9.3, −0.3)
TRT	63 (4.3)	63 (3.5)	(−1.8, 3.2)
Reading words	0.49 (0.09)	0.48 (0.07)	(−0.06, 0.05)
Reading non-words	0.74 (0.14)	0.72 (0.15)	(−0.1, 0.08)

Descriptive statistics (mean and SD) for the young and older adults separately as well as confidence intervals for the group differences are provided for all measures.

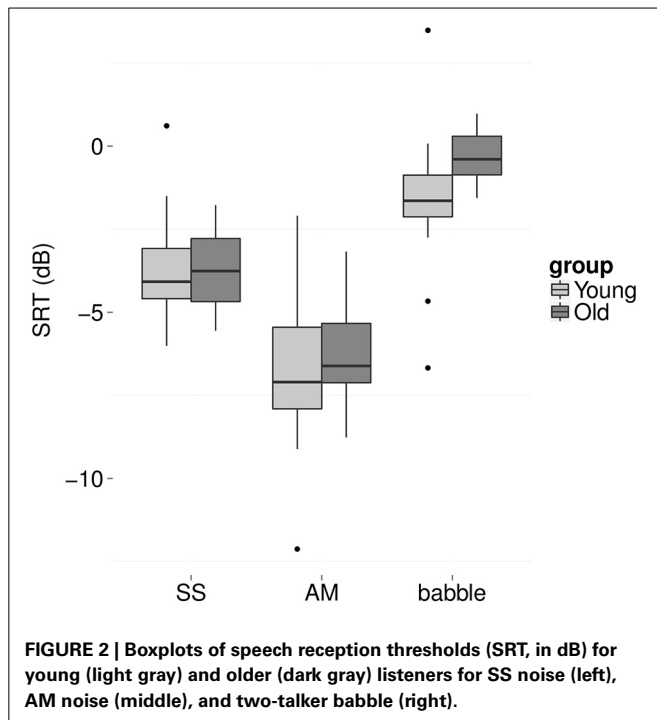
[$t_{(36)} = -6.4$, $p < 0.001$]. This could potentially contribute to any group differences that might exist for the auditory tasks (SRT, gap detection, AM detection, and FM detection; see Section 3.6).

Analyses were conducted on a PTA across 0.5–4 kHz since the auditory tasks in this study, with the exception of the SRT task, did not have energy above 4 kHz. While the materials in the SRT task did contain energy above 4 kHz, stimuli for six older adults were spectrally shaped using the NAL formula to account for audibility differences.

3.2. SPEECH PERCEPTION IN NOISE

Older adults were expected to perform more poorly (i.e., higher SRTs) in all three background noises. However, the older adults had higher SRTs only in the presence of two-talker babble (**Figure 2**). A mixed effects model with condition (AM, SS, babble) and group (young, old) as fixed factors and participant and sentence list as random factors showed a significant interaction between condition and group [$F_{(2, 186)} = 5.6$, $p = 0.004$]. *Post-hoc* independent *t*-tests revealed a significant difference between the two age groups for babble only, with young listeners performing better than older listeners by 1.4 dB [$t_{(36)} = 2.8$, $p = 0.008$, Cohen's $d = 0.9$; all other $p > 0.6$].

Overall, SRTs in AM noise were expected to be lower (i.e., better) compared to SRTs in SS noise, indicative of dip listening. Furthermore, SRTs in babble were expected to be higher (i.e., worse) compared to the two noise maskers (c.f. Rosen et al., 2013). *Post-hoc* independent *t*-tests indeed revealed a significant dip listening effect, with lower SRTs in AM compared to SS noise [$t_{(37)} = 12.9$, $p < 0.001$, Cohen's $d = 1.4$, mean difference = 2.7 dB]. In addition, SRTs in babble were significantly higher



compared to the two noise maskers [SS: $t_{(37)} = 8.5$, $p < 0.001$, Cohen's $d = 2.5$, mean difference = 2.6 dB; AM: $t_{(37)} = 16.3$, $p < 0.001$, Cohen's $d = 1.4$, mean difference = 5.3 dB].

While there may be no group differences in SRTs in SS or AM noise and only a small difference in babble, it may be the case that particular older adults experience increased difficulties with one or more of the maskers. To explore these individual differences, we performed a deviance analysis (c.f. Ramus et al., 2003). The SRT scores were converted to z-scores and the deviance threshold was set to 1.65 SD above the mean SRT of the young group. Thus, participants were identified who performed more poorly than the poorest 5% of a young population.

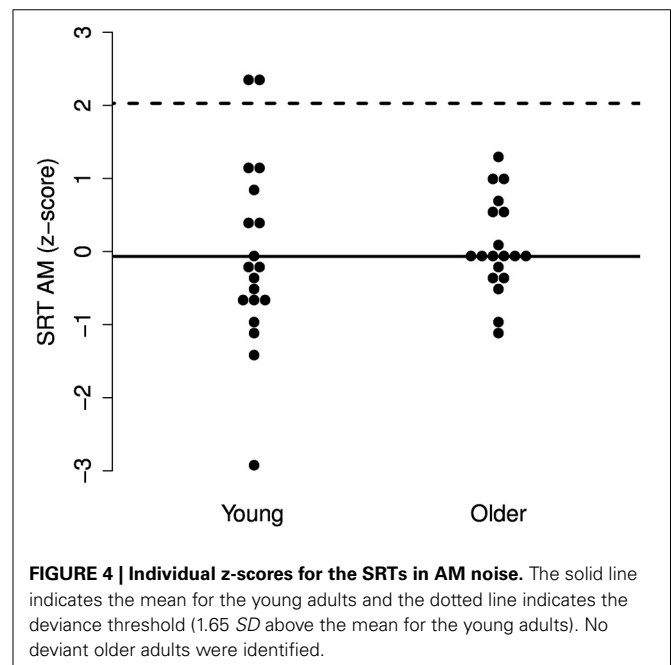
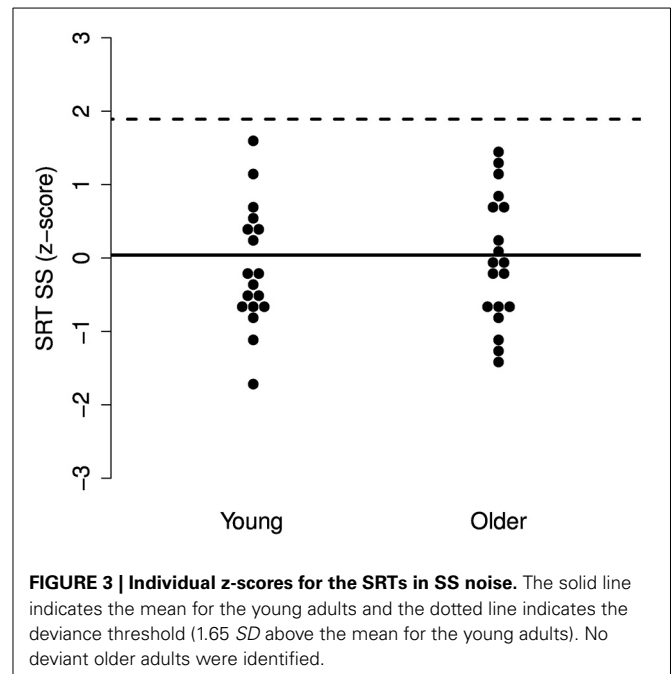
The results, illustrated in Figures 3–5, indicate that none of the older adults performed particularly poorly in any of the maskers. This supports the idea that normal-hearing older adults do not necessarily experience increased difficulties understanding speech in noise.

3.3. SUBJECTIVE MEASURE OF SPEECH PERCEPTION IN NOISE

While the SRT data showed some group differences (in the presence of two-talker babble only), older adults did not report increased difficulties understanding speech in noise. An independent t -test on the subjective measure of speech perception in noise (SSQ questionnaire) did not reveal a significant difference between the two age groups [$t_{(36)} = 1.3$, $p = 0.2$]. It should be pointed out, however, that the difference in SRTs in two-talker babble was small (1.4 dB) and that older adults did not perform more poorly in AM and SS noise compared to the young adults.

3.4. AUDITORY TEMPORAL PROCESSING

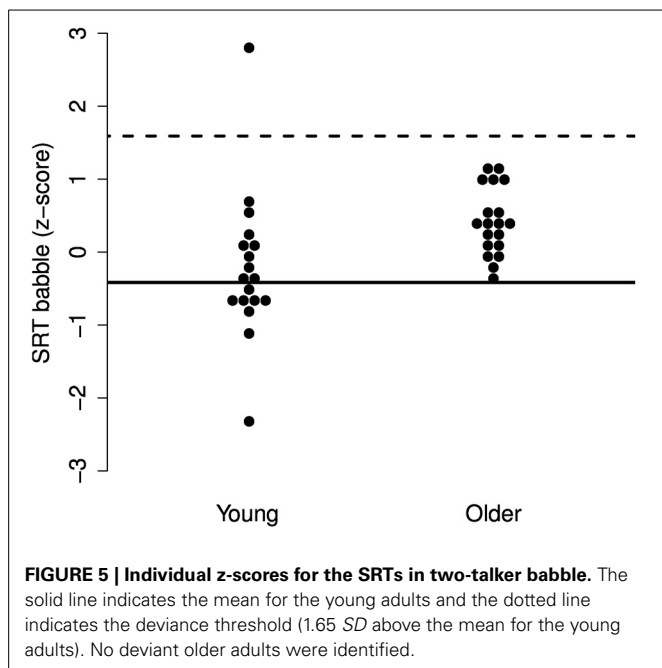
While previous studies have reported age-related declines in auditory temporal processing (Pichora-Fuller and Schneider, 1992;



Snell, 1997; Vongpaisal and Pichora-Fuller, 2007; He et al., 2008; Füllgrabe, 2013), no support for such a deficit was found in this study. AM, FM, and gap detection thresholds did not differ significantly between the young and older adults.

Independent t -tests on the two measures derived from the TMTF (AM efficiency and TMTF cut-off frequency) revealed no significant group differences [AM efficiency: $t_{(35)} = -0.23$, $p = 0.8$; TMTF cut-off: $t_{(35)} = -0.07$, $p = 0.9$].

These findings were supported by a mixed effects model on the AM detection thresholds with rate (10, 20, 40, 80, and 160)



and group (young, old) as fixed factors and participant as a random factor. The analysis revealed a significant main effect of rate [$F_{(1, 148)} = 220, p < 0.001$], due to the fact that the shape of the TMTF resembles a low-pass filter. However, no group or interaction effects were found [group $F_{(1, 36)} = 0.4, p = 0.5$; interaction $F_{(1, 148)} = 1.2, p = 0.27$], which means that the AM detection thresholds at the five different rates did not differ between the young and older adults.

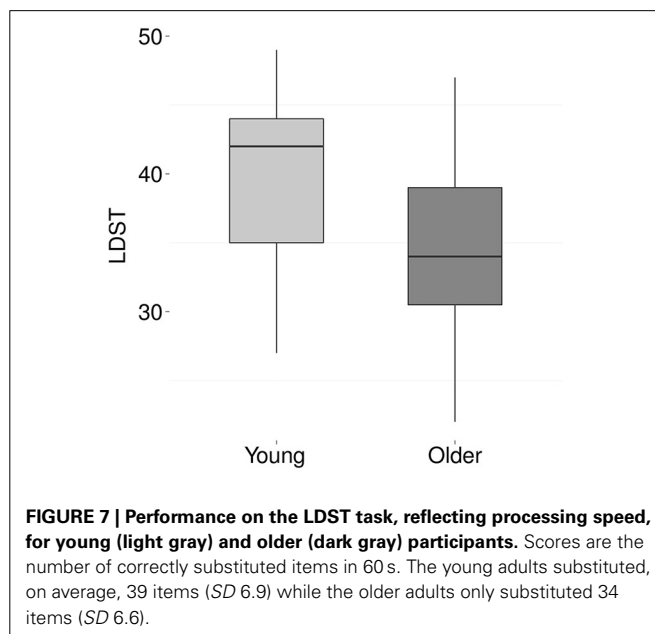
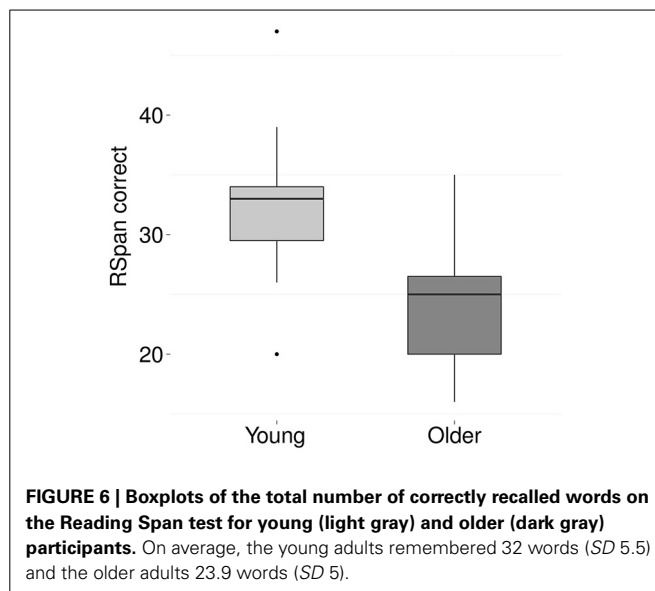
Similarly, independent *t*-tests did not reveal significant differences between the two age groups in terms of FM and gap detection thresholds [FM $t_{(36)} = 0.6, p = 0.5$; gap $t_{(36)} = 0.7, p = 0.4$].

3.5. COGNITIVE PROCESSING

Figures 6, 7 show the results for the different cognitive processing tasks. Five independent *t*-tests were carried out to examine the effect of age on various cognitive skills. The analyses revealed an age-related decline in working memory, as indicated by fewer correctly remembered items on the Reading Span task [$t_{(36)} = 4.7, p < 0.001$, Cohen's $d = 1.5$]. In addition, a significant age-effect was found for processing speed, with older adults performing fewer substitutions on the letter-digit-substitution task [$t_{(36)} = 2.2, p = 0.04$, Cohen's $d = 0.7$]. No age-effects were found for attention [$t_{(32)} = -1.3, p = 0.2$], TRT [$t_{(35)} = -0.6, p = 0.59$], or reading skills [words $t_{(36)} = 0.3, p = 0.8$; non-words $t_{(35)} = 0.4, p = 0.7$].

3.6. PREDICTING SPEECH PERCEPTION IN NOISE

Of primary interest was the extent to which the various auditory and cognitive measures could predict listeners abilities to perceive speech in the three noises. The results have so far indicated age-related declines in speech perception in babble (but not SS and AM noise), working memory, and processing speed. In addition, while both groups had near-normal hearing, thresholds for the



older adults were significantly higher. These findings indicate that the normal-hearing older adults had no problems understanding speech in SS and AM noise, despite some age-related cognitive declines and slightly higher audiometric thresholds. One of the questions that remains, however, is whether these age-related declines can account for the group difference in SRTs in babble.

Furthermore, the fact that the older adults only experienced increased difficulties understanding speech in two-talker babble, but not in the two noise maskers (SS and AM noise), suggests that the relative contribution of the various auditory and cognitive processes involved in the perception of speech in noise differs depending on the masker type. A question to be answered, then, is which of the auditory and cognitive measures can account for

the inter-individual differences in the perception of speech in the presence of babble and noise maskers.

To determine which of the auditory or cognitive measures was predictive of speech understanding in babble and noise maskers, best subsets regression analyses were conducted (Hastie et al., 2009). Since the SRTs in AM and SS noise were highly correlated ($r = 0.736$, $p < 0.001$, $R^2 = 0.54$, **Figure 8**), the regression was performed on the average of the two.

3.6.1. Data reduction

Due to the relatively large number of possible predictors (twelve) given our sample size (38 participants), a principal components analysis (PCA) using varimax rotation with Kaiser normalization was performed on the cognitive and temporal processing tasks separately to reduce the number of predictors for the regression analysis. Missing data points (see Section 3) were replaced by the mean. The resulting principal components (PC) were saved as Anderson-Rubin scores to ensure uncorrelated PC scores.

PCA on the six cognitive measures (LDST, RSpan, TRT, TEA, and TOWRE words and non-words) resulted in the extraction of two components, following the Kaiser criterion (eigenvalues > 1). Together they explained 63% of the variance in the data, with PC1 accounting for 34% and PC2 for 29% (see **Table 2**). The first PC was interpreted as an overall measure of linguistic closure

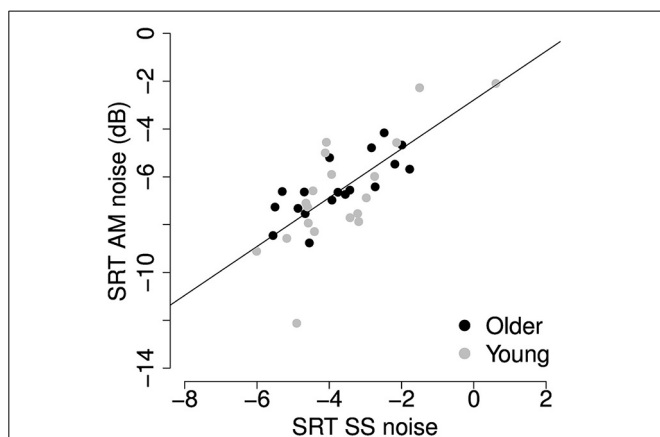


FIGURE 8 | Scatter plot of SRTs in SS and AM noise reveals a strong correlation ($r = 0.736$, $p < 0.001$) between performance in the two noise maskers.

Table 2 | PCA Factor loadings: Cognitive processing.

	Linguistic closure	Memory and processing speed
LDST	-0.187	0.763
Reading span	0.068	0.857
TEA	0.356	-0.518
TRT	0.635	-0.394
TOWRE words	0.899	0.005
TOWRE non-words	0.811	-0.117

Factor loadings for each of the cognitive processing measures. Factor loadings > 0.4 are highlighted in bold font.

(c.f. Zekveld et al., 2007) as it mainly reflected the TRT and the two measures of reading ability (TOWRE). The second PC primarily reflected processing speed (LDST) and working memory (RSpan). Note that the measure of attention (TEA) did not group clearly with either of the two components.

An initial PCA on the four temporal processing measures (TMTF cut-off frequency, AM efficiency, FM, and gap detection thresholds) suggested the extraction of three PCs; the two AM detection measures grouped together, but the FM and gap detection scores loaded significantly onto separate components (see **Table 3**). Since the latter two components were dominated by a single temporal processing measure, the raw FM and gap detection thresholds were entered into the regression model instead. A subsequent PCA was performed on the two AM detection measures (TMTF cut-off frequency and AM efficiency), which resulted in the extraction of a single component that explained 66% of the variance in the AM detection data (**Table 3**).

3.6.2. Regression

Following data reduction, the seven possible predictors that were entered into the regression models were; age group, PTA across 0.5–4 kHz, PC linguistic closure, PC memory and processing speed, PC AM detection, FM detection, and gap detection. Note that while individual differences in audiometric thresholds above 4 kHz could also have contributed to differences in SRTs, especially since the stimuli were filtered with a relatively shallow filter, a PTA across 6–8 kHz was not included in the regression models as a possible predictor. This is because the NAL-shaping that was applied for some older adults from 6 kHz upwards means the audiometric thresholds do not accurately reflect audibility differences in this region. Best subsets linear regressions were performed for SRTs in babble and noise (averaged across AM and SS) separately. The final models were selected based on the Bayesian Information Criterion (BIC; Schwarz, 1978).

The analyses indicated that SRTs in babble were best predicted by PTA across 0.5–4 kHz and FM detection thresholds [$R^2 = 0.32$, $F_{(2, 35)} = 8.3$, $p = 0.001$; see **Table 4**]. Thus, age-related cognitive declines in working memory and processing speed did not in fact predict SRTs in babble. Instead, when audiometric

Table 3 | PCA factor loadings: Temporal processing.

PCA across all measures	PC1	PC2	PC3
FM detection	0.004	-0.068	0.968
Gap detection	0.011	0.953	-0.076
TMTF cutoff frequency	0.797	-0.360	-0.295
AM efficiency	0.824	0.332	0.264
PCA across AM measures	PC AM detection		
TMTF cutoff frequency	0.811		
AM efficiency	0.811		

Factor loadings for each of the temporal processing measures (top) and for the amplitude-modulation detection measures only (bottom). Factor loadings > 0.4 are highlighted in bold font.

thresholds were accounted for, FM detection thresholds were the primary predictor of SRTs in babble. This would imply that TFS processing in part determines speech understanding in the presence of competing talkers.

SRTs in noise, by contrast, were best predicted by a model with PTA across 0.5–4 kHz, linguistic closure, and memory and processing speed [$R^2 = 0.32$, $F_{(3, 34)} = 5.48$, $p = 0.004$; see **Table 4**]. The fact that, after controlling for audiometric thresholds, the two cognitive measures, rather than FM detection thresholds, were significant predictors of SRTs in noise suggests that TFS processing might be less important for the perception of speech in noise maskers than in the periodic two-talker babble.

While the results from the best subsets regression analyses appear to suggest that the underlying processes accounting for individual differences in speech perception in two-talker babble and noise maskers is different, this may in fact not be the case. Even though the regression coefficients may be significant in one model but not the other, these differences in significance are in themselves not necessarily significant (Gelman and Stern, 2006).

To assess whether the slopes of the predictors in the two models were indeed significantly different, a linear regression with the four predictors that were significant in either of the two best subsets regression models (PTA 0.5–4 kHz, FM, PC linguistic closure, PC memory and processing speed) was performed on both SRTs in babble and noise separately (see **Table 5**). The results of this regression model are in line with the results of the best subsets regressions, with the same predictors coming out as significant [SRT babble: $R^2 = 0.36$, $F_{(4, 33)} = 4.7$, $p = 0.004$; SRT noise maskers: $R^2 = 0.37$, $F_{(4, 33)} = 4.839$, $p = 0.003$]. Since both models now contained the same predictors, the regression coefficients could be compared. In order to do so, a subsequent linear regression was conducted on both SRTs, with an additional dummy-coded predictor indicating the type of background noise (i.e., babble or noise maskers). The interaction between the dummy variable and the original predictors indicated whether the slopes of the predictors differed depending on the type of background noise. The results did not reveal any significant interactions (see **Table 5**), suggesting that even though some measures

Table 4 | Best subsets regression.

Dependent variable	Predictors	<i>b</i>	β	<i>SE</i>	<i>p</i>	R^2 change
SRT babble	PTA 0.5–4 kHz	0.17	0.51	0.047	<0.001***	0.2
	FM	1.57	0.35	0.64	0.02*	0.12
SRT noise maskers	PTA 0.5–4 kHz	0.09	0.31	0.045	0.045*	0.06
	Linguistic closure	0.59	0.4	0.21	0.008**	0.16
	Memory and processing speed	0.5	0.34	0.22	0.03*	0.1

Results of the best subsets regression analyses on SRTs in babble and the average SRTs across the two noise maskers (i.e., AM and SS noise; *significant at $\alpha = 0.05$, **significant at $\alpha = 0.01$, ***significant at $\alpha = 0.001$). Note that β refers to the standardized regression coefficient. The R^2 change reflects the proportion of the variance accounted for as predictors are added to the model.

Table 5 | Full regression model.

Dependent variable	Predictors	<i>b</i>	β	<i>SE</i>	<i>p</i>	R^2 change
SRT babble	PTA 0.5–4 kHz	0.16	0.49	0.05	0.003**	0.2
	Linguistic closure	0.35	0.22	0.23	0.14	0.06
	Memory and processing speed	0.001	0.0008	0.26	0.99	0.02
	FM	1.46	0.32	0.7	0.045*	0.08
SRT noise maskers	PTA 0.5–4 kHz	0.115	0.379	0.046	0.02*	0.06
	Linguistic closure	0.55	0.37	0.21	0.01*	0.16
	Memory and processing speed	0.65	0.44	0.24	0.01*	0.1
	FM	0.97	0.24	0.64	0.14	0.04
Interaction with SRT	PTA 0.5–4 kHz	0.023	0.1	0.034	0.5	0.005
	Linguistic closure	−0.1	−0.038	0.16	0.53	0.001
	Memory and processing speed	−0.33	−0.12	0.18	0.07	0.017
	FM	0.24	0.1	0.48	0.61	0.001

Top: Results of the regression analyses on the SRTs in babble and the average SRTs across the two noise maskers (i.e., AM and SS noise) with the four predictors that were significant in either of the two best subsets regression models (PTA 0.5–4 kHz, FM, PC linguistic closure, PC memory and processing speed). Bottom: Results of the regression analysis on both SRT measures with an additional dummy-coded predictor indicating the type of background noise (i.e., babble or noise maskers) assessing whether the slopes of the predictors differed depending on the type of background noise. Significant results are highlighted in bold font (*significant at $\alpha = 0.05$, **significant at $\alpha = 0.01$). Note that β refers to the standardized regression coefficient. The R^2 change reflects the proportion of the variance accounted for as predictors are added to the model.

significantly predicted SRTs in one type of background noise but not the other, the regression coefficients across the models were themselves not significantly different. In other words, there is no support for the claim that the underlying processes involved in the perception of speech in babble and noise maskers are different.

4. DISCUSSION

The aim of this study was to assess why older adults, even in the absence of hearing impairment, typically experience increased difficulties understanding speech in noise. These difficulties are typically attributed to an age-related decline in central auditory processing, particularly in the time domain, and/or a decline in cognitive function (CHABA, 1988). This study examined the relative contribution of age-related declines in both auditory temporal and cognitive processing on the perception of speech in the presence of different noise maskers.

First, it is important to note that the data in fact suggest that older adults with fairly good hearing do not necessarily perform more poorly on a speech in noise task when ecologically valid stimuli are used. Group differences were found only in the presence of two-talker babble but not in steady-state (SS) or fluctuating (AM) noise maskers. These findings are in line with the idea that competing speech is particularly detrimental for older adults (Tun and Wingfield, 1999; Helfer and Freyman, 2008; Rajan and Cainer, 2008). The fact that the older adults performed more poorly only in the presence of two-talker babble, but not the two noise maskers, suggests that these difficulties may be due to increased susceptibility to informational masking (c.f. Freyman et al., 2004). However, it may similarly be attributable to a reduced ability to make use of periodicity cues to successfully segregate the target and masker.

Contrary to expectations, the data suggest that normal hearing older adults do not have reduced glimpsing abilities (c.f. Stuart and Phillips, 1996; Peters et al., 1998; Dubno et al., 2002, 2003; Gifford et al., 2007; Grose et al., 2009). It should be noted, however, that the idea that older adults have impaired glimpsing abilities is perhaps somewhat controversial since age-related declines in FMB reported in the literature may in part have been the result of group differences that also became apparent in SS noise (c.f. Stuart and Phillips, 1996; Dubno et al., 2002, 2003; Bernstein and Grant, 2009).

It is perhaps surprising that the older adults did not perform more poorly on the speech in noise task compared to the younger listeners. One might argue that the tasks were not challenging enough. However, it is important to remember that the task was adaptive and therefore always got difficult. Moreover, while studies in the past have often used simple stimuli, such as syllables (e.g., Stuart and Phillips, 1996; Dubno et al., 2002, 2003) or simple BKB or HINT sentences (e.g., Gifford et al., 2007; Rajan and Cainer, 2008), this study used the more challenging IEEE sentences (see also Grose et al., 2009). It should be noted, however, that it remains possible that the older adults had to expend greater listening effort to perform on a par with the younger listeners.

Given that older adults are relatively unimpaired in their perception of speech in noise, could it be that the older adults are similarly unimpaired in terms of auditory temporal and cognitive processing? While an age-related decline in temporal auditory

processing is well documented in normal-hearing older adults (e.g., CHABA, 1988; Frisina and Frisina, 1997; Pichora-Fuller and Souza, 2003; Gordon-Salant, 2005; Pichora-Fuller et al., 2007), this study found no decline in either ENV or TFS processing. However, the fact that AM detection thresholds were not different between young and older adults is likely because age effects only become apparent at higher modulation rates than those assessed in the present study (above about 200 Hz, Purcell et al., 2004; Grose et al., 2009). Furthermore, the lack of an age-related increase in gap detection thresholds may be related to the temporal location of the gap. He et al. (1999) only found large age-related declines when the gap was located close to the stimulus onset or offset (at 5 or 95% of the stimulus duration), and when the gap location was random from trial to trial. Consistent with our findings, gaps in the central region of a noise burst were equally detectable by younger and older listeners, even when randomly located. Whatever the exact nature of the deficit in the older listeners found by He et al. (1999) is, it is certainly not a simple deficit in ENV processing. Instead, the importance of gap uncertainty suggests a cognitive component. What is perhaps most surprising is the absence of a decline in TFS processing as this has been found using a variety of psychophysical measures (He et al., 2007; Grose and Mamo, 2010; Füllgrabe, 2013). While aging has been shown to negatively affect frequency modulation (FM) detection using low carrier frequencies (≤ 4 kHz) and low modulation rates (≤ 5 Hz) (He et al., 2007), which is thought to be primarily dependent on the neural phase-locking (Moore and Sek, 1995, 1996), we did not replicate this finding.

Similarly, aging has often been associated with declines in cognitive abilities thought to be important for the perception of speech in noise, such as working memory, attention, and processing speed (Craik and Byrd, 1982; Kausler, 1994; Salthouse, 1996). The current data indeed show declines in both working memory and processing speed. By contrast, however, attentional switching, as measured by the Visual Elevator task (Robertson et al., 1996), was not affected by age. This is somewhat surprising since this task is thought to be similar to the Wisconsin Card Sorting Test (Nelson, 1976; Robertson et al., 1996), which has repeatedly been shown to be negatively affected by age (Rhodes, 2004). Another factor thought to be important for the perception of speech in noise is linguistic closure, which was assessed by the TRT task (Zekveld et al., 2007). The literature is inconclusive as to whether linguistic closure is negatively affected by age. The results from the present study suggest that older adults do not have problems reconstructing partially masked text. This may be because linguistic closure is representative of crystallized intelligence, which does not decline with age, as opposed to fluid intelligence, which does decline with age (Horn and Cattell, 1967).

It should be noted that the absence of any age-related declines in attention, linguistic closure, and perhaps even auditory temporal processing, could in part be attributed to the fact that the older adults who participated in this study were exceptional, if only in the sense that they had good hearing. Given that cognitive declines have been linked to hearing loss (c.f. Lin et al., 2013), it may not be surprising that the normal hearing older adults who participated in this study were relatively unimpaired in the cognitive domain. This means, however, that while this study may

tell us something about normal hearing older adults, the findings cannot be generalized to a more typical hearing impaired older population.

Despite the declines in working memory and processing speed, normal hearing older adults did not have increased difficulties understanding speech in SS and AM noise. This suggests that cognitive declines associated with aging do not inevitably lead to speech in noise problems. Furthermore, while the older adults performed worse on the speech perception task in the presence of two-talker babble, this could not be explained by age-related cognitive declines in working memory or processing speed when accounting for differences in audiometric thresholds. This lack of association may in part be attributed by the fact that the inter-individual variability in the data set was relatively small. Instead, however, individual differences in SRTs in babble were best predicted by audiometric thresholds and TFS processing, as measured by the FM detection task. It should be noted, however, that since the older adults had higher audiometric thresholds, it is difficult to distinguish between an explanation based on age, and one based on hearing status. The fact that TFS processing, second to audiometric thresholds, was predictive of speech perception in the presence of competing talkers suggests that variability in performance was largely due to differences in abilities to use periodicity cues. However, whether the difficulties in the presence of babble are in fact due to a reduced ability to use periodicity cues in the masker, informational masking, or even reduced glimpsing abilities remains unclear.

While it is tempting to conclude that the underlying processes involved in the perception of speech in babble and noise maskers is different, the current study did not provide sufficient support for this idea. In fact, TFS processing may be equally important for the perception of speech in noise maskers as in the presence of competing speech. Similarly, while cognitive processing was found to be predictive of SRTs in noise maskers, they may be equally important in the presence of babble. Since the predictor coefficients across the two regression models (SRTs in babble and noise maskers) were not significantly different, no conclusions can be drawn regarding differences in underlying processes involved in speech perception in the two interferer types.

In sum, this study set out to determine the relative contribution of age-related declines in auditory temporal and cognitive processing on the perception of speech in different maskers for normal-hearing older adults. The findings can be summarized as follows:

1. Older adults meeting a relatively stringent criterion for normal hearing experienced increased difficulties understanding speech only in the presence of two-talker babble.
2. Glimpsing abilities in 10-Hz sinusoidal amplitude-modulated noise were not reduced for the older adults.
3. While age-related declines in temporal auditory processing are well documented for older adults, even in the absence of hearing loss, this study failed to identify a decline in either envelope or temporal fine structure processing.
4. Older adults showed cognitive declines in working memory capacity and processing speed. Despite these declines,

however, speech perception in steady-state and amplitude-modulated noise was not impaired. Moreover, reduced working memory capacity and processing speed could not explain SRTs in babble beyond differences in audiometric thresholds.

AUTHOR CONTRIBUTIONS

This work is part of Tim Schoof's PhD project, supervised by Stuart Rosen.

FUNDING

This work was supported by a PhD studentship grant funded jointly by Action on Hearing Loss and Age UK (grant S19).

ACKNOWLEDGMENTS

The authors would like to thank Rebecca Oyekan for her help with data collection, Steve Nevard for technical support, Adriana Zekveld and J. H. M. Van Beek for sharing the TRT test, and Jerker Rönnerberg for sharing the Reading Span test.

REFERENCES

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). Aging affects neural precision of speech encoding. *J. Neurosci.* 32, 14156–14164. doi: 10.1523/JNEUROSCI.2176-12.2012
- Baddeley, A. (1986). *Working Memory*. Cambridge: Oxford University Press.
- Bernstein, J. G. W., and Brungart, D. S. (2011). Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *J. Acoust. Soc. Am.* 130, 473–488. doi: 10.1121/1.3589440
- Bernstein, J. G. W., and Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 125, 3358–3572. doi: 10.1121/1.3110132
- Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459
- Besser, J., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2012). New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J. Speech Lang. Hear. Res.* 55, 194–209. doi: 10.1044/1092-4388(2011/11-0008)
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.
- Byrne, D., and Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear.* 7, 257–265. doi: 10.1097/00003446-198608000-00007
- CHABA. (1988). Speech understanding and aging. *J. Acoust. Soc. Am.* 83, 859–895. doi: 10.1121/1.395965
- Cohen, G. (1987). Speech comprehension in the elderly: the effects of cognitive changes. *Br. J. Audiol.* 21, 221–226. doi: 10.3109/03005368709076408
- Craik, F. I. M., and Byrd, M. (1982). "Aging and cognitive deficits: the role of attentional resources," in *Aging and Cognitive Processes*, eds F. I. M. Craik and S. Treub (New York, NY: Plenum), 191–211.
- Craik, F. I. M., and Jennings, J. (1992). "Human memory," in *Handbook of Aging and Cognition*, eds F. I. M. Craik and T. A. Salthouse (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.), 51–110.
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Darwin, C. J., and Carlyon, R. P. (1995). *Auditory Grouping*. London: Academic Press.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 111, 2897–2907. doi: 10.1121/1.1480421

- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 113, 2084–2094. doi: 10.1121/1.1555611
- Eddins, D. A. (1993). Amplitude modulation detection of narrow-band noise: effects of absolute bandwidth and frequency region. *J. Acoust. Soc. Am.* 93, 470–479. doi: 10.1121/1.405627
- Eddins, D. A. (1999). Amplitude-modulation detection at low- and high-audio frequencies. *J. Acoust. Soc. Am.* 105, 829–837. doi: 10.1121/1.426272
- Eddins, D. A., Hall, J. W., and Grose, J. H. (1992). The detection of temporal gaps as a function of frequency region and absolute noise bandwidth. *J. Acoust. Soc. Am.* 91, 1069–1077. doi: 10.1121/1.402633
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J. Acoust. Soc. Am.* 115, 2246–2256. doi: 10.1121/1.1689343
- Frisina, D. R., and Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear. Res.* 106, 95–104. doi: 10.1016/S0378-5955(97)00006-3
- Füllgrabe, C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 22, 313–315. doi: 10.1044/1059-0889(2013)12-0070
- Gatehouse, S., and Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int. J. Audiol.* 43, 85–99. doi: 10.1080/14992020400050014
- Gelman, A., and Stern, H. (2006). The difference between significant and not significant is not itself statistically significant. *Am. Stat.* 60, 328–331. doi: 10.1198/000313006X152649
- Gifford, R. H., Bacon, S. P., and Williams, E. J. (2007). An examination of speech recognition in a modulated background and forward masking in young and older listeners. *J. Speech Lang. Hear. Res.* 50, 857–864. doi: 10.1044/1092-4388(2007)060
- Gordon-Salant, S. (2005). Hearing loss and aging: new research findings and clinical implications. *J. Rehabil. Res. Dev.* 42:9. doi: 10.1682/JRRD.2005.01.0006
- Grose, J. H., and Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear Hear.* 31, 755–760. doi: 10.1097/AUD.0b013e3181e627e7
- Grose, J. H., Mamo, S. K., and Hall III, J. W. (2009). Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 30, 568–575. doi: 10.1097/AUD.0b013e3181ac128f
- Hasher, L., and Zacks, R. T. (1988). “Working memory, comprehension, and aging: a review and a new view,” in *The Psychology of Learning and Motivation*, 22 Edn., ed G. Bower (New York, NY: Academic Press), 193–225.
- Hasher, L., Zacks, R. T., and May, C. P. (1999). “Inhibitory control, circadian arousal, and age,” in *Attention and Performance XVII*, eds D. Gopher and A. Koriati (Cambridge, MA: MIT Press), 653–675.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer-Verlag. doi: 10.1007/978-0-387-84858-7
- He, N., Horwitz, A. R., Dubno, J. R., and Mills, J. H. (1999). Psychometric functions for gap detection in noise measured from young and aged subjects. *J. Acoust. Soc. Am.* 106, 966–978. doi: 10.1121/1.427109
- He, N., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J. Acoust. Soc. Am.* 124, 3841–3849. doi: 10.1121/1.2998779
- He, N., Mills, J. H., and Dubno, J. R. (2007). Frequency modulation detection: effects of age, psychophysical method, and modulation waveform. *J. Acoust. Soc. Am.* 122, 467–477. doi: 10.1121/1.2741208
- Helfer, K. S., and Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear Hear.* 29, 87–98. doi: 10.1097/AUD.0b013e31815d638b
- Horn, J. I., and Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychol.* 26, 107–129. doi: 10.1016/0001-6918(67)90011-X
- Houtgast, T., and Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *Int. J. Audiol.* 47, 287–295. doi: 10.1080/14992020802127109
- Humes, L. E., Burk, M. H., Coughlin, M. P., Busey, T. A., and Strauser, L. E. (2007). Auditory speech recognition and visual text recognition in younger and older adults: similarities and differences between modalities and the effects of presentation rate. *J. Speech Lang. Hear. Res.* 50, 283–303. doi: 10.1044/1092-4388(2007)021
- Humes, L. E., and Dubno, J. R. (2010). “Factors affecting speech understanding in older adults,” in *The Aging Auditory System*, eds S. Gordon-Salant, R. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 211–258.
- Kausler, D. H. (1994). *Learning and Memory in Normal Aging*. San Diego, CA: Academic Press.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375
- Lin, F., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Int. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18866–18869. doi: 10.1073/pnas.0607364103
- Moore, B. C., and Sek, A. (1996). Detection of frequency modulation at low modulation rates: evidence for a mechanism based on phase locking. *J. Acoust. Soc. Am.* 100(4 Pt 1), 2320–2331. doi: 10.1121/1.417941
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Otolaryngol.* 9, 399–406. doi: 10.1007/s10162-008-0143-x
- Moore, B. C. J. (2012). “The importance of temporal fine structure for the intelligibility of speech in complex backgrounds,” in *Speech Perception and Auditory Disorders*, eds T. Dau, J. C. Dalsgaard, and T. Poulsen (Centertryk A/S), 21–32.
- Moore, B. C. J., and Sek, A. (1995). Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *J. Acoust. Soc. Am.* 97, 2468–2478. doi: 10.1121/1.411967
- Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex* 12, 313–324. doi: 10.1016/S0010-9452(76)80035-4
- Peters, R. W., Moore, B. C., and Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 103, 577–587. doi: 10.1121/1.421128
- Pichora-Fuller, M. K., and Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *J. Acoust. Soc. Am.* 91(4 Pt 1), 2129–2135. doi: 10.1121/1.403673
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Pichora-Fuller, M. K., Schneider, B. A., Macdonald, E., Pass, H. E., and Brown, S. (2007). Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. *Hear. Res.* 223, 114–121. doi: 10.1016/j.heares.2006.10.009
- Pichora-Fuller, M. K., and Souza, P. E. (2003). Effects of aging on auditory processing of speech. *Int. J. Audiol.* 42, S11–S16. doi: 10.3109/14992020309074638
- Plomp, R., and Mimpfen, A. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18, 43–52. doi: 10.3109/00206097909072618
- Purcell, D. W., John, S. M., Schneider, B. A., and Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *J. Acoust. Soc. Am.* 116, 3581–3593. doi: 10.1121/1.1798354
- Rajan, R., and Cainer, K. E. (2008). Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. *Neuroscience* 154, 784–795. doi: 10.1016/j.neuroscience.2008.03.067
- Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865. doi: 10.1093/brain/awg076
- Rhodes, M. G. (2004). Age-related differences in performance on the Wisconsin card sorting test: a meta-analytic review. *Psychol. Aging* 19, 482–494. doi: 10.1037/0882-7974.19.3.482
- Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1996). The structure of normal human attention: the test of everyday attention. *J. Int. Neuropsychol. Soc.* 2, 525–534. doi: 10.1017/S1355617700001697
- Roccaforte, W. H., Burke, W. J., Bayer, B. L., and Wengel, S. P. (1992). Validation of a telephone version of the mini-mental state examination. *J. Am. Geriatr. Soc.* 40, 697–702.
- Rönnerberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42(Suppl. 1), S68–S76. doi: 10.3109/14992020309074626
- Rönnerberg, J., Lyxell, B., Arlinger, S., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725

- Rönneberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 26326–26329. doi: 10.4103/1463-1741.70505
- Rosen, S., Souza, P., Ekelund, C., and Majeed, A. A. (2013). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J. Acoust. Soc. Am.* 133, 2431–2443. doi: 10.1121/1.4794379
- Rothauser, E. H., Chapman, N. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., et al. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.* 17, 225–246. doi: 10.1109/TAU.1969.1162058
- Salthouse, T. A. (1985). *A Theory of Cognitive Aging*. Amsterdam: North Holland.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403
- Schneider, B. A., and Hamstra, S. J. (1999). Gap detection thresholds as a function of tonal duration for younger and older listeners. *J. Acoust. Soc. Am.* 106, 371–380. doi: 10.1121/1.427062
- Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (2010). “The effects of senescent changes in audition and cognition on spoken language comprehension,” in *The Aging Auditory System, Springer Handbook of Auditory Research*, eds S. Gordon-Salant, R. D. Frisina, A. Popper, and D. Fay (Berlin: Springer), 167–210.
- Schooneveldt, G. P., and Moore, B. C. J. (1987). Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotonic versus dichotic presentation of the flanking band. *J. Acoust. Soc. Am.* 82, 1944–1956. doi: 10.1121/1.395639
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *J. Neurosci.* 33, 13686–13694. doi: 10.1523/JNEUROSCI.1783-13.2013
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303
- Snell, K. B. (1997). Age-related changes in temporal gap detection. *J. Acoust. Soc. Am.* 101, 2214–2220. doi: 10.1121/1.418205
- Stuart, A., and Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear Hear.* 17, 478–489. doi: 10.1097/00003446-199612000-00004
- Takahashi, G. A., and Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear. Res.* 35, 1410–1421. doi: 10.1044/jshr.3506.1410
- Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: PRO-ED Publishing, Inc.
- Tun, P. A., and Wingfield, A. (1999). One voice too many: adult age differences in language processing with different types of distracting sounds. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 54, P317–P327. doi: 10.1093/geronb/54B.5.P317
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., and Jolles, J. (2006). The Letter Digit Substitution Test: normative data for 1,858 healthy participants aged 24–81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex. *J. Clin. Exp. Neuropsychol.* 28, 998–1009. doi: 10.1080/13803390591004428
- Van der Linden, M., Brédart, S., and Beerten, A. (1994). Age-related differences in updating working memory. *Br. J. Psychol.* 85(Pt 1), 145–152. doi: 10.1111/j.2044-8295.1994.tb02514.x
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 66, 1364–1380. doi: 10.1121/1.383531
- Vongpaisal, T., and Pichora-Fuller, M. K. (2007). Effect of age on F0 difference limen and concurrent vowel identification. *J. Speech Lang. Hear. Res.* 50, 1139–1156. doi: 10.1044/1092-4388(2007/079)
- Zekveld, A. A., George, E. L. J., Kramer, S. E., Goverts, S. T., and Houtgast, T. (2007). The development of the text reception threshold test: a visual analogue of the speech reception threshold test. *J. Speech Lang. Hear. Res.* 50, 576–584. doi: 10.1044/1092-4388(2007/040)

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 May 2014; accepted: 21 October 2014; published online: 12 November 2014.

Citation: Schoof T and Rosen S (2014) The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Front. Aging Neurosci.* 6:307. doi: 10.3389/fnagi.2014.00307

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Schoof and Rosen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cognitive spare capacity in older adults with hearing loss

Sushmit Mishra¹, Stefan Stenfelt^{1,2}, Thomas Lunner^{1,2,3}, Jerker Rönnerberg¹ and Mary Rudner^{1*}

¹ Department of Behavioural Sciences and Learning, Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

² Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

³ Eriksholm Research Centre, Oticon A/S, Snekkærsten, Denmark

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Bruce A. Schneider, University of Toronto at Mississauga, Canada
Chad Rogers, Volen National Center for Complex Systems, USA

*Correspondence:

Mary Rudner, Department of Behavioural Sciences and Learning, Linköping University, SE-581 83 Linköping, Sweden
e-mail: mary.rudner@liu.se

Individual differences in working memory capacity (WMC) are associated with speech recognition in adverse conditions, reflecting the need to maintain and process speech fragments until lexical access can be achieved. When working memory resources are engaged in unlocking the lexicon, there is less Cognitive Spare Capacity (CSC) available for higher level processing of speech. CSC is essential for interpreting the linguistic content of speech input and preparing an appropriate response, that is, engaging in conversation. Previously, we showed, using a Cognitive Spare Capacity Test (CSCT) that in young adults with normal hearing, CSC was not generally related to WMC and that when CSC decreased in noise it could be restored by visual cues. In the present study, we investigated CSC in 24 older adults with age-related hearing loss, by administering the CSCT and a battery of cognitive tests. We found generally reduced CSC in older adults with hearing loss compared to the younger group in our previous study, probably because they had poorer cognitive skills and deployed them differently. Importantly, CSC was not reduced in the older group when listening conditions were optimal. Visual cues improved CSC more for this group than for the younger group in our previous study. CSC of older adults with hearing loss was not generally related to WMC but it was consistently related to episodic long term memory, suggesting that the efficiency of this processing bottleneck is important for executive processing of speech in this group.

Keywords: working memory, cognitive spare capacity, updating, inhibition, episodic long-term memory

INTRODUCTION

Communication is vital for social participation but may be hampered by adverse listening conditions. Such conditions arise when target speech is accompanied by background noise, possibly including other talkers, and when the listener has a hearing impairment (Mattys et al., 2012). Listening in adverse conditions is associated with individual working memory capacity (WMC, for reviews see Akeroyd, 2008; Besser et al., 2013). Working memory is the ability to maintain and process task-relevant information on line and is necessary for a wide range of complex cognitive activities including speech understanding (Baddeley, 2003). While listening under adverse conditions, individuals may maintain in working memory fragments of spoken information that are not masked by noise and process them to achieve speech understanding (Rönnerberg et al., 2013). This processing may involve executive functions including updating to add relevant information to working memory and inhibition to exclude irrelevant information (McCabe et al., 2010; Sörqvist et al., 2010; Rudner et al., 2011; Rudner and Lunner, 2013). There is also evidence that linguistic closure ability supports listening in adverse conditions (Besser et al., 2013; Zekveld et al., 2013) and this may be another process that is recruited into working memory during speech understanding. Because WMC is limited (Baddeley and Hitch, 1974; Just and Carpenter, 1992), it is likely that when cognitive processes are recruited during listening under adverse conditions, fewer resources are available for higher cognitive processing of heard speech (Pichora-Fuller, 2003; Arehart et al.,

2013; Mishra et al., 2013a,b; Rudner and Lunner, 2013). In other words, cognitive spare capacity (CSC) is reduced. WMC varies substantially between individuals (Just and Carpenter, 1992) and decreases with age (Nilsson et al., 1997; Nyberg et al., 2012). Thus, it is not surprising that there are individual differences in CSC (Mishra et al., 2013a,b) and there may also be age-related differences. However, work to date has shown that CSC is not reliably associated with WMC (Mishra et al., 2013a,b).

The role of working memory in speech understanding has been conceptualized in the Ease of Language Understanding model (ELU; Rönnerberg, 2003; Rönnerberg et al., 2008, 2013). The ELU model postulates that in optimal listening situations, when the listener is young and healthy and speech is clear, non-accented and in their native language without background noise or reverberation (Mattys et al., 2012), understanding is implicit or effortless as the incoming signal can be smoothly and rapidly matched with the lexical and phonological representations stored in the mental lexicon in long term memory (LTM, Luce and Pisoni, 1998). In adverse listening conditions, on the other hand, a mismatch may occur between the incoming speech signal and stored representations. This leads to explicit or conscious cognitive processing of the signal in order for language understanding to occur. Such processing may include inhibition, to keep working memory clear of irrelevant information such as speech produced by a non-target talker or other background noise, but also inappropriate inferences about individual words or the gist of the conversation (Pichora-Fuller, 2007). Updating skills are required to correctly

prioritize maintenance of information held in working memory in relation to new, incoming information and old information in the form of episodic or semantic representations in LTM. Linguistic closure skills are required to make efficient inferences concerning the lexical and semantic identity of spoken information on the basis of the ongoing processing of fragments of speech information held in working memory (Besser et al., 2013).

Aging is associated with both sensory and cognitive decline. Older adults usually have raised hearing thresholds and reduced spectral and temporal resolution (Gordon-Salant, 2005). In addition, they may have temporal auditory processing deficits (Pichora-Fuller and Souza, 2003). Further, reduced cognitive resources may make speech recognition more difficult for older adults compared to young (Pichora-Fuller and Singh, 2006; Mattys et al., 2012), due to impoverished encoding of the target stimuli into episodic memory, despite adequate recognition (Pichora-Fuller et al., 1995; Heinrich and Schneider, 2011; Sörqvist and Rönnerberg, 2012). When target stimuli are presented against a speech-like background, older individuals encounter additional problems because they are not as efficient at processing target information present in the gaps in the speech-like masker (George et al., 2007), nor at suppressing automatic linguistic processing of the irrelevant speech (Ben-David et al., 2012). Thus, factors relating to both sensory and cognitive decline are likely to increase the risk of mismatch between speech input and representations in LTM, leading to a relative decrease in CSC for older adults with hearing loss compared to younger adults with normal hearing under similar environmental circumstances. Further, there are differences in the way younger and older adults deploy their cognitive resources during speech understanding (Pichora-Fuller et al., 1995; Murphy et al., 2000; Wong et al., 2009) and these are likely to be compounded by working memory load or the amount of information that needs to be held in working memory at any one time (Reuter-Lorenz and Cappell, 2008). In particular, older adults may make full use of all available cognitive resources when load is still relatively low, while younger adults still have spare capacity to cope with higher load.

Recent work has demonstrated that a greater degree of hearing loss in older adults is associated with poorer LTM (Lin et al., 2011; Rönnerberg et al., 2011). It has been suggested that the mechanism behind this may be related to the mismatch function as described by the ELU model (Rönnerberg et al., 2013). In particular, if mismatch occurs regularly over an extended period of time, LTM may be accessed less frequently leading to disuse and less efficient LTM function (Rönnerberg et al., 2011; Classon et al., 2013). As we have seen, mismatch frequency in any given situation may be related to both sensory and cognitive factors and in older adults, general cognitive slowing may make it harder to resolve mismatch (Pichora-Fuller, 2003). Thus, relatively well-preserved cognitive processing speed and LTM efficiency are likely to enhance CSC in older adults with hearing loss.

While noise, hearing impairment and increasing age make it harder to understand speech, viewing the speaker's face may enhance the listener's ability to segregate the speech signal from the noise, thereby making it easier to understand (Campbell, 2009). It has been suggested that the presence of visual cues helps the listener, especially older adults, to attend to the incoming

signal at the most critical time for encoding (Helfer and Freyman, 2005). This results in less signal uncertainty and fewer cognitive demands in anticipating target stimuli than in the absence of visual cues (Besle et al., 2004; Moradi et al., 2013). Thus, visual cues seem to reduce the cognitive demands of listening in noise, especially for older adults. Seeing the talker's face during encoding improves WMC and in older adults has been shown to reduce neural activation, indicating a processing benefit (Frtusova et al., 2013).

To investigate CSC and assess the effect of noise, memory load and visual cues on executive processing of speech, we developed a test of CSC, CSCT. In the CSCT, participants listen to lists of spoken two-digit numbers presented serially with (AV) and without (A-only) a video of the talker's face. Presentation may take place in quiet or in background noise adjusted to a level at which good stimulus intelligibility is maintained. At the end of each list, participants recall two (low memory load) or three numbers (high memory load) depending on instructions eliciting one of two different executive functions (updating and inhibition). In young adults with normal hearing thresholds, CSCT performance is generally better for inhibition than updating, and low than high memory load (Mishra et al., 2013a,b). Steady-state background noise, which may be described as a rushing sound, reduces CSCT scores even when intelligibility is high (Mishra et al., 2013b). However, at the same SNR, speech-like background noise that consists of unintelligible speech fragments pieced together, does not seem to reduce CSCT scores in young adults with normal hearing thresholds, possibly because executive skills allowing dynamic tracking of target speech and simultaneous suppression of non-target speech leading to richer memory representation (Mishra et al., 2013b; Zion Golumbic et al., 2013). Although visual cues restore performance in steady-state background noise, probably by facilitating segregation of target speech from noise (Mishra et al., 2013b) they reduce performance in quiet, probably by causing distraction when the auditory signal provides adequate information to solve the CSC task (Mishra et al., 2013a,b). CSCT performance in young adults with normal hearing is predicted by updating skills but not, generally speaking, by WMC (Mishra et al., 2013a,b).

In the present study, we investigated CSC in adults with hearing loss who were older than the participants in our previous CSC studies (Mishra et al., 2013a,b). The CSCT and a cognitive test battery were administered. In the CSCT, audibility was restored by individualized amplification. We expected to find an overall pattern of results generally similar to that in our previous CSC studies but with stronger effects of load and noise manipulations due to an overall decrease in CSC attributable to age and hearing loss. Specifically, we expected better performance in inhibition than updating conditions and in low than high memory load conditions, as well as better performance in quiet than in noise. However, we also expected speech-like as well as steady-state noise to reduce CSCT scores due to age-related reduction in executive skills (Nyberg et al., 2012) and hearing-loss-related reduction in the ability to segregate target from background speech-like noise (Festen and Plomp, 1990; George et al., 2006, 2007; Lorenzi et al., 2006; Ben-David et al., 2012). We expected visual cues to restore performance in noise

conditions (Frtusova et al., 2013). Because younger and older adults use different strategies to deploy cognitive resources, we expected to find a different pattern of associations between CSCT performance and the cognitive test battery (Pichora-Fuller et al., 1995; Murphy et al., 2000; Reuter-Lorenz and Cappell, 2008; Wong et al., 2009; Avivi-Reich et al., 2014). In particular we expected better CSC to be associated with faster cognitive processing speed (Pichora-Fuller, 2003) and better LTM (Rönnberg et al., 2011).

METHODS

PARTICIPANTS

Twenty-seven adults with mild-to-moderate hearing loss and no reported tinnitus consented to participate in the study. They were all recruited from the hearing clinic at Linköping University Hospital, Sweden. Two participants opted to drop out of the testing and one participant was excluded due to poor vision. Thus, 24 participants (61–75 years of age, $M = 69$, $SD = 4.7$), 14 males and 10 females, completed the testing. All participants had sensorineural hearing loss (Air-Bone gap < 10 dB HL) and the average pure-tone hearing threshold (PTA₄) across 0.5, 1, 2, and 4 kHz was 34.5 dB HL ($SD = 3.6$), see **Figure 1**. An epidemiological study covering the same area showed that 73.1% of the population in the age range of 70–80 years and 42.1% in the age range of 60–70 years had at least a mild hearing loss (Johansson and Arlinger, 2003). In the present study, all participants had mild (PTA₄: 26–40 dB HL; WHO, 2013) hearing loss, except for one participant aged 74 years who had moderate (PTA₄: 41–60 dB HL; WHO, 2013) hearing loss. Hearing thresholds of all participants at all four frequencies were within one standard deviation of population means for the age group reported by Cruickshanks et al. (1998). Thus, hearing status was representative for their age group. The participants reported that their hearing loss was acquired post-lingually and that they did not have any otological, psychological or neurological problems. Three participants reported that they used hearing aids occasionally while the others were non-users. Visual acuity after correction was normal as measured using the Jaeger eye chart (Weatherly, 2002). Ethical approval for the study was obtained from the regional ethical review board.

COGNITIVE SPARE CAPACITY TEST (CSCT)

The CSCT is an auditory working memory task that systematically manipulates storage and executive processing demands along with modality of presentation and noise conditions (Mishra et al., 2013a,b).

Material

The CSCT stimuli consisted of AV and A-only recordings of Swedish two digit numbers 13–99 (Mishra et al., 2013a,b). The numbers were spoken by two native Swedish speakers, one male and one female, with no distinctive dialect. The levels of the numbers were equated for equal intelligibility in steady-state noise (Mishra et al., 2013b). The two-digit numbers were arranged serially in 48 lists of 13 numbers each. Numbers were never repeated within lists or condition and numbers spoken by same speaker were repeated between two and eight times across all the lists. Half

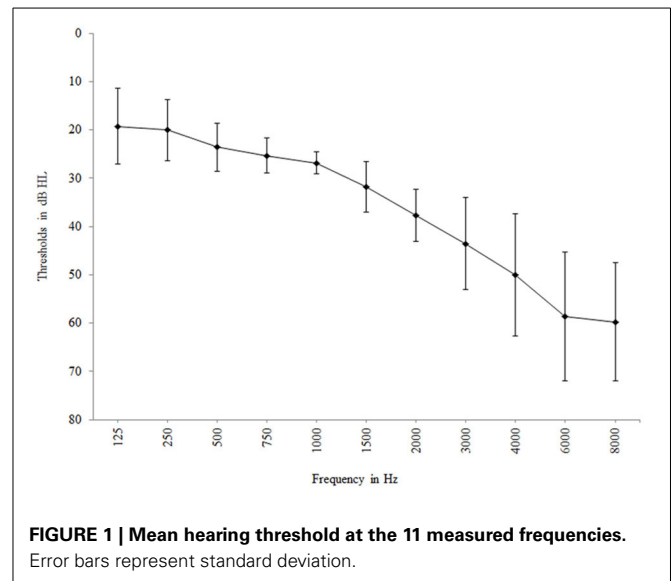


FIGURE 1 | Mean hearing threshold at the 11 measured frequencies. Error bars represent standard deviation.

of the lists included AV stimuli and the other half included A-only stimuli. Within each modality, 12 lists were used for each of the two CSCT tasks. The serial position of the target numbers within the lists and contingent task demands were balanced across the lists. For more details of CSCT materials, see Mishra et al. (2013a).

Noise

Stationary and modulated noises were used. The stationary noise was a steady-state speech-weighted (SSSW) noise, having the same long term average spectrum as the stimuli (numbers). The modulated noise was the International Speech Testing Signal (ISTS), which contains concatenated short segments of speech in six different languages (Holube et al., 2010) and is thus speech-like but unintelligible.

Individualizing SNR and amplification

Audibility is a key factor for speech understanding (Humes, 2007) and thus for optimizing the task in CSCT. Therefore, the CSCT lists were presented with amplification compensating for the hearing loss of the participants and at individualized SNR ensuring an intelligibility level of around 90% for the SSSW noise. An adaptive procedure implemented in MATLAB (Version 2009b) was used to determine the individualized SNR for presentation of the CSCT. This procedure was implemented in two steps and was based on the stimulus materials (numbers) in the A-only modality and the SSSW noise (Mishra et al., 2013b). In the first step, a number was presented at an SNR of 5 dB and the participant was instructed to repeat the number he or she heard. Then the noise was increased by steps of 3 dB each time the number was repeated correctly. When the participant made an incorrect response, the SNR was improved by 1 dB and a new number was presented. Thirty such randomly selected numbers, were presented consecutively with a step-size of 1 dB for the level of noise to determine the 84% intelligibility level in a four up-one down procedure (Levitt, 1971). In the second step, in order to achieve an intelligibility level of approximately 90% in SSSW noise, the SNR obtained for 84% intelligibility was

increased by 0.5 dB for each individual. To verify the individual intelligibility at this new SNR in both SSSW and ISTS noises, 60 randomly selected numbers were presented at the set SNR in SSSW and another 60 numbers in ISTS noise to measure the intelligibility levels for both noise types. In the CSCT, the ISTS noise was presented at exactly the same SNR as the SSSW noise.

The signal (numbers and noise) for all speech in noise tests with auditory presentation was amplified using the Cambridge prescriptive formula (Cameq) for linear hearing aids (Moore and Glasberg, 1998). This amplification was implemented in a master hearing aid (MHA) system (Grimm et al., 2006). The participant's audiogram was used to set the gain according to the Cameq fitting rule giving individual amplification for each participant.

Tasks

There are two different CSCT tasks, updating and inhibition which are designed to engage the corresponding executive functions. In the updating task, the participants are asked to recall either the highest or the lowest value item spoken by the male and female speaker in the particular list. Thus, each time an item is presented that meets the criterion, the participant has to encode this item into working memory and discard any previous item which it replaces. In the inhibition task, the participants are asked to recall either two odd or even value items spoken by a particular speaker. Thus, the participant has to inhibit encoding of items produced by a particular talker while monitoring items of the desired parity. These tasks are performed with either AV or A-only stimulus presentation. After each list, the participant is requested to report two specified list items, depending on the task to be performed. In half of the trials, which were the low memory load trials, only these two numbers are reported. The two specified numbers never include the first item in the list. In the other half of the trials, the participant is requested to report the first number in the list along with the two specified items, i.e., three numbers in total need to be held in working memory but only two of them are subject to executive processing. These are the high memory load trials. For these trials, the first number (dummy item) is not included in the scoring. Thus, all scoring in the CSCT is based on correct report, in any order, of two numbers.

Experimental design

All participants performed the CSCT with stimulus presentation in quiet (no noise), SSSW noise and ISTS noise. Thus, there were a total of 24 conditions of presentation in the CSCT with two executive tasks (Updating, inhibition), two memory loads (High, low), two modalities (AV, A-only) of presentation and three noise conditions (quiet, SSSW and ISTS noise) in a $2 \times 2 \times 2 \times 3$ design.

Administration of CSCT

The CSCT was administered using DMDX software (Forster and Forster, 2003; Mishra et al., 2013a,b). The participants performed the CSCT under 12 different conditions per executive task in separate blocks and hence two lists per condition were tested. The order of the conditions was pseudo-randomized within the two task blocks and balanced across the participants. For the noisy conditions, the noise sound files were played together with the

AV and A-only stimulus files in DMDX. The noise onset was 1 s prior to onset of stimulus and the noise offset was at least 1 s after the stimulus offset. The lists of numbers were always presented at 65 dB SPL and the level of the noise was varied depending upon the individualized SNR level before individualized amplification for hearing loss. The same individualized SNR was used for all noisy trials. Across all the conditions (noisy or quiet), the duration of presentation of each number list was 33 s in AV and A-only modality. The visual stimuli were presented using a computer with screen size of 14.1 inches and the amplified auditory stimuli were presented through Sennheiser HDA 200 headphones.

The participants were provided with written instructions for the particular executive task before each of the blocks and the instructions were also elaborated orally. In addition to this, before each list the participant was prompted on the computer screen as to which version of the executive task was to be performed, what the modality was and whether to remember two or three numbers (high or low load). The task prompt remained on screen until the participant pressed a button to continue to the test. At the end of each list, an instruction "Respond now" appeared on the screen and the participant was required to say the target numbers. Corrections to reported numbers were allowed and responses were audio recorded. The participant then pressed another button when they were ready to continue. All the participants practiced each task with two lists before doing the test. The participants were specifically instructed to keep looking at the screen during stimulus presentation. This applied even during presentation in the A-only modality where a fixation cross was provided at center screen. If they looked away from the screen, the test was stopped after presentation of the list and the participants were reinstructed to keep looking at the screen.

COGNITIVE TEST BATTERY

Reading span

The participants read series of sentences which appeared on the computer screen one at a time (Daneman and Carpenter, 1980; Rönnberg et al., 1989). Each series consisted of three to six sentences presented in increasing series length while each sentence consisted of three words. There was an interval of 50 ms between words and each word was shown for 800 ms. Half of the sentences were coherent and half were absurd. After each sentence, the participant was given 1.75 s to judge the semantic coherence of the sentence before the next once appeared. The participant responded "yes" (if the sentence was coherent) or "no" (if the sentence was absurd). At the end of each series of sentences, the participants were prompted by an instruction on the screen to recall either the first or the last word of all the sentences in the series in the order in which they appeared on the screen. All participants practiced with a series of three sentences before the actual testing and the practice was repeated if necessary. There were a total of 54 sentences in the actual test. The dependent measure was the total number of words correctly recalled in any order.

Text reception threshold (TRT)

The TRT (Zekveld et al., 2007) provides a measure of linguistic closure (Besser et al., 2013). In TRT, sentences in red appear

word by word on a computer screen partially masked by black bars. The participants are asked to guess the sentence correctly. A Swedish version of the TRT, using the Hearing in Noise Test (HINT) sentences (Hällgren et al., 2006) was used in the present study. First, a practice list with 20 sentences was presented which was followed by the actual testing where two similar lists were presented. All words remained on the screen until the sentence was completed and after presentation of the last word the sentence remained visible for 3.5 s. The presentation rate of the words in each sentence was equal to the speaking rate in a corresponding speaker file. If the participants were unsuccessful in reading the sentence, feedback was provided. A one-up-one-down adaptive procedure with a step-size of 6% was applied to target percentage of unmasked text required to read 50% of the sentences entirely correctly. The average percentage of unmasked text from the two lists of sentences was used as dependent variable.

Letter memory

The letter memory task (Morris and Jones, 1990; Miyake et al., 2000) was presented using a DMDX platform. Series of consonants were presented at the center of the computer screen. The participants were asked to hold the four most recent letters in mind and then prompted to say them at the end of each series. Responses were audio-recorded. In order to ensure that the participant followed the instructed strategy and continuously updated working memory until the end of the trial, series length was randomized across trials. Two series consisting of seven and nine letters were presented as practice and the actual testing consisted of 12 series varying in length between five and 11 items. The practice sequences were repeated until participants followed the instructed strategy. The number of consonants correctly recalled irrespective of order was the independent measure of updating.

Simon

A visual analog of the Simon task (Simon, 1969; adapted from Pratte et al., 2010), consisting of presentation of red and blue rectangular blocks on a computer screen, was used to provide a measure of the inhibition. The blocks appeared on the left or the right of the computer screen successively at intervals of 2 s. The participants were instructed to respond as quickly as possible while maintaining accuracy by pressing a button on the right hand side of the screen when they saw a red block and when they saw a blue block they pressed a button on the left hand side of the screen. A total of 16 blocks were presented using DMDX. No practice item was provided. When the spatial position of the stimulus and correct response key coincided, the trial was termed congruent otherwise incongruent. The participant had to ignore the spatial position in which the block appeared in the task. The difference in reaction time between the incongruent and congruent trials was taken as the dependent variable. The mean reaction time obtained on the congruent trial of the Simon task for each participant was taken as a measure of processing speed.

Delayed recall of reading span

A delayed free recall of the reading span test (Mishra et al., 2013b) was used to measure the episodic LTM of the participants. In this test the participants were asked to recall words or

sentences remembered from the reading span test after approximately 60 min, without forewarning. During the 60 min, the participants performed the other tests in the cognitive test battery. The score in the delayed free recall of the reading span test was the total number of words recalled by the participant, irrespective of the order and the performance in the reading span test. The participants did not have any time restriction to recall the words or sentences.

PROCEDURE

The testing was conducted in two sessions. All auditory testing took place in a sound-treated booth with the participants facing the computer screen. Each session took approximately 90 min. The participants, on arriving for the testing, were fully briefed about the study and a consent form was signed. The participants were provided with written instructions about the test and instructions were verbally elaborated if needed. All the participants underwent vision screening and audiometric testing in the audiometric booth. In a separate room, the reading span test was administered followed by the Simon task, the letter memory test and the TRT test. Individual SNRs for the CSCT were determined and the delayed recall of the reading span test was the last test of the first session. In the second session, CSCT was conducted. The participants were allowed to take breaks between the tests.

DATA ANALYSIS

To ensure that the performance of the three occasional hearing-aid users did not differ from that of the other participants, it was checked that their scores in the various tests were within one standard deviation of the mean score of the participants who did not use hearing aids. An overall repeated measures analysis of variance (ANOVA) on the CSCT scores was conducted. The inter-correlation among the cognitive tests and the association between cognitive functions and CSCT was assessed using Pearson's correlations.

RESULTS

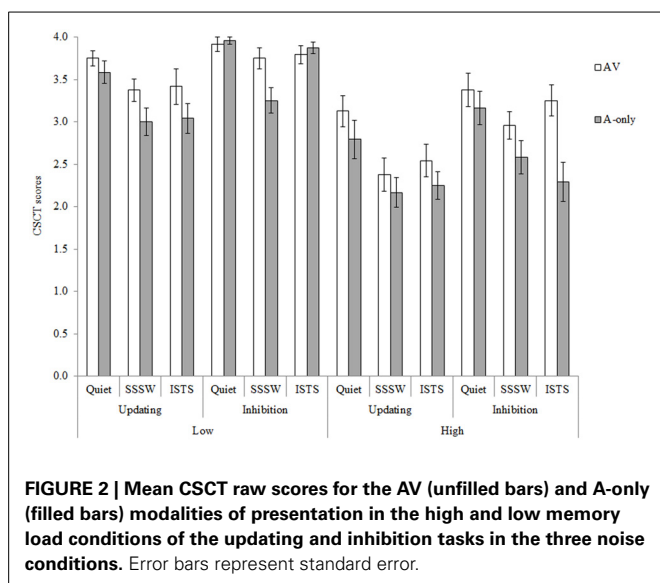
INTELLIGIBILITY

The mean SNR for CSCT presentation in noise was -0.17 dB ($SD = 1.39$). The mean intelligibility levels were 94.5% ($SD = 3.0$) and 88.3% ($SD = 3.0$) for the SSSW and ISTS noise, respectively. The difference between these levels was statistically significant, $t_{(46)} = 7.05$, $p < 0.01$.

COGNITIVE SPARE CAPACITY TEST (CSCT)

Mean raw scores are shown in **Figure 2**. The maximum possible score per condition was four, as two lists were presented per condition. Performance in the inhibition task in the low memory load for quiet and ISTS noise conditions approached ceiling. Hence, all analyses of CSCT data were conducted on the rationalized arcsine-transformed scores (Studebaker, 1985) to counteract data skewing. Performance on the updating and inhibition subsets correlated significantly, ($r = 0.57$, $p < 0.01$), confirming internal and construct validity.

The overall repeated measures ANOVA revealed main effects of all four factors: executive function, $F_{(1, 23)} = 30.00$, $MSE = 0.23$, $p < 0.001$, showing higher CSCT scores in inhibition



than updating conditions; memory load, $F_{(1, 23)} = 71.71$, $MSE = 0.35$, $p < 0.001$, showing higher CSCT scores in low than high memory load conditions; modality, $F_{(1, 23)} = 26.23$, $MSE = 0.14$, $p < 0.001$, showing CSCT scores were higher in AV than A-only conditions and noise, $F_{(2, 46)} = 23.78$, $MSE = 0.18$, $p < 0.001$. Pair-wise comparisons with Bonferroni adjustment for multiple comparisons were conducted in order to identify statistically significant differences in performance between the three noise types. They showed that CSCT performance in quiet was better than in both ISTS and SSSW noise ($p < 0.05$), but there was no difference between performance in ISTS and SSSW noise ($p = 0.13$). All main effects were in line with our prediction. It should be borne in mind here that intelligibility was significantly higher in SSSW noise than in ISTS noise. To test whether this difference in intelligibility influenced memory performance, we examined whether there was a difference in recall of the first list item in the high memory load conditions when items were presented in SSSW noise compared to ISTS noise. There was no statistically significant difference, $t_{(46)} = 0.01$, $p > 0.05$. This suggests that the lack of difference in CSCT performance in SSSW and ISTS noise is not an artifact of intelligibility differences. There were no statistically significant Two-Way or Three-Way interactions.

COGNITIVE TEST BATTERY

Table 1 shows the mean performance and standard deviation in the cognitive test battery. In the reading span semantic judgment task, the mean score was 50.5 ($SD = 3.20$) out of 54 possible responses, demonstrating adherence to instructions. We excluded the delayed recall of reading span score of one participant that was more than two standard deviations above the mean score.

Table 2 shows the correlations among the cognitive tests, PTA₄ threshold and age. PTA₄ threshold was associated with age, and reading span was associated with letter memory and delayed recall of reading span, see **Table 2**.

Table 3 shows the overall and factor-wise association between CSCT performance and cognitive skills. Delayed recall of reading

Table 1 | Mean performance and standard deviation (SD) in the cognitive test battery and results of two-tailed independent sample *t*-tests with young adults with normal hearing thresholds included in the reanalysis.

Cognitive test	Units	Mean	SD	<i>t</i> , <i>p</i>
Reading span	Words recalled (max 54)	21.38	4.61	4.8, 0.00
Letter memory	Letters recalled (max 48)	37.79	4.09	3.65, 0.00
Simon	Difference in reaction time in ms between incongruent and congruent trials	133.46	82.84	2.25, 0.03
Simon congruent trials	Reaction time in ms	628.29	140.12	2.11, 0.04
TRT	Percentage unmasked text	52.53	4.62	3.23, 0.00
Delayed recall of reading span	Words recalled	5.90, <i>n</i> = 23	2.90	4.91, 0.00

span was associated with CSCT irrespective of how scores were split. Letter memory was associated with the overall CSCT score as well as performance in CSCT updating, A-only, high memory load, quiet and ISTS noise conditions. There was no statistically significant correlation between Simon and the inhibition conditions of the CSCT ($p = 0.63$). However, TRT was associated with performance in inhibition conditions. A higher score in TRT indicates poorer performance. Therefore, the negative correlation shows that better TRT performance is associated with better CSCT performance in the inhibition condition. Working memory as measured by the reading span test correlated significantly with the performance of CSCT in ISTS noise conditions. Reaction times on the congruent trials of the Simon task, which was our measure of processing speed, did not correlate significantly with CSCT performance. We did not include a measure of motor processing speed in the present study and thus we cannot exclude the possibility that cognitive processing speed was confounded by differences in motor skills.

DISCUSSION

In the present study we investigated CSC for speech heard in quiet and in noise in adults with hearing loss in AV and A-only modality of presentation. We did this by administering the CSCT (Mishra et al., 2013a,b), a test of CSC that measures individuals' ability to perform executive processing of heard material at different memory loads. The CSCT was presented with individualized amplification in SSSW and ISTS noise as well as in quiet. In the two noise conditions, it was presented at an estimated speech intelligibility level of approximately 90%.

CSCT PERFORMANCE

In line with predictions, performance was better when memory load was low compared to high and when the task involved

Table 2 | Coefficients of correlations (Pearson's *r*) between age, average pure tone thresholds across the four frequencies 0.5, 1, 2, and 4 kHz (PTA₄) and cognitive test scores.

	PTA ₄	Reading span	Letter memory	Simon	TRT	Delayed recall of reading span	Simon congruent trials
Age	0.51*	−0.12	−0.11	−0.28	0.34	−0.19	0.35
PTA ₄		−0.11	−0.18	−0.29	0.23	−0.27	−0.03
Reading span			0.51*	−0.27	0.02	0.44*	−0.03
Letter memory				0.21	−0.36	0.36	−0.12
Simon					−0.40	−0.23	−0.05
TRT						−0.07	0.37
Delayed recall of reading span							−0.19

*Correlation is significant at the 0.05 level (two-tailed).

Table 3 | Coefficients of correlations (Pearson's *r*) between factorwise CSCT scores and cognitive test scores.

CSCT	Reading span	Letter memory	Simon	TRT	Delayed recall of reading span	Simon congruent trials
Overall	0.29	0.51*	−0.05	−0.36	0.64**	−0.29
Updating	0.33	0.50*	−0.16	−0.23	0.69**	−0.18
Inhibition	0.18	0.40	0.11	−0.42*	0.41*	−0.33
AV	0.29	0.38	0.04	−0.36	0.61**	−0.33
A-only	0.25	0.57**	−0.12	−0.31	0.60**	−0.22
Low load	0.21	0.26	−0.29	−0.13	0.42*	−0.35
High load	0.17	0.50*	0.03	−0.39	0.58**	−0.22
Quiet	0.27	0.46*	−0.03	−0.29	0.77**	−0.26
SSSW	−0.09	0.18	−0.07	−0.24	0.66**	−0.24
ISTS	0.44*	0.54**	−0.15	−0.32	0.77**	−0.25

*Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed).

inhibitory processing rather than updating. Also in line with expectation, performance was better in quiet than in noise and in the AV than A-only modality in noise and quiet. This was contrary to our previous study (Mishra et al., 2013a,b) in which visual cues hindered performance in the quiet conditions. In Mishra et al. (2013b) when young adults performed CSCT in quiet and in noise, the Two-Way interaction between modality and noise was significant, revealing higher CSCT scores in the A-only compared to the AV modality in quiet and the opposite in noise. Although the finding of poorer performance with visual cues is unexpected in relation to much of the perceptual and cognitive literature, it is in line with the results of recent studies showing that superfluous information carried in the visual stream may reduce performance on a dual task paradigm (Fraser et al., 2010; Gosselin and Gagné, 2011). This phenomenon may arise when executive demands make it difficult to prioritize task-related processing in the presence of low priority stimuli (Lavie, 2005). We proposed that in conditions where all the information needed to solve the task was available in the auditory signal, assuming optimum speech intelligibility for participants with normal hearing listening in quiet, the visual cues constituted a distraction (Mishra et al., 2013a,b). It has been shown that the presence of visual cues reduces the cognitive demands for perception of speech (Besle et al., 2004; Moradi et al., 2013). The reduction in cognitive demands leads to better representation of the target signal in memory (Pichora-Fuller et al., 1995; Heinrich and Schneider, 2011). Even in quiet, the signal is degraded for older adults with hearing loss due to

receiver limitation (Mattys et al., 2012). Hence, we suggest that even in quiet conditions, seeing the talker's face helps older individuals with hearing loss to form better cognitive representations of spoken words leading to higher performance in AV modality, possibly by viseme and phoneme information working together (Feld and Sommers, 2009).

CSCT performance was poorer in both types of noise than in quiet, whereas in our previous study, only SSSW reduced CSCT performance for young adults (Mishra et al., 2013b). It has been demonstrated that being older and having a hearing loss are associated with poorer speech segregation, especially when noise is speech-like (Festen and Plomp, 1990; George et al., 2006, 2007; Ben-David et al., 2012), probably because relevant cognitive functions are less efficient and deployed differently (Pichora-Fuller et al., 1995; Murphy et al., 2000; Reuter-Lorenz and Cappell, 2008; Wong et al., 2009). We suggest that this is also the cause of lower CSC in older adults with hearing loss.

EXECUTIVE PROCESSING IN CSCT

To check that the CSCT tasks tapped into the intended executive functions, we investigated correlations with the cognitive test battery. The performance on the updating task of CSCT collapsed across other factors correlated with performance on the letter memory task, confirming previous results (Mishra et al., 2013a,b) and showing that the updating task of the CSCT does tap updating skills. However, as we found in our previous study using the same paradigm (Mishra et al., 2013b) the correlation between

performance on the inhibition task of the CSCT and the Simon task was not statistically significant. In both studies, noise was introduced for two out of three lists in an unpredictable manner. Thus, inhibition skills available for solving the executive task may have been reduced, even for the lists presented in quiet. An alternative explanation could be that the Simon task does not measure inhibition. However, we find this explanation implausible as we found a significant correlation ($r = -0.46$) between Simon and CSCT performance under inhibition conditions in a study in which noise was not presented in any condition (Mishra et al., 2013a). Depletion of inhibition skills was probably compounded in the present study by the reduced executive skills of the participants (c.f. Ben-David et al., 2012). The statistically significant association between CSCT performance in inhibition conditions and TRT performance, which provides a measure of linguistic closure, found in the present study for adults with hearing loss as well as in our previous study for adults without hearing loss (Mishra et al., 2013b), suggests that linguistic closure may compensate for depleted executive skills. Zekveld et al. (2012) pointed out that TRT predicted speech perception in noise when the irrelevant cues for speech understanding had to be disregarded or inhibited. Similarly in the inhibition task of CSCT, the numbers of same parity spoken by the opposite gender had to be disregarded. Further studies should investigate the interplay of executive function and linguistic closure ability during higher processing of speech.

The main effects of executive function, memory load and modality revealed that CSCT performance was lower when the task was updating, memory load was high and the visual cues were absent. Under these conditions, CSCT performance consistently correlated with updating skills. Thus, good updating skills seem to be particularly important for higher level processing of speech when task demands are particularly high for older individuals with hearing loss. This finding contrasts with the finding of our previous study using the same paradigm (Mishra et al., 2013b) in which good updating skills were associated with good CSCT performance in almost all conditions. In the present study, consistent associations between CSCT performance and delayed recall of reading span were found across the board.

ROLE OF WORKING MEMORY AND EPISODIC LONG-TERM MEMORY (LTM) IN CSCT PERFORMANCE

Performance in CSCT was not significantly associated with performance in the reading span test except in the ISTS noise conditions. Previous work has shown that speech recognition in modulated noise, especially speech noise, is associated with WMC (Zekveld et al., 2013). This may be because listening in modulated noise involves integrating fragments of information available in the dips in the noise (Lunner, 2003). Hence, it is not surprising to find that CSCT performance in ISTS noise but not in SSSW noise was associated with WMC as measured by reading span performance. In line with our prediction, better episodic LTM, as measured by delayed recall of reading span, was consistently associated with better CSCT performance. Recent work has demonstrated that older adults with hearing loss may have a limited LTM (Lin et al., 2011; Rönnberg et al., 2011). Thus, LTM may form a processing bottleneck for this group and individuals

with more efficient LTM are likely to be able to process speech with fewer demands on cognitive resources (Rönnberg et al., 2013), resulting in larger CSC. This interpretation is also in line with notion that there are age-related changes in depth of processing of heard material (Craig and Rose, 2012). In older adults, a general cognitive slowing makes matching of the incoming signal with representations stored in LTM more effortful and susceptible to errors (Pichora-Fuller, 2003) and hence faster processing speed may lead to higher scores in CSCT. However, in the present study we did not find any such evidence.

To compare the performance of the participants with hearing impairment in the present study with that of participants without hearing impairment in a previous study (Mishra et al., 2013b), a reanalysis was performed. We expected that the CSCT and cognitive test scores of the participants in the present study would be lower than those of younger adults with normal hearing in our previous study. Further, we expected the CSCT scores of the participants in the present study would be more influenced by high memory load, noise (Pichora-Fuller et al., 1995; Heinrich and Schneider, 2011) and the absence of visual cues (Frtusova et al., 2013).

REANALYSIS

In the present study, the participants were older adults with hearing loss. The hearing thresholds of these participants were similar to those of similar age cohorts in epidemiological studies (e.g., Cruickshanks et al., 1998; Johansson and Arlinger, 2003), suggesting that the hearing loss was age-related. Older adults with normal hearing were not selected in this study because such a group is not representative of the population of older adults. Thus, the reanalysis explores the effect of aging and concomitant auditory decline on CSCT performance. This is achieved by comparing the data of the present study with those of a previous study (Mishra et al., 2013b) where CSCT was administered to young adults with normal hearing.

METHODS

Participants

The reanalysis included the participants in the present study and the 20 young adult participants in a previous study (Mishra et al., 2013b). The young adults were 19–35 years of age ($M = 25.9$; $SD = 4.4$) and the mean PTA₄ 3.2 dB HL ($SD = 3.2$). There were statistically significant differences in age ($t = 30.77$, $p < 0.01$) and PTA₄ ($t = 23.25$, $p < 0.01$) with the participants in the present study being older and having higher hearing thresholds.

Material and noise

The material consisted of lists of two digits numbers presented in AV and A-only modality which were prepared into lists, each containing 13 numbers. The lists were presented in quiet and in SSSW and ISTS noise at intelligibility levels approximating 90%. The material and noise were identical in both the studies.

Individualizing SNR and amplification

The methods used for individualizing SNR were identical for the two groups. However, because the younger participants did not have any hearing loss they were not provided with amplification.

Tasks, experimental design, and procedure

In both the studies, the testing was conducted in two sessions. In the first session the audiometric testing, vision screening, cognitive test battery and individualized SNR for CSCT presentation was determined. The cognitive test battery consisted of reading span, Simon, letter memory, TRT, and delayed recall of reading span. The order of testing was also identical in both studies. For CSCT, in both studies, order of the conditions was pseudo-randomized within the two executive task blocks and was balanced across the participants in the same manner. Hence, the task, experimental design and procedure were identical in both studies.

Data analysis

The data were analyzed using a mixed repeated measures ANOVA on CSCT scores with the two groups of participants as a between subjects variable. Where significant interactions were obtained in ANOVAs, the simple main effects observed were investigated using *post-hoc* Tukey's Honestly Significant Difference (HSD) test. In order to test simple main effects in accordance with our a-priori hypothesis, planned comparisons were carried out. Independent sample *t*-tests were used to compare performance across groups on the cognitive test battery.

RESULTS

Intelligibility

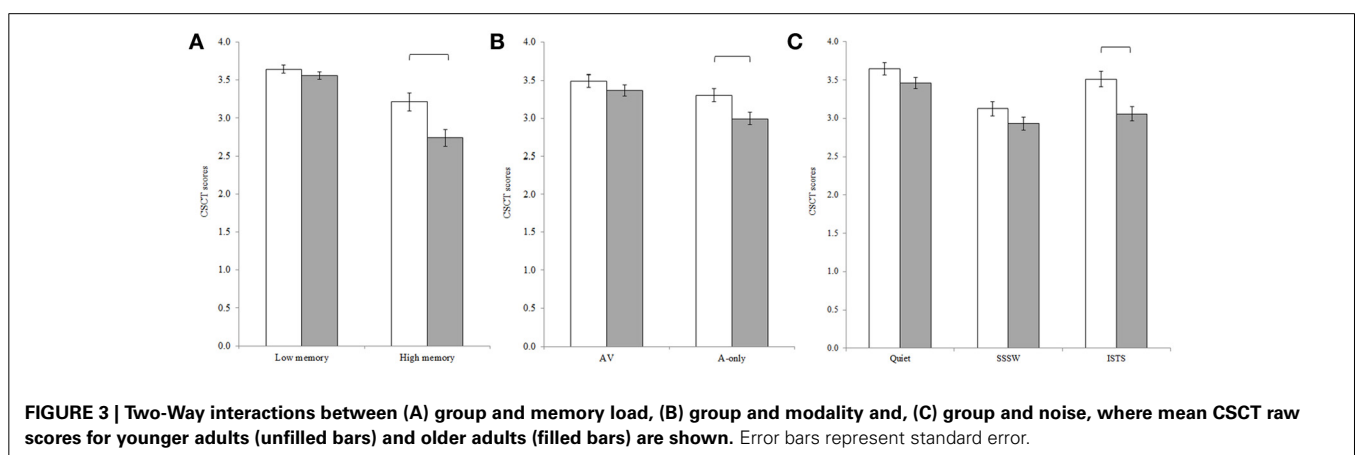
The mean SNR for CSCT presentation in noise for the older adults with hearing loss was -0.17 dB ($SD = 1.39$) and for the young adults, the mean SNR of presentation was -2.17 dB ($SD = 0.85$; Mishra et al., 2013b). The mean intelligibility levels for older adults with hearing loss were 94.5% ($SD = 3.0$) and 88.3% ($SD = 3.0$) for the SSSW and ISTS noise respectively. For the young adults, the mean intelligibility level was 93.8% ($SD = 3.0$) and 92.3% ($SD = 2.9$) for the SSSW and ISTS noise, respectively, (Mishra et al., 2013b). A mixed repeated measures ANOVA conducted on the actual intelligibility levels in SSSW and ISTS noise showed that there was a main effect of noise, $F_{(1, 42)} = 43.86$, $MSE = 0.01$, $p < 0.001$, indicating that the intelligibility levels were higher in SSSW compared to ISTS noise. The Two-Way interaction between group and noise was also statistically significant, $F_{(1, 42)} = 12.49$, $MSE = 0.01$, $p = 0.001$. *Post-hoc* Tukey HSD tests assessing this interaction revealed that in the SSSW

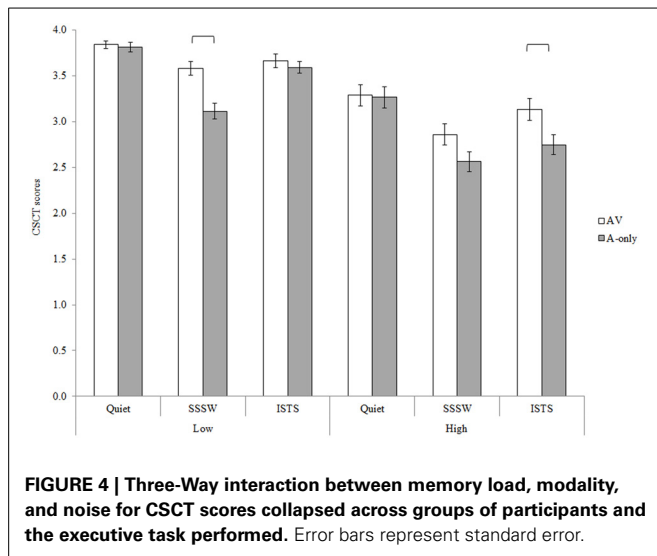
noise there was no statistically significant difference in intelligibility levels between groups, but in ISTS noise, the intelligibility level for the participants in the present study was significantly lower than that for the young adults with normal hearing, statistically.

CSCT

The repeated measures ANOVA revealed a main effect of group, $F_{(1, 42)} = 7.78$, $MSE = 0.02$, $p < 0.01$, showing that the participants in the present study had lower CSCT scores compared to young adults with normal hearing thresholds. In line with results for the two groups separately, main effects of executive function, $F_{(1, 42)} = 35.80$, $MSE = 0.25$, $p < 0.001$; memory load, $F_{(1, 42)} = 93.44$, $MSE = 0.29$, $p < 0.001$; modality, $F_{(1, 42)} = 26.31$, $MSE = 0.12$, $p < 0.001$, and noise, $F_{(2, 84)} = 37.84$, $MSE = 0.18$, $p < 0.001$, were observed. Pair-wise comparisons with Bonferroni adjustment for multiple comparisons revealed that the CSCT scores in quiet were significantly higher than the scores in ISTS noise ($p = 0.001$) which in turn were significantly higher than the scores in SSSW noise ($p < 0.001$). The Two-Way interactions between group and memory load [$F_{(1, 42)} = 7.52$, $MSE = 0.29$, $p < 0.01$], group and modality [$F_{(1, 42)} = 5.07$, $MSE = 0.12$, $p < 0.05$], and group and noise [$F_{(2, 84)} = 3.64$, $MSE = 0.19$, $p < 0.05$] were statistically significant, see Figure 3. *Post-hoc* Tukey HSD tests assessing these Two-Way interactions revealed, in line with our predictions, that although the participants in the present study had significantly lower CSCT scores compared to the younger adults in high memory load conditions, there was no statistically significant difference in performance between groups in low load conditions. Further, although participants in the present study had significantly lower CSCT scores in the A-only modality than the younger adults, there was no statistically significant difference in performance between groups in AV conditions. Participants in the present study had lower CSCT scores than younger adults in ISTS noise but there was no statistically significant difference in performance between groups in quiet or in SSSW noise.

The Two-Way interaction between modality and noise [$F_{(2, 84)} = 9.25$, $MSE = 0.11$, $p < 0.01$] and the Three-Way interaction between memory load, modality and noise [$F_{(2, 84)} = 3.48$, $MSE = 0.13$, $p < 0.05$], see Figure 4, were also statistically





significant. Neither of these interactions interacted further with group.

Post-hoc Tukey's HSD tests investigating the Two-Way interaction between modality and noise revealed that visual cues significantly enhanced performance in SSSW noise ($p < 0.01$) and ISTS noise ($p < 0.01$) but not in quiet. Investigation of the Three-Way interaction revealed that the findings of the Two-Way interaction were modulated by memory load. In particular, although visual cues enhanced CSCT scores in ISTS noise ($p < 0.05$) when memory load was high, this was not the case when memory load was low. Visual cues enhanced CSCT scores in SSSW noise ($p < 0.01$) in low memory load conditions, but in high memory load conditions the difference in scores did not reach significance with *post-hoc* testing.

Cognitive test battery

Independent sample *t*-tests showed that, in line with our prediction, the performance of the participants in the present study was significantly poorer in all the cognitive tests than that of the young adults with normal hearing thresholds (Mishra et al., 2013b, see Table 1). We also found that the mean reaction time for congruent trials in the Simon task in the present study was significantly longer than that found for the younger adults statistically, showing that the participants in the present study had a slower cognitive processing speed than the young adults.

DISCUSSION

On combining the data of CSCT performance by the older adults with hearing loss in the present study and the young adults without hearing loss in the previous study (Mishra et al., 2013b), we found a main effect of group revealing lower CSCT scores for the older adults with hearing loss compared to the younger adults without hearing loss. Even though intelligibility was held relatively constant, it is likely that the background noise placed a greater burden on cognitive resources in the older adults with hearing loss because of low level auditory processing deficits (Festen and Plomp, 1990; George et al., 2006, 2007; Ben-David et al., 2012). As expected, the older group also performed worse

on the cognitive test battery (c.f. Salthouse, 1980; Rönnberg, 1990; Pichora-Fuller and Singh, 2006). Thus, lower CSC for older adults with hearing loss is probably due both to poorer fundamental cognitive abilities and more pressure on those abilities while listening in noise. Examination of the Two-Way interactions with the group factor revealed that the poorer performance of the older adults was driven mainly by performance differences in more challenging conditions: high memory load, A-only modality of presentation and in ISTS noise, in line with our prediction. Across groups, the benefit of visual cues was evident only in noise. It is likely that the visual cues make it easier to distinguish target speech from background noise thus reducing the demand for executive resources during segregation (Besle et al., 2004; Helfer and Freyman, 2005; Mishra et al., 2013b; Moradi et al., 2013). When executive resources are spared during listening, more are likely to be available for solving the CSCT, leading to better performance. In other words, visual cues enhance CSC during listening in noise by freeing executive resources. However, the benefit of visual cues in noise was modulated by memory load and thus further work is needed to investigate the interplay of visual cues, memory load and noise.

GENERAL DISCUSSION

The findings of the present study further our understanding of CSC. Using the CSCT, we found lower CSC in older adults with hearing loss compared to the younger adults with normal hearing whom we had tested with the same experimental paradigm in a previous study (Mishra et al., 2013b). This was despite the fact that amplification was provided to compensate for hearing loss and that SNR was individualized to ensure intelligibility. Performance was also significantly poorer on all of the tests in the cognitive test battery. Thus, we suggest that poorer CSC in older adults with hearing loss compared to younger adults with normal hearing is probably due both to poorer cognitive abilities and more pressure on those abilities, rather than differences in hearing thresholds.

As we had predicted, factors previously found to decrease CSC in young adults with normal hearing (Mishra et al., 2013b) had an even stronger effect on the older individuals with hearing loss. In particular, increasing memory load and removing visual cues reduced CSC more for the older group. Further, whereas, steady state noise, but not speech-like, noise reduced CSC for the younger participants in our previous study (Mishra et al., 2013b), both kinds of noise reduced CSC for the older participants in the present study. Lower CSC for older adults with hearing loss in speech-like noise is likely to be related to the poorer ability to segregate target speech from non-target speech, that is well-attested in the literature. This poorer ability probably has several causes at a number of levels including less efficient processing of target information present in the gaps in the speech-like masker (George et al., 2007) poorer inhibition of irrelevant speech (Ben-David et al., 2012), impoverished encoding of target stimuli (Pichora-Fuller et al., 1995; Heinrich and Schneider, 2011; Sörqvist and Rönnberg, 2012) lower WMC (Nyberg et al., 2012) and differences in deployment of cognitive resources during speech understanding (Pichora-Fuller et al., 1995; Murphy et al., 2000; Wong et al., 2009). Further research is needed to tease apart these effects.

The finding of a stronger effect of increasing memory load on CSC for older adults with hearing loss compared to younger adults with normal hearing is in tune with work showing differences in the deployment of cognitive resources for younger and older adults in response to changes in memory load. In particular, imaging studies have shown that older adults display a greater increase in brain activity in response to cognitive load than younger adults, and that brain activity reaches a plateau at a lower level in older adults than in younger adults (Grady, 2012).

Importantly, visual cues enhanced CSC for older adults with hearing loss. Further, this effect was stronger for this group than for young adults with normal hearing. This finding demonstrates that for older adults with hearing loss, visual cues can support the kind of executive processing of speech that may be used in everyday conversation (c.f. Frtusova et al., 2013), possibly because of viseme and phoneme information working together (Feld and Sommers, 2009). Because intelligibility was held relatively constant between groups while cognitive skills were poorer for the older group, it is likely that the mechanism behind this phenomenon is related to cognitive processes.

It is worth noticing that although the older adults with hearing loss generally had reduced CSC compared to young adults; this did not apply in all conditions. When CSCT tasks were performed in quiet, with low memory load and in presence of visual cues, there was no significant difference across groups. This suggests that when listening conditions are optimized, there is no difference in CSC between older adults with hearing loss and younger adults with normal hearing.

In the present study, we found that better CSCT performance in older adults with hearing loss was associated with better updating skills, but only in those conditions in which the participants performed worse than the younger adults with normal hearing who took part in our previous study using the same experimental paradigm (Mishra et al., 2013b). In that study, we found that updating skills predicted CSCT performance in virtually all conditions and suggested that updating skills became particularly important in CSC when inhibition resources are depleted by constantly being prepared to cope with background noise. In the present study, we found that better episodic LTM was associated with better CSCT performance. This pattern of findings further supports the notion that younger adults with normal hearing and older adults with hearing loss deploy cognitive resources differently, especially in relation to changes in task difficulty (Pichora-Fuller et al., 1995; Murphy et al., 2000; Wong et al., 2009; Grady, 2012). In particular, it suggests that LTM may form a processing bottleneck for this group (c.f. Lin et al., 2011) and that more efficient LTM may allow more efficient executive processing of speech with less depletion of CSC (Rönnberg et al., 2013).

CONCLUSION

Older adults with hearing loss have lower CSC than young adults without hearing loss, probably because they have poorer cognitive skills and deploy them differently. However, visual cues and efficient episodic LTM enhance CSC more for the older group.

ACKNOWLEDGMENTS

This study was supported by grant number 2007-0788 to Mary Rudner from the Swedish Council for Working Life and Social Research. We thank Mathias Hällgren for technical support.

REFERENCES

- Akeroyd, M. A. (2008). Are individual differences in speech perception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142
- Arehart, K. H., Souza, P., Baca, R., and Kates, J. M. (2013). Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear.* 34, 251–260. doi: 10.1097/AUD.0b013e318271aa5e
- Avivi-Reich, M., Daneman, M., and Schneider, B. A. (2014). How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment. *Front. Syst. Neurosci.* 8:21. doi: 10.3389/fnsys.2014.00021
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201
- Baddeley, A., and Hitch, G. (1974). "Working memory," in *Recent Advances in Learning and Motivation*, Vol. 8, ed G. A. Bower (New York, NY: Academic Press), 47–90.
- Ben-David, B. M., Tse, V. Y., and Schneider, B. A. (2012). Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker? *Hear. Res.* 290, 55–63. doi: 10.1016/j.heares.2012.04.022
- Besle, J., Fort, A., Delpuech, C., and Giard, M. H. (2004). Bimodal speech: early suppressive visual effects in the human auditory cortex. *Eur. J. Neurosci.* 20, 2225–2234. doi: 10.1111/j.1460-9568.2004.03670.x
- Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459
- Campbell, R. (2009). "The processing of audiovisual speech: empirical and neural bases," in *The Perception of Speech: From Sound to Meaning*, 1st Edn., eds B. C. J. Moore, L. K. Tyler, and W. Marslen-Wilson (Oxford, NY: Oxford University Press), 133–150.
- Classon, E., Rudner, M., and Rönnberg, J. (2013). Working memory compensates for hearing related phonological processing deficit. *J. Commun. Disord.* 46, 17–29. doi: 10.1016/j.jcomdis.2012.10.001
- Craik, F. I., and Rose, N. S. (2012). Memory encoding and aging: a neurocognitive perspective. *Neurosci. Biobehav. Rev.* 36, 1729–1739. doi: 10.1016/j.neubiorev.2011.11.007
- Cruikshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., et al. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The epidemiology of hearing loss study. *Am. J. Epidemiol.* 148, 879–886. doi: 10.1093/oxfordjournals.aje.a009713
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Feld, J., and Sommers, M. (2009). Lipreading, processing speed, and working memory in younger and older adults. *J. Speech Lang. Hear. Res.* 52, 1555–1565. doi: 10.1044/1092-4388(2009/08-0137)
- Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736. doi: 10.1121/1.400247
- Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124. doi: 10.3758/BF03195503
- Fraser, S., Gagné, J.-P., Alepins, M., and Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech-cues. *J. Speech Lang. Hear. Res.* 53, 18–33. doi: 10.1044/1092-4388(2009/08-0140)
- Frtusova, J. B., Winneke, A. H., and Phillips, N. A. (2013). ERP evidence that auditory-visual speech facilitates working memory in younger and older adults. *Psychol. Aging* 28, 481–494. doi: 10.1037/a0031243
- George, E. L. J., Festen, J. M., and Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing impaired listeners. *J. Acoust. Soc. Am.* 120, 2295–2311. doi: 10.1121/1.2266530

- George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *J. Acoust. Soc. Am.* 121, 2362–2375. doi: 10.1121/1.2642072
- Gordon-Salant, S. (2005). Hearing loss and aging: new research findings and clinical implications. *J. Rehabil. Res. Dev.* 42, 9–24. doi: 10.1682/JRRD.2005.01.0006
- Gosselin, P. A., and Gagné, J.-P. (2011). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int. J. Audiol.* 50, 786–792. doi: 10.3109/14992027.2011.599870
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505. doi: 10.1038/nrn3256
- Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (2006). The master hearing aid: a PC-based platform for algorithm development and evaluation. *Acta Acust. United Ac.* 92, 618–628.
- Hällgren, M., Larsby, B., and Arlinger, S. (2006). A Swedish version of the Hearing in Noise Test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–237. doi: 10.1080/14992020500429583
- Heinrich, A., and Schneider, B. A. (2011). The effect of presentation level on memory performance. *Ear Hear.* 32, 524–532. doi: 10.1097/AUD.0b013e31820a0281
- Helfer, K. S., and Freyman, R. L. (2005). The role of visual speech cues in reducing energetic and informational masking. *J. Acoust. Soc. Am.* 117, 842–849. doi: 10.1121/1.1836832
- Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). Development and analysis of an international speech test signal (ISTS). *Int. J. Audiol.* 49, 891–903. doi: 10.3109/14992027.2010.506889
- Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *J. Am. Acad. Audiol.* 18, 590–603. doi: 10.3766/jaaa.18.7.6
- Johansson, M. S. K., and Arlinger, S. (2003). Prevalence of hearing impairment in population in Sweden. *Int. J. Audiol.* 42, 18–28. doi: 10.3109/14992020309056081
- Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295X.99.1.122
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477.
- Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., and Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore longitudinal study of aging. *Neuropsychology* 25, 763. doi:10.1037/a0024238
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18866–18869. doi: 10.1073/pnas.0607364103
- Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19, 1–36. doi: 10.1097/00003446-199802000-00001
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, S49–S58. doi: 10.3109/14992020309074624
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006
- McCabe, D. P., Roediger, H. L. III, McDaniel, M. A., Balota, D. A., and Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology* 24, 222–243. doi: 10.1037/a0017619
- Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013a). Visual information can hinder working memory processing of speech. *J. Speech Lang. Hear. Res.* 56, 1120–1132. doi: 10.1044/1092-4388(2012/12-0033)
- Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013b). Seeing the talker's face supports executive processing of speech in steady state noise. *Front. Syst. Neurosci.* 7:96. doi: 10.3389/fnsys.2013.00096
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. (2000). The unity and diversity of executive functions and their contribution to complex frontal lobe tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi:10.1006/cogp.1999.0734
- Moore, B. C., and Glasberg, B. R. (1998). Use of a loudness model for hearing-aid fitting. I. Linear hearing aids. *Br. J. Audiol.* 32, 317–335. doi: 10.3109/03005364000000083
- Moradi, S., Lidestam, B., and Rönnberg, R. (2013). Gated audiovisual speech identification in silence versus noise: effect on time and accuracy. *Front. Psychol.* 4:359. doi: 10.3389/fpsyg.2013.00359
- Morris, N., and Jones, D. M. (1990). Memory updating in working memory: the role of the central executive. *Br. J. Psychol.* 81, 111–121. doi: 10.1111/j.2044-8295.1990.tb02349.x
- Murphy, D. R., Craik, F. I. M., Li, K. Z. H., and Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323
- Nilsson, L. G., Backman, L., Erngrund, K., Nyberg, L., Adolfsson, R., Bucht, G., et al. (1997). The Betula prospective cohort study: memory, health, and aging. *Aging Neuropsychol. Cogn.* 4, 1–32. doi: 10.1080/13825589708256633
- Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., and Bäckman, L. (2012). Memory aging and brain maintenance. *Trends Cogn. Sci.* 16, 292–305. doi: 10.1016/j.tics.2012.04.005
- Pichora-Fuller, M. K. (2003). Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *Int. J. Audiol.* 42, S59–S67. doi: 10.3109/14992020309074625
- Pichora-Fuller, M. K. (2007). “Audition and cognition: what audiologists need to know about listening,” in *Hearing Care for Adults*, eds C. Palmer and R. Seewald (Stäfa: Phonak), 71–85.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103
- Pichora-Fuller, M. K., and Souza, P. E. (2003). Effects of aging on auditory processing of speech. *Int. J. Audiol.* 42, S11–S16. doi: 10.3109/14992020309074638
- Pratte, M. S., Rouder, J. N., and Morey, R. D., and Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Atten. Percept. Psychophys.* 72, 2013–2025. doi:10.3758/APP.72.7.2013
- Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Rönnberg, J. (1990). Cognitive and communication function: the effects of chronological age and “handicap age.” *Eur. J. Cogn. Psychol.* 2, 253–273. doi: 10.1080/09541449008406207
- Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, S68–S76. doi: 10.3109/14992020309074626
- Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Lang. Hear. Res.* 32, 725–735.
- Rönnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternäng, O., Wahlin, A., et al. (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *J. Speech Lang. Hear. Res.* 54, 705–726. doi: 10.1044/1092-4388(2010/09-0088)
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167
- Rudner, M., Foo, C., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4
- Rudner, M., and Lunner, T. (2013). Cognitive spare capacity as a window on hearing aid benefit. *Semin. Hear.* 34, 297–306. doi: 10.1055/s-0033-1356642
- Salthouse, T. (1980). “Age and memory: strategies for localizing the loss,” in *New Direction in Memory and Aging*, eds L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, and L. W. Thompson (Hillsdale, NJ: Lawrence Erlbaum Associates Inc), 47–65.
- Simon, J. R. (1969). Reactions towards the source of stimulation. *J. Exp. Psychol.* 81, 174–176. doi: 10.1037/h0027448

- Sörqvist, P., Ljungberg, J. K., and Ljung, R. (2010). A sub-process view of working memory capacity: evidence from effects of speech on prose memory. *Memory* 18, 310–326. doi: 10.1080/09658211003601530
- Sörqvist, P., and Rönnerberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what role for working memory capacity? *J. Speech Lang. Hear. Res.* 55, 210–218. doi: 10.1044/1092-4388(2011/10-0353)
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *J. Speech Lang. Hear. Res.* 28, 455–462.
- Weatherly, S. L. (2002). Back to basics: testing visual acuity with the Jaeger eye chart. *Mater. Eval.* 60, 928–929.
- WHO. (2013). Available online at: http://www.who.int/pbd/deafness/hearing_impairment_grades/en/
- Wong, P. C., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., and Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* 47, 693–703. doi: 10.1016/j.neuropsychologia.2008.11.032
- Zekveld, A. A., George, E. L. J., Kramer, S. E., Govers, S. T., and Houtgast, T. (2007). The development of the text reception threshold test: a visual analogue of the Speech reception threshold test. *J. Speech Lang. Hear. Res.* 50, 576–584. doi: 10.1044/1092-4388(2007/040)
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., Heslenfeld, D. J., and Rönnerberg, J. (2012). Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain Lang.* 122, 103–113. doi: 10.1016/j.bandl.2012.05.006
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., and Rönnerberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *J. Acoust. Soc. Am.* 134, 2225–2234. doi: 10.1121/1.4817926
- Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, S., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanism underlying selective neuronal tracking of attended speech at a “cocktail party.” *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 December 2013; accepted: 29 April 2014; published online: 19 May 2014.
Citation: Mishra S, Stenfelt S, Lunner T, Rönnerberg J and Rudner M (2014) Cognitive spare capacity in older adults with hearing loss. *Front. Aging Neurosci.* 6:96. doi: 10.3389/fnagi.2014.00096

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Mishra, Stenfelt, Lunner, Rönnerberg and Rudner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The effect of functional hearing loss and age on long- and short-term visuospatial memory: evidence from the UK biobank resource

Jerker Rönnerberg^{1*}, Staffan Hygge², Gitte Keidser³ and Mary Rudner¹

¹ Linnaeus Centre HEAD, Department of Behavioural Sciences and Learning, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

² Environmental Psychology, Faculty of Engineering and Sustainable Development, University of Gävle, Gävle, Sweden

³ National Acoustic Laboratories, Sydney, Australia

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Steve Majerus, Université de Liège, Belgium

David R. Moore, University of Cincinnati College of Medicine, USA

*Correspondence:

Jerker Rönnerberg, Linnaeus Centre HEAD, Department of Behavioral Sciences and Learning, Swedish Institute for Disability Research, Linköping University, SE-581 83 Linköping, Sweden
e-mail: jerker.ronnerberg@liu.se

The UK Biobank offers cross-sectional epidemiological data collected on >500,000 individuals in the UK between 40 and 70 years of age. Using the UK Biobank data, the aim of this study was to investigate the effects of functional hearing loss and hearing aid usage on visuospatial memory function. This selection of variables resulted in a sub-sample of 138,098 participants after discarding extreme values. A digit triplets functional hearing test was used to divide the participants into three groups: poor, insufficient and normal hearers. We found negative relationships between functional hearing loss and both visuospatial working memory (i.e., a card pair matching task) and visuospatial, episodic long-term memory (i.e., a prospective memory task), with the strongest association for episodic long-term memory. The use of hearing aids showed a small positive effect for working memory performance for the poor hearers, but did not have any influence on episodic long-term memory. Age also showed strong main effects for both memory tasks and interacted with gender and education for the long-term memory task. Broader theoretical implications based on a memory systems approach will be discussed and compared to theoretical alternatives.

Keywords: visuospatial tasks, memory systems, functional hearing loss, age, hearing aids

INTRODUCTION

There is sufficient evidence to conclude that there is a connection between sensory decline and cognitive decline. Decline in one function is associated with decline in the other and the strength of the association has been empirically shown to increase with increasing age (Lindenberger and Baltes, 1994; Baltes and Lindenberger, 1997; Valentijn et al., 2005). This may suggest that there is some kind of common cause (e.g., neural degeneration) that explains the association, but more recent longitudinal evidence does not unequivocally support this hypothesis (Lindenberger and Ghisletta, 2009). Another explanation is that the sensory loss actually causes the cognitive decline (called the sensory deprivation hypothesis), and a third alternative is that cognitive decline drives sensory loss (Baltes and Lindenberger, 1997).

In this paper we focus on what might be dubbed the interactive hypothesis. Under this hypothesis, research has targeted mechanisms that underlie the online interaction (e.g., during speech understanding) between different hearing-related perceptual aspects on the one hand and cognitive aspects on the other. One such mechanism is perceptual stress or perceptual degradation, where it is typically assumed that even when stimuli are audible, the hearing loss affects the quality of encoding of

memory items (e.g., Pichora-Fuller, 2003; McCoy et al., 2005). Another mechanism relates to the attentional costs that may be involved, implying that even a mild hearing loss draws on central attention resources, hence affecting memory encoding negatively (e.g., Sarampalis et al., 2009; Tun et al., 2009; Heinrich and Schneider, 2011). Still another possibility is that the long-term cognitive consequences of hearing loss strike selectively at different memory systems, even when audibility is high at testing (Rönnerberg et al., 2011; Ng et al., 2013), and even when the to-be-remembered items are encoded in modalities other than audition (e.g., motor encoding, Rönnerberg et al., 2011).

In this study, we pursue this memory systems approach with strictly non-auditory encoding conditions so as to minimize hearing-related perceptual encoding problems, hence making a conservative test of the set of hypotheses that hearing loss affects encoding more generally (i.e., independently of encoding conditions), that the locus of the effect is at the level of memory systems, and that there is selectivity in terms of which system is most affected. We outline the reasons for the predictions below:

In Rönnerberg et al. (2011) it was found that hearing loss had a negative effect on both episodic and semantic long-term memory, but not on short-term/working memory. This

held true even when chronological age was statistically controlled for and for tasks that did not rely solely on auditory encoding, thus minimizing the reliance on potential perceptual degradation (e.g., Schneider et al., 2002) or attentional effort (e.g., Tun et al., 2009). Using linear Structural Equation Models (SEM), Rönnerberg et al. (2011) demonstrated that models that *combined* the degree of hearing loss with the degree of visual acuity did not make satisfactory predictions of memory decline for *any* memory system. Thus, the results suggest that relative decline in a memory system is tightly connected specifically to hearing loss rather than to sensory decline in general.

Rönnerberg et al. (2011) explained their findings on the basis of relative use/disuse of memory systems, essentially stating that working or short-term memory is often occupied with storage of heard words and with reconstruction and repair of misheard words or sentences, whereas episodic long-term memory will become relatively less used in individuals with hearing loss because of the higher probability of mismatches (or no-matches) between input phonology and stored phonological representations of words in semantic long-term memory. Therefore, unlocking of the lexicon, and hence, episodic memory encoding/retrieval, will occur to a lesser extent for individuals with hearing loss than for individuals with normal hearing while working or short-term memory will be engaged to the same extent, if not more.

The prediction regarding semantic long-term memory based on a use/disuse concept is less clear because it could be argued that semantic and contextual knowledge would have to be used more than episodic memory to compensate for misheard or non-matching words (Rönnerberg et al., 2008, 2011, 2013). This is evident e.g., in studies of false hearing, where older adults rely to a larger extent on context (Rogers et al., 2012). However, the data suggest a decline due to hearing loss even for semantic memory, especially for phonologically sensitive fluency tasks (Rönnerberg et al., 2011) and for nonword recall tasks (Janse and Newman, 2013).

Testing the short-term/working memory system in more detail, Verhaegen et al. (2014) have recently shown that, especially in auditory short-term memory tasks that rely on serial recall of words, there is an effect of hearing loss that is not related to age (see also Pichora-Fuller et al., 1995; van Boxtel et al., 2000; Schneider et al., 2010). This effect occurs even when the hearing loss of the study sample was mild (25–30 dB). They also argued that the results did not support the neural degeneration hypothesis (i.e., an example of a common cause) since young and old participants with hearing loss performed on a par, thus leaving most of the explanatory power to hearing status and not to age, as both groups were outperformed by a third group of young individuals with normal hearing. It was further reasoned that because speeded non-word repetition was intact even in the hearing-impaired groups, the actual perceptual processes were intact. It was proposed, in line with several other studies (cf. McCoy et al., 2005; Wingfield et al., 2005; Tun et al., 2009; Piquado et al., 2010), that increased demands on attention may instead be a plausible hypothesis regarding the mechanism involved (Verhaegen et al., 2014).

In the current study, based on a large sample ($N = 138,098$) of people not using hearing aids from the much larger UK Biobank Resource ($N > 500,000$), we therefore focused on the effects of hearing loss and age on memory tasks that were *not* confounded by possible auditory perceptual degradation, or by attentional demands related to hearing difficulties, strictly testing the memory systems hypothesis.

Testing the memory systems hypothesis, we used two types of memory tasks, tapping visuo-spatial working memory and visuo-spatial episodic long-term memory, respectively. The working memory task was a card-pair matching game in which participants had to remember cards that were the same (pictures of ordinary animals/objects like e.g., cat/ball) after having had a short inspection time. Two versions of the task were employed, an easy one with three pairs (which was considered to be a warm-up task), and a more difficult one with six pairs (loading highly on visuospatial memory). Thus, we opted for the six pair version in our analysis to maximise the demands on working memory.

As a proxy for episodic long-term memory function and to determine whether we could replicate the negative effect of hearing loss on episodic long-term memory (Rönnerberg et al., 2011), we used a prospective long-term memory task, a task that has a clear episodic long-term memory component (Burgess and Shallice, 1997). At the beginning of the session, subjects were given instructions (written on the computer screen) stating that they were to touch a colored shape when prompted at the end of the session. Crucially, they were also informed that the prompt on the screen would say *blue square*, but as a prospective memory test, they should instead touch the *orange circle*.

Although short-term memory has been shown to be affected by hearing loss (Verhaegen et al., 2014), it should be noted that the data by Rönnerberg et al. (2011) suggest that working memory/short-term memory is *relatively* less affected by hearing loss than episodic long-term memory. This is the central hypothesis in the present study. Thus, by using the two visuospatial memory indices briefly described above, we were able to make a very conservative test of the hypothesis that functional hearing loss is more strongly related to episodic long-term memory decline than to short-term or working memory decline and that these declines are not caused by perceptual degradation or lack of attention resources. Semantic memory measures were not included in the present study.

In a separate sample from the UK Biobank resource ($N = 3751$, see “Additional Analyses” section below), we also checked for the effects of hearing aid usage, with the hypothesis that this may have a protective effect against memory decline (Rönnerberg et al., 2011). This has not been examined in detail in previous studies: for example, in Rönnerberg et al. (2011) we only used data from individuals with hearing loss who were also users of hearing aids, in the seminal studies reported by Baltes and Lindenberger (1997) hearing aid usage was not separately accounted for (see Arlinger, 2003), and in the Verhaegen et al. (2014) study, the participant sample did *not* use hearing aids.

Finally, as we used visuospatial memory tests, we also deemed it appropriate to use two simple measures of visual acuity/vision problems as another sensory-specific possibility to

explain any hearing loss-related decline. In this way we can cast more light on the influential Baltes-Lindenberger common-cause hypothesis.

The sample from the UK Biobank resource used in the present study is extremely large compared to that used in any other study in the literature on this topic. It will guarantee statistical power and generalizability.

METHODS

OVERALL SAMPLE

The UK Biobank resource consists of data obtained from more than 500,000 participants. In the present study, we excluded participants who were born outside of the UK and the Republic of Ireland, as unknown language and cultural differences may significantly affect their cognitive abilities. We also excluded participants whose data sets were incomplete across measures of hearing and cognition. In addition, in the first main analyses we did not include hearing aid users (HAUse). This resulted in a study sample of 138,098 participants. Among these, 75,065 were females and 63,033 were male; giving a slightly skewed ratio of 54/46 (%). Age ranged from 39–70 years reflecting the UK Biobank population as a whole.

SUBJECTIVE REPORTS

The UK Biobank population also answered yes/no questions about “difficulty with hearing in general” ($N = 439,510$) and “difficulty following a conversation if there is background noise (such as TV, radio, children playing)” ($N = 448,416$). Among the UK population, 114,717 (25%) reported having general difficulty with hearing, 169,055 (37%) had difficulty hearing in noise, and 14,010 (3%) wore hearing aids. In our sample of 138,098 persons, we had data for 130,206 on reported general difficulty with hearing, and 24% reported such difficulty. For hearing in noise, we had data for 134,673 persons and 34% reported difficulties with that. With respect to hearing aid usage, 3751 persons (2.6%) in our sample reported wearing a hearing aid.

Furthermore, participants were asked whether they wore glasses (no/yes) and whether they had diagnosed eye problems/disorders other than those corrected for by the use of glasses. In our sub-sample of 138,098, 89% (of 137,978) reported having eye-glasses and 88% (of 101,845) reported having no additional eye-problems.

Participants were also asked which of six qualifications they had obtained. To simplify further analyses, a new highest level of qualification variable was created that assumes that a College or University degree (rated 1) > A levels/AS levels (rated 2) > O levels/GSEs (rated 3) > CSEs (rated 4) > NVQ or HND or HNC (rated 5) > Other professional qualifications; e.g., nursing or teaching (rated 6). In our sub-sample, we had valid values for 116,947 on qualification and the distribution across qualification levels 1–6 was 38.8%, 13.8%, 26.6%, 7.1%, 7.7%, and 6.0%, respectively.

The study presented here is covered by a Research Tissue Bank approval obtained by UK Biobank from its governing Research Ethics Committee, as recommended by the National Research Ethics Service.

TESTS

Participants attended 1 of 22 assessment centers spread throughout the UK. All test data used in this study were obtained through a self-administered program running of a computer with a touch screen that collected responses to questionnaires and tests on hearing in noise and cognition. Incomplete data sets were collected as it was possible for participants to be selective in which questionnaires and tests they responded to.

The digit triplets test (DTT)

The participants completed a functional hearing test in which they were presented with digit triplets in a steady state, speech-shaped noise (Smits et al., 2004) and had to enter (on a numberpad shown on the touchscreen) which three digits they had heard (forced choice). The speech reception threshold in noise (SRTn) was the SNR arrived at after 15 presentations, during which noise was adaptively changed after each presentation depending on whether the three digits were correctly identified or not. These SNR could vary between -12 and $+8$ dB, where a high and positive score indicated worse hearing. Each ear was tested separately (unaided) under headphones. As a first step a best ear SRTn variable was created to be used in further analyses. One reason for choosing the best ear is that it dominates auditory function in daily life, and is typically used in insurance compensation for assessment of e.g., occupational hearing loss (Dobie, 1996, see also Dawes et al. (2014)). For those who only completed the test on one ear, it is assumed that this was the better ear, and this result is recorded. As a second step, we classified the participants on the basis of the criteria used by Dawes et al. (2014), where “normal” hearing was assumed for SRTn values below -5.5 dB, “insufficient” hearing as -5.5 to -3.5 dB, and “poor” hearing as a threshold above -3.5 dB (variable was denoted Hear). This classification, in turn, was based on earlier work within the HearCom project (Smits et al., 2004; Vlaming et al., 2011).

Smits et al. (2004) found a relatively high correlation between the Dutch DTT and pure tone audiometry of $r = 0.77$. One reason for a lack of perfect correlation is that people with similar audiograms can have different psychoacoustic profiles (e.g., individual differences in frequency and temporal resolution) and hence perform differently when listening to speech in noise. Therefore, it seems reasonable that DTT also has been found to correlate highly with speech-in noise-recognition measures (such as with Plomp and Mimpen (1979) Sentences in Noise; $r = 0.85$; Smits et al., 2004). Together, the DTT can be considered as a functional hearing test (Dawes et al., 2014); see also the General Discussion section below.

Cognitive tests

Four tests of cognitive function were performed in the following order: (1) Prospective Memory test: Shape—Part 1; (2) Pairs memory test; (3) Verbal Reasoning test; (4) Reaction time: Snap; (5) Prospective Memory test: Shape—Part 2. We here describe the pairs matching and the prospective memory tests, as they are used for the short-term—long-term memory distinction relevant to this paper. Data on reverse digit span were also available from the UK Biobank resource but were not used in the present study with its focus on visuo-spatial memory function.

Pairs memory test: visuospatial working memory (VSWM). VSWM was measured with a pairs matching game. Participants were presented first with a round of three pairs of cards depicting different designs of objects and then, twice, with a round of six pairs of cards. The layout was purely random each time. There were no specific selection criteria applied to choosing the designs of the pictures other than that they should look reasonably distinct. Thus, there were no systematic phonological or semantic relationships between the English lexical labels of the pairs of objects. During each round, the pictures were turned over after a short inspection period. The 2×3 layout was shown for 3 s before pointing and the 2×6 layout was shown for 5 s. The participants were asked to identify as many pairs as possible with the fewest attempts by touching “pairs” of the same object on the screen. When the participant made an error, this was indicated by the word “miss” appearing at the center of the screen. When the participant gave a correct answer, the word “pair” would appear on the screen. For each correctly identified pair, the cards were removed and two blank spaces were left in the position where they had previously been placed. The participants could continue until they had identified all pairs. Time allowed for matching of pairs was unrestricted. The participants were allowed to continue until they had discovered all pairs correctly. The dependent variable is thus the number of errors made before all the pairs had been matched. We considered the three-pairs round a warm-up trial for the six pairs round, which constituted the dependent variable.

Prospective long-term memory (PLTM). PLTM consisted of two parts:

Part 1. The initial instruction to the participant was the following: “At the end of the games we will show you four colored shapes and ask you to touch the Blue Square. However, to test your memory, we want you to actually touch the Orange Circle instead. Once the “Next” button was touched, a hidden timer was started to record the delay interval until the answer to this question (asked after the reaction time test) was requested. Then the Pairs matching test, the Fluid intelligence test and the Reaction time (Snap) test were performed.

Part 2. After the Reaction time (Snap) test was finished, the following text was shown to the participant: “That’s the last game. Just one more thing left to do...”. The participant then selected “Next”; and the Shapes screen appeared with the text: “Please touch the Blue Square then touch the “Next” button” was presented. At this point the delay interval timing ended. If the participant touched any of the symbols it was highlighted by surrounding it in a yellow box. If the participant touched the Next button without having highlighted a symbol they were shown the message: “Please touch a symbol (a colored shape) before touching the “Next” button” If the participant then touched any symbol other than the Blue Square, then Next, the test ended. If the participant touched the Blue Square, they were prompted with the message: “At the start of the games we asked you to remember to touch a different symbol when this screen appeared. Please try to remember which symbol it was and touch it now”. If the participant touched the Blue Square again then this message was repeated (*ad infinitum*), otherwise the program

accepted their new selection and the test ended. The dependent variable was scored in three steps: correct at first attempt, correct at a subsequent attempt, and not correct at first or following attempts (which were given the scores 1, 2, and 3, respectively).

RATIONALE FOR THE STATISTICAL ANALYSES

For the memory measures, logarithmic transformation of the number of errors made in VSWM and the errors scores in PLTM were computed (for both measures: natural logarithm of $x + 1$) to counteract the skewed distribution of the raw scores. Also, for the analyses of the VSWM and PLTM tasks, individuals with values above the 99th percentile on the six pairs matching tasks were excluded to build in a safeguard against outliers. Our initial analyses were also restricted to participants who did not use hearing aids.

To be able to compare error rates on the dependent variables VSWM and PLTM in the ANOVAs, rather than in regression analyses with dummy coding of the interactions, the age and the hearing variables were divided into sub-groups. Our aim was to have at least about 100 observations for each combination of age and hearing status. With the functional hearing status variable already divided into three groups (Good, Insufficient, and Poor), as suggested by Dawes et al. (2014); see also Smits et al. (2004)), and outlined above in The Digit-Triplets Test (DTT) section, a choice had to be made about age-group spans.

We preferred four age spans, and that the two middle spans would be 10 years. With hearing status groups already defined, the pragmatic solution was to move the two middle 10 year age spans down from the maximum age of 70 years in our sample, and ensure that the N in the smallest Age \times Hearing status groups were ≈ 100 or more. With these criteria our oldest group was defined as >67 years, and the youngest as <48 years, with two 10-year age spans in between.

RESULTS

The Age by Hearing status distribution is shown in **Table 1** of our $N = 138,098$ in our subsample. **Table 1** also shows the defining criteria for the three hearing status groups: Normal, Insufficient, and Poor.

Table 2 shows the dichotomized fractions of men and people with an education other than University, College, A level, AS level in the Age \times Hearing status groups. These fractions do not vary substantially between sub-groups, but the means in the groups were statistically evaluated in our subsequent analyses (see below under Section Additional Analyses).

We also decided to take a parametric approach to how to treat the logarithmic error scores for VSWM and PLTM. The basic issue is whether it can be justified to treat the scores as being on an interval scale, and reanalyze them with parametric tests, such as ANOVA, or whether the data only meet ordinal scale properties and should thus be subjected to non-parametric tests. We concluded that an ANOVA approach is justified, but we will discuss the pros and cons of that at the end of the Results Section and also provide non-parametric analyses of our data to support the parametric statistical analyses.

Table 1 | Number of persons in Age-groups and the three-step functional hearing status groups.

	Age	Hear			Total
		Normal < -5.5	Insuff -5.5 to -3.5	Poor > -3.5	
1	<48	23147	881	90	24118
2	48–57	38724	2369	197	41290
3	58–67	55617	7340	835	63792
4	>67	7113	1567	218	8898
Total		124601	12157	1340	138098

Table 2 | Proportions of men (1st fraction in each cell of the table) and proportions of persons with an education other than University, College, A level, AS level (2nd fraction) in the Age by Hearing status groups.

Age		Hear						Total	
		Normal < −5.5		Insuff −5.5 to −3.5		Poor > −3.5			
		Men	LoEduc	Men	LoEduc	Men	LoEduc	Men	LoEduc
1	<48	0.45	0.44	0.42	0.54	0.37	0.66	0.45	0.45
2	48–57	0.43	0.45	0.40	0.50	0.46	0.55	0.43	0.45
3	58–67	0.47	0.49	0.47	0.54	0.54	0.61	0.47	0.59
4	>67	0.49	0.54	0.49	0.58	0.57	0.64	0.49	0.55
Total		0.46	0.47	0.46	0.53	0.52	0.61	0.46	0.47

Note. Fractions (0.0–1.0) of men and persons with an education other than University, College, A level, AS level in the Age by Hearing groups. For the fraction of men there are valid observations for the same 138,098 persons as in our standard sub-sample, but for education the total number is 116,947.

EFFECTS OF HEARING LOSS AND AGE ON PERFORMANCE IN THE TWO MEMORY TESTS

Figure 1 presents the mean error scores ($\ln(1+x)$) plotted as a function of age and hearing according to Dawes et al. (2014), called Hear, with categories in SRT dB: Normal = <−5.5 Insuff = −5.5 to −3.5, Poor > −3.5). The left panel presents the data for VSWM and the right panel gives the data for PLTM. The ANOVAs were computed separately for VSWM and PLTM with Hear and Age as independent between-person factors. As can be seen from **Figure 1** and as confirmed by the ANOVAs (see **Table 3**) there are significant effects of both Hear and Age. The Age effect is about equal in terms of F -values for the two memory tests, but the effect of Hear for PLTM appears to be stronger than it is for VSWM. Also, there is a significant interaction Hear \times Age for VSWM, but not for PLTM.

Thus, PLTM seems to be more sensitive to functional hearing status and judging from **Figure 1**, the dominating difference is between the poor and the insufficient hearers. To statistically corroborate this difference we conducted follow-up ANOVAs on the 12,157 insufficient hearers and compared them with the 1340 poor hearers. For PLTM, there was a marked difference between the poor and insufficient hearers, $F_{(1,13489)} = 68.9$, $p < 0.000$, between the normal and insufficient hearers, $F_{(1,136750)} = 256.6$, $p < 0.000$, and a significant effect of Age, $F_{(3,13489)} = 12.18$, $p < 0.000$, but no significant effect of their interaction ($F < 1$). For VSWM, there was no significant difference between the poor and insufficient hearers, ($F < 1$), a main effect of Age, $F_{(3,13489)} = 12.81$, $p < 0.000$, and no significant interaction ($F < 1$).

Thus, the ANOVAs and the pattern of simple main effects results strongly support the conclusion that there is a crucial difference in the pattern of age-related performance between PLTM and VSWM, especially when comparing the poor and insufficient hearers. Poor compared to insufficient hearing is markedly more deleterious to PLTM than it is to VSWM.

POWER AND EFFECT SIZE

In **Table 3** it can also be noted that the observed power is very high because of the large samples. Effect sizes (Cohen's d) were calculated for pairwise comparisons between levels of Age and Hear for VSWM and PLTM, respectively, and are shown in **Table 4**.

As shown in **Table 4**, the effect sizes are mostly small (<0.20), but the effect of Hear is systematically greater and in the medium range for PLTM than VSWM. Particularly, the effect size of the comparison between normal and poor hearers for PLTM exceeds medium (>0.50), which is quite impressive with such a large sample. However, the effect sizes for the comparisons normal vs. insufficient hearers and insufficient and poor hearers were 0.25 and 0.32, respectively, which is closer to the small effect size.

Therefore, effects sizes are quite in line with the results from the separate ANOVAs, which showed large effects of both Hear and Age, the Age effect being about equal for the VSWM and PLTM, but also that the Hear effects were larger for PLTM than for VSWM.

ADDITIONAL ANALYSES

To assess whether using a hearing aid modulated memory decline, we computed separate ANOVAs on the following sub-sample: for

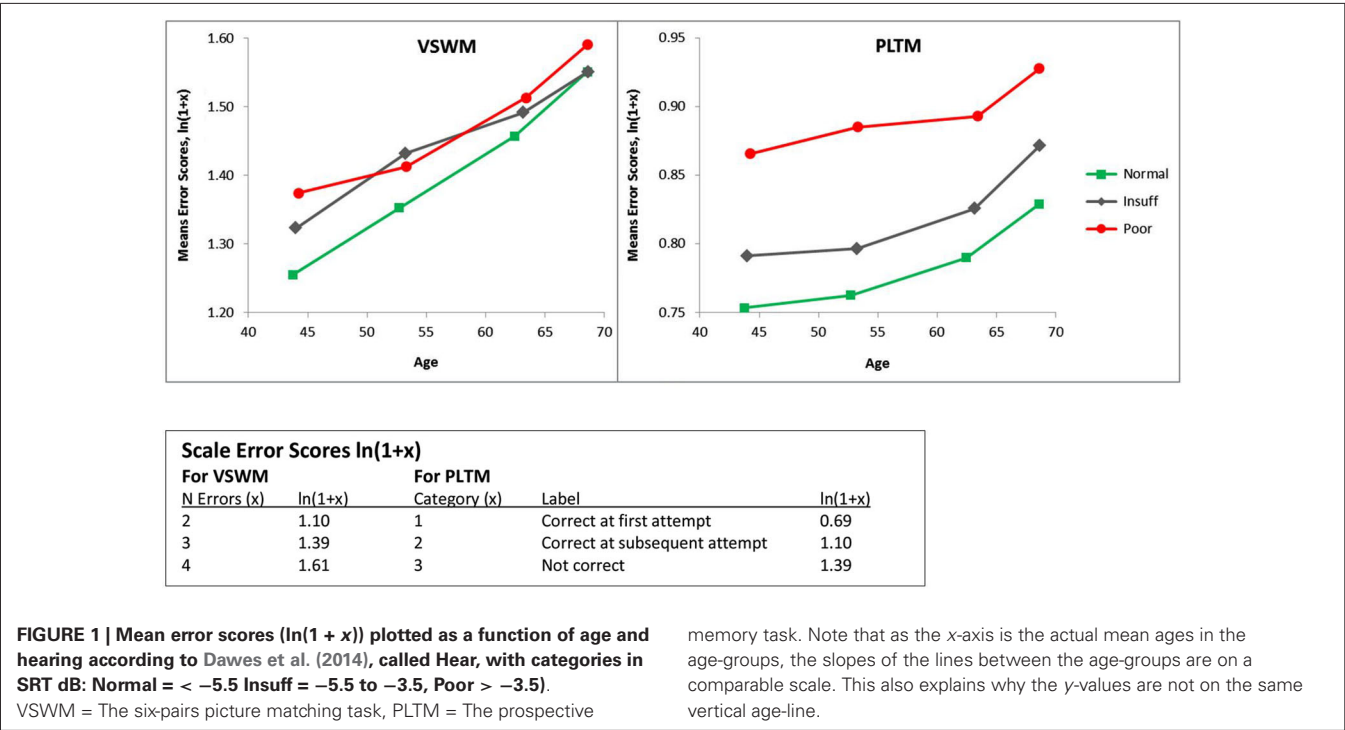


Table 3 | F-tables for VSWM (upper panel) and PLTM (lower panel) by Hear and Age for the values given in Figure 1.

	Sum of squares	df	Mean square	F	Sign. p =	Observed power
VSWM source						
Hear	15.822	2	7.911	21.085	0.000	1.000
Age	41.860	3	13.953	37.189	0.000	1.000
Hear*Age	6.659	6	1.110	2.958	0.007	0.906
Error	51810.502	138086	0.375			
PLTM source						
Hear	16.861	2	8.431	243.940	0.000	1.000
Age	4.186	3	1.395	40.376	0.000	1.000
Hear*Age	0.146	6	0.024	0.705	0.645	0.285
Error	4772.280	138086	0.035			

a total of 3751 of HAUse we had data on their Age and Hearing status, as well as on their scores within the 99th percentile on the memory tasks. Of these, 2139 were normal hearers (57%, out of 3751 HAUse), 1080 insufficient hearers (29%), and 532 were poor hearers (14%).

When adding HAUse as a separate third variable to Age and Hear in our separate ANOVAs, we noted a beneficial main effect of HAUse, shown as a reduction in the number of errors for VSWM for HAUse compared to non-users ($F_{(1,141825)} = 4.86, p < 0.05$). For VSWM there was also a significant interaction Hear \times HAUse, $F_{(2,141825)} = 4.20, p < 0.05$, see **Figure 2**. A test of the simple main effects of HAUse indicated at significant difference between HA-users and No HA-user with poor hearing, $F_{(1,141825)} = 7.10, p < 0.01$ (with a Cohen's d effect size of = 0.185) but not at the other two levels of hearing ($F < 1$). Thus, for VSWM the results indicated that for the normal hearers there was not much of a

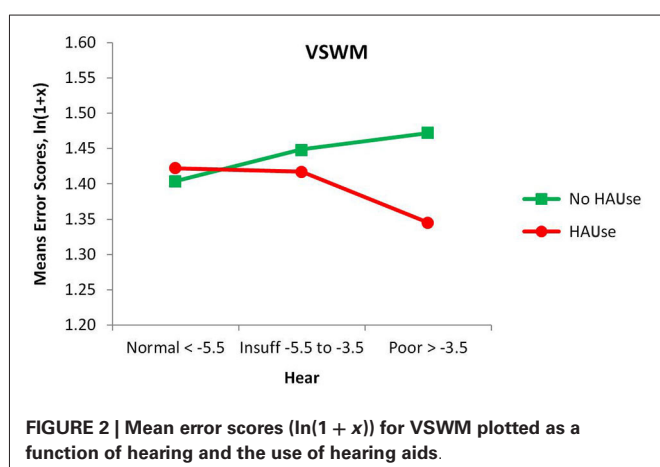
difference between those with and without hearing aids, but with increasing hearing loss the degree of “protection” against memory errors afforded by wearing hearing aids increased (see **Figure 2**). However, the effect size is relatively low, but inspecting the 95% confidence intervals for the means of the three levels of Hear in **Figure 2** for the HA-users indicated that the mean for the poor hearers was outside the lower bounds of the means for the normal and insufficient hearers.

For PLTM, there was no main effect of HAUse, ($F < 1$), and no significant interaction Hear \times HAUse ($F < 1$), but there was a significant interaction Age \times HAUse, $F_{(3,141825)} = 6.05, p < 0.000$, which was specified by the interaction Hear \times Age \times HAUse, $F_{(6,141825)} = 3.06, p < 0.01$, (not given in any figure) showing that the poor hearers with hearing aids in the youngest group have markedly higher error scores than was the case for the other HAUse (their value of 0.967 is far above the upper 95% confidence

Table 4 | Effect sizes (Cohen's *d*) for VSWM and PLTM between adjacent levels and the highest vs. lowest levels of Age and Hear, for the same analyses shown in Table 3 and Figure 1.

Age, years	Cohen's <i>d</i>	
	VSWM	PLTM
<48 vs. 48–57	0.162	0.059
48–57 vs. 58–67	0.171	0.162
58–67 vs. >67	0.146	0.214
<48 vs. >67	0.478	0.461
Hear		
Normal vs. Insufficient	0.134	0.250
Insufficient vs. Poor	0.041	0.324
Normal vs. Poor	0.175	0.646

Note. The values in the Table can be compared to Cohen's (1988) proposed rules of thumb for interpreting effect sizes: a "small" effect size is 0.20, a "medium" effect size is 0.50, and a "large" effect size is 0.80.



limits for all of the other 11 Age \times Hear groups with hearing aids. However, a warning is in place for this group, as it has the lowest *N* in that analysis, only 22). Thus, generally speaking, PLTM was not positively affected by the use of hearing aids, but for VSWM we could observe some more "protection" against making errors, as is suggested from the HAUse \times Hear interaction in Figure 2. However, two points should be noted about this interaction: one is that we had so called normal hearers who used hearing aids. The fact that they seek treatment with presumably very mild or non-existent functional hearing loss is usually because of some other kind of communication difficulties. If the cochlear function does not contribute to these problems, we suggest that there are some underlying central processing or cognitive defects that contribute to the person's experiences of having difficulties with communication. Second, we cannot be sure about causality (see more under Section General Discussion).

To eliminate Gender and Education (dichotomized as in Table 2) as confounders (cf. Table 2), we added these two independent variables to Age and Hear in a MANOVA, ending up with $N = 116,947$, as in Table 2. For VSWM there were no significant main effects or interactions involving Gender and/or Education. For PLTM there was a main effect of Education, $F_{(1,116899)} = 85.19$,

$p < 0.001$, and an interaction Hear \times Education, $F_{(2,116899)} = 7.73$, $p < 0.001$. These effects indicated that the persons with a lower education made more errors, and that this disadvantage was more marked for those with poor hearing. The 95% confidence interval for the poor group included the insufficient group for those with a higher education, but for those with a lower education, the insufficient group was lower by far in errors and outside the 95% confidence interval for the poor group. However, we cannot be conclusive about education causing better episodic long-term memory, but there are studies that suggest that schooling affects brain function and cognition many decades after schooling has terminated (Glymour et al., 2008; Nyberg et al., 2012).

For PLTM, there was also an interaction Age \times Gender \times Education, $F_{(3,116899)} = 2.74$, $p < 0.05$, meaning that males with lower education and in the age range 48–57 years, made more errors than women in the same group. However, caution should be observed when interpreting these results as the number of persons in 4 of the 48 ($= 4 \times 3 \times 2 \times 2$) cells come as low as $n < 30$, particularly for the youngest and oldest poor hearers with high education.

Furthermore, replacing Hearing status in the original ANOVAs with the binary scored subjective reports of hearing difficulty and hearing difficulty in noise, did not yield any significant main effects or interaction (all F s < 1.97).

We also tested whether using eyeglasses or having reported eye problems had any association with the memory data but found no such relationships. Thus, it is mainly the objectively measured functional hearing loss (the SRTn for the DTT) that accounts for the observed memory declines.

PROBING THE CATEGORIZATION OF HEARING STATUS

To safeguard against missing some more delicate and detailed effects when a rather crude hearing criterion like the three-step Hear-distinction was employed, an analysis with a four-step hearing criterion (Hear4) was also performed. In this four-step criterion the extreme groups were the same as in the original Hear4 criterion, but the former middle-group (Insuff) was split into two groups, Insuff1 (SRT -5.5 to -5.0) and Insuff2 (SRT > -5.0 to 3.5). The number of persons are shown in Table 5.

The results of the Hear4 grouping is depicted in Figure 3, which has the same y-axis as Figure 1, to make a visual inspection easy. However, the Hear4 grouping did not change the pattern of significant effects in the overall ANOVA already reported above in Table 3.

As can be seen when comparing Figure 1 and Figure 3, the splitting of the Hear insufficient group into two groups, did not reveal that Insuff2 approached the group with the poorest hearers. Insuff2 remained close to Insuff1 in performance on the two memory measures. This indicates that the pronounced problems with memory are mainly restricted to the 1% of the sample that has the worst hearing.

In a similar vein, we also probed what would happen to the scores for VSWM and PLTM when the group with poor hearers ($N = 1\,340$) was divided into three poor hearing groups (Bad, Worse, Worst, see Figure 4 for hearing criteria, N s = 369, 549, 422 respectively). The results are shown in Figure 4, and the corresponding ANOVAs indicated that the only significant

Table 5 | Number of persons in Age-groups and the four-step Hearing status groups.

	Age	Hear4				Total
		Normal < -5.5	Insuff1 -5.5--3.5	Insuff2 -5.5--3.5	Poor > -3.5	
1	< 48	23147	447	434	90	24118
2	48–57	38724	1119	1250	197	41290
3	58–67	55617	3175	4165	835	63792
4	> 67	7113	574	993	218	8898
Total		124601	5315	6842	1340	138098

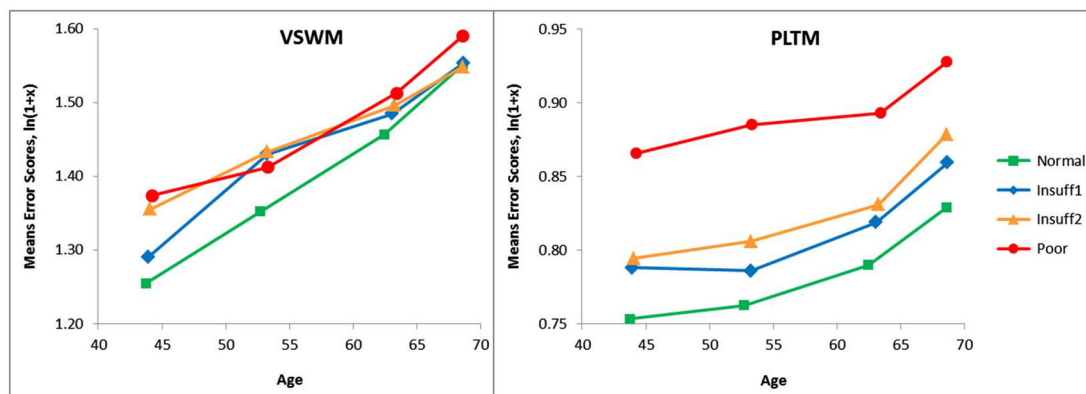


FIGURE 3 | Same as Figure 1 except that a four-step (Hear4) hearing grouping was employed. Categories in SRT dB: Normal = < -5.5 Insuff1 = -5.5 to -5.0, Insuff2 = -5.0 to -3.5

Poor > -3.5. The Poor and Normal hearing group are the same as in Figure 1, and their lines have the same legends.

effect for VSWM was as a main effect of Age, $F_{(2,1328)} = 3.25$, $p < 0.05$. For PLTM there was no significant effect of Age ($p > 0.10$), but as indicated in Figure 4, the average errors in the worst sub-group of the poor hearers were higher than in the bad group. This difference was significant in a one-tailed t -test, $t_{(789)} = 1.78$, $p < 0.05$, but Cohen's d was low (0.127).

Thus, a more fine-grained sub-grouping of the poor hearers pinpoints the extremely poor hearers, the worst group, as responsible for a significant share of the increase in error scores for PLTM, but not to the same extent for VSWM.

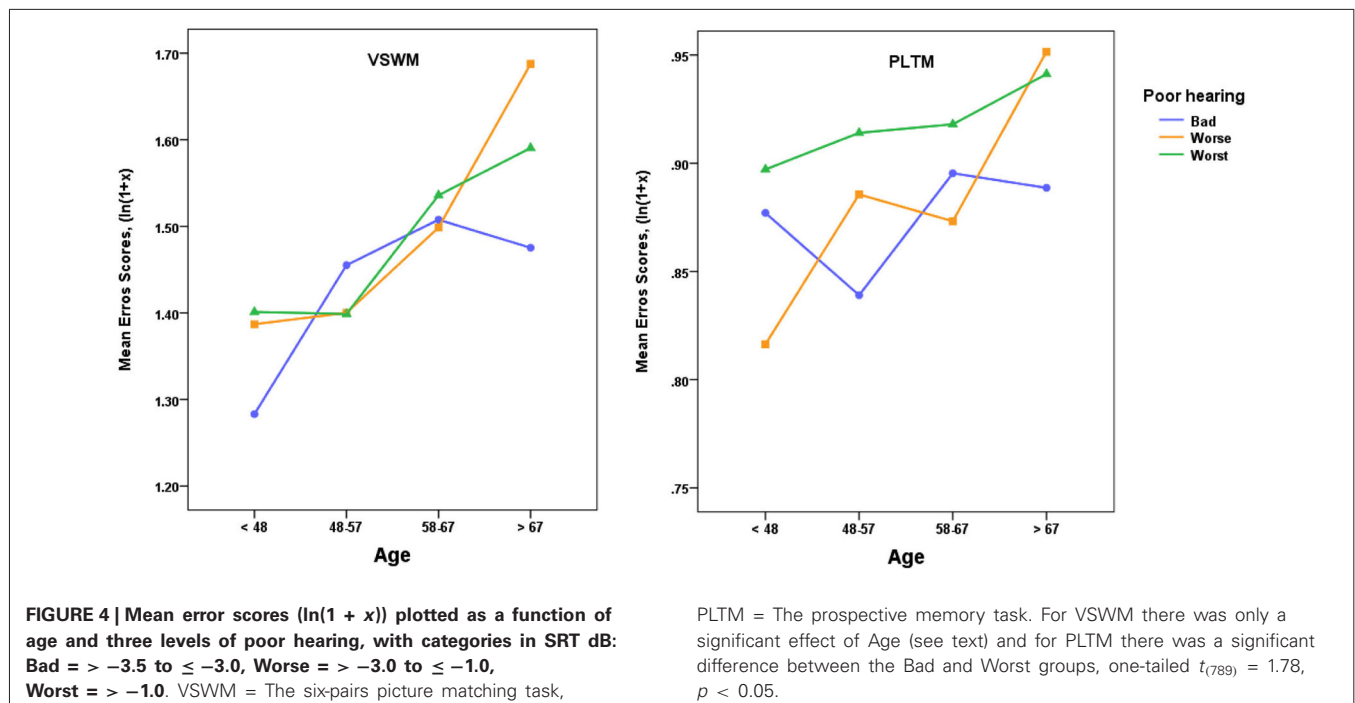
The general result of these analyses is that the effects of functional hearing loss are robust and prominent for episodic long-term memory, and driven by extremely poor hearers. Being a hearing aid user had no effect on the association between hearing and episodic long-term memory, but did influence the association between hearing and working memory; hearing aid users among poor hearers performed better than non-users. Education and gender modulated the episodic long-term memory decline but not working memory. Age affected both memory systems negatively, but interacted with gender and education only for episodic long-term memory.

PARAMETRIC AND NON-PARAMETRIC TESTING OF VSWM AND PLTM

There may be some doubt as to whether the scale properties of our measures of VSWM and PLTM meet the assumptions for a parametric ANOVA-test.

However, ANOVAs are known to be robust against violations of the underlying assumptions (discussed in several elementary text books in statistics, e.g., Howell, 2007). A normal distribution is not necessary, and testing skewed distribution against each other may be acceptable if the distributions have the same kind of skewness. Histograms of our VSWM scores showed a unimodal symmetric distribution. The PLTM measure showed a skewed distribution with more observation at the lower end of the scale. The VSWM measure showed a unimodal symmetric distribution, if the interval band width was set to 0.5.

We also made analyses of VSWM and PLTM with the SPSS Generalized Linear Model, which do not make any assumptions about the distributions of the scores. Analyses with VSWM and PLTM as ordinal scale dependent measures, and with Age and Hear as independent variables, in the same way as for the data in Figure 1, Table 3, showed exactly the same pattern of significant effects as the ANOVA analyses. For VSWM the effects of Age and Hear were significant with $ps < 0.000$ and the p -value of their interaction was 0.025. For PLTM the effects of Age and Hear were also significant with $ps < 0.000$, but the p -value of their interaction was > 0.10 . It was also the case in this SPSS Generalized Linear Model that the effect of Age was about equal for VSWM and PLTM. However, for VSWM the effect of Hear was much weaker than that of Age, while for PLTM the effect of Hear was more substantial than for Age. Thus, in the non-parametric tests we show the same relative effects as those reported from the separate parametric ANOVA analyses as well as from the effect



sizes reported. Finally, there is a notable difference in the basic original scales for PLTM and VSWM.

Prospective long-term memory is based on a trichotomization (correct on first attempt, correct at a subsequent attempt, not correct at first or following attempts), while the scale for VSWM was number of errors on and interval scale from 0 to 15. Thus, there was a substantial underestimation of the actual number of errors made in the PLTM task. In spite of this underestimation, poor functional hearing turned out to be significantly related to PLTM, which makes the result even more striking in light of the main hypothesis of the present paper.

GENERAL DISCUSSION

The focal finding of this study is that functional hearing loss is clearly related to visuospatial episodic long-term memory (PLTM). This result is important for several reasons.

First, it shows that the negative effect of functional hearing loss is not restricted to mechanisms coupled to auditory perceptual degradation (Schneider et al., 2002, 2010) or to consumption of attention resources due to a compromised auditory signal (Tun et al., 2009; Verhaegen et al., 2014). Although the results in the Rönnerberg et al. (2011) study already generalized to verbal tasks with alternative kinds of encoding than the purely auditory or audiovisual (i.e., using motor encoding, Nyberg et al., 1992), the present study has taken a further significant step: here, we demonstrate a robust effect of hearing loss that generalizes to visuo-spatial encoding and subsequent memory retrieval of these kinds of stimuli. Therefore, the negative effects are more pervasive in terms of encoding modality than previously imagined or documented (cf. Rönnerberg et al., 2011).

Second, the results replicate the Rönnerberg et al. (2011) result showing a stronger impact of hearing loss on episodic long-term

memory function rather than on short-term/working memory. The effect size for the poor hearers compared to the normal hearers is substantial (in between medium and large) for PLTM but not for VSWM. Subsequent analyses of subgroups of the poor hearers also showed that the worst subgroup differed from the bad subgroup, but at this level of detail the effect size is relatively low.

Third, the analysis of VSWM revealed a negative effect of functional hearing status, but the relative effects are small and much smaller than for the PLTM task. This finding fits with the overall picture of results from Verhaegen et al. (2014), who also found (significant) negative effects of mild hearing loss on certain short-term memory tasks. Nevertheless, this is also in line with the claim (Rönnerberg et al., 2011) that there should be a relatively stronger effect of hearing loss on episodic long-term memory compared to short-term or working memory, mainly because mismatches would reduce the number of times the episodic long-term memory system would be used for encoding, storage and retrieval (Rönnerberg et al., 2013).

Fourth, as the effect of using a hearing aid had a relatively positive (error-reducing) effect on the visuospatial working memory task but not on the episodic long-term memory task, the results mimic the Rönnerberg et al. (2011) data in that all participants wore hearing aids in that sample—and the negative effect of hearing loss only persisted for semantic and episodic long-term memory. Thus, one more general interpretation of the two sets of results is that there is an effect of hearing loss on short-term memory and long-term memory, the effect is smaller for short-term memory or working memory, and can be at least potentially be compensated for by the use of hearing aids for the poor hearers. This pattern of results agrees with the recent data by Verhaegen et al. (2014) where negative effects of hearing loss

were found even in short-term memory tasks, but note that hearing aids were not used by the participants in that study sample.

A counterargument against the positive effect being due to the use of hearing aids as such would be to reverse causality as follows: if good memory were causing people to get and use hearing aids, the group with normal functional hearing who used hearing aids would have better memory. However, since this was not the case (cf. **Figure 2**) and the poor hearers with hearing aids do have better working memory, then it is likely that the hearing aid is reducing the effect of hearing loss on working memory, and possibly also compensating for the loss as shown by the relative improvement seen for the poor hearers compared to normal hearers.

Understanding the benefit provided by hearing aids (although constrained and small) rests on the fact that functional hearing loss affects PLTM and hearing aid benefit VSWM, i.e., both variables affect the two memory systems selectively. In this study, it happened with a visuospatial VSWM task, but similar results could have been found with an auditory WM task, that is, the general picture that is emerging is that of multimodal processing. The important aspect is the difference in basic cognitive mechanisms underpinning the two tasks, and how other variables latch on to the different properties of those two memory systems.

However, it is also important to note that there could be some initial selection bias relating to individual stages of acceptance of the hearing loss, with the motivation to change and to actively seek help (Manchaiah et al., 2014). Furthermore, yet another interpretation is that the persons who were poor hearers had worn their hearing aids for longer periods of time than the other groups (as hearing loss is usually progressive), and therefore they had developed compensatory skills. However, since the use of a hearing aid did not improve episodic long-term memory, the potential benefit from wearing a hearing aid is relatively *restricted* to VSWM and the effect size was also low. This is also in general agreement with Rönnberg et al. (2011), where we also observed negative effects of hearing loss on episodic long-term memory despite the fact that all participants wore hearing aids. Finally, it is also possible that some hidden cognitive capacity that is not tested in the UK Biobank data set is responsible for the observed interaction. Future research may be more hypothesis-driven in this respect.

Fifth, background variables such as education and gender interact with age for the PLTM task suggesting that the long-term component demonstrates qualitatively different properties compared to working memory. This generally shows that it is important to consider the type of memory system when we are evaluating background variables. It is suggested here that episodic long-term memory is more dependent on crystallized knowledge such as linguistic competence, which is mediated by education (Nyberg et al., 2012) and gender expectations (Lundervold et al., 2014). That kind of competence can also help in decoding the visuospatially presented objects.

Sixth, the negative effect of aging is pervasive across memory systems in the current study, i.e., for both VSWM and PLTM. What we found in Rönnberg et al. (2011) was that hearing loss displayed a negative effect on episodic long-term memory, even

when age was statistically controlled for. This is also what we find here: poor hearers are especially prone to error in the PLTM task.

Seventh, the details of the results also show that the relative weighting of the impact of age and hearing loss plays out differently for the two memory tasks. Age is relatively more important for VSWM than for PLTM while hearing loss has a relatively more adverse effect on PLTM than on VSWM. Thus, age and poor hearing play at least partially different roles and may also rely on different mechanisms (Rönnberg et al., 2011).

Eighth, Pelle et al. (2011) have shown that individual differences in hearing acuity (pure tone thresholds) predict activation of bilateral superior temporal regions during auditory sentence comprehension, and that the loss of gray matter is proportional to the degree of audiometric hearing loss, especially in the right auditory cortex. A recent study by Lin et al. (2014) shows that declines in regional brain volumes over 6.4 years are associated with hearing loss, especially in the *right* temporal lobe (superior temporal gyrus, middle temporal gyrus and inferior temporal gyrus), and that this decline is comparable to loss of brain volume in participants with diagnosed mild cognitive impairment (Driscoll et al., 2009). This result is also in line with the previous study by Lin et al. (2011), using a follow-up period that was twice as long, and showing that the risk of developing Alzheimer's disease is related to hearing loss. However, with our current state of knowledge, it may be too speculative to assume that atrophy in the temporal lobe *also directly* affects visuospatial processing, especially for the PLTM task. Thus, the challenge for future research is to address the many kinds of functional and multimodal brain compensations that may occur due to temporal lobe atrophy, and which also lead to selectivity at the memory systems level.

Ninth, the important aspect here is that we replicate the selectivity *predicted* by the Ease of Language Understanding (ELU) model in the relationship between hearing loss and working memory on the one hand, and episodic long-term memory on the other for different types of tasks (cf. Rönnberg et al., 2011). Again, this effect occurs despite the fact that the underlying scale for PLTM is more conservative (but see more under Section Methodological Issues). This kind of selectivity is not predicted by a common cause account. Also, the association between hearing loss and memory system must be considered to be more central, as our peripheral measures of visual acuity (i.e., wearing eye glasses) did not show any distinctive contribution to memory performance, which is perhaps less surprising than the fact that reported eye problems (which may include more central deficits such as amblyopia) did not show any relationship either. If this line of reasoning is correct, then we may argue for a hearing loss-related central and multimodal mechanism that explains the PLTM decline (Rönnberg et al., 2013) rather than a hypothesis claiming that neural degeneration in general affects both vision and audition in tandem with a general cognitive decline (i.e., the common cause hypothesis, see e.g., Lindenberger and Ghisletta, 2009). However, our claim of a central mechanism should be considered with due caution. One point is that there was no fine-grained or advanced measure of visual acuity/spatial resolution in the UK Biobank database, hence potential associations with visual processing may

be underestimated (cf. Humes et al., 2013). Another related point is about causality: even if our hypothesis is about hearing as the independent variable, it is in principle possible that a degradation of visuospatial functions (affecting visuo-spatial memory) may have caused a functional hearing loss. However, the literature on brain tissue degeneration (e.g., Peelle et al., 2011; Lin et al., 2014) suggests that there are right-hemisphere effects that are caused by hearing loss and related to its severity, and again, at least in this study, we do not see any signs of a reversed causality.

Tenth, summarizing across the findings of the current and the Rönnerberg et al. (2011) study, functional hearing loss seems to affect episodic long-term memory in general, irrespective of encoding modality, which is why we see effects in visuospatial tasks in the present study, and in Rönnerberg et al. (2011) for motor, visual and auditory encoding. The causal nature of the effects needs, however, to be verified in longitudinal studies.

Overall, the large sample in the current study has been helpful in detecting substantial effect sizes related to functional hearing losses. Importantly, it should be noted that these effects apply to non hearing-aid users in the main analyses, suggesting that even relatively mild functional hearing losses do indeed suggest early deterioration of episodic long-term memory function in particular. Altogether, considering the current state of knowledge, including our previous finding that hearing aid wearers show episodic long-term memory deficits related to degree of hearing loss (e.g., Rönnerberg et al., 2011), as well as the fact that decline in memory functions represents an important and integral part of dementia and that hearing impairment is related to a substantially increased risk of dementia of Alzheimer type (e.g., Lin et al., 2011), we suggest that the current result is very important from a public health perspective.

METHODOLOGICAL ISSUES

It could be argued that the DTT is confounded by a *short-term memory* component (as perception and recall of digit triplets are required). If the short-term or working memory component was crucial, one would then predict that DTT performance should co-vary with VSWM and not with PLTM. Digit triplet test performance did not co-vary with VSWM. The reason for the lack of an association with VSWM could be that a “load” of a digit triplet is clearly below what is typically given as the normal digit span size (i.e., 7 ± 2). Instead, the DTT variable predicted a decline in PLTM. This kind of double dissociation represents evidence in favor of an interpretation of the present results in terms of a negative effect of functional hearing loss on episodic long-term memory, as outlined by the ELU model (Rönnerberg et al., 2011).

It is also clear that there is little reason to believe that the DTT is confounded specifically by semantic *long-term memory* processes (Moore et al., 2014). The DTT has been found to be correlated highly with both an adaptive speech-in-noise test and audiometric testing: the primary interpretation is that it is an auditory speech component that is shared, not a cognitive or linguistic component (cf. Smits et al., 2004). Second, the DTT calls on stored knowledge of a small set of overlearned

phonologically dissimilar items with *limited* semantic content whose representation is unlikely to change as a function of either hearing loss or age-related cognitive change. Third, the response format (a touch pad on the screen with the digits laid out) acts as a reminder of the set of available items. Fourth, it is currently unknown how central and peripheral auditory factors play out in the DTT. Further research is needed (cf. Moore et al., 2014), and it would be of interest in the future to investigate the association between hearing and memory using both threshold and functional hearing data.

Another concern that may be raised against the selectivity in the effect of hearing loss on memory systems is the possibility that the results may be confounded by *task difficulty*. However, the PLTM-task was *less* difficult than the VSWM-task in terms of how many percent of the participants produced a correct response on the first trial (80.6 for the PLTM and 7.1% for the VSWM-task). Also, the range of the raw values of number of errors were three for PLTM (0, 1 and 2, or more) and 16 for VSWM (0–15). The logarithmic ranges and means were: VSWM range 0.00–2.77, mean 1.40–0 errors = 7.1%; PLTM range 0.69–1.39, mean –0.78–0 errors = 80.6%. Again, the PLTM task was less difficult than the VSWM task, had fewer steps, and was less sensitive, but still produced significant differences with substantial effect sizes due to functional hearing loss. Reliability estimates are not available from the UK Biobank resource. If we had observed the opposite pattern, viz. that functional hearing loss was associated with larger effects for the VSWM task, then it could have been argued that the effect (at least partially) was due to greater task difficulty that provoked the negative memory effect. In all, it seems unlikely that aspects related to task difficulty could explain the results obtained in the current study.

Finally, visuospatial memory function was *not* related to subjective ratings of hearing disability collected in the UK Biobank database, which suggests that the obtained effects may be based on the loss and objectively determined by an audiogram or by an objective test such as the DTT (Rönnerberg et al., 2011; Dawes et al., 2014). Likewise, recent data show that perceived effort in quiet and noise in work-related tasks is hardly ever related to a whole range of cognitive capacities relevant for speech understanding in noise (Hua et al., 2014). This may point to a more general issue regarding ratings of hearing problems and/or effort ratings as predictors of memory or perceptual functions. Several factors may play a role here: it may be the case that the ratings must involve an explicit component of the function under scrutiny and that the function *per se* is explicit (see Rudner et al., 2012; Ng et al., 2013). In the current case, the rating of hearing disability may be too coarse (binary) to measure the explicit functions tapped by VSWM and PLTM. It may also be the case that these types of tasks are less representative of everyday memory problems involved in subjective experiences of hearing problems.

CONCLUSION

In all, connecting the memory systems hypothesis with the demands of the visuospatial processing in the memory tasks, the putative negative long-term effect of functional hearing loss is more pronounced for episodic long-term memory (i.e., for

PLTM) than for working memory or short-term memory (i.e., for VSWM). This is in line with the ELU prediction about mismatch and relative use/disuse of memory systems (Rönnerberg et al., 2011). There may also be a biological basis for a transfer effect from functional hearing loss to episodic long-term memory, including visuospatial and other kinds of multimodal memory encoding formats. It remains for future research to show how e.g., hearing loss-related brain atrophy in the right temporal lobe is associated with general episodic memory deficits.

ACKNOWLEDGMENTS

This research has been conducted using the UK Biobank Resource. It was partly supported by the grant to the Linnaeus Centre HEAD, Linköping university, Sweden, from the Swedish Research Council (349- 2007-8654), and was partly supported by the Department of Health and Aging in Australia.

REFERENCES

- Arlinger, S. (2003). Negative consequences of uncorrected hearing loss—a review. *Int. J. Audiol.* 42, S17–S20. doi: 10.3109/14992020309074639
- Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12
- Burgess, P. W., and Shallice, T. (1997). “The relationship between prospective and retrospective memory: neuropsychological evidence,” in *Cognitive Models of Memory*, ed M. A. Conway (London: Psychology Press), 247–272.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edn. Hillsdale, NJ: Erlbaum.
- Dawes, P., Fortnum, H., Moore, D. R., Emsley, R., Norman, P., Cruickshanks, K. J., et al. (2014). Hearing in middle age: a population snapshot of 40–69 year olds in the UK. *Ear. Hear.* 35, e44–e51. doi: 10.1097/aud.0000000000000010
- Dobie, R. A. (1996). Compensation for hearing loss. *Audiology* 35, 1–7. doi: 10.3109/00206099609071925
- Driscoll, I., Davatzikos, C., An, Y., Wu, X., Shen, D., Kraut, M., et al. (2009). Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. *Neurology* 72, 1906–1913. doi: 10.1212/WNL.0b013e3181a82634
- Glymour, M. M., Kawachi, I., Jencks, C. S., and Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *J. Epidemiol. Community Health* 62, 532–537. doi: 10.1136/jech.2006.059469
- Heinrich, A., and Schneider, B. A. (2011). Elucidating the effects of ageing on remembering perceptually distorted word pairs. *Q. J. Exp. Psychol. (Hove)* 64, 186–205. doi: 10.1080/17470218.2010.492621
- Howell, D. C. (2007). *Statistical Methods for Psychology*. 6th Edn. Belmont, CA: Thomson Wadsworth.
- Hua, H., Emilsson, M., Ellis, R., Widén, S., Möller, C., and Lyxell, B. (2014). Cognitive skills and the effect of noise on perceived effort in employees with aided hearing impairment and normal hearing. *Noise Health* 16, 79–88. doi: 10.4103/1463-1741.132085
- Humes, L. E., Busey, T. A., Craig, J., and Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten. Percept. Psychophys.* 75, 508–524. doi: 10.3758/s13414-012-0406-9
- Janse, E., and Newman, R. S. (2013). Identifying nonwords: effects of lexical neighborhoods, phonotactic probability and listener characteristics. *Lang. Speech* 56, 421–441. doi: 10.1177/0023830912447914
- Lin, F. R., Ferrucci, L., An, Y., Goh, J. O., Doshi, J., Metter, E. J., et al. (2014). Association of hearing impairment with brain volume changes in older adults. *Neuroimage* 90, 84–92. doi: 10.1016/j.neuroimage.2013.12.059
- Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B., and Ferrucci, L. (2011). Hearing loss and incident dementia. *Arch. Neurol.* 68, 214–220. doi: 10.1001/archneurol.2010.362
- Lindenberger, U., and Baltes, P. B. (1994). Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339–355. doi: 10.1037/0882-7974.9.3.339
- Lindenberger, U., and Ghisletta, P. (2009). Cognitive and sensory declines in old age: Gauging the evidence for a common cause. *Psychol. Aging* 24, 1–16. doi: 10.1037/a0014986
- Lundervold, A. J., Wollschläger, D., and Wehling, E. (2014). Age and sex related changes in episodic memory function in middle aged and older adults. *Scand. J. Psychol.* 55, 225–232. doi: 10.1111/sjop.12114
- Manchaiah, V. K. C., Molander, P., Rönnerberg, J., Andersson, G., and Lunner, T. (2014). The acceptance of hearing disability among adults experiencing hearing difficulties: a cross-sectional study. *BMJ Open* 4:e004066. doi: 10.1136/bmjopen-2013-004066
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A* 58, 22–33. doi: 10.1080/02724980443000151
- Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., et al. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS One* 9:e107720. doi: 10.1371/journal.pone.0107720
- Ng, E. H. N., Rudner, M., Lunner, T., and Rönnerberg, J. (2013). Relationships between self-report and cognitive measures of hearing aid outcome. *Speech Lang. Hear.* 16, 197–207. doi: 10.1179/205057113x13782848890774
- Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., and Bäckman, L. (2012). Memory aging and brain maintenance. *Trends Cogn. Sci.* 16, 292–305. doi: 10.1016/j.tics.2012.04.005
- Nyberg, L., Nilsson, L.-G., and Bäckman, L. (1992). Recall of actions, sentences and nouns—influences of adult age and passage of time. *Acta Psychol.* 79, 245–254. doi: 10.1016/0001-6918(92)90060-q
- Peelle, J. E., Troiani, V., Grossman, M., and Wingfield, A. (2011). Hearing loss in older adults affects neural systems supporting speech comprehension. *J. Neurosci.* 31, 12638–12643. doi: 10.1523/JNEUROSCI.2559-11.2011
- Pichora-Fuller, M. K. (2003). Processing speed: psychoacoustics, speech perception and comprehension. *Int. J. Audiol.* 42, S59–S67. doi: 10.3109/14992020309074625
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. K. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Piquado, T., Cousins, K. A. Q., Wingfield, A., and Miller, P. (2010). Effects of degraded sensory input on memory for speech: behavioral data and a test of biologically constrained computational models. *Brain Res.* 1365, 48–65. doi: 10.1016/j.brainres.2010.09.070
- Plomp, R., and Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *J. Acoust. Soc. Am.* 66, 1333–1342. doi: 10.1121/1.383554
- Rogers, C. S., Jacoby, L. L., and Sommers, M. S. (2012). Frequent false hearing by older adults: the role of age differences in metacognition. *Psychol. Aging* 27, 33–45. doi: 10.1037/a0026231
- Rönnerberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternäng, O., Wahlin, Å., et al. (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *J. Speech Lang. Hear. Res.* 54, 705–726. doi: 10.1044/1092-4388(2010/09-0088)
- Rönnerberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Rönnerberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167
- Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., and Rönnerberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *J. Am. Acad. Audiol.* 23, 577–589. doi: 10.3766/jaaa.23.7.7
- Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *J. Speech Lang. Hear. Res.* 52, 1230–1240. doi: 10.1044/1092-4388(2009/08-0111)
- Schneider, B. A., Daneman, M., and Pichora-Fuller, M. K. (2002). Listening in aging adults: from discourse comprehension to psychoacoustics. *Can. J. Exp. Psychol.* 56, 139–152. doi: 10.1037/h0087392

- Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (2010). "The effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis. Handbook of Auditory Research*, eds R. D. Gordon-Salant, A. Frisina, A. Popper and D. Fay (Berlin: Springer), 167–210.
- Smits, C., Kapteyn, T. S., and Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int. J. Audiol.* 43, 15–28. doi: 10.1080/14992020400050004
- Tun, P. A., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity and the attentional costs of effortful listening. *Psychol. Aging* 24, 761–766. doi: 10.1037/a0014802
- Valentijn, S. A. M., van Boxtel, M. P. J., van Hooren, S. A. H., Bosma, H., Beckers, H. J. M., Ponds, R. W. H. M., et al. (2005). Change in sensory functioning predicts change in cognitive functioning: results from a 6-year follow-up in the Maastricht aging study. *J. Am. Geriatr. Soc.* 53, 374–380. doi: 10.1111/j.1532-5415.2005.53152.x
- van Boxtel, M. P. J., van Beijsterveldt, C. E. M., Houx, P. J., Anteunis, L. J. C., Metsemakers, J. F. M., and Jolles, J. (2000). Mild hearing impairment can reduce verbal memory performance in a healthy adult population. *J. Clin. Exp. Neuropsychol.* 22, 147–154. doi: 10.1076/1380-3395(200002)22:1;1-8;ft147
- Verhaegen, C., Collette, F., and Majerus, S. (2014). The impact of aging and hearing status on verbal short-term memory. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 21, 464–482. doi: 10.1080/13825585.2013.832725
- Vlaming, M. S. M. G., Kollmeier, B., Dreschler, W. A., Rainer, M., Wouters, J., Gover, B., et al. (2011). HearCom: hearing in the communication society. *Acta Acust. United Acust.* 97, 175–192. doi: 10.3813/aaa.918397
- Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 May 2014; accepted: 07 November 2014; published online: 09 December 2014.

Citation: Rönnberg J, Hygge S, Keidser G and Rudner M (2014) The effect of functional hearing loss and age on long- and short-term visuospatial memory: evidence from the UK biobank resource. *Front. Aging Neurosci.* 6:326. doi: 10.3389/fnagi.2014.00326

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Rönnberg, Hygge, Keidser and Rudner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



What's on TV? Detecting age-related neurodegenerative eye disease using eye movement scanpaths

David P. Crabb*, Nicholas D. Smith and Haogang Zhu

Department of Optometry and Visual Science, School of Health Sciences, City University London, London, UK

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Jonathan Marotta, University of Manitoba, Canada

Michael Wall, University of Iowa, USA

*Correspondence:

David P. Crabb, Department of Optometry and Visual Science, School of Health Sciences, City University London, Northampton Square, London EC1V 0HB, UK
e-mail: d.crabb@city.ac.uk

Purpose: We test the hypothesis that age-related neurodegenerative eye disease can be detected by examining patterns of eye movement recorded whilst a person naturally watches a movie.

Methods: Thirty-two elderly people with healthy vision (median age: 70, interquartile range [IQR] 64–75 years) and 44 patients with a clinical diagnosis of glaucoma (median age: 69, IQR 63–77 years) had standard vision examinations including automated perimetry. Disease severity was measured using a standard clinical measure (visual field mean deviation; MD). All study participants viewed three unmodified TV and film clips on a computer set up incorporating the Eyelink 1000 eyetracker (SR Research, Ontario, Canada). Eye movement scanpaths were plotted using novel methods that first filtered the data and then generated saccade density maps. Maps were then subjected to a feature extraction analysis using kernel principal component analysis (KPCA). Features from the KPCA were then classified using a standard machine based classifier trained and tested by a 10-fold cross validation which was repeated 100 times to estimate the confidence interval (CI) of classification sensitivity and specificity.

Results: Patients had a range of disease severity from early to advanced (median [IQR] right eye and left eye MD was -7 [-13 to -5] dB and -9 [-15 to -4] dB, respectively). Average sensitivity for correctly identifying a glaucoma patient at a fixed specificity of 90% was 79% (95% CI: 58–86%). The area under the Receiver Operating Characteristic curve was 0.84 (95% CI: 0.82–0.87).

Conclusions: Huge data from scanpaths of eye movements recorded whilst people freely watch TV type films can be processed into maps that contain a signature of vision loss. In this *proof of principle* study we have demonstrated that a group of patients with age-related neurodegenerative eye disease can be reasonably well separated from a group of healthy peers by considering these eye movement signatures alone.

Keywords: eye movements, scanpaths, glaucoma, perimetry, eye tracking, KPCA, perception, diagnosis procedures

INTRODUCTION

The ever increasing elderly population will cause an “epidemic” of age-related neurological disease in the 21st century. At present, healthcare detection and monitoring of patients with sensory impairments resulting from chronic age-related neurodegenerative disease is done, mainly inadequately, in a clinic; a system that is likely unsustainable in the future. Instead of relying on infrequent tests in a clinic, focus should shift to capturing health-related data acquired as part of a person's ordinary daily activities.

Eye movements are a continuous and ubiquitous part of sensory perception. Whenever we interact with the visual environment we generate saccadic eye movements. Saccades move the eyes in a ballistic fashion from one point to another, interspersed by fixations where the eye is stable. Scanpaths, revealing the sequence of fixations and saccades, collected non-invasively during a period of time that a person is, for example, simply engaged in watching a TV program could give an “eye movement

signature.” The main idea reported in this paper is to show how these signatures could contain features that can be used to detect if a person has a chronic neurodegenerative condition.

Glaucoma is a generic term for age-related disease of the optic nerve which can lead to irreversible loss of the visual field: the area which can be seen when the eye is directed forward, including both central and peripheral vision. Medical treatment to control the condition is largely successful, but once diagnosed all patients with glaucoma normally need lifelong treatment, and lifelong monitoring within hospitals and clinics, so that any worsening of visual damage can be detected. Therefore, people with glaucoma represent a major workload of eye services, with an estimated one million outpatient appointments per year in the UK for example (National Institute for Health and Clinical Excellence, 2009). The visual field tests (perimetry) are often difficult for patients to do and they are not done with sufficient frequency to adequately monitor most patients (Chauhan et al., 2008; Fung et al., 2013;

Glen et al., 2014). Moreover detection rates for the disease are poor with an estimated 50% of all cases undetected (Rudnicka et al., 2006).

In this work we use glaucomatous optic neuropathy as an example of an age-related neurodegenerative disease (Yücel and Gupta, 2008; Caprioli, 2013). The pathogenesis of glaucoma shares many features with other chronic age-related neurodegenerative disease: there is, for example, ample evidence linking the etiology and disease process in glaucoma to Alzheimer's disease (AD) (Bayer et al., 2002; Sivak, 2013). The epidemiology and impact of glaucoma is well known but the pathogenesis of the disease is multifaceted and not well understood. Optic neuropathy is characterized in the clinic by changes in the optic nerve head (ONH) and thinning of the nerve fiber layer. This is almost certainly a result of a non-specific gradual reactive change of glial cells resulting in chronic retinal ganglion cell death and then loss of visual function (Tezel and Fourth ARVO/Pfizer Ophthalmics Research Institute Conference Working Group, 2009; Sivak, 2013).

There are at least two theories to explain eye gaze. In short, eye movement can be driven by factors that purposely direct fixations toward task-driven locations or in the absence of such task demands eyes are likely directed to salient regions. Investigators have already used scanpaths to gain insights into what an observer is doing or their mental state. For example, a large variety of studies have confirmed that eye movements contain rich signatures about an observer. An excellent up-to-date review of the literature is given elsewhere (Borji and Itti, 2014).

Investigators have revealed data showing that simple viewing patterns in controlled experiments can detect eye-movement abnormalities that can discriminate schizophrenia cases from control subjects with good accuracy (Benson et al., 2012). Other workers, extracting salient features from a series of films, have used eye movements to classify patients with attention deficit hyperactivity disorder and Parkinson's disease (Tseng et al., 2013). The work of this type has attempted to demonstrate the classification of clinical populations from natural viewing. Errors in the ability to make anti-saccades (an eye movement purposely directed in the opposite direction from a target) has been repeatedly implicated in AD (Crawford et al., 2013) and patients with AD have also been shown to display irregular eye movements when reading and in other tasks (Lueck et al., 2000; Mosimann et al., 2004). Our laboratory has published preliminary evidence that patients with eye disease make different types of eye movements when compared to age-related control subjects, when performing different types of task (Crabb et al., 2010; Smith et al., 2012; Glen et al., 2013). Other research regarding the nature and consistency of the types of eye movement patterns shown by groups of individuals as they view scenes have been considered (Castelhano and Henderson, 2008; Cristino and Baddeley, 2009; Dorr et al., 2010). Data analyses in these studies typically, and inadequately, rely on simple counts and averages of, for example, number of fixations, saccade amplitude, and region of interest measures. We propose computational approaches to analyze eye-tracking data not used before, considering sequences of saccades within the scanpaths. We also apply machine classifiers to learn combinations of multidimensional features extracted from the

scanpaths in order to discover patterns that belong to groups of patients.

Therefore, in this study we test the hypothesis that age-related neurodegenerative eye disease can be detected by examining patterns of eye movement recorded whilst a person naturally watches a TV program. We do this with a case-control study with the aim of providing evidence that patients with a clinical diagnosis of glaucoma can be reasonably well separated from age-related healthy people using data from their scanpaths alone.

MATERIALS AND METHODS

PARTICIPANTS

People with glaucoma were recruited from clinics at Moorfields Eye Hospital NHS Foundation Trust, London. All patients had an established clinical diagnosis of chronic open angle glaucoma (COAG) for at least 2 years and were between 50 and 80 years of age. COAG is defined, following clinical guidelines, by the presence of reproducible visual field defects in at least one eye with corresponding damage to the ONH and an open iridocorneal drainage angle on gonioscopy (National Institute for Health and Clinical Excellence, 2009). The diagnosis was made by a glaucoma specialist. A deliberate attempt was made to recruit a sample of patients with a range of disease severity according to visual field loss. Patients were purposely not recruited if they had any ocular disease other than glaucoma (except for an uncomplicated lens replacement cataract surgery). In addition, at the point of recruitment, patients had slit lamp biomicroscopy performed by an ophthalmologist to further exclude any other concomitant macular pathology, ocular surface disease or any significant problems with dry eye.

Healthy people (controls), of a similar age to the patients, were recruited from the City University London Optometry Clinic; this is a primary care center where people routinely receive a full eye examination, which includes measurement of visual acuity, refraction, binocular vision assessment, pupil reactions, slit-lamp assessment of the anterior eye, measurement of intraocular pressure, visual field assessment and indirect ophthalmoscopy of the macula, ONH, and peripheral retina. The City University London Optometry Clinic is located just meters from the hospital meaning all participants were drawn from an identical geodemographic area.

All participants (patients and controls) were required to have a corrected visual acuity of at least 0.18 logMAR (Snellen equivalent 6/9) in each eye at the time of their most recent examination. Astigmatic error was less than ± 2.5 Dioptres in all those recruited. Participants were only recruited if they had no significant health problems (other than their glaucoma for patients), meaning no difficulty with self-care, mobility, pain, anxiety, and depression; this was ascertained at recruitment by self-report to questions based on the EQ-5D instrument (Rabin and de Charro, 2001) added to the participation information sheet. Participants were not enrolled if they were taking any significant medication other than that for their glaucoma. ("Significant medication" included anti-depressants or treatment for diabetes or significant use of β -blocker medication, all of which were deliberately mentioned.) Recruitment of patients and controls was made

simultaneously over a period of about 9 months with an effort to make the two groups age-related.

The study was approved by the Moorfields and Whittington Research Ethics Committee, London and the School of Health Sciences Research and Ethics Committee, City University London. Written informed consent, according to the tenets of the Declaration of Helsinki, was obtained prior to examination from each participant. Data was anonymized and stored in a secure database.

SUPPLEMENTARY VISION TESTING AND COGNITIVE SCREENING

All participants underwent vision testing on the day of the study. Visual fields were measured in both eyes with automated perimetry using the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, CA, USA) employing the central standard 24-2 Swedish Interactive Testing Algorithm. This test is a clinical gold-standard for measuring the severity of functional damage caused by glaucoma. The Glaucoma Hemifield Test (GHT) on the HFA is an established statistical test for early glaucomatous visual field defects (Asman and Heijl, 1992). The GHT was “outside normal limits” for all eyes in all patients and “within normal limits” for all eyes in all controls. If any of the visual fields were flagged by the HFA output as “unreliable,” as assessed by the false positive, false negatives or poor fixation measures, then the test was repeated. The HFA mean deviation (MD) is a standard clinical measure of the overall severity of a visual field defect, relative to healthy age-matched observers, with more negative values indicating greater visual field loss (Flammer, 1986; Artes et al., 2011). MD values were used as a measure of overall glaucomatous disease severity in the patient group. Patients with MD better than -6 dB in both eyes were classified as having early glaucomatous damage whilst those with MD worse than -12 dB in both eyes were considered to have advanced disease; these patients would be typically symptomatic and would, for example, likely fail the visual field component for fitness to drive in the UK (Saunders et al., 2012). All other patients would be considered to have moderate disease severity. These values were taken from a widely used and well established criterion for summarizing disease stages in glaucoma (Hodapp et al., 1993; Mills et al., 2006).

Three other vision tests were performed on the day of the experiment. Corrected binocular visual acuity (BVA) was measured using an Early Treatment Diabetic Retinopathy Study (ETDRS) chart. All participants recorded BVA of at least 0.18 logMAR (Snellen equivalent of 6/9). We chose to restrict the study to patients with preserved VA to allow for the impact of glaucoma to be better isolated. Contrast sensitivity (CS) was measured in log units with a Pelli-Robson chart. The Oculus C-Quant straylight meter (Oculus GmbH, Wetzlar, Germany) was used to measure abnormal light scattering in the eye media, in order to eliminate significant media opacity and other lens type artifacts as confounding ocular conditions; all participants were required to be within “normal limits” for this test.

All participants were examined with a modified version of the Middlesex Elderly Assessment of Mental State (MEAMS, Pearson, London, UK), a psychometric test designed to detect gross impairment of specific cognitive skills such as memory and

object recognition in an elderly population. Each section of the MEAMS is scored independently with lower scores indicating a more significant cognitive impairment. The use of MEAMS for screening of impairments has been validated by a number of research studies (Morris et al., 2000). All participants passed the MEAMS test. Individual scores (percent) were also recorded to compare patients with controls.

MAIN EXPERIMENT

Participants viewed three separate unmodified TV and film clips with sound on a 54 cm monitor (Iiyama Vision Master PRO 514, Iiyama Corporation, Tokyo, Japan) at a resolution of 1600 by 1200 pixels (refresh rate 100 Hz). One clip was an excerpt from an entertainment program (309 sec; *Dads Army*, BBC Television) which covered the full screen (subtending a half-angle of 20.3° by 14.9°). The other two clips were taken from a feature film (200 sec; *The History Boys*, 20th Century Fox) and a sports program (436 sec; *2010 Vancouver Winter Olympics Men's Ski Cross*, BBC Television); both these clips were recorded at a 16:9 ratio, therefore they contained black rectangles at the top and bottom of the screen (subtended a half-angle of 17.3° by 10.6°). Participants were positioned, using a chin rest, at a viewing distance of 60 cm. The volume was kept at the same level throughout all the trials and the films were in color. All participants wore trial frames with a refractive correction suitable for the viewing distance of 60 cm to ensure that any obstruction to the field of view caused by spectacle frames would be equivalent for everyone.

Monocular eye movements were recorded at a temporal resolution of 1000 HZ during the task using an SR Research EyeLink 1000 (SR Research Ltd., Ontario, Canada). A purpose-written application developed in C++ using Microsoft DirectShow was used to display the video and link with the SR Research EyeLink API. The eye giving the best quality pupil detection and corneal reflection was chosen for tracking. The EyeLink proprietary algorithm was used to calibrate and verify the participants' point of regard in relation to the correct location on the display. After measuring fixation the default calibration technique monitors eye movement to a stimulus presented at nine points on the monitor. This process is repeated giving a measure of the validation of the calibration. This process only took a few minutes at most. Calibration accuracy flagged by the system to be of a “good” level was a prerequisite before each trial. A drift correction was also performed before each of the three films was displayed, and in the case where a large drift (greater than 5°) was detected, a recalibration performed. The films were shown in the same order for all participants.

EYE MOVEMENT AND DATA ANALYSIS

The EyeLink 1000 gives average eye position accuracy of better than 0.5° and uses velocity and acceleration thresholds of $30^\circ/\text{s}$ and $8000^\circ/\text{s}^2$, respectively to identify saccades. This simple definition means smooth pursuits are excluded. Since the recordings were done over a relatively long uninterrupted period of time then it was not unusual for the eye tracker to lose the position of the point of regard because of blinks and loss of pupil position. In these instances the data is not useful. An automated filtering

technique was applied to the collected data. A sliding window, passing over the entire temporal eye movement trace associated with each film for each person, counts the number of saccades made per second. The percentage of 1 sec regions containing one or more saccades is delineated per video clips per person and labeled as “good” data. Any clip that contains less than 60% of these regions of “good” data is excluded from the analysis. This filtering was only used to exclude whole film clips. If a film clip was included then all the available measured saccades made whilst viewing it were used.

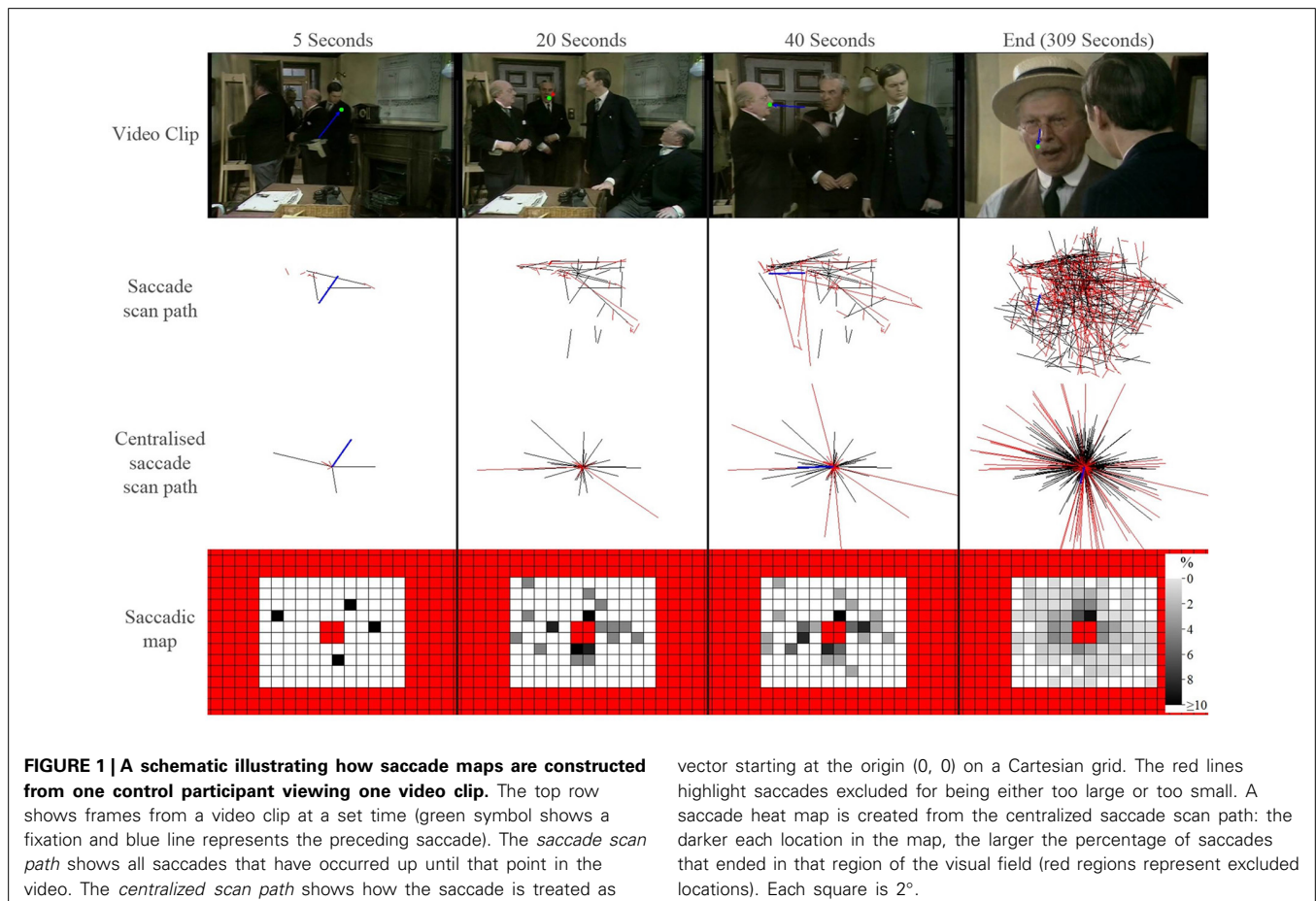
Next our software application builds a scanpath of saccades and fixations for each person viewing each clip. From this a saccade map is constructed on a grid of size 12° by 10° half angle subdivided into 2° regions. This grid is populated with the end position of every single saccade after it is assumed to be a vector starting at the origin (0, 0) on a Cartesian grid. Saccades with end points falling outside this region are excluded. In addition, the central 4 regions are excluded from the map thus removing indeterminable saccades of small amplitude. The frequency distribution of the saccade endpoints across the entire map is then recorded. This methodology is illustrated with the schematic shown in **Figure 1** and in a movie given in Supplementary Material. For the next stage of the analysis we have a matrix of 116 values (10×20 with 4 central excluded values) for each video clip for each person.

We first use kernel principal component analysis (KPCA) to extract classifiable features from the saccade maps. KPCA is widely used in pattern recognition and image processing problems (Scholkopf et al., 1999; Kwang In Kim et al., 2002; Jian Yang et al., 2005). In short, KPCA is similar to PCA but it handles non-linearities in the data by implicitly transforming data into high dimensional feature space via the kernel function and then performing a linear analysis in that space. The technique extracts features in an unsupervised fashion meaning, from a practical point of view, that none of the saccade maps are labeled as a case or a control. The feature space in KPCA is defined by “kernels” that quantifies the “distance” measure between every pair of participants. Each saccade map is serialized to a 116 point vector and the difference between participants on one video trial is defined as the Euclidean distance between the two corresponding vectors.

The kernel k_{ij} between two participants i and j is then defined as a non-normalized Gaussian distribution (1) where *meanDiff* and *maxDiff* are the mean and maximum difference of all video trials between the two participants:

$$k_{ij} = e^{-\frac{1}{2} \frac{(\text{meanDiff} + \text{maxDiff})^2}{0.2^2}} \quad (1)$$

The kernel k_{ij} forms the KPCA feature Gram matrix **K** which is then normalized and decomposed into principal eigenvectors, the



importance of these is evaluated by ranking the corresponding eigenvalues. The features for each person in our study are then calculated as the projection onto these principal axes. Crucially we then hypothesize that these “mathematical” features carry characteristics that will allow us to efficiently separate the patients from the controls on these data alone. In order to do this, significant features with the highest eigenvalues are input into a Naïve Bayes linear classification algorithm. The classifier is trained and tested by a 10 fold cross validation on all the data, repeated 100 times. In particular, with each iteration of the cross validation, 90% of the participants are randomly sampled to train a Naïve Bayes model; the diagnostic performance of this model is then tested on the remaining 10% participants. The average sensitivity of the technique is then estimated (with 95% confidence intervals) at fixed specificity using all iterations for the entire sample of the data. These values are used to construct a Receiver Operating Characteristic (ROC) curve summarizing the classification potential of the methodology.

The analytical methods described were implemented in purpose written programs using MATLAB R2013a (MathWorks Inc., Natick, MA).

RESULTS

Seventy-eight people were recruited and took part in the study. We failed to acquire sufficient eye tracking data in two individuals (both patients). Therefore, our study sample comprised 44 patients and 32 healthy people (controls) with mean age of 69 (standard deviation [SD]: 8) and 68 (SD: 9) years, respectively. These means were not significantly different (two sample *t*-test; $P = 0.58$) and the distribution of the ages were also equivalent (*F*-test of variances; $P = 0.35$) meaning the groups represent age-similar populations. Twenty-two of the patients (50%) and 17 controls (53%) were female. Glaucoma disease severity in the patient group, as described by HFA MD in both eyes, is shown in **Figure 2**; 9 (20%) and 11 (25%) of the patients had early and advanced glaucomatous disease, respectively. Median (interquartile range) right eye and left eye MD was -7 (-13 to -5) dB and -9 (-15 to -4) dB, respectively. Summary statistics for the other vision and supplementary tests are shown in **Table 1**. As expected CS differed between groups. The difference between group BVA was statistically significant ($P = 0.002$) but the actual size of the average difference was clinically small (95% CI for the mean difference of 0.03 – 0.13), representing about 4 letters on the chart, reflecting our minimum acuity inclusion criterion.

Saccade maps for every film clip for every participant are given in **Figure 3**. In total 205 out of 234 (88%) film clips were wholly included in the analysis. There was no statistically significant difference (Chi-Squared Test; $P = 0.43$) between the average inclusion rates for the controls (92%) and patients (86%).

An illustration of the results from the feature extraction, as applied to the saccade maps, is given in **Figure 4**. Here the two most significant features (eigenvectors, in arbitrary units) are plotted against each other. This visualization hints at the good separation achieved by considering the saccade maps alone when the data is reduced to just two feature axes. KPCA revealed that the five feature axes with the largest eigenvalues accounts for 35% variance in the data. These five feature axes were then used in

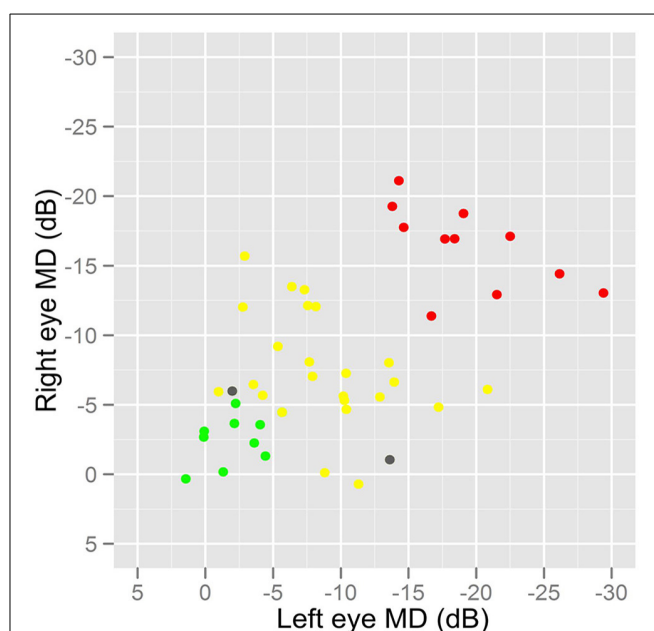


FIGURE 2 | A graph showing the distribution of disease severity in the patients. HFA MD values were used as a measure of overall glaucomatous disease severity in the patient group. Patients with MD better than -6 dB in both eyes were classified as having early glaucomatous damage (green) whilst those with MD worse than -12 dB in both eyes were considered to have advanced disease (red). All other patients would be considered to have moderate disease severity (yellow). Two patients, excluded because of sufficient eye tracking data, are shown as gray symbols.

Table 1 | Group summary statistics (mean [SD]) and comparison (two-sample *t*-test) for ETDRS corrected binocular LogMAR visual acuity (BVA), Pelli-Robson contrast sensitivity and MEAMS.

	Patients <i>n</i> = 44	Controls <i>n</i> = 32	<i>p</i> -value
BVA (LogMAR)	0.02 (0.13)	-0.06 (0.09)	0.002
CS (Log)	1.84 (0.18)	1.95 (0.00)	<0.001
MEAMS (%)	97.9 (3.7)	98.1 (3.9)	0.82

the Naïve Bayes linear classification algorithm. An ROC curve summarizing the “diagnostic precision” of the classifier is shown in **Figure 5**. The area under the ROC curve is 0.85 (95% confidence interval of 0.82 – 0.87). One point is highlighted on the ROC curve since this illustrates that the technique has a sensitivity (hit rate) of 76% (95% confidence interval of 58 – 86%) at 90% specificity.

DISCUSSION

A whole series of visual and neurological processes coalesce in order to allocate gaze efficiently. It therefore seems reasonable that gaze patterns might be inhibited or be altered by visual and neurological disorders. This simple notion underpins our driving hypothesis that scanpaths from gaze patterns might be idiosyncratic of neurodegenerative conditions. It might be that these patterns can only be realized after examining an extensive period

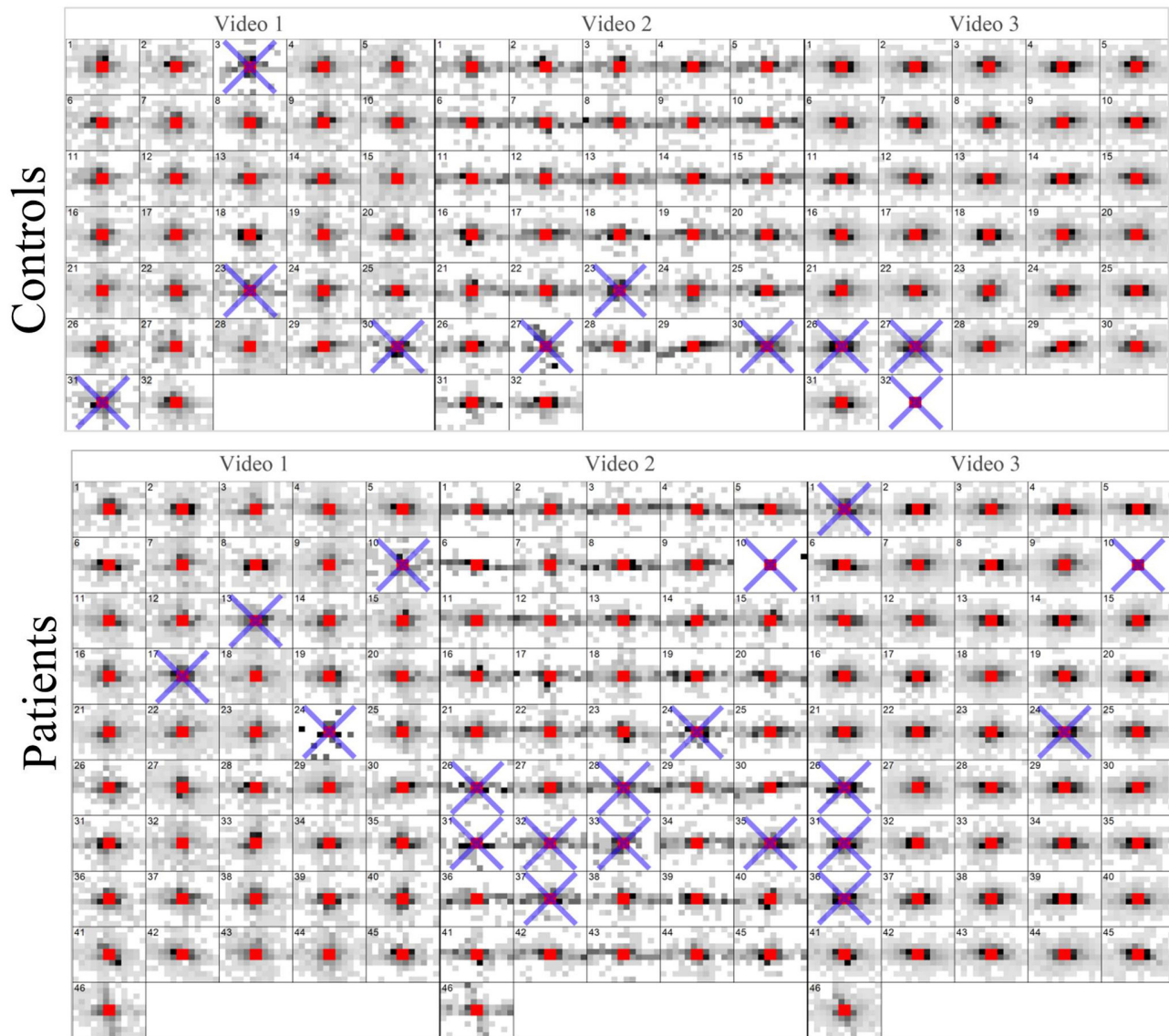
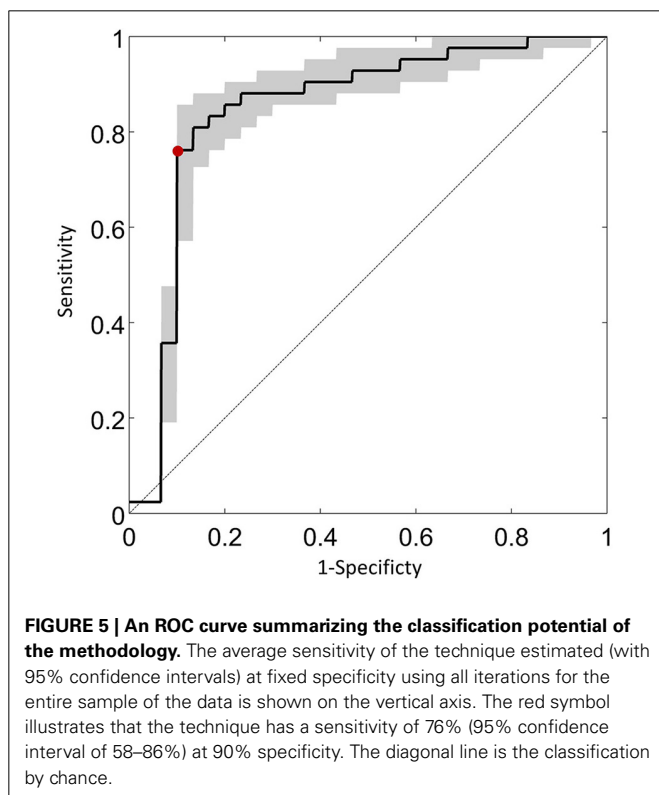
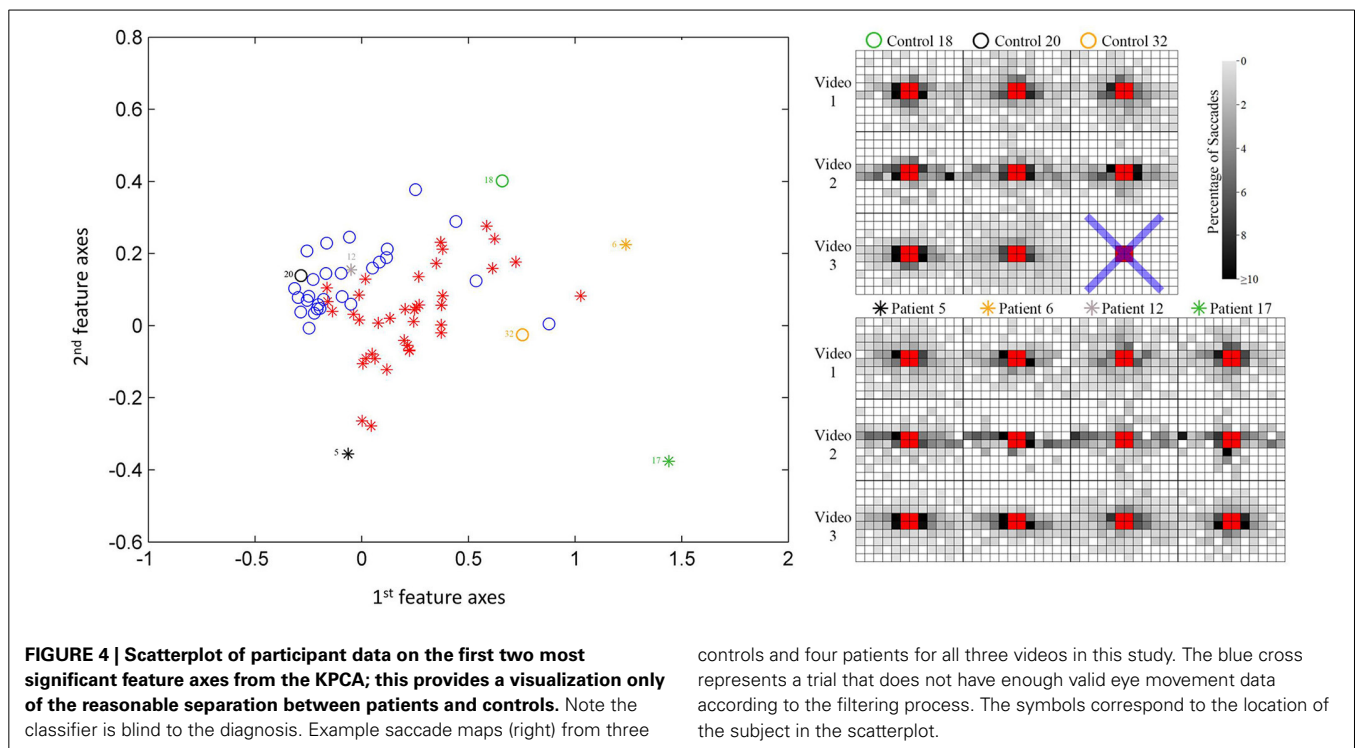


FIGURE 3 | Saccade maps for every video clip for every participant. A blue cross indicates a trial (video) that did not have enough valid eye movement data according to the filtering process. These maps represent the entire data processed by the KPCA.

of recorded eye movements, but this is now possible with modern eye tracking equipment and statistical methods that can interpret the mass of scanpath data that is yielded by them. In this case-control study we chose patients with one particular neurodegenerative eye disease. Gaze patterns were examined simply whilst viewing everyday TV type films in a quite uncontrolled fashion. The results from this study demonstrate proof of principle for the effectiveness of our approach. Using novel methodology we show that features extracted from extensive maps of saccades made whilst watching TV clips can be quantified to correctly differentiate a group of patients with a clinical diagnosis of glaucoma from a group of age-similar healthy people. For example, at a relatively high specificity (90%) we classified a patient in our sample with sensitivity of 76% by using the saccade maps alone.

Some context for this estimate of “diagnostic accuracy” is needed. There are three well-established index tests used to detect glaucoma: intraocular pressure (IOP) measurements, visual field tests and clinical assessments of the optic nerve. A systematic review of the clinical effectiveness of detecting glaucoma reported an enormous range for the estimates of diagnostic performance in these tests (Burr et al., 2007; Mowatt et al., 2008). The pooled estimates of the sensitivity of the clinically used tests in detecting COAG ranged from as low as 42–92%, whilst specificity ranged from 75 to 95%. Elevated IOP is only a risk factor for disease but tonometry, the instrumentation used to assess high IOP, is routinely used in primary care and case finding. The lower estimates for diagnostic accuracy (~ 40–50%) were for IOP measures. Unsurprisingly those index tests reported to have



better performance tended to be perimetric tests sharing direct attributes with the reference standard. Methods for assessing the ONH, either by direct observation by expert ophthalmologist or an imaging device fare no better in detecting glaucoma, yielding

only modest accuracy when used in isolation. In one well reported study, even expert ophthalmologists were shown to only classify ONH photographs moderately well for detecting glaucoma with diagnostic accuracy similar to what we report in our study (Reus et al., 2010).

Our findings need to be discussed in the context of current research. A catalog of research, extensively reviewed elsewhere, suggests that laboratory recordings of eye movements can provide valuable information about neurodegenerative disease, and hold promise as biomarkers for characterizing the efficacy of neuroprotective and neurorestorative therapies (Anderson and MacAskill, 2013). At the same time research findings about the nature and consistency of gaze patterns shown by groups of individuals as they view scenes or films are somewhat mixed (Borji and Itti, 2014). On the one hand, studies have shown that individuals within a group tend to produce similar eye movements for the same movie of natural scenes; a finding thought to be influenced by scene driven saliency effects whereby certain properties of the image cause individuals to produce a certain type of eye movement (Itti, 2006; Cristino and Baddeley, 2009; Tatler et al., 2011). Conversely, some studies have reported high levels of variability of eye movements between subjects in response to viewing scenes, which could be indicative of idiosyncratic viewing patterns within an individual (Andrews and Coppola, 1999; Castelhamo and Henderson, 2008), whilst others, for example, have shown eye movements to natural images vary enormously as a function of personality (Mercer Moss et al., 2012). More recently other workers have pursued the effect of age in processing video (Kirkorian et al., 2012). Others did not find evidence for idiosyncratic viewing patterns of the same subject across different movies (Dorr et al., 2010). Our study is most closely

aligned with the attempts made to classify patients with attention deficit hyperactivity disorder, fetal alcohol spectrum disorder and Parkinson's disease using natural viewing eye movements by Tseng et al. (2013). A novel computational model of visual attention based on saliency properties of each frame of specifically constructed scene-shuffled videos was used as a benchmark to predict gaze. By using a case-control study, similar in design to ours, the investigators managed to classify a small group of patients with acceptable levels of diagnostic accuracy. The investigators also used machine learning methods to analyse the eye movement data and to classify the differences between predicted gaze and actual gaze in the individuals (Tseng et al., 2013). Our work differs because we did not build a model predicting where people should look. Neither did we specifically select dynamic content to be used. The novelty of our hypothesis rested on the principal of demonstrating that patients could be separated from controls where the dynamic content being viewed was not extensively controlled. Our experimental effect was surprisingly large given this condition.

The results from this study hint at important clinical applications and we speculate on these briefly now. Potential tests of age-related neurological conditions, like glaucoma, based on the concept formulated in this study would require little patient action beyond the passive viewing of movies. Such a procedure would have the potential to provide a continuous assessment of changes either as the disease developed, or during treatment, within a more realistic visual environment. Moreover, eye tracking will likely become more affordable, practical, and robust in the near future, driven not by scientific research but by demands of computer gaming, mobile technology, and developments in human computer interaction. Our contention is that anomalies in eye movements tracked during viewing of naturalistic stimuli have the potential to be developed into a rigorous test that could be incorporated into an everyday activity, like something as simple as watching a movie.

The experimental design of our study had several strengths. The sampling was done carefully to ensure that the cases had the same age, sex, and general health profile as the controls. All participants passed the MEAMS test for cognitive ability and average scores from this test did not differ between the two groups. Our experimental design tried to eliminate other visual factors affecting the results. For example, we chose to restrict the study to patients with preserved visual acuity and healthy optical media (assessed by straylight measures) to allow for the contribution of functional loss due to glaucoma to be better isolated. In addition, the patients in this study had a range of disease severity (Figure 2) adding to the strength of our findings. (Tests with diagnostic promise become more sensitive as the disease becomes more severe; a study including participants with advanced disease only would by default report better sensitivity.) Of course, at this stage, it is very important to acknowledge results from this preliminary case-control report do not even remotely suggest that our method would translate into clinically significant gains in the diagnostic precision of the disease; this must be the subject of a study that must follow appropriate standards (Bossuyt et al., 2003).

Our study had other notable novel attributes. For example, the majority of eye movement studies reported in the literature have been done on trained observers or young healthy volunteers: our study clearly demonstrates that it is possible to collect extensive eye movement data on elderly people; more than one quarter of our participants were older than 75 years. Moreover, we took advantage of modern statistical methods (KPCA) to assess the high dimensional data yielded by the eye tracking. These methods are unsupervised and “learn” the discriminatory features without training on data that is labeled as “case” or “control.”

There are several limitations to our study. This proof of concept study was designed to develop our method and test it on the same material. It is well established that if modeling and testing is done on the same data then model estimates of effects will be overly optimistic. A far better experiment would have used one sample of participants to develop the classifier and another sample for testing. Our sample size was relatively small, although large enough to demonstrate an effect. All the testing was done on the same experimental set up; the repeatability of these results on another experimental set up is unknown. The study included no tests for attention deficits and so we cannot be sure that the two groups would have had the same level of attention. Furthermore, assessment for cognition was based on a modified screening test only. Therefore, we cannot be certain that the two groups would be the same had a more thorough cognitive examination been carried out. Two patients were excluded because we could not extract meaningful scanpath data from them, and this would be important if clinically meaningful estimates sensitivity and specificity were being reported. Moreover, our results and the paucity of data offer no real hint about what particular characteristics of the saccade maps are suggestive of abnormality and this awaits further study. Moreover, only patients with glaucoma have been considered and we have no idea if these results will translate to other age-related neurodegenerative conditions. Furthermore, the pathogenesis of glaucoma shares many features with other chronic age-related neurodegenerative disease but it is not typically or primarily classified as such. The site of damage histologically appears to be at the level of the ONH and is thought to be due to the interaction of IOP, cerebrospinal fluid pressure, ONH blood flow and changes in lamina cribrosa anatomy with retinal ganglion cell changes being secondary.

In conclusion we have shown scanpaths of eye movements recorded whilst people freely watch TV type films can be processed into maps that contain a signature of age-related neurological disease. In this *proof of principle* study we have demonstrated that a group of patients with glaucoma can be reasonably well separated from a group of healthy peers by considering these eye movement signatures alone. Future studies will consider larger groups of patients and other age-related neurological disorders.

ACKNOWLEDGMENTS

We thank Fiona Glen and Robyn Burton for study recruitment and helping to carry out the experiments. We also thank Ryo Asaoka (now at University of Tokyo Graduate School of Medicine) for help with study recruitment and carrying out some of the ophthalmological examinations. This study was funded by

a project grant awarded by *Fight For Sight* (United Kingdom) www.fightforsight.org.uk.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnagi.2014.00312/abstract>

REFERENCES

- Anderson, T. J., and MacAskill, M. R. (2013). Eye movements in patients with neurodegenerative disorders. *Nat. Rev. Neurol.* 9, 74–85. doi: 10.1038/nrneuro.2012.273
- Andrews, T. J., and Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Res.* 39, 2947–2953. doi: 10.1016/S0042-6989(99)00019-X
- Artes, P. H., O'Leary, N., Hutchison, D. M., Heckler, L., Sharpe, G. P., Nicoletta, M. T., et al. (2011). Properties of the statpac visual field index. *Invest. Ophthalmol. Vis. Sci.* 52, 4030–4038. doi: 10.1167/iov.10-6905
- Asman, P., and Heijl, A. (1992). Glaucoma hemifield test. Automated visual field evaluation. *Arch. Ophthalmol.* 110, 812–819. doi: 10.1001/archophth.1992.01080180084033
- Bayer, A. U., Ferrarri, F., and Erb, C. (2002). High occurrence rate of glaucoma among patients with Alzheimer's disease. *Eur. Neurol.* 47, 165–168. doi: 10.1159/000047976
- Benson, P. J., Beedie, S. A., Shephard, E., Giegling, I., Rujescu, D., and St Clair, D. (2012). Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. *Biol. Psychiatry* 72, 716–724. doi: 10.1016/j.biopsych.2012.04.019
- Borji, A., and Itti, L. (2014). Defending Yarbus: eye movements reveal observers' task. *J. Vis.* 14, 29. doi: 10.1167/14.3.29
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., et al. (2003). The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann. Intern. Med.* 138, W1–W12. doi: 10.7326/0003-4819-138-1-200301070-00010
- Burr, J. M., Mowatt, G., Hernández, R., Siddiqui, M. A. R., Cook, J., Lourenco, T., et al. (2007). The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol. Assess.* 11, iii–iv, ix–x, 1–190. doi: 10.3310/hta11410
- Caprioli, J. (2013). Glaucoma: a disease of early cellular senescence. *Invest. Ophthalmol. Vis. Sci.* 54, ORSF60–ORSF6067. doi: 10.1167/iov.13-12716
- Castellano, M. S., and Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Can. J. Exp. Psychol.* 62, 1–14. doi: 10.1037/1196-1961.62.1.1
- Chauhan, B. C., Garway-Heath, D. F., Goñi, F. J., Rossetti, L., Bengtsson, B., Viswanathan, A. C., et al. (2008). Practical recommendations for measuring rates of visual field change in glaucoma. *Br. J. Ophthalmol.* 92, 569–573. doi: 10.1136/bjo.2007.135012
- Crabb, D. P., Smith, N. D., Rauscher, F. G., Chisholm, C. M., Barbur, J. L., Edgar, D. F., et al. (2010). Exploring eye movements in patients with glaucoma when viewing a driving scene. *PLoS ONE* 5:e9710. doi: 10.1371/journal.pone.0009710
- Crawford, T. J., Higham, S., Mayes, J., Dale, M., Shaunak, S., and Lekwuwa, G. (2013). The role of working memory and attentional disengagement on inhibitory control: effects of aging and Alzheimer's disease. *Age (Dordr)* 35, 1637–1650. doi: 10.1007/s11357-012-9466-y
- Cristino, F., and Baddeley, R. (2009). The nature of the visual representations involved in eye movements when walking down the street. *Vis. cogn.* 17, 880–903. doi: 10.1080/13506280902834696
- Dorr, M., Martinetz, T., Gegenfurtner, K. R., and Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *J. Vis.* 10:28. doi: 10.1167/10.10.28
- Flammer, J. (1986). The concept of visual field indices. *Graefes Arch. Clin. Exp. Ophthalmol.* 224, 389–392. doi: 10.1007/BF02173350
- Fung, S. S. M., Lemer, C., Russell, R. A., Malik, R., and Crabb, D. P. (2013). Are practical recommendations practiced? A national multi-centre cross-sectional study on frequency of visual field testing in glaucoma. *Br. J. Ophthalmol.* 97, 843–847. doi: 10.1136/bjophthalmol-2012-302903
- Glen, F. C., Baker, H., and Crabb, D. P. (2014). A qualitative investigation into patients' views on visual field testing for glaucoma monitoring. *BMJ Open* 4:e003996. doi: 10.1136/bmjopen-2013-003996
- Glen, F. C., Smith, N. D., and Crabb, D. P. (2013). Saccadic eye movements and face recognition performance in patients with central glaucomatous visual field defects. *Vision Res.* 82, 42–51. doi: 10.1016/j.visres.2013.02.010
- Hodapp, E., Parrish, R. K. I., and Anderson, D. R. (1993). *Clinical Decisions in Glaucoma*. St. Louis, MO: The CV Mosby Co.
- Itti, L. (2006). Quantitative modelling of perceptual salience at human eye position. *Vis. cogn.* 14, 959–984. doi: 10.1080/13506280500195672
- Kim, K. I., Jung, K., and Kim, H. J. (2002). Face recognition using kernel principal component analysis. *IEEE Signal Process. Lett.* 9, 40–42. doi: 10.1109/97.991133
- Kirkorian, H. L., Anderson, D. R., and Keen, R. (2012). Age differences in online processing of video: an eye movement study. *Child Dev.* 83, 497–507. doi: 10.1111/j.1467-8624.2011.01719.x
- Lueck, K. L., Mendez, M. F., and Perryman, K. M. (2000). Eye movement abnormalities during reading in patients with Alzheimer disease. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 13, 77–82.
- Mercer Moss, F. J., Baddeley, R., and Canagarajah, N. (2012). Eye movements to natural images as a function of sex and personality. *PLoS ONE* 7:e47870. doi: 10.1371/journal.pone.0047870
- Mills, R. P., Budenz, D. L., Lee, P. P., Noecker, R. J., Walt, J. G., Siegartel, L. R., et al. (2006). Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. *Am. J. Ophthalmol.* 141, 24–30. doi: 10.1016/j.ajo.2005.07.044
- Morris, R. G., Worsley, C., and Matthews, D. (2000). Neuropsychological assessment in older people: old principles and new directions. *Adv. Psychiatr. Treat* 6, 362–370. doi: 10.1192/apt.6.5.362
- Mosimann, U. P., Fellinger, J., Ballinari, P., Hess, C. W., and Müri, R. M. (2004). Visual exploration behaviour during clock reading in Alzheimer's disease. *Brain* 127, 431–438. doi: 10.1093/brain/awh051
- Mowatt, G., Burr, J. M., Cook, J. A., Siddiqui, M. A. R., Ramsay, C., Fraser, C., et al. (2008). Screening tests for detecting open-angle glaucoma: systematic review and meta-analysis. *Invest. Ophthalmol. Vis. Sci.* 49, 5373–5385. doi: 10.1167/iov.07-1501
- National Institute for Health and Clinical Excellence, (2009). *Glaucoma: Diagnosis and Management of Chronic Open Angle Glaucoma and Ocular Hypertension. Clinical Guidelines* CG85. (London: National Institute for Health and Clinical Excellence MidCity Place). Available online at: www.nice.org.uk
- Rabin, R., and de Charro, F. (2001). EQ-5D: a measure of health status from the EuroQol Group. *Ann. Med.* 33, 337–343. doi: 10.3109/07853890109002087
- Reus, N. J., Lemij, H. G., Garway-Heath, D. F., Airaksinen, P. J., Anton, A., Bron, A. M., et al. (2010). Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European optic disc assessment trial. *Ophthalmology* 117, 717–723. doi: 10.1016/j.ophtha.2009.09.026
- Rudnicka, A. R., Mt-Isa, S., Owen, C. G., Cook, D. G., and Ashby, D. (2006). Variations in primary open-angle glaucoma prevalence by age, gender, and race: a Bayesian meta-analysis. *Invest. Ophthalmol. Vis. Sci.* 47, 4254–4261. doi: 10.1167/iov.06-0299
- Saunders, L. J., Russell, R. A., and Crabb, D. P. (2012). Practical landmarks for visual field disability in glaucoma. *Br. J. Ophthalmol.* 96, 1185–1189. doi: 10.1136/bjophthalmol-2012-301827
- Scholkopf, B., Smola, A., and Müller, K.-R. (1999). "Kernel principal component analysis," in *Advances in Kernel Methods - Support Vector Learning*, eds B. Schölkopf, C. J. C. Burges and A. J. Smola (Massachusetts, MA: MIT Press), 327–352.
- Sivak, J. M. (2013). The aging eye: common degenerative mechanisms between the Alzheimer's brain and retinal disease. *Invest. Ophthalmol. Vis. Sci.* 54, 871–880. doi: 10.1167/iov.12-10827
- Smith, N. D., Glen, F. C., and Crabb, D. P. (2012). Eye movements during visual search in patients with glaucoma. *BMC Ophthalmol.* 12:45. doi: 10.1186/1471-2415-12-45
- Tatler, B. W., Hayhoe, M. M., Land, M. F., and Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *J. Vis.* 11:5. doi: 10.1167/11.5.5
- Tezel, G., and Fourth ARVO/Pfizer Ophthalmics Research Institute Conference Working Group (2009). The role of glia, mitochondria, and the immune system in glaucoma. *Invest. Ophthalmol. Vis. Sci.* 50, 1001–1012. doi: 10.1167/iov.08-2717
- Tseng, P.-H., Cameron, I. G. M., Pari, G., Reynolds, J. N., Munoz, D. P., and Itti, L. (2013). High-throughput classification of clinical populations from natural

- viewing eye movements. *J. Neurol.* 260, 275–284. doi: 10.1007/s00415-012-6631-2
- Yang, J., Frangi, A. F., Yang, J.-Y., Zhang, D., and Jin, Z. (2005). KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 230–244. doi: 10.1109/TPAMI.2005.33
- Yücel, Y., and Gupta, N. (2008). Glaucoma of the brain: a disease model for the study of transsynaptic neural degeneration. *Prog. Brain Res.* 173, 465–478. doi: 10.1016/S0079-6123(08)01132-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 June 2014; accepted: 23 October 2014; published online: 11 November 2014.

Citation: Crabb DP, Smith ND and Zhu H (2014) What's on TV? Detecting age-related neurodegenerative eye disease using eye movement scanpaths. *Front. Aging Neurosci.* 6:312. doi: 10.3389/fnagi.2014.00312

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Crabb, Smith and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition

Christian Füllgrabe^{1*}, Brian C. J. Moore² and Michael A. Stone^{3,4}

¹ MRC Institute of Hearing Research, Nottingham, UK

² Department of Psychology, University of Cambridge, Cambridge, UK

³ School of Psychological Sciences, University of Manchester, Manchester, UK

⁴ Central Manchester NHS Hospitals Foundation Trust, Manchester, UK

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Jerker Rönnberg, Linköping University, Sweden

Larry E. Humes, Indiana University, USA

Tim Schoof, University College London, UK

*Correspondence:

Christian Füllgrabe, MRC Institute of Hearing Research, Science Road, Nottingham, NG7 2RD, UK
e-mail: christian.fullgrabe@ihr.mrc.ac.uk

Hearing loss with increasing age adversely affects the ability to understand speech, an effect that results partly from reduced audibility. The aims of this study were to establish whether aging reduces speech intelligibility for listeners with normal audiograms, and, if so, to assess the relative contributions of auditory temporal and cognitive processing. Twenty-one older normal-hearing (ONH; 60–79 years) participants with bilateral audiometric thresholds ≤ 20 dB HL at 0.125–6 kHz were matched to nine young (YNH; 18–27 years) participants in terms of mean audiograms, years of education, and performance IQ. Measures included: (1) identification of consonants in quiet and in noise that was unmodulated or modulated at 5 or 80 Hz; (2) identification of sentences in quiet and in co-located or spatially separated two-talker babble; (3) detection of modulation of the temporal envelope (TE) at frequencies 5–180 Hz; (4) monaural and binaural sensitivity to temporal fine structure (TFS); (5) various cognitive tests. Speech identification was worse for ONH than YNH participants in all types of background. This deficit was not reflected in self-ratings of hearing ability. Modulation masking release (the improvement in speech identification obtained by amplitude modulating a noise background) and spatial masking release (the benefit obtained from spatially separating masker and target speech) were not affected by age. Sensitivity to TE and TFS was lower for ONH than YNH participants, and was correlated positively with speech-in-noise (SiN) identification. Many cognitive abilities were lower for ONH than YNH participants, and generally were correlated positively with SiN identification scores. The best predictors of the intelligibility of SiN were composite measures of cognition and TFS sensitivity. These results suggest that declines in speech perception in older persons are partly caused by cognitive and perceptual changes separate from age-related changes in audiometric sensitivity.

Keywords: aging, normal hearing, speech identification, temporal envelope, temporal fine structure, cognition

INTRODUCTION

Aging in adults is associated with deterioration and increased effortfulness of all levels of speech processing (from identification to comprehension), especially in noisy and reverberant conditions (e.g., CHABA, 1988). It has long been known that hearing sensitivity declines with increasing age (Bunch, 1929; Corso, 1963) and this is associated with poorer speech identification (Harris et al., 1956; Delk et al., 1957). More recently, it has also become apparent that hearing-impaired people report a lower quality of life (Dalton et al., 2003), experience more social isolation (Weinstein and Ventry, 1982; Strawbridge et al., 2000) and depression (Gopinath et al., 2009; Huang et al., 2010), and show poorer cognitive functioning and accelerated cognitive decline (Lin et al., 2011, 2013) than normal-hearing people. This suggests that speech communication difficulties not only constitute a socio-psychological handicap for the affected person

(Arlinger, 2003) but also represent an important financial burden for society in terms of social and health care provision (Mohr et al., 2000; Hjalte et al., 2012; Foley et al., 2014).

Modern digital hearing aids, which provide frequency-specific amplification, at least partially restore audibility of those sounds that would not otherwise be perceived by the hearing-impaired person. These aids are the standard treatment in most cases of hearing loss. While aided speech identification in quiet and background noise generally improves with increasing audibility (e.g., Humes, 2002), the observed benefit often falls short of what would be expected based on audibility (Humes and Dubno, 2010). One possible explanation for this is that age-related changes in supra-threshold auditory processing and cognition—that are not captured by an audiometric assessment—contribute to the speech-identification difficulties of older people (e.g., Humes et al., 2013b; Moore et al., 2014; Schoof and Rosen, 2014).

To study the existence of age effects unrelated to audibility, most previous research adopted a cross-sectional design, in which a group of older participants (generally somewhat arbitrarily taken as ≥ 60 years) was compared to a group of young controls. Given the high prevalence of hearing loss in the older population (Davis, 1995; Cruickshanks et al., 1998; Agrawal et al., 2008), establishing audiometric equality between these age groups to control for the effect of audibility is not easy. Alternative solutions have been sought to matching the age groups, for example by: (1) spectrally shaping the speech signal to equate audibility across groups; (2) distorting the speech signals delivered to the young normal-hearing (YNH) participants (e.g., by adding noise) to simulate the hearing loss of the older participants; or (3) statistically partialling out the effect of hearing loss. None of these approaches controls for possible “central effects of peripheral pathology” (Willott, 1996) in the older participants, i.e., physiological and anatomical changes in the central auditory system induced by peripheral pathology (Robertson and Irvine, 1989; Ison et al., 2010). The approach using mathematical adjustments has the additional disadvantage that, since age and audiometric thresholds are not statistically independent, partialling out the effect of hearing sensitivity also removes some of the age effect, resulting in an underestimation of the effect of the latter (Martin et al., 1991). In the present study, the older participants were selected to have hearing sensitivity matching that of a YNH control group over a wide frequency range and in both ears. In addition, a relatively large number of older normal-hearing (ONH) participants was recruited, to allow calculation of correlations across measures within the ONH group.

Many previous studies of aging focussed on either perceptual or cognitive processes involved in speech processing, frequently employing a single measure of the process under study. Here, we attempted to study the interplay and relative contribution of both of these processes in the case of speech-in-noise (SiN) identification, using multiple indices of perceptual, cognitive, and speech processing.

The choice of perceptual tasks was motivated by our knowledge of how sounds are represented or “coded” in the auditory system. Acoustic broadband signals, such as speech, are decomposed in the cochlea into a series of bandpass-filtered signals, each corresponding to a specific position on the basilar membrane. The response at each place can be considered as a temporal envelope (TE; corresponding to the slow variations in overall amplitude over time) imposed on a time-varying carrier, the temporal fine structure (TFS; faster variations corresponding to the rapid oscillations in the filtered waveform). Both types of temporal information are represented in the auditory system by the timing of neural discharges (phase locking) to the TE (e.g., Frisina, 2001; Sayles et al., 2013) or TFS (e.g., Young and Sachs, 1979). In the healthy auditory system, both TE and TFS cues, and their comparison across different places on the basilar membrane, are used for speech identification (for a review, see Moore, 2014).

Aging in the absence of elevated audiometric thresholds does not seem to have a significant negative effect on frequency selectivity, as measured using psychophysical tuning curves or the notched-noise procedure (Lutman et al., 1991; Peters and Moore, 1992; Sommers and Humes, 1993; Gifford and Bacon, 2005).

Although some studies reported a widening of the auditory filters with increasing age (Patterson et al., 1982; Glasberg et al., 1984), the older participants in those studies were either not audiometrically screened or had higher audiometric thresholds than the younger participants. Since elevated audiometric thresholds have been shown to be associated with greater auditory filter bandwidths (Moore, 2007), hearing loss most likely confounded the results. Given that the aim of the present study was to compare young and older participants with matched audiograms, measures of frequency selectivity were not included. Rather we focussed on measures of sensitivity to TE and TFS, based on behavioral (Pichora-Fuller and MacDonald, 2008; Moore, 2014) and neurophysiological (Walton et al., 1998; Clinard et al., 2010) data suggesting that aging negatively affects the processing of TE and TFS information.

Several studies of speech identification have used a signal-processing technique called vocoding (Dudley, 1939) to disrupt TFS information and reduce spectral cues, while substantially preserving information in the TE. These studies have shown that TE information in a few spectral bands can be sufficient for good identification of speech in quiet (Van Tasell et al., 1987; Shannon et al., 1995; Lorenzi et al., 2000). Modulation frequencies in the range 4–16 Hz seem to be especially important for the identification of speech in quiet (Drullman et al., 1994a,b). However, when speech is presented against interfering speech maskers, both slower and faster TE cues, associated respectively with prosodic (Füllgrabe et al., 2009) and fundamental frequency (Stone et al., 2008) information, become important for identification. Older listeners seem less able to use these complex TE patterns across different places on the basilar membrane to achieve speech identification (Souza and Boike, 2006; Schvartz et al., 2008; Sheldon et al., 2008), possibly due to reduced sensitivity to TE cues. Such a reduction should not be due to the presence of reduced hearing sensitivity in some of those listeners since, when the audibility of the stimuli is controlled for, hearing-impaired listeners have either similar (Moore and Glasberg, 2001) or better (Füllgrabe et al., 2003) TE sensitivity than normal-hearing listeners. Also, several studies using older listeners with nearly normal audiograms reported significant age-related decrements in the detection of sinusoidal amplitude modulation (SAM) imposed on pure-tone (He et al., 2008) or noise carriers (Takahashi and Bacon, 1992; Kumar and Sangamanatha, 2011). However, the results of the studies using noise carriers might have been affected by higher audiometric thresholds (especially in the high-frequency range) for the older than the younger participants, resulting in a smaller audible carrier bandwidth, which negatively affects SAM detection (Eddins, 1993). Here, TE sensitivity was assessed by measuring thresholds for detection of SAM presented over a range of modulation frequencies.

TFS information does not seem to be critical for the identification of speech in quiet. It may be more important when background sounds are present, perhaps by providing cues for auditory scene analysis (segregation of target and background sounds), such as sound-lateralization and voice-pitch cues (for an overview, see Moore, 2014). It has been argued that people with hearing loss have reduced TFS sensitivity (Smoski and Trahiotis, 1986; Hopkins and Moore, 2011), resulting in lower

speech intelligibility (Lorenzi et al., 2006). An increasing number of studies (Pichora-Fuller and Schneider, 1992; Ross et al., 2007; Grose and Mamo, 2010; Moore et al., 2012a; Füllgrabe, 2013; Whitmer et al., 2014) indicate that age *per se* may also negatively affect TFS sensitivity. However, most studies used young and older participants whose audiograms were not matched, which could have led to the observed differences. Here, TFS sensitivity was assessed monaurally and binaurally for audiometrically matched YNH and ONH participants.

The decision to conduct a cognitive assessment, in addition to psychoacoustic tasks, was motivated by the general assumption that top-down cognitive processes are involved in speech processing (Eysenck and Keane, 2000) and empirical evidence that many cognitive functions decline with age (e.g., Baltes and Lindenberger, 1997; Park et al., 2002). Akeroyd (2008) reviewed 20 studies investigating the link between performance on SiN and cognitive tasks. He concluded that, while cognition was generally linked to SiN identification, there was no single cognitive test that consistently showed such an association. Across-study differences in sample characteristics (age, hearing status, general cognitive functioning), speech material (syllables, words, sentences), and listening conditions (interfering noise or babble), as well as their interactions, might account for the observed discrepancies. Here, we used a battery of cognitive tests to investigate the role of particular cognitive abilities (such as memory, attention, and processing speed) and general cognitive functioning in SiN processing.

Finally, the choice of two types of speech tasks (closed-set phoneme identification without semantic or syntactic context vs. open-set sentence identification with linguistic context) reflects an attempt to capture different levels of speech processing (Pickett, 1999). Varying the listening conditions (e.g., in quiet, in reverberation, in the presence of different types of maskers) was meant to modulate the perceptual and cognitive load (Mattys et al., 2009). Here, identification performance was assessed for maskers producing little informational masking (Durlach et al., 2003) in the absence of reverberation, and for maskers producing considerable informational masking in the presence of reverberation.

In summary, the present study aimed to measure possible deficits in the ability to identify speech in quiet and in background sounds that occur with increasing age, in spite of the absence of hearing loss as measured by the audiogram. The aims were to establish: (1) the existence and magnitude of such deficits; (2) the degree of awareness of the deficits; and (3) the extent to which the deficits were associated with declines in auditory and cognitive processing.

MATERIALS AND METHODS

A discussion of methodological issues related to this study is provided in the Supplementary Material: *Methodological issues*.

PARTICIPANTS

Potential participants were recruited from the Cambridge (UK) area through age-targeted (18–29 years or ≥ 60 years) advertisements posted in public spaces (e.g., doctors' surgeries) and appeals to social and community clubs. Nine younger (six

females) and 21 older (20 females) native English speakers were retained for this study based on them having normal hearing sensitivity as defined by the audiometric criteria given below. The mean age of the YNH participants was 23 years (standard deviation, $SD = 3$; range = 18–27) and that of the ONH participants was 67 years ($SD = 5$; range = 60–79). All ONH participants completed the Mini Mental State Examination (Folstein et al., 1975) to screen for cognitive impairment, generally taken as indexed by scores $< 24/30$ points. All obtained full marks, bar one, who scored 29; this observation is consistent with population-based norms for 65–69-year olds with at least some university education (Crum et al., 1993). The number of years of formal education was, on average, 16.2 ($SD = 2.0$) and 16.8 ($SD = 1.9$) for the YNH and ONH groups, respectively. An independent-samples *t*-test showed that the age-group difference was not significant [$t_{(28)} = 0.712$, $p = 0.482$; two-tailed]. However, given that this proxy measure of cognitive ability is likely biased by cohort effects (ONH participants could have been prevented by historical circumstances and societal attitude toward education from attaining further education, while some YNH participants still had not completed their education), the two non-verbal sub-tests of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), Block Design and Matrix Reasoning, were also used to confirm the equivalence of the groups in terms of general cognitive functioning. Performance on the two tests can be combined into a performance IQ (see WASI manual). While the mean raw scores for the two tests differed across age groups (see Result section), the corresponding performance IQ scores (incorporating an age correction), were 123 ($SD = 7$) for the YNH and 122 ($SD = 11$) for the ONH group. This corresponds to the 92nd and 88th percentiles, respectively. According to an independent-samples *t*-test, the difference in age-corrected performance IQ was not significant [$t_{(28)} = -0.441$, $p = 0.663$; two-tailed]. Individual differences in mental functions show a high stability across the human lifespan (Deary et al., 2000; Gow et al., 2011) and the inter-individual variability in various cognitive abilities does not seem to increase with age (e.g., Salthouse, 2004, 2012). Under these circumstances, it is reasonable to assume that both age groups were sampled from the same cognitively high-functioning stratum of the underlying young and older populations.

All participants were fully informed about the aims of the study (approved by the local Cambridge University Ethics committee), provided written consent, and received monetary compensation for their participation.

Audiological assessment of hearing

Following a clinical interview (including questions about difficult listening situations), pure-tone air-conduction audiometry was conducted using a Grason-Stadler GSI61 Clinical Audiometer with TDH-50P headphones, following the procedure recommended by the British Society of Audiology (BSA, 2004). In this study, normal hearing sensitivity was defined as audiometric thresholds ≤ 20 dB Hearing Level (HL) in both ears at octave frequencies between 0.125 and 4 kHz, as well as at 3 and 6 kHz. Audiograms for the YNH and ONH participants are shown in **Figure 1** (for a comparison with audiometric thresholds found

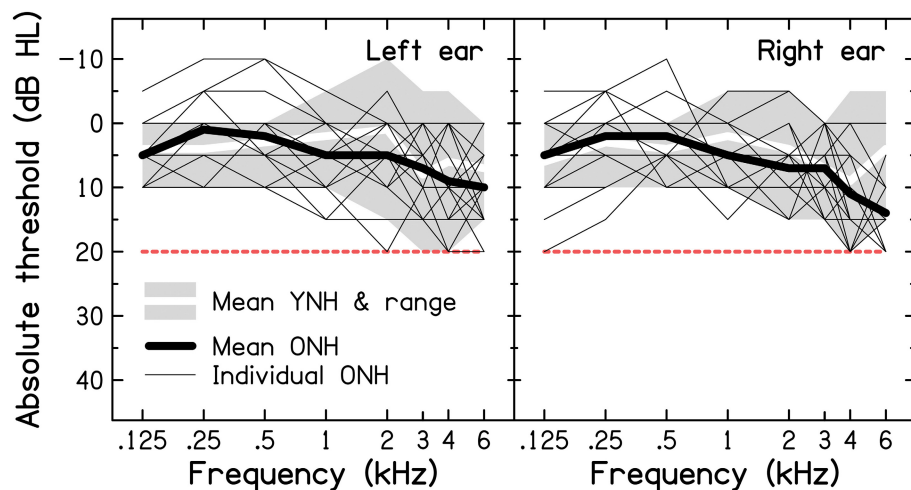


FIGURE 1 | Results of pure-tone air-conduction audiometry for the left (left panel) and right ears (right panel) of the nine YNH and 21 ONH participants. The thin and thick black lines represent the individual and mean audiograms of the ONH participants. The thick

white lines and associated light-gray shaded areas represent the mean audiograms and ranges of audiometric thresholds for the YNH participants, respectively. The dashed red line indicates the audiometric inclusion criteria used in the present study.

in a sample of older volunteers with self-reported normal hearing, see Supplementary Material: *Audiometric screening results for older volunteers with self-reported normal hearing*). Mean audiometric thresholds for the two age groups (thick lines) were very similar at all frequencies in both ears (the grand pure-tone average, $PTA_{0.125-6\text{ kHz}}$, was 5.1 and 6.1 dB HL for the YNH and ONH groups, respectively), except for the right ear at 6 kHz, where the threshold for the ONH group was higher by 8.5 dB. The mean audiometric threshold did not differ significantly across groups, as shown by an independent-samples t -test [$t_{(28)} = 0.808$, $p = 0.426$; two-tailed].

Subjective assessment of hearing

Paper-and-pencil versions of two self-report inventories, routinely used for the assessment of hearing-aid benefit, were administered to all participants to assess their hearing abilities in various everyday listening conditions. At the time of questionnaire completion, none of the participants was aware of the outcome of the audiometric assessment.

The Abbreviated Profile of Hearing Aid Benefit. The Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox and Alexander, 1995) is a questionnaire composed of 24 short statements (e.g., “I can understand conversations even when several people are talking.”). Respondents are asked to estimate how frequently they experience problems in the described situation, by selecting one of seven ordinal response alternatives, ranging from “Always” (=99%) to “Never” (=1%). Sub-scale scores are computed for Ease of Communication (EC), Reverberation (RV), Background Noise (BN), and Aversiveness (AV) by averaging across the responses to six statements for each sub-scale. The average APHAB scores for the two age groups are shown in the left panel of **Figure 2**. The frequency of experiencing problems was very similar for the two groups, as confirmed by a

mixed-design repeated-measures analysis of variance (ANOVA) with Age group as the between-subjects factor and APHAB sub-scale as the within-subjects factor. The effect of Age group was not significant [$F_{(1, 28)} = 0.008$, $p = 0.930$], nor was the Age group*APHAB sub-scale interaction [$F_{(1.229, 34.417)} = 0.222$, $p = 0.691$]¹.

The Speech, Spatial, and Qualities of hearing scale. The Speech, Spatial, and Qualities of hearing scale (SSQ; Gatehouse and Noble, 2004) is a 50-item questionnaire developed to assess how effectively auditory information is being processed in a variety of everyday listening situations. Unlike the APHAB, it includes situations explicitly involving auditory scene analysis and cognitive abilities, such as focusing on one sound source in the presence of others, or attending to multiple sound sources simultaneously. One item was excluded from the original questionnaire since it was only applicable to hearing-aid users. For each of the remaining items, respondents are asked to estimate their (dis)ability in performing an auditory-based “activity” by selecting a number on an 11-point response scale, ranging from “0” (= complete disability) to “10” (= no disability). Each item is associated with one of three sub-scales: Speech hearing (14 items; e.g., “Can you have a conversation in the presence of someone whose voice is the same pitch as that of the person you’re talking to?”), Spatial hearing (17 items; e.g., “Do you have the impression of sounds being exactly where you would expect them to be?”), and other Qualities of hearing (18 items; e.g., “Do you find it easy to recognize different people you know by the sound of each one’s voice?”). Average ability scores for the two groups are shown in the right panel of **Figure 2** for each of the sub-scales. Ratings were very similar for the two groups, as confirmed by a repeated-measures ANOVA

¹Here and for subsequent ANOVAs, when the assumption of sphericity was not met, the Greenhouse-Geisser correction was used.

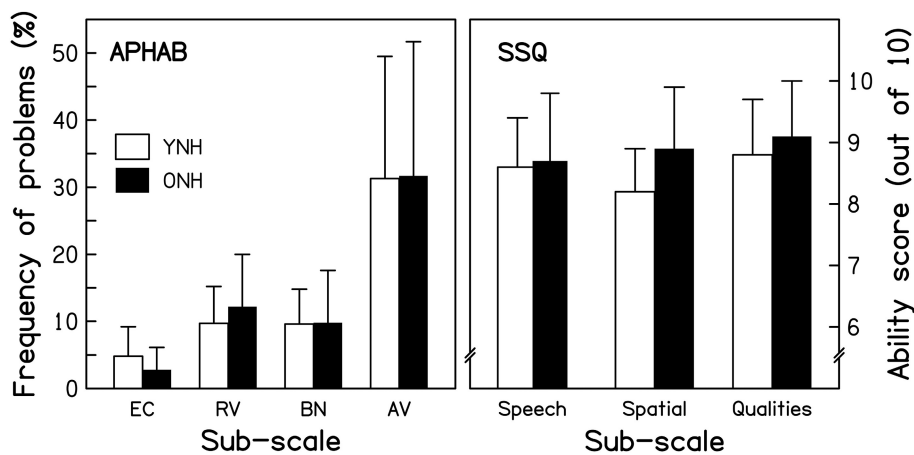


FIGURE 2 | Scores for the YNH (open bars) and ONH (filled bars) participants for two questionnaires. For the Abbreviated Profile of Hearing Aid Benefit (APHAB; left panel), responses in terms of frequency of experiencing the described problems are averaged for each of four sub-categories: Ease of communication (EC), Reverberation (RV), Background noise (BN), and Aversiveness (AV). For the Speech, Spatial,

and Qualities of hearing scale (SSQ; right panel), responses on an 11-point scale (0–10, with greater scores reflecting less disability) are averaged for the sub-categories of Speech hearing (14 questions), Spatial hearing (17 questions), and Qualities of hearing (19 questions). Note that more hearing difficulties are indicated by taller and smaller bars in the left and right panels, respectively.

that showed a non-significant effect of Age group [$F_{(1, 28)} = 1.097, p = 0.304$] and a non-significant Age group*SSQ sub-scale interaction [$F_{(2, 56)} = 1.506, p = 0.231$].

EQUIPMENT

For all auditory tasks (unless otherwise stated; see Section Assessment of sensitivity to TE information), stimuli were played with 16-bit precision through a Lynx L22 soundcard hosted in a PC, under control of custom-written software in Matlab or VisualBasic. The sampling frequency was dependent on the task but was at least 16 kHz. The soundcard output signal was buffered by a Mackie 1202-VLZ PRO mixing desk, and delivered over Sennheiser HDA580 headphones to the participants, who were seated in a sound-attenuating booth. Depending on the task, response entry was made via either a mouse click on virtual buttons displayed on a computer screen, a manual button press, or orally.

For the cognitive tests, the experimenter sat with the participant in a large sound-attenuating booth or a quiet room. Depending on the test administered, participants gave their responses either orally or manually.

SPEECH TASKS

Consonant identification

Bisyllabic vowel-consonant-vowel (VCV) stimuli with 21 different consonants were used. An /a/ was used for the initial and final vowels. The consonants were /p, t, k, b, d, g, f, θ, s, ʃ, h, v, z, r, l, j, w, t, dj, n, m/. Four utterances of each VCV were spoken by a female talker with a standard British accent, with the emphasis on the second syllable. Recordings were made in an anechoic room with 16-bit quantization and a 44.1-kHz sampling rate, later digitally down-sampled to 16 kHz.

Consonant identification was assessed using a 1-interval, 21-alternative forced-choice procedure. In each run, all 21 VCVs were

presented once in random order. Following the presentation of a VCV, the participant indicated which consonant had been heard by selecting one of 21 virtual buttons, each labeled with the orthographical representation of one of the consonants in a meaningful CV word.

VCVs were presented in quiet and in three types of background masker whose long-term average spectrum was shaped to be the same as that of the VCVs: (1) unmodulated² noise; (2) noise with 100%, 5-Hz SAM³ applied to its TE; and (3) noise with 100%, 80-Hz SAM applied to its TE. Masked speech testing, for example during speech audiometry (see Katz et al., 2009), is traditionally performed using unmodulated speech-shaped noise. Here, performance was also assessed with modulated noises, because it has been suggested that age-related speech-identification deficits might be exacerbated when the background has such fluctuations (Takahashi and Bacon, 1992; Stuart and Phillips, 1996; Dubno et al., 2002). The choice of the 5-Hz SAM frequency was motivated by the finding of Füllgrabe et al. (2006) that, compared to consonant identification in the presence of an unmodulated noise, the largest improvement in performance was observed for a noise with an SAM frequency between 4 and 16 Hz. This release from masking is believed to reflect the ability to take advantage of the minima in the fluctuating noise to detect speech cues, a phenomenon referred to as “dip listening” (e.g., Cooke, 2006; Füllgrabe et al., 2006). A higher modulation frequency

²We deliberately refrain from using the qualifiers “steady” or “steady-state” that are generally used in the literature to refer to a noise on which no amplitude modulation is impressed. Recent work by Stone et al. (2011, 2012) indicates that the random amplitude fluctuations intrinsic to a notionally steady noise impede speech perception.

³The TE of such a SAM stimulus is given by the equation: $E(t) = 1 + m \sin(2\pi f_m t + \phi)$, where m is the modulation depth, here fixed at 1 (100% modulation), f_m is the modulation frequency, and ϕ is the starting phase of the modulation, which was randomized for each presentation.

was also used to test the hypothesis that older listeners show less masking release when the masker has only short temporal dips, due to decreased temporal resolution (Takahashi and Bacon, 1992; He et al., 2008; Kumar and Sangamanatha, 2011) or increased susceptibility to forward masking (Dubno et al., 2002; Gifford et al., 2007). Noises were ramped on and off using a 50-ms raised-cosine function, and started and ended synchronously with the VCVs. VCVs were presented at 65 dB Sound Pressure Level (SPL), approximating normal conversational speech levels (Olsen, 1998). The noise level was varied in 4-dB steps to give signal-to-noise ratios (SNRs) of -2 to -14 dB for the unmodulated noise, and -6 to -18 dB for the SAM noises. All stimuli were lowpass filtered at 6 kHz (with frequency components above 6.125 kHz attenuated by at least 100 dB) to produce zero audibility above that frequency for both age groups. Stimuli were presented diotically.

Practice was given prior to data collection using four training runs, drawn from each of the four conditions (quiet, unmodulated noise at -2 -dB SNR, 5-Hz SAM noise at -6 -dB SNR, and 80-Hz SAM noise at -6 -dB SNR). Visual feedback and the possibility of repeating a given VCV were provided, and participants were encouraged to use the repeat option whenever necessary. No feedback was provided during the test phase, in which the 13 experimental conditions were presented twice, once in each of two test blocks separated by a break. Each test block started with the speech-in-quiet condition. In the first block, identification was then assessed using the unmodulated, the 5-Hz SAM, and the 80-Hz SAM noise conditions; for each noise type, the SNR conditions were presented in descending order. In the second block, the noise conditions were presented in reverse order to balance possible learning and fatigue effects.

Sentence identification

Target sentences were taken from the main corpus of the Adaptive Sentence Lists (MacLeod and Summerfield, 1990) and comprised 18 lists, plus four “trash” lists (the sentences in these were not as well matched for difficulty as for the other lists). Each list contained 15 sentences. Sentences (mean duration = 1510 ms) had three key words and a simple syntactic structure, and were somewhat predictable (e.g., “*They moved the furniture*”). Sentences were spoken by a male talker with a standard British accent, and presented either in quiet or against two interfering male talkers (one with a British and one with a soft Australian accent), reading from prose passages in a normal conversational manner (Moore et al., 2008). Pauses exceeding 300 ms were truncated “by hand,” and the two interfering talkers were added together at the same root-mean-square (rms) level. To simulate real-world listening conditions (containing reverberation and spatially separate sound sources), the target and interfering speech were played through one of two Tannoy Precision 8D self-powered loudspeakers to a KEMAR head-and-torso manikin in a moderately reverberant lecture theater ($RT_{60} = 0.67, 0.67, 0.54, 0.56, 0.53, 0.53$, and 0.53 s for 1/3-octave-wide bands centered at 0.125, 0.25, 0.5, 1, 2, 4, and 8 kHz, respectively; Moore et al., 2010). The loudspeakers were positioned at $\pm 60^\circ$ relative to KEMAR’s sagittal plane, and at a distance of 1.5 m. Recordings with 16-bit quantization and a 44.1-kHz sampling rate were obtained separately from the two

ears, and then processed off-line (including an inverse diffuse field correction for KEMAR’s meatal response). For the masked conditions, left- and right-ear recordings of the target speech played through one loudspeaker were combined with left- and right-ear recordings of the interfering speech, respectively, when played through the same loudspeaker (giving rise to a co-located percept of the talkers) or through the other loudspeaker (giving rise to spatially separate percepts of the talkers). Target sentences were inserted into randomly selected 3-s excerpts of the interfering-talker mixture. The onset of the target sentences varied randomly from 0 to 500 ms relative to the onset of the interfering speech. The level of the target speech was fixed at 65 dB SPL and the level of the interfering speech was varied in 4-dB steps to give SNRs of -2 to -18 dB. All stimuli were lowpass filtered at 6 kHz (with attenuation of at least 100 dB above 6.125 kHz).

The task was to repeat orally as many words as possible from each target sentence. Response time was unlimited. The trash lists were used to present six practice conditions: quiet, co-located at -2 and -6 dB SNR, and spatially separate at -6 , -10 , and -14 dB SNR. Lists 1–18 from the main corpus were used for the 16 experimental conditions obtained by combining the three factors (Masker location, SNR, and Target position), plus two quiet conditions, one for each target position. The order of presentation of conditions was counterbalanced using a Latin-square design.

SUPRA-THRESHOLD PSYCHOACOUSTIC TASKS

Assessment of sensitivity to TE information

The threshold for detecting SAM imposed on a 4-kHz sinusoidal carrier was measured using a 3-interval, 3-alternative forced-choice procedure with feedback. On each trial, three consecutive 1-s observation intervals were presented, separated by 415-ms silences. One interval, selected at random, contained the SAM tone (the “target”) and the other two intervals (the “standards”) contained the unmodulated carrier. All stimuli had the same rms level. The task was to indicate the interval containing the target.

Modulation frequencies (f_m) of 5, 30, 90, and 180 Hz were used to characterize the temporal-modulation-transfer function (TMTF; Viemeister, 1979), covering the three types of TE-based percepts, namely loudness fluctuations, roughness, and residue pitch (see Figure 2 in Joris et al., 2004). The modulation depth (m) at the start of a run was set to 0.5, 0.6, 0.6, and 0.7 for the four values of f_m , respectively. The value of m was changed adaptively using a 3-down, 1-up stepping rule, estimating the 79%-correct point on the psychometric function (Levitt, 1971). The initial step size was a factor of 1.78, and the step size was reduced to a factor of 1.26 after the first two reversals. After a total of 70 trials, the run was terminated, and the geometric mean of the values of m at the last eight reversals was taken as the threshold estimate.

A 4-kHz carrier was used to ensure that TE and not spectral cues were used to perform the task; spectral sidebands produced by the SAM would not have been resolved even for $f_m = 180$ Hz (Kohlrausch et al., 2000). The level of the carrier was set to 30 dB Sensation Level (SL), based on the participant’s absolute thresholds for a 4-kHz pure tone, measured at the beginning of the test. This limited the spread of the excitation pattern, thus minimizing

“off-frequency listening,” which has been shown to affect TE processing (Füllgrabe et al., 2005).

Prior to data collection, participants received practice in the form of one threshold run for each value of f_m . The test phase consisted of two repeated measures for each f_m , administered first in one order (30, 180, 5, and 90 Hz), and, after a break, in the reverse order.

All stimuli were digitally generated using a PC-controlled Tucker-Davis-Technologies (TDT) system with a 16-bit digital-to-analog converter (DD1, 50-kHz sampling rate), lowpass filtered at 20 kHz (Kemo VBF8, mark 4), attenuated (TDT P4), passed through a headphone buffer (TDT HB6), and delivered diotically at 65 dB SPL.

Assessment of sensitivity to TFS information

The ability to detect changes in the temporal fine structure (TFS) of tones, within the same ear and across ears, was assessed using two tests developed by Moore and colleagues (Moore and Sek, 2009; Hopkins and Moore, 2010). In both tests, a 2-alternative forced-choice procedure with feedback was used. On each trial, two consecutive intervals were presented, separated by 500 ms. Each interval contained four consecutive 400-ms tones, separated by 100 ms. All tones were shaped using a 20-ms raised-cosine function. In one interval, selected at random, the TFS of all tones was identical (the standard). In the other interval (the target), the first and third tones were the same tones as in the standard interval while the second and fourth tones differed in their TFS. Listeners with “normal” TFS sensitivity perceive the change in TFS as a variation either in pitch (in the monaural TFS test; see below) or in lateralization (in the binaural TFS test; see below), and thus can identify the interval containing the changing tone sequence when large TFS differences are used. Initially, the difference in TFS between tones was set to the maximum value possible, without producing ambiguous percepts. The manipulated variable was adaptively adjusted, using a 2-down, 1-up stepping rule to estimate the 71%-correct point on the psychometric function (Levitt, 1971). The value of the manipulated variable was changed by a factor of 1.25³ until the first reversal, then by a factor of 1.25² until the next reversal, and by a factor of 1.25 thereafter. After eight reversals, the run was terminated and the geometric mean of the values at the last six reversals was taken as the threshold estimate. When the SD of the log values at the last six reversals exceeded 0.2, indicating high variability, the estimate was discarded and a new run was conducted. If the adaptive procedure called for values exceeding the maximum more than twice during a run, the adaptive procedure was terminated, and 40 constant-stimuli trials were presented with the value fixed at its maximum. Two valid threshold estimates were obtained for each condition. Practice in the form of at least one threshold run for each of the four monaural and two binaural conditions was provided prior to data collection.

Monaural TFS test. Monaural TFS sensitivity was assessed using the TFS1 test (Moore and Sek, 2009). Participants were asked to discriminate harmonic complex tones, with a fundamental frequency F_0 , from similar tones in which all components were shifted up by the same amount in Hz, resulting in inharmonic complex tones. The frequency shift was the manipulated variable,

and it was initially set to $0.5F_0$. The tones had the same envelope repetition rate, but different TFS. The starting phases of the components in each tone were random, resulting in random differences in the shape of the TE of the complex tones and preventing TE shape from being used as a cue. All tones were passed through a bandpass filter (with a bandwidth of $1F_0$ and slopes of 30 dB/octave), centered on $11F_0$. Since the auditory system does not resolve harmonics above the 8th (Plomp, 1964; Moore and Ohgushi, 1993), all components in the passband were unresolved and, consequently, differences in excitation pattern for the two tones were minimal (Hopkins and Moore, 2007). Two F_0 s, 91 and 182 Hz, were used, corresponding to filter center frequencies of 1 and 2 kHz, respectively. To mask combination tones and components falling on the skirts of the bandpass filter, threshold equalizing noise (TEN; Moore et al., 2000) was presented. Its level/ERB_N (Moore, 2012) was 15 dB below the overall level of the tones. The TEN was gated on and off with 20-ms raised-cosine ramps and started 300 ms before the first tone in the first interval and ended 300 ms after the last tone in the second interval. The overall level of the complex tones was set to 30 dB SL, based on absolute-threshold measurements for pure tones at the two filter center frequencies. Each ear was tested separately.

Binaural TFS test. Binaural TFS sensitivity was assessed using the TFS-LF test (Hopkins and Moore, 2010). Participants were asked to discriminate binaurally presented pure tones with identical phases at the two ears (perceived as emanating from a central position inside the head) from tones with a phase shift between the ears (perceived as being lateralized toward one ear). The interaural phase shift ($\Delta\phi$) was the manipulated variable. Two frequencies, 0.5 and 0.75 kHz, were used. For both, $\Delta\phi$ was initially set to 180°. All tones were gated on and off synchronously in the two ears to avoid the use of interaural differences in TE to perform the task. The level of the tones was set to 50 dB SPL in each ear.

COGNITIVE TASKS

With the advent of cognitive hearing science (e.g., Arlinger et al., 2009; Rönnberg et al., 2011) new interest has been sparked concerning the role of cognition in normal and pathological speech perception. An increasing number of studies have included some form of cognitive assessment. However, the number and diversity of cognitive tests have generally been small, and their choices have not always been explicitly motivated. In keeping with past attempts to assess more systematically the relationship between cognitive functioning and speech perception (e.g., Van Rooij et al., 1989; Van Rooij and Plomp, 1990; Jerger et al., 1991; Humes et al., 1994), a large number of cognitive abilities was investigated in the present study. The reasons for selecting the specific cognitive tasks in terms of their relationship with age and speech processing are discussed in the Supplementary Material: *Relationship between cognitive-task performance, age, and speech intelligibility*.

Digit Span test

The Digit Span (DS) test, taken from the Wechsler Adult Intelligence Scale—Third Edition (WAIS-IIIUK; Wechsler, 1997), is assumed to assess short-term-memory (STM) capacity (i.e., the

temporary storage of information) and working-memory (WM) capacity (i.e., storage plus processing of information), using the Digits Forward (DS-F) and Digits Backward (DS-B) tests, respectively. In the former, digit sequences of increasing length (from 2 to 9 digits) are presented verbally at one digit per second for immediate verbal recall. Two trials for each sequence length are presented. The task is discontinued after two incorrect answers for a given sequence length. The final DS-F score used here corresponded to the sum of recalled digits for all entirely correctly reported sequences; the maximum total score was 88. In the DS-B test, digit sequences of increasing length (containing 2 to 8 digits) are presented, but the digits have to be recalled in reverse order (i.e., from last to first). The final DS-B score was computed in the same way as the DS-F score; the maximum total score was 70. An initial practice trial was given for each test.

Reading Span test

The Reading Span (RS) test, originally developed by Daneman and Carpenter (1980), is one implementation of a complex span test (Conway et al., 2005), designed to assess the key properties of the limited-capacity working-memory (WM) system, namely memory storage and information processing (Baddeley, 1992). Here, a computerized version (Rönnberg et al., 1989) of the RS test of Baddeley et al. (1985) was used, in which short, grammatically correct sentences were displayed in a word-by-word fashion on a computer screen (e.g., “The ball—bounced—away”) at a rate of one word every 800 ms. A 1750-ms silent interval separated the end of one sentence from the beginning of the next sentence. Half of the sentences were sensible while the others were absurd (e.g., “The pear—drove—the bus”). Sentences were arranged in three sets of three, four, five, and six sentences, and presented in order of increasing length. All sets were administered, irrespective of the participant’s performance. The task was to read aloud each sentence and then to indicate by a verbal “yes/no” response if the sentence made sense or not (processing component of WM). At the end of each set, the participant was instructed to recall in correct serial order either the first or the last word of each sentence (storage component of WM). The requested recall position (first or last) varied pseudo-randomly (with first-word recalls in half of the sets) but was identical for all participants. Prior to testing, practice was given in the form of one three-sentence set, which was repeated if necessary until the instructions were clearly understood. To assess whether participants traded performance on the semantic-judgment task in favor of the recall task in an age-dependent manner, the number of errors on the semantic-judgment task was analyzed: there was no significant difference between the two age groups [$t_{(27)} = -0.528$, $p = 0.602$; two-tailed]. Following others (Lunner, 2003; Sörqvist and Rönnberg, 2012), the percentage of first and last words correctly recalled in any order out of the total number of words to be recalled (i.e., 54) was taken as an indicator of WM capacity.

Test of Everyday Attention

The Test of Everyday Attention (TEA; Robertson et al., 1994) is a neuropsychological test designed to assess the integrity of different, functionally independent attentional systems (Posner and Petersen, 1990). Using principal-component analysis and

cross-validating with established tests of attention, Robertson et al. (1996) identified four putative cognitive processes probed by the following eight sub-tests of the TEA:

The *Map Search* and *Telephone Search* tests require the participant to visually search, as quickly as possible, for predetermined symbols, either on a map or in a telephone directory. The number of identified symbols and the time per symbol are recorded in the first and second tests, respectively. Performance on both tests indexes selective attention. In the *Elevator Counting* and *Lottery* tests, participants count the number of tones in sequences of varying length, and monitor a 10-min recording of lottery ticket numbers for winning numbers, respectively. Performance on these tests, and the *Telephone Search while Counting* test (in which participants perform the Telephone Search test while simultaneously counting the number of tones in a sequence), assesses sustained attention. In the *Visual Elevator* test, the participant counts the number of visual symbols in an ascending and descending order, following visual instructions. Time to completion for correct trials assesses attentional switching. The auditory analog of this test is the *Elevator Counting with Reversals* test, using tones of different pitches as items to be counted and also as instructions to count up or down. Performance on this test and on the *Elevator Counting with Distraction* test (which requires counting the number of tones in a sequence while ignoring interleaved distractor tones of a different frequency) indexes audio-verbal WM. Practice was provided for all sub-tests prior to testing according to the TEA instructions.

Trail Making test

The two parts of the paper-and-pencil version of the neuropsychological Trail Making (TM) test (Reitan, 1955) were administered, following the protocol described by Bowie and Harvey (2006). In Part A, 25 encircled Arabic numerals (1–25), randomly distributed on a white sheet of paper, had to be connected in ascending order. The participants were instructed to complete the task as quickly and as accurately as possible, and time to completion was recorded. It is generally assumed that this part assesses psycho-motor speed and visual search (e.g., Crowe, 1998). In Part B, 12 Arabic numerals (1–12) and 12 letters (A–L) had to be connected in ascending order, alternating between numerals and letters (A-1-B-2-C-3, etc.). Keeping two mental sets in memory and switching between them requires additional executive control (Arbuthnott and Frank, 2000). Prior to test administration, participants completed shorter practice versions of each part.

Derived measures, for example the difference between completion times for the two parts (Part B—Part A; e.g., Sanchez-Cubillo et al., 2009) or the ratio of the two completion times (Part B/Part A; e.g., Lamberty et al., 1994) have been used to provide a “purer” estimate of executive control abilities. However, as pointed out by Verhaeghen and De Meersman (1998), age-group differences in difference scores are still confounded by age-related deficits in processing speed. Hence, we computed the normalized derived measure [(Part B—Part A)/Part A] to assess executive control.

Block Design test

Block Design (BD) constitutes a standard measure of performance IQ in many test batteries of intelligence. It is assumed

to measure spatial perception, visual abstract processing, and problem solving. Here, we used the BD version from the WASI (Wechsler, 1999), in which participants had to manually rearrange four or nine two-color blocks to replicate 13 target “designs” displayed on a series of test cards and presented in order of increasing difficulty. The two easiest designs were used as practice. Time to completion for each design was measured and transformed to a point score (from 0 to 7); designs completed after predefined cutoff times were scored as zero. The maximum total score was 71.

Matrix Reasoning test

Matrix Reasoning (MR) is another standard test for measuring non-verbal abstract reasoning. Here, we used the version taken from the WASI (Wechsler, 1999), comprising 35 items, organized in order of increasing difficulty. Each item was composed of a matrix of geometric patterns with one element missing. The task was to choose from five response alternatives the one that best completes the matrix. The two easiest designs were used as practice. No time limit was imposed. The maximum total score was 35.

RESULTS AND DISCUSSION

Age-group differences in sensitivity and performance were assessed, using independent-samples *t*-tests, and, in cases of the simultaneous manipulation of within-subjects factors, mixed-design repeated-measures analyses of variance (ANOVAs) with Age group as the between-subjects factor. To assess the strength of association between the various measures of supra-threshold auditory processing, cognitive abilities, and speech identification, Pearson product-moment correlation coefficients were computed for the entire group of participants (see table of all correlations in the Supplementary Material: *Grand correlation matrix for the combined group of young and older normal-hearing participants*), for the ONH participants alone, and for all participants with the effect of age partialled out. Finally, multiple regression analyses were conducted to quantify the relative contribution of different processing abilities to consonant and sentence identification.

SPEECH TASKS

Several authors (e.g., Dubno and Ahlstrom, 1997; Demeester, 2011) have highlighted the possibility that changes in the audiogram of a few dB may be associated with changes in speech perception in noise. This motivated the matching of audiograms for the two age groups used in the present study. The results presented in this section will mainly be compared to those for studies using fairly stringent definitions of normal audiograms (e.g., thresholds ≤ 25 dB HL over a wide range of frequencies) and using lowpass-filtered target speech to restrict the spectrum of the stimuli to the frequency range where audiometric thresholds were normal, or to studies where age groups were audiometrically matched.

Individual identification scores were transformed into rationalized arcsine units (RAUs, Studebaker, 1985) for statistical analyses. To ease interpretation, the averaged transformed data were transformed back to percentages for the presentation of the results in the figures and the text.

Consonant identification

Intelligibility. Group-mean consonant-identification scores are plotted in **Figure 3** for the YNH (open symbols) and ONH (filled symbols) participants for speech in quiet (left-most symbols) and in the three noise types (different panels) as a function of SNR.

Consonant identification in quiet was near-perfect for both age groups, but, consistent with previous results (Gelfand et al., 1985; Gordon-Salant, 1986), the ONH participants made slightly but significantly more confusions than the YNH participants [group difference of 1.7% points; $t_{(28)} = -2.051$, $p = 0.05$; two-tailed].

Consistent with Gelfand et al. (1986), the addition of background noise resulted in a larger decrease in identification scores for the ONH than for the YNH participants. However, the effect of decreasing SNR was similar for the two groups. Contrary to the assumption that the effect of age is greater for temporally fluctuating backgrounds than for unmodulated backgrounds (Takahashi and Bacon, 1992; Stuart and Phillips, 1996; Dubno et al., 2002), the three background noises yielded age-associated decrements of similar sizes: age differences averaged across SNRs were 9.7, 11.6, and 10.2% points for the unmodulated, 5-Hz SAM, and 80-Hz SAM noise, respectively. Given the use of different SNR ranges, a separate ANOVA was conducted for each noise type. In all three cases, there were significant effects of Age group [$F_{(1, 28)} = 16.027$, 21.865, and 23.413, respectively, all $p < 0.001$] and SNR [$F_{(3, 28)} = 603.483$, 68.312, and 391.025, respectively, all $p < 0.001$], but the Age group*SNR interaction was not significant [$F_{(3, 28)} = 1.002$, 0.270, and 1.533, respectively, $p = 0.396$, 0.847, and 0.212, respectively]. Scores for the different noise types were correlated moderately to strongly across all participants (all $r \geq 0.612$, all $p < 0.001$), indicating that participants tended to perform consistently poorly or well.

Modulation masking release. The benefit derived from modulation of the background noise was computed as the difference in identification scores obtained at a given SNR with either of the two SAM noises and the unmodulated noise, and will be referred to as modulation masking release (MMR). The mean MMR values for 5-Hz and 80-Hz SAM are shown in **Figure 4** for the two age groups.

Consistent with previous results for YNH participants (Füllgrabe et al., 2006), all participants showed more MMR for the lower than the higher SAM frequency. For the 5-Hz SAM noise, the amount of MMR increased monotonically with decreasing SNR, while it remained roughly constant for the 80-Hz SAM noise. Across all participants, identification performance for the unmodulated noise was negatively correlated with the amount of MMR ($r = -0.464$, all $p = 0.010$); participants who had low scores for unmodulated noise tended to show high MMR. However, this could be due to the fact that performance for the unmodulated noise enters into both quantities that were correlated.

Age-group differences in MMR were very small, barely exceeding 3% points. The main effects of SAM frequency [$F_{(1, 28)} = 61.810$, $p < 0.001$] and SNR [$F_{(1.629, 45.613)} = 28.347$, $p < 0.001$]

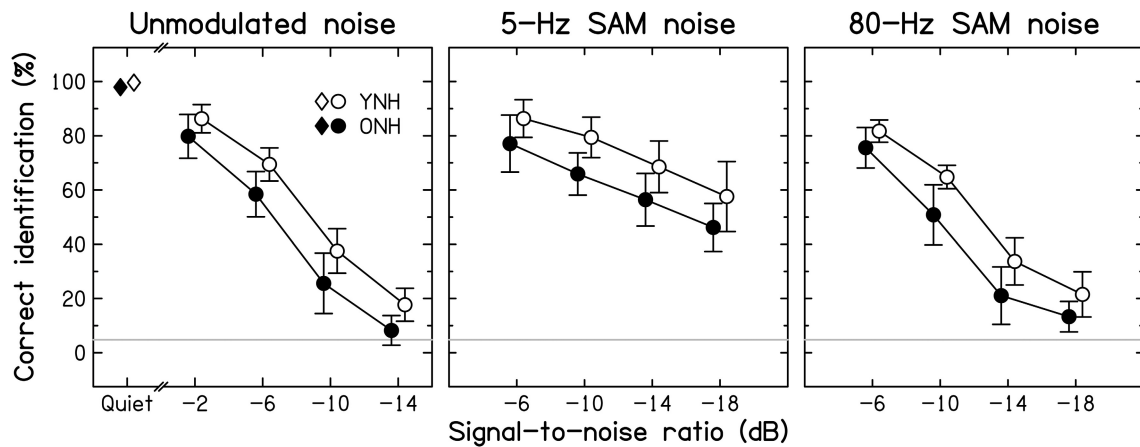


FIGURE 3 | Average consonant-identification performance in different listening conditions for YNH (open symbols) and ONH (filled symbols) participants. Identification scores are given for the quiet condition (diamonds) and as a function of the signal-to-noise ratio (SNR) for the unmodulated, 5-Hz SAM, and

80-Hz SAM noise conditions (left, middle, and right panels, respectively). Here, and in the following figures, data points for the two groups are slightly displaced horizontally to aid visibility. Chance-level performance is indicated by the gray horizontal lines. Error bars represent ± 1 SD.

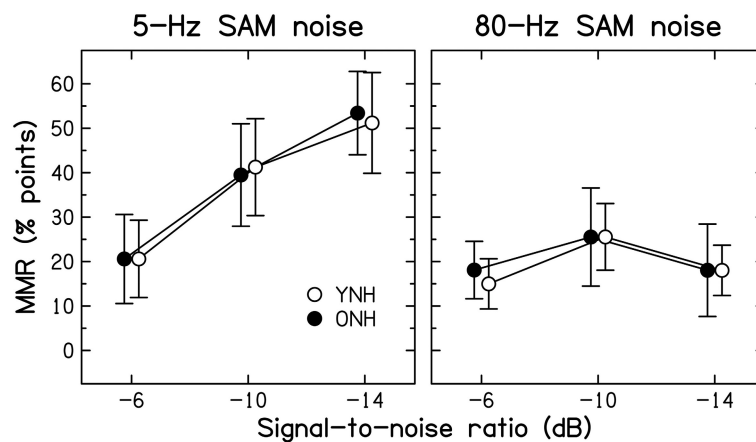


FIGURE 4 | Average amount of modulation masking release (MMR, in percentage points) for YNH (open symbols) and ONH (filled symbols) participants. MMR is the difference in scores obtained using an SAM noise [SAM frequency = 5 Hz (left panel) or 80 Hz (right panel)] and an unmodulated noise.

were significant, as was the interaction between these two factors [$F_{(1.836, 51.398)} = 18.363$, $p < 0.001$]. However, neither the main effect of Age group [$F_{(1, 28)} = 0.001$, $p = 0.972$] nor any of the two- or three-way interactions involving this factor were significant (all $F < 1$, all $p \geq 0.607$). These results indicate that the ability to “listen in the dips” does not decrease with increasing age, at least when peripheral hearing sensitivity is normal and matched across age groups. Some earlier investigations (Dubno et al., 2002, 2003; Grose et al., 2009) reported age-group differences in MMR. However, the older participants in those studies had higher audiometric thresholds than the younger participants, especially in the high-frequency range, and the bandwidth of the speech signals was not limited to the audiometrically normal range, which might explain the discrepancy between the present and previous results.

Sentence identification

Intelligibility. Figure 5 presents group-mean scores for speech in quiet and in two-talker babble presented from the same spatial location as the target speech (“co-located”) or from a different spatial location (“separate”). The position of the target speech (localized toward the left or the right) was counterbalanced across conditions. However, a paired-samples t -test for the quiet conditions [$t_{(29)} = -1.000$, $p = 0.326$; two-tailed] and separate ANOVAs for the co-located [$F_{(1, 28)} < 1$, $p = 0.957$] and separate [$F_{(1, 28)} = 3.609$, $p = 0.068$] conditions revealed no significant effect of this factor. Hence, scores were pooled across the two target positions.

Unmasked speech identification was at ceiling and almost the same for the two age groups [$t_{(28)} = -0.648$, $p = 0.522$; two-tailed]. For the co-located condition, the ONH participants

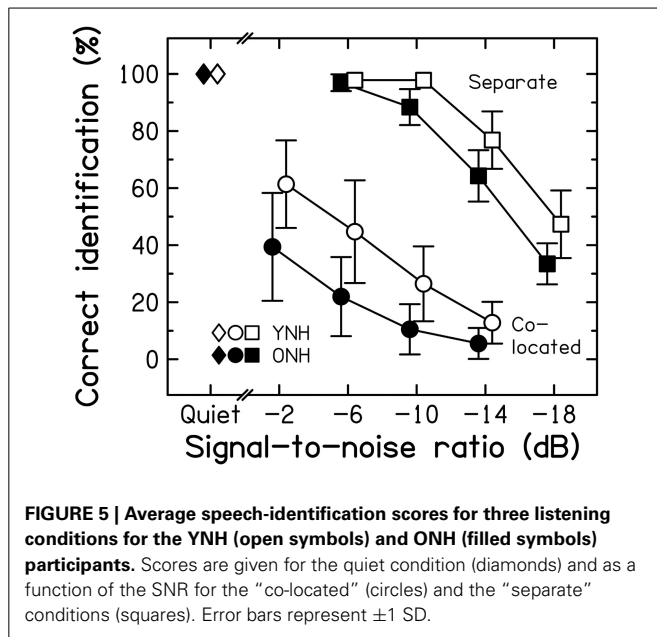


FIGURE 5 | Average speech-identification scores for three listening conditions for the YNH (open symbols) and ONH (filled symbols) participants. Scores are given for the quiet condition (diamonds) and as a function of the SNR for the “co-located” (circles) and the “separate” conditions (squares). Error bars represent ± 1 SD.

performed more poorly than the YNH participants over the entire range of SNRs. The age-group difference was 22% points at the highest SNR, and dropped to 7% points at the lowest SNR, most likely due to a floor effect. The main effects of SNR [$F_{(3, 84)} = 97.151, p < 0.001$] and Age group [$F_{(1, 28)} = 17.154, p < 0.001$] were significant but the SNR*Age group interaction was not [$F_{(3, 84)} < 1, p = 0.433$]. Performance was better for the separate than for the co-located condition for both groups. Scores were lower for the ONH than for the YNH participants for the three lowest SNRs; at the most favorable SNR, a ceiling effect was most likely responsible for the very similar scores for the two groups. There were significant main effects of SNR [$F_{(3, 84)} = 592.247, p < 0.001$] and Age group [$F_{(1, 28)} = 19.200, p < 0.001$] and a significant interaction [$F_{(3, 84)} = 7.594, p < 0.001$]. Performance in the two masking conditions was correlated strongly across all participants ($r = 0.753, p < 0.001$).

Spatial masking release. The improvement in speech identification produced by a difference in the spatial locations of target and masker signals (compared to the co-located case) is referred to as spatial masking release (SMR; e.g., Freyman et al., 1999). We quantified SMR by calculating the difference in scores for the separate and co-located conditions for the SNR that did not yield a floor effect for the former and a ceiling effect for the latter. For an SNR of -10-dB, the SMR values for the YNH and ONH participants were 84.6 and 86% points, respectively. This difference across age groups was not significant [$t_{(28)} = 0.369, p = 0.715$; two-tailed]. Thus, consistent with earlier investigations (Gelfand et al., 1988; Li et al., 2004; Singh et al., 2008; Cameron et al., 2011), these results provide no evidence to support the idea that the ability to use spatial separation between target and interfering speech declines with age when the audiogram is normal. This is surprising given that ONH participants have been shown to be less sensitive than YNH participants to inter-aural time differences (ITDs; Ross et al., 2007; Grose and Mamo, 2010; Füllgrabe,

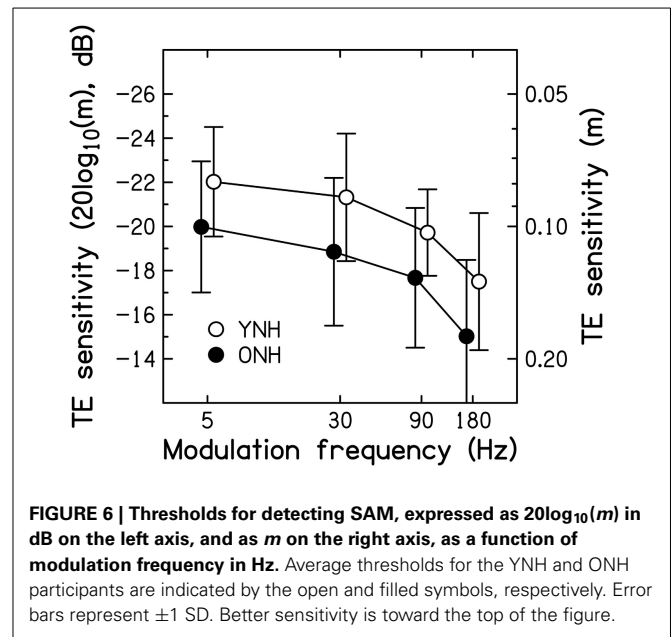


FIGURE 6 | Thresholds for detecting SAM, expressed as $20\log_{10}(m)$ in dB on the left axis, and as m on the right axis, as a function of modulation frequency in Hz. Average thresholds for the YNH and ONH participants are indicated by the open and filled symbols, respectively. Error bars represent ± 1 SD. Better sensitivity is toward the top of the figure.

2013; see also Section *Assessment of sensitivity to TFS information*). However, the potency of ITD cues in the physiological range in inducing sequential stream segregation does not seem to be affected in those listeners (Füllgrabe and Moore, 2014), and the listening conditions used here afforded additional cues (e.g., monaural spectral cues and interaural intensity differences) that contributed to SMR (Singh et al., 2008). The processing of these cues seems to be relatively unaffected by aging (Herman et al., 1977; Babkoff et al., 2002).

SUPRA-THRESHOLD PSYCHOACOUSTIC TASKS

Assessment of sensitivity to TE information

Mean SAM detection thresholds for the two age groups⁴ are shown as a function of modulation frequency in Figure 6. To ease comparison with previously published results, the modulation depth at threshold (m , right axis) is expressed in dB, as $20\log_{10}(m)$, on the left axis. The TMTFs for both groups are similar in shape to those reported in previous studies for pure-tone carriers (Kohlrausch et al., 2000; Füllgrabe and Lorenzi, 2003).

On average, thresholds were 2–2.5 dB higher (worse) for the ONH than for the YNH participants. The effects of SAM frequency [$F_{(2.353, 61.190)} = 20.132, p < 0.001$] and Age group [$F_{(1, 26)} = 4.208, p = 0.050$] were significant, but the interaction was not [$F_{(2.353, 61.190)} < 1, p < 0.946$]. These results are generally consistent with previous studies reporting significant age-related decrements in the TMTF by 2–7 dB, as measured using pure-tone (He et al., 2008) and noise carriers (Takahashi and Bacon, 1992; Kumar and Sangamanatha, 2011), although in those studies the older participants had higher audiometric thresholds than the younger participants. Also, those studies showed the largest decrements for higher modulation frequencies, whereas here the decrement was independent of modulation frequency,

⁴Only seven YNH participants completed the SAM detection task.

suggestive of a deficit in processing efficiency and not temporal resolution (e.g., Hill et al., 2004). In other words, as for very young participants (i.e., normal-hearing children aged 4–7 years; Hall and Grose, 1994), the peripheral encoding of TE information seems young-adult-like but the processing of this information is less efficient.

Assessment of sensitivity to TFS information

To allow comparison of results from the adaptive and constant-stimulus procedures, thresholds (in Hz for the monaural test, and in degrees for the binaural test) and percent-correct scores were transformed into the value of the sensitivity index d' that would be obtained for the largest possible value of the manipulated variable, that is $0.5F_0$ for the TFS1 test and 180° for the TFS-LF test (for further details, see Hopkins and Moore, 2007, 2010). The d' values obtained in this way were sometimes very large. The utility of this conversion is that scores for both the adaptive and constant-stimulus procedures are transformed into a single scale, and values on this scale increase monotonically with improving performance. Each ear was tested separately in the TFS1 test. However, paired t -tests revealed no significant differences between the d' values for the left and right ears [$t_{(28)} = -0.044$, $p = 0.965$ and $t_{(28)} = -0.747$, $p = 0.461$ for the filter center frequencies of 1 and 2 kHz, respectively; both two-tailed]. Hence, results were pooled across the two ears for further analysis and presentation. Average d' values for the two age groups⁵ in the two monaural and the two binaural test conditions are shown in Figure 7.

Monaural TFS d' values for YNH participants were in good agreement with published data but d' values for the binaural TFS test were considerably higher (better) than previously observed (Moore and Sek, 2009; Hopkins and Moore, 2010, 2011; Moore et al., 2012b), possibly due to more protracted practice in the present study, to the longer tone duration, or to the longer interval between the two sets of four stimuli in each trial that was used here. Mean d' scores were higher for YNH than ONH participants. According to independent-samples t -tests, the differences between the age groups were significant [1 kHz: $t_{(27)} = -2.427$, $p = 0.011$; 2 kHz: $t_{(8,256)} = -2.971$, $p = 0.009$ ⁶; 0.5 kHz: $t_{(27)} = -3.306$, $p = 0.002$; 0.75 kHz: $t_{(27)} = -3.703$, $p < 0.001$; all one-tailed] and remained so after applying a Holm-Bonferroni correction for multiple comparisons. This confirms previous evidence for an age-related TFS processing deficit for smaller and/or audiometrically normal but unmatched participant groups (Hopkins and Moore, 2011; Moore et al., 2012b; Füllgrabe, 2013).

When measured at the same frequency (0.5, 1, or 2 kHz), audiometric thresholds (for each ear or averaged across the two ears) and TFS d' values (for each ear or for binaural processing) were not significantly correlated (r between 0.064 and -0.321 ; all $p \geq 0.090$, uncorrected). Hence, TFS sensitivity for our normal-hearing participants was not associated with absolute

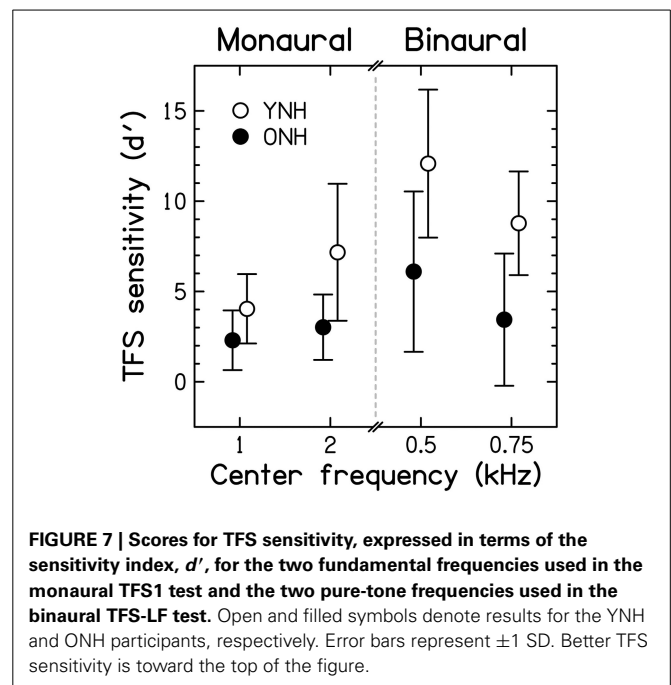


FIGURE 7 | Scores for TFS sensitivity, expressed in terms of the sensitivity index, d' , for the two fundamental frequencies used in the monaural TFS1 test and the two pure-tone frequencies used in the binaural TFS-LF test. Open and filled symbols denote results for the YNH and ONH participants, respectively. Error bars represent ± 1 SD. Better TFS sensitivity is toward the top of the figure.

threshold at the test frequency. Results from previous studies using audiometrically unmatched young and older (Hopkins and Moore, 2011; Moore et al., 2012b) or older participants with a range of ages (Moore et al., 2012a) generally agree with the present finding for binaural TFS sensitivity, but showed significant correlations between absolute threshold and monaural TFS sensitivity.

Surprisingly, the correlation between d' values for the two center frequencies used for the TFS1 test failed to reach significance ($r = 0.322$, $p = 0.088$), perhaps because individual differences in TFS sensitivity were relatively small at 1 kHz, or because TFS sensitivity might show idiosyncratic variations across frequency, even for audiometrically normal ears. However, d' values for the two center frequencies used for the TFS-LF test were highly correlated ($r = 0.763$; $p < 0.001$), and the correlation remained significant after partialling out the effect of age ($r_{\text{age}} = 0.663$; $p < 0.001$). The d' value averaged over the two frequencies of the TFS1 test was moderately correlated with the d' value averaged over the two frequencies of the TFS-LF test ($r = 0.541$, $p = 0.002$), but the correlation became non-significant after partialling out the effect of age ($r = 0.251$, $p = 0.197$). This is consistent with previous suggestions that the TFS1 and TFS-LF tests tap partially different abilities (Hopkins and Moore, 2011; Moore et al., 2012b), perhaps because the latter involves additional binaural processing occurring in the brainstem.

COGNITIVE TASKS

To facilitate comparison across cognitive tests and with findings of previous cognitive-aging studies (e.g., Park et al., 2002; Salthouse, 2009), the data were transformed into z -scores, using the mean and the SD of the entire group (YNH and ONH combined), prior to statistical analyses. Reaction-time data were multiplied

⁵Results from one YNH participant were not included due to incomplete data.

⁶Levene's test indicated inequality of variance for the two groups. Hence here (and whenever else applicable), corrected degrees of freedom and t -values were used.

by -1 after being transformed into z-scores so that better performance was represented by higher z-scores across all tests. Group means and SDs for the seven tests (plus the derived measure for the TM test) are shown in **Figure 8** for the YNH (open symbols) and ONH (filled symbols) participants. Performance for the TEA was computed as the average of the unit-weighted z-scores for the eight sub-tests. For each cognitive measure, the effect size, expressed as Cohen's d ⁷, is given at the bottom of the panel. Gray and black panel frames denote non-significant ($p > 0.05$) and significant ($p \leq 0.05$) group differences, respectively. Bold panel frames indicate differences that remain significant after applying a Holm-Bonferroni correction.

For all tests, mean scores were higher for the YNH than for the ONH participants, but the effect size varied from small ($d \sim 0.2$) for the two DS tests, to large ($d > \sim 0.8$) for the remaining tests. Performance on each of the two DS tests and the derived measure for the TM test did not differ significantly for the two age groups

(all $p \geq 0.461$; two-tailed; uncorrected), but all other tests showed significant effects of age group (all $p \leq 0.011$; two-tailed; uncorrected) which remained significant after correcting for multiple comparisons. The group means of the raw scores for the eight cognitive measures and the results of independent-samples t -tests are given in the Supplementary Material: *Raw scores and statistical results for cognitive measures*.

It is often assumed that the DS-B and RS require both information storage and processing, while the DS-F involves only information storage. However, performance on the RS test, but not on the two DS tests, was significantly affected by age, suggesting that the "re-ordering task" (DS-B) is more closely related to STM tests (such as DS-F) than to complex WM tests (for a discussion of this point, see Bopp and Verhaeghen, 2005). This interpretation is supported by a significant correlation between scores for the two DS tests ($r = 0.622$, $p < 0.001$; two-tailed) but non-significant correlations between scores for either of these tests and RS scores (both $r \leq 0.271$, both $p \geq 0.155$; two-tailed).

Figure 9 gives the scores for each of the eight sub-tests of the TEA, grouped by the putatively assessed attentional process

⁷Due to the unequal size of the two age groups, Cohen's d was calculated using the square root of the pooled variance rather than the mean variance (Howell, 2002).

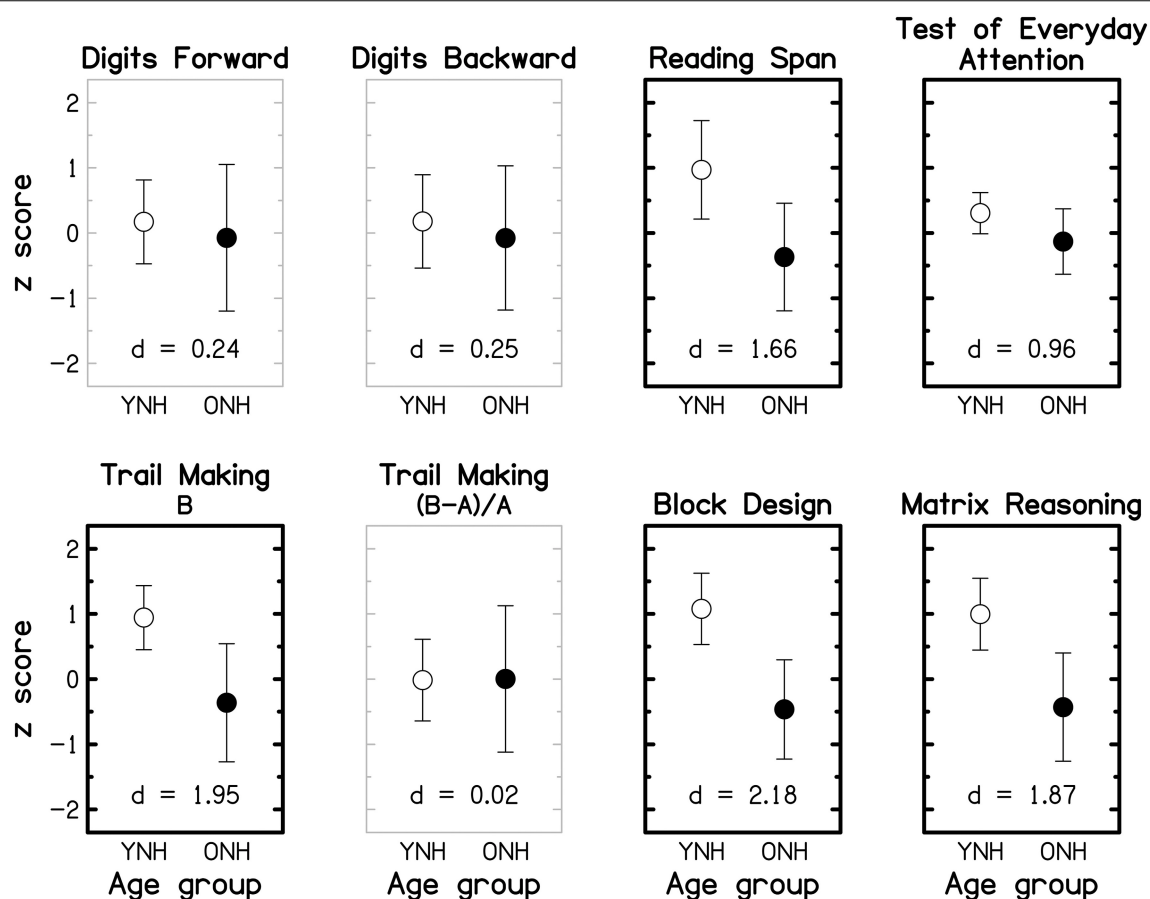


FIGURE 8 | Group-mean performance (in z-scores) for YNH (open symbols) and ONH (filled symbols) participants on different cognitive tasks. Error bars represent ± 1 SD. Gray panel frames indicate non-significant

group differences ($p > 0.05$). Bold black panel frames denote significant results at $p \leq 0.05$ that remained significant after applying a Holm-Bonferroni correction. The effect size is given by Cohen's d at the bottom of each panel.

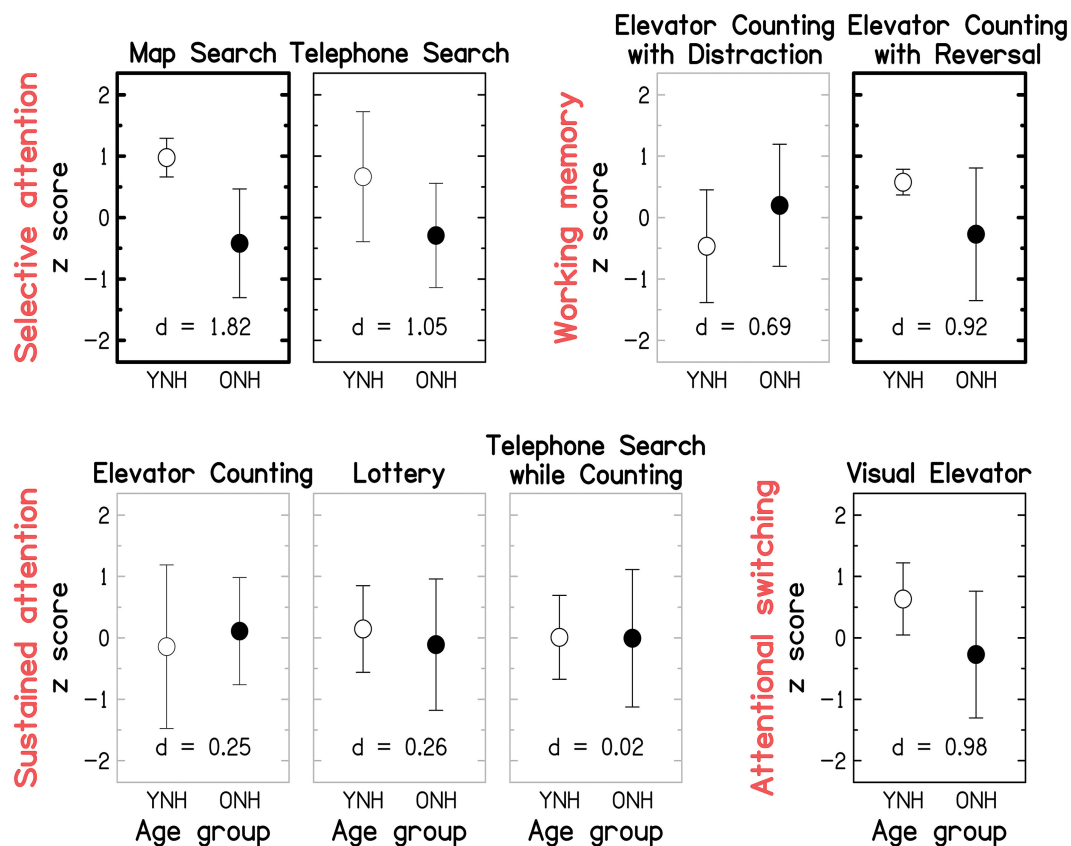


FIGURE 9 | Group-mean performance (in z-scores) for YNH (open symbols) and ONH (filled symbols) participants on each of the eight sub-tests of the Test of Everyday Attention. Sub-tests are grouped by the underlying attentional processes (see red labels) they are assumed to assess according to Robertson et al. (1996): Selective attention (Map Search, Telephone Search), Audio-verbal working memory (Elevator Counting with

Distraction, Elevator Counting with Reversal), Sustained attention (Elevator Counting, Lottery, Telephone Search while Counting), and Attentional switching (Visual Elevator). Gray panel frames indicate non-significant group differences ($p > 0.05$). Black panel frames denote significant results at $p \leq 0.05$. Bold panel frames indicate significant results after applying a Holm-Bonferroni correction. Otherwise as **Figure 8**.

identified by (Robertson et al., 1996); note that subsequent factor analyses only partially confirmed these groupings (Chan, 2000; Bate et al., 2001). Effect size and statistical significance are indicated for each sub-test, as for **Figure 8**. The raw mean scores and statistical results are given in the Supplementary Material: *Raw scores and statistical results for cognitive measures*. The pattern of results is broadly consistent with the nomenclature (see red labels in **Figure 9**) suggested by Robertson et al. (1996): large and significant age effects (all $p \leq 0.021$; two-tailed; uncorrected) were observed for both of the selective-attention tests, one of the WM tests, and the attentional-switching test, although the effects became non-significant for the Telephone Search and Visual Elevator tests after correction for multiple comparisons. All three tests of sustained attention yielded small and non-significant age effects (all $p \geq 0.521$).

Given that the TEA was designed as a neuropsychological screening tool, it is not surprising that ceiling effects were observed for some of the sub-tests (Robertson et al., 1996). In our “healthy” sample, most participants performed perfectly on the Elevator Counting test and many YNH participants scored close to ceiling on the Map Search test. At least for the latter test,

administering version B of the test might overcome this problem in the future; indeed, a group of 31 YNH participants tested on that version as part of an unrelated study yielded a lower mean score of 65.3/80 (compared to 75.4/80 in the current study).

CORRELATION AND REGRESSION ANALYSES

The strength of the association between supra-threshold auditory processing, various cognitive abilities, and SiN identification was evaluated by conducting correlation and regression analyses. However, the analysis of data from “extreme” age groups using these statistical tools can be problematic (Hofer et al., 2003). As demonstrated in the previous section, TE and TFS sensitivity, cognitive processing, and speech perception were all generally poorer for the ONH than for the YNH group. Even if no association between psychoacoustic, cognitive, and speech measures existed within each age group, use of the combined scores across all participants could reveal a significant relationship between the measures. To avoid this pitfall, we followed the example of Grassi and Borella (2013), and computed correlations not only across all participants, but also restricting the analyses to the data for the ONH group, and also after partialling out the effect of age.

Relationship between auditory temporal processing and speech perception

To reduce the effect of errors of measurement, masked speech-identification scores for each participant were averaged across the different SNRs and masker types, to give a single composite score for consonants and a single composite score for sentences. Similarly, a composite score for TE sensitivity was obtained by averaging detection thresholds for the four modulation frequencies, and a composite score for TFS sensitivity was obtained by averaging d' values across the two TFS1 and the two TFS-LF conditions.

Figure 10 shows individual composite consonant and sentence identification scores for the YNH and ONH participants plotted against the composite measures of TE sensitivity (left column) and TFS sensitivity (right column). In each panel, significant correlation coefficients are given (r for the entire group, r_{ONH} for the ONH group only, and $r_{\text{-age}}$ for the entire group when age was partialled out). The boldness of the font increases with increasing significance (from $p \leq 0.05$ to 0.001). For the entire group,

speech scores were strongly and significantly associated with TFS sensitivity, and were somewhat more weakly associated with TE sensitivity. When only the ONH group was considered, or participant age (alone or together with composite cognition; see Section *Relationship between cognitive abilities and speech perception*) was partialled out, the strength of the correlation between TFS sensitivity and performance on both speech tasks was somewhat reduced but remained significant. In contrast, TE sensitivity was no longer significantly associated with sentence identification and its correlation with consonant identification, while still significant at $p \leq 0.05$, was only moderate.

Given the small number of YNH participants in this study no detailed correlational analysis for this group is presented. However, it is noteworthy that for our YNH sample TFS sensitivity was correlated strongly with sentence identification in noise ($r = 0.839$, $p = 0.009$). Neher et al. (2011) did not find a correlation between TFS sensitivity and a measure of speech perception for a similarly sized “youngish” normal-hearing group. However, they only assessed the relationship for binaural TFS sensitivity

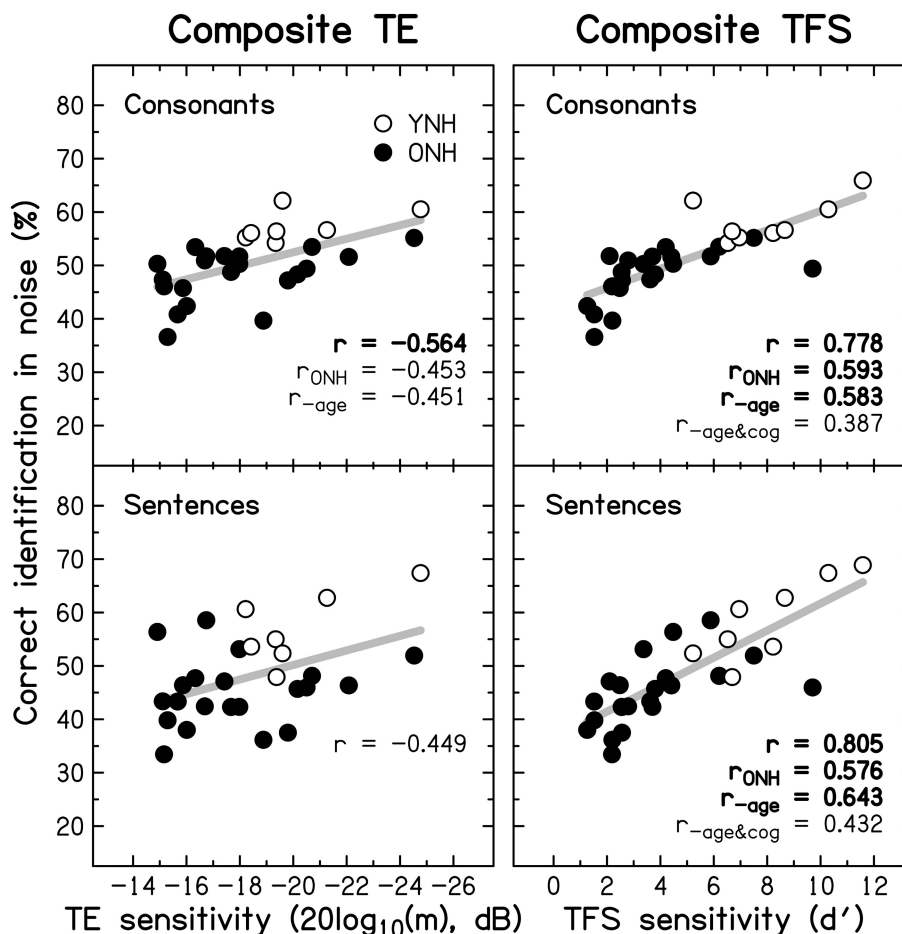


FIGURE 10 | Scatter plots of composite sensitivity to TE (left panel) and TFS (right panel) vs. composite consonant (top row) and sentence (bottom row) identification in noise. The thick gray line represents the best linear fit to the data from the entire group composed of YNH (open symbols) and ONH participants (filled

symbols). Significant (at $p \leq 0.05$; uncorrected) correlation coefficients for all participants (r), for the ONH participants only (r_{ONH}), and for all participants with age ($r_{\text{-age}}$) or with age and composite cognition ($r_{\text{-age\&cog}}$) partialled out, are given in each panel. Bold font indicates significance at $p \leq 0.001$.

and for target speech presented at a different azimuth from the speech maskers.

Based on the evidence that sensorineural hearing loss is associated with a reduced ability to process TFS information (e.g., Buss et al., 2004; Lacher-Fougère and Demany, 2005; Santurette and Dau, 2007; for an overview, see Moore, 2014), some authors (Lorenzi and Moore, 2008; Moore, 2008; Hopkins and Moore, 2009) have suggested that the large speech-perception deficit experienced by hearing-impaired listeners in the presence of modulated noise could be a consequence of their inability to use TFS information to take advantage of the minima in the noise. Similarly, it is often assumed that dip listening requires a certain degree of temporal resolution (Festen, 1993; Stuart and Phillips, 1996; Füllgrabe et al., 2006; George et al., 2006; Grose et al., 2009). To test the role of TE and TFS sensitivity in MMR, and its dependence on age, a composite measure of MMR was calculated for the consonant-identification task, by averaging individual scores across the different SNRs and two SAM frequencies. The scatter plots in **Figure 11** indicate that MMR was not significantly associated with the composite measures of TE sensitivity (left panel; $r = 0.280$, $p = 0.148$) or TFS sensitivity (middle panel; $r = -0.233$, $p = 0.224$). Also, MMR was not significantly correlated with composite sentence identification in the presence of co-located speech interference (right panel; $r = -0.204$, $p = 0.279$).

Relationship between cognitive abilities and speech perception

The association between cognitive measures and identification of consonants and sentences in noise was assessed for the ONH group and for the entire group after partialling out the effect of age (see **Table 1**). Correlation coefficients significant at $p < 0.05$ are shown in black.

Somewhat similar patterns of results were observed for the two correlational analyses for the two speech tasks. Considering only those results that remained significant after applying a Holm-Bonferroni correction (values in boldface), mainly scores for DS-F, DS-B, TM-B, and BD were correlated consistently with speech

identification scores. Performance on the RS test was not significantly associated with speech perception (even though there was a significant moderate and positive correlation when young and older participants were considered together, consistent with results reported by Besser et al., 2012). This finding contrasts with a growing body of evidence that WM capacity, as measured by the RS test, is correlated with speech perception for hearing-impaired listeners (e.g., Rudner et al., 2011), and does not support the notion that “WM capacity also seems to play an important role when people with normal hearing must understand language spoken in acoustically adverse conditions” (Rönnberg et al., 2013). A survey of previous studies administering the RS test and a measure of SiN perception to YNH participants revealed a mixed pattern of results: while Moradi et al. (2014) reported a significant (but uncorrected) moderate positive correlation between scores on the two tasks, others either found significant results only for a sub-set of the tested SiN conditions (Zekveld et al., 2011; Besser et al., 2012; Kilman et al., 2014) which, contrary to predictions, did not always include the most adverse conditions, or failed to find any evidence for a relationship between WM capacity and SiN performance (Zekveld et al., 2014). The significant correlations (but uncorrected for multiple comparisons) found in studies including adults from a wider age range with no or only partial audiometric confirmation of normal hearing were possibly confounded by age-related changes in audibility, supra-threshold auditory processing, and/or cognition (Besser et al., 2012; Ellis and Munro, 2013). Consistent with this, Besser et al. (2012) reported that the moderate correlation between performance on the RS and speech test was no longer significant after partialling out the effect of age. In summary, it is currently unclear if individual differences in WM capacity in the audiometrically normal-hearing young or older population are the main contributor to the observed variability in SiN perception. Further studies are warranted to explicitly address this issue, including the questions: (1) is the RS test the most appropriate measure of WM (Besser et al., 2012; Sörqvist and Rönnberg, 2012); (2) which of the sub-processes of WM does the RS test probe

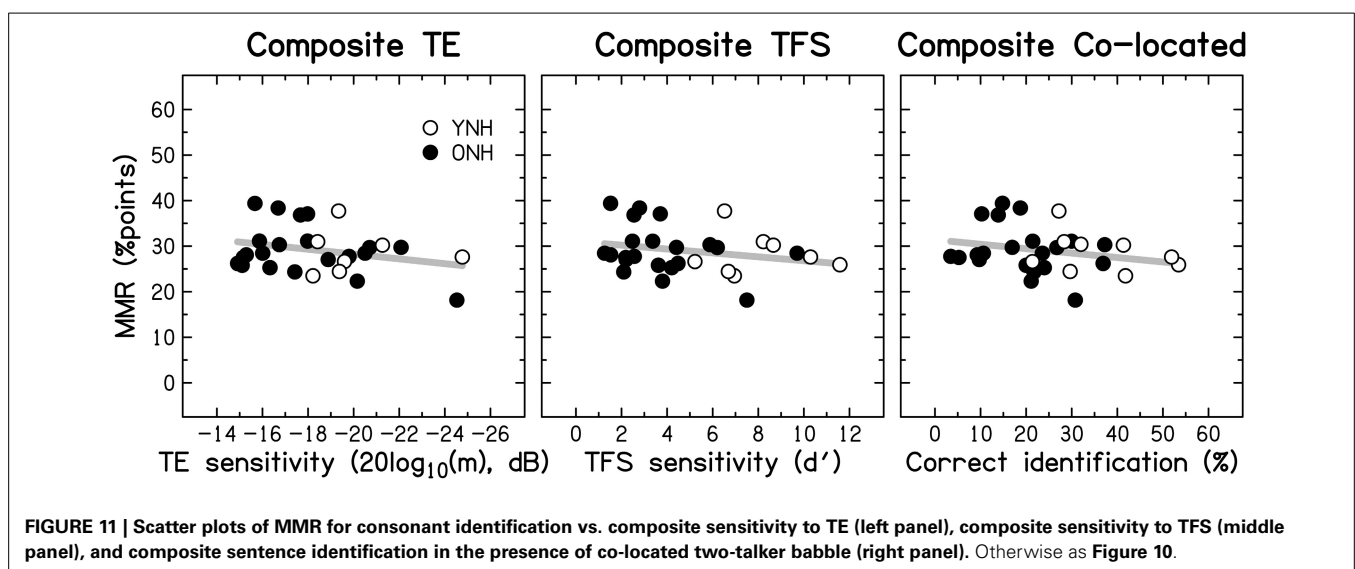


Table 1 | (A) Pearson product-moment correlation coefficients for results on eight cognitive measures vs. consonant- (first and third result columns) and sentence-identification performance in noise (second and fourth result columns). Results for the ONH group only and for the entire group after partialling out participant age are given in result columns 1–2 and 3–4, respectively. Gray values indicate non-significant correlations ($p > 0.05$). Values in black indicate significant results at $p \leq 0.05$. Values in boldface indicate significant results after applying a Holm-Bonferroni correction. (B) Correlation coefficients for performance on the eight sub-tests of the TEA vs. speech-identification performance in noise. Otherwise as (A).

	ONH		Age partialled out	
	Consonants	Sentences	Consonants	Sentences
(A) COGNITIVE MEASURES				
Digits Forward	0.546	0.761	0.516	0.654
Digits Backward	0.260	0.597	0.276	0.446
Reading Span	0.037	0.174	0.127	0.266
Test of Everyday Attention	0.577	0.525	0.477	0.408
Trail Making B	0.386	0.643	0.401	0.607
Trail Making [(B-A)/A]	0.072	−0.199	−0.008	−0.295
Block Design	0.671	0.612	0.673	0.558
Matrix Reasoning	0.222	0.361	0.204	0.245
(B) TEA SUB-TESTS				
Map Search	0.431	0.368	0.387	0.360
Telephone Search	0.532	0.024	0.408	0.017
Elevator Counting with Distraction	0.559	0.479	0.308	0.311
Elevator Counting with Reversal	0.532	0.483	0.454	0.387
Elevator Counting	−0.108	0.107	−0.086	0.082
Lottery	0.241	−0.096	0.293	−0.033
Telephone Search while Counting	−0.023	0.422	−0.043	0.285
Visual Elevator	0.195	0.298	0.100	0.155

(Unsworth and Engle, 2007; Sörqvist et al., 2010); and (3) what constitutes an acoustically adverse condition?

In an attempt to characterize the relationship between general cognitive functioning and speech perception, a composite cognition score was computed by averaging the unit-weighted z -scores (Salthouse, 1991; Lindenberger et al., 2001) from all eight cognitive measures, independently of whether or not they were associated with speech perception. Such an all-inclusive approach was meant to avoid “cherry picking” the cognitive tests yielding the strongest correlations with speech perception. Figure 12 shows the scatter plots of scores for identification of consonants and sentences in noise against the composite cognition scores.

Considering the entire participant group, scores for both speech tasks were strongly associated with cognition. Limiting the analysis to the ONH group, or partialling out the effect of age, reduced the strength of the correlation but it remained moderate (for consonant identification) to strong (for sentence identification). All analyses yielded larger correlation coefficients and

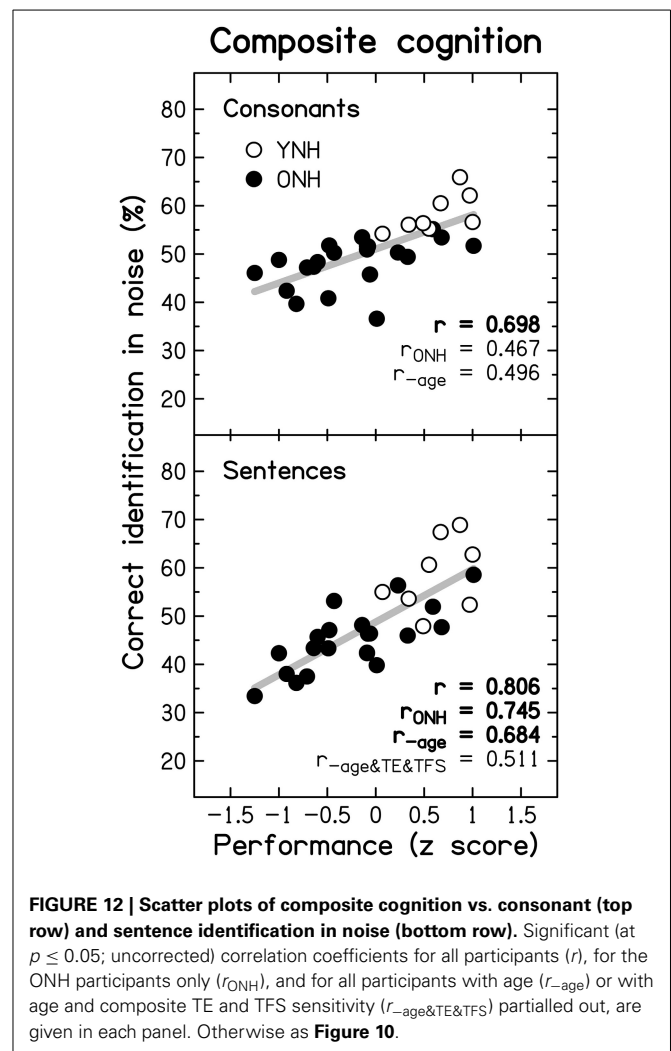


FIGURE 12 | Scatter plots of composite cognition vs. consonant (top row) and sentence identification in noise (bottom row). Significant (at $p \leq 0.05$; uncorrected) correlation coefficients for all participants (r), for the ONH participants only (r_{ONH}), and for all participants with age (r_{age}) or with age and composite TE and TFS sensitivity ($r_{\text{age\&TE\&TFS}}$) partialled out, are given in each panel. Otherwise as Figure 10.

smaller p values for sentence than for consonant identification. Interestingly, the correlation between cognition and sentence identification was still moderate and significant ($p = 0.009$) after partialling out the effects of age, composite TE sensitivity, and composite TFS sensitivity.

Despite the link of cognition with consonant identification, neither form of release from masking was associated with cognitive functioning: MMR, $r = -0.198$, $p = 0.303$; SMR, $r = -0.125$, $p = 0.519$. In other words, the ability to benefit from temporal dips in a masker or spatial separation between a target and masker was not related to cognition.

Relationship between cognitive abilities and temporal auditory processing

Since a link between cognitive abilities (especially WM capacity) and temporal processing has recently been suggested (Troche and Rammsayer, 2009; Broadway and Engle, 2011), correlations were computed between composite sensitivity scores for TE and TFS and the eight cognitive measures (see Table 2A), and the eight sub-tests of the TEA (see Table 2B). No evidence was found that performance on the RS test, assumed to index WM capacity, was

linked to temporal processing abilities. However, for several other cognitive tests (DS-F, BD, TEA) there were significant positive moderate correlations, mainly with TFS sensitivity, even when the effect of age was partialled out. Amongst the TEA sub-tests, scores for one selective-attention test (Map Search) and the two WM tests (Elevator Counting with Distraction and Elevator Counting with Reversal) were significantly correlated with TFS sensitivity. The relationship between cognition and TFS processing might occur because some level of cognitive ability is required to perform well on the TFS tests. Alternatively, or in addition, it may occur because both are linked to the integrity and precision of neural processing.

Multiple regression analysis

To explore the relative importance of the factors contributing to the variance in consonant and sentence identification, multiple regression analyses (using the stepwise method) were carried out separately for the two speech tasks, using composite scores for TE sensitivity, TFS sensitivity, and cognition as predictor variables.

For consonant identification, the most parsimonious significant model that emerged was based on the single predictor TFS sensitivity [$F_{(1, 26)} = 28.826$, $p < 0.001$]. The model explained 50.8% of the variance. The standardized regression coefficient for the TFS variable was 0.725 ($p < 0.001$).

Table 2 | (A) Pearson product-moment correlation coefficients for results on eight cognitive measures vs. composite TE (first and third results columns) and composite TFS sensitivity (second and forth result columns). Otherwise as Table 1. (B) Correlation coefficients for performance on the eight sub-tests of the TEA vs. composite temporal sensitivity. Otherwise as (A).

	ONH		Age partialled out	
	TE	TFS	TE	TFS
(A) COGNITIVE MEASURES				
Digits Forward	-0.136	0.596	-0.147	0.493
Digits Backward	-0.163	0.476	-0.153	0.397
Reading Span	0.004	0.008	-0.088	0.045
Test of Everyday Attention	-0.517	0.735	-0.440	0.579
Trail Making B	-0.263	0.588	-0.258	0.541
Trail Making [(B-A)/A]	-0.139	-0.058	-0.106	-0.104
Block Design	-0.286*	0.639	-0.288	0.615
Matrix Reasoning	-0.070	0.431	-0.034	0.327
(B) TEA SUB-TESTS				
Map Search	-0.515	0.582	-0.476	0.563
Telephone Search	-0.522	0.328	-0.432	0.316
Elevator Counting with Distraction	-0.298	0.535	-0.213	0.414
Elevator Counting with Reversal	-0.459	0.639	-0.376	0.564
Elevator Counting	-0.105	0.100	-0.127	-0.068
Lottery	0.156	0.041	0.037	0.044
Telephone Search while Counting	-0.090	0.215	0.028	0.078
Visual Elevator	-0.352	0.535	-0.199	0.371

For sentence identification, the significant model that explained the most (68%) of the variance (adjusted $R^2 = 0.68$) was based on Cognition and TFS sensitivity [$F_{(2, 25)} = 29.679$, $p < 0.001$]. The standardized regression coefficients were 0.509 ($p = 0.004$) for cognition and 0.392 ($p = 0.020$) for TFS sensitivity.

SUMMARY AND GENERAL DISCUSSION

Increasing acknowledgement of speech identification and comprehension problems among older people (for a review, see Gordon-Salant et al., 2010), combined with awareness of the increasing proportion of older people in most Western countries (e.g., Christensen et al., 2009), has spawned a considerable number of studies investigating the age-related auditory and cognitive changes that underlie speech perception problems. In the absence of gross cognitive dysfunction, elevated audiometric thresholds in older listeners have been identified as the major contributor to the reduction in speech intelligibility (Van Rooij and Plomp, 1992; Humes, 1996, 2007; Dubno et al., 2008). However, audibility generally did not explain all of the variance in identification performance.

STUDY AIMS, FINDINGS, AND IMPLICATIONS

The aims of the present study were to confirm the existence of SiN identification difficulties in the older population and to investigate the nature and relative importance of the associated age-related changes in supra-threshold auditory and cognitive processing. The following steps were taken to control for the roles of audibility and cochlear status in SiN identification: (1) only young and older participants with normal audiograms over a wide frequency range (0.125–6 kHz) were included; (2) the average audiograms of the two groups were matched; and (3) high-frequency information from the signals was removed to ensure zero audibility of frequency components above 6 kHz for both age groups. Two speech tasks were administered to capture different levels of processing and complexity: consonant identification in unmodulated or modulated speech-shaped noise under “dry” conditions, and sentence identification in spatially co-located or separate speech maskers under reverberant conditions. To clarify the factors contributing to individual and age-group performance on these tasks, participants were also characterized in terms of: (1) their sensitivity to TE and TFS cues, which are known to be important for speech intelligibility and auditory scene analysis; and (2) their performance on a battery of cognitive tests probing cognitive abilities such as memory, attention, and processing speed. The main findings are summarized below:

1. Identification scores for consonants in quiet and in speech-shaped noise were lower for ONH than for YNH participants. The size of this effect was independent of masker type (unmodulated vs. modulated), masker modulation frequency (5 vs. 80 Hz), and SNR.
2. Modulation masking release did not differ for the two age groups.
3. Identification scores for sentences in quiet were identical for the two age groups, but identification of sentences in the presence of spatially co-located or separate interfering two-talker babble was worse for the ONH group. The size of this

age effect was similar across SNRs within the same masker-location condition for SNRs at which performance was not affected by a floor or ceiling effect.

4. Spatial masking release did not differ for the two age groups.
5. The lower speech-perception performance of the ONH participants was not associated with higher subjective ratings of hearing disabilities.
6. Sensitivity to TE information was reduced for ONH participants; modulation detection thresholds were higher by 2–2.5 dB across all modulation frequencies used ($f_m = 5$ –180 Hz).
7. Sensitivity to monaural and binaural TFS information was reduced for ONH participants.
8. Performance on most, but not all, cognitive tests was worse for the ONH than for the YNH participants. Cognitive abilities spared by aging included short-term memory and sustained attention.
9. Composite sensitivity to temporal information was correlated positively and moderately (for TE sensitivity) or strongly (for TFS sensitivity) with consonant and/or sentence identification in noise. After partialling out the effects of age and composite cognition, a moderate correlation with TFS sensitivity remained, and was significant.
10. Composite sensitivity to TE or TFS was not correlated significantly with modulation masking release. Neither was sentence identification in co-located two-talker babble.
11. Performance on some cognitive tests was correlated positively and moderately to strongly with TFS sensitivity. Correlations were moderate, but remained significant, after partialling out the effect of age.
12. The measure of composite cognition was correlated positively and strongly with consonant and sentence identification in noise. After partialling out the effects of age, TE sensitivity, and TFS sensitivity, a moderate correlation with sentence identification remained, and was significant.
13. Composite cognition and sensitivity to TFS explained 68% of the variance in sentence-in-speech identification. Composite TFS sensitivity explained 51% of the variance in consonant-in-noise identification.

Most of the age-group differences in auditory and cognitive processing and their associations with intelligibility in noise were statistically significant. However, despite these deficits in test performance, the ONH participants did not report more hearing disabilities than the YNH participants on either of the two questionnaires used (for possible explanations, see the Supplementary Material: *Discrepancy between measured and self-assessed hearing difficulties*). Thus, the practical significance of these age effects for speech processing in everyday life remains to be determined. It is likely that age-related deficits are even more pronounced in more variable listening conditions (for effects of stimulus variability, see Sommers, 1997; Golomb et al., 2007) and when the speech material has greater syntactic complexity (Wingfield et al., 2006) than used here, even when compensatory mechanisms (Wingfield and Grossman, 2006; Reuter-Lorenz and Cappell, 2008) and changes in cognitive strategies (Lemaire, 2010) might help to offset the deleterious effects of variability and complexity (e.g., through the

enhanced use of contextual knowledge; Pichora-Fuller et al., 1995; Wingfield, 1996; Dubno et al., 2000; Pichora-Fuller, 2008). Also, the impact of aging on speech communication might manifest itself in other ways than in a decrement in identification performance (e.g., changes in conversational discourse pragmatics; Kiessling et al., 2003; McKellin et al., 2007).

In a recent review of the literature on the topic of age-related central factors in presbycusis, Humes et al. (2012) concluded that, given the lack of control of confounding variables such as hearing sensitivity and cognition in most studies, there is insufficient evidence to support a “pure” form of central presbycusis (i.e., age-related central auditory decline in the absence of peripheral and/or cognitive changes). Our data showed a significant negative correlation between age and composite TFS (but not TE) sensitivity in our audiometrically and performance-IQ-matched normal-hearing participants, even after partialling out the effect of composite cognition ($r_{\text{cog}} = -0.450$, $p = 0.016$; two-tailed). Also, the correlation between age and the identification of meaningless VCVs in noise (a condition in which participants mainly had to rely on acoustic cues) was significant, even after partialling out the effect of composite cognition ($r_{\text{cog}} = -0.475$, $p = 0.011$; two-tailed)⁸. These findings support the idea of a form of presbycusis that is not confounded by age-related cognitive changes and that is not related to changes in the cochlea that lead to elevated audiometric thresholds. The presbycusis could be a result of neural changes anywhere from the auditory nerve up to higher centers in the auditory system (e.g., Sergeyenko et al., 2013).

Despite an increasing body of evidence showing age deficits affecting various aspects and levels of auditory and cognitive processing, many experimental studies investigating the perceptual consequences of hearing loss on auditory perception did not use age-matched experimental groups, but compared *young* normal-hearing to *older* hearing-impaired participants (for an example from previous work by the authors, see the Supplementary Material: *Confounding age effect in a study of hearing loss*). Consequently, it is likely that many published results overestimate the effects of hearing loss as measured by the audiogram, and need to be “corrected” for the effect of age. A similar word of caution applies to many studies of aging, in which audibility across age groups was generally not matched and often only loosely controlled for, in spite of clear evidence that differences in audiometric thresholds can result in differences in speech identification (Humes, 1996; Dubno and Ahlstrom, 1997; Demeester, 2011).

FUTURE DIRECTIONS

The use of the audiogram as the main clinical measure of hearing status, and the use of the diagnostic term presbycusis (literally “elderly hearing”) to refer to “age-related hearing loss” both reflect the common assumption that the speech processing difficulties of older persons are mainly related to, and are predictable from, their audiogram. However, the data reported

⁸In these correlational analyses, the use of extreme groups (young vs. old) might have increased the power of the significance test but this is considered acceptable for the purpose of proof-of-concept studies (Preacher et al., 2005).

here confirm and extend evidence that has been accumulating over several decades, showing that, even when the audiogram is normal, deficits in central-auditory and cognitive processing are ubiquitous in the older population and are associated with poorer speech identification. This highlights the need to expand audiological assessment beyond tests of pure-tone audibility (Kricos, 2006) and to devise effective rehabilitative interventions for speech-perception difficulties in older listeners that target not only peripheral dysfunction through frequency-specific amplification *via* hearing aids but also age-related changes in central auditory and cognitive functions through the provision of auditory-based perceptual training programs (Dubno, 2013; Ferguson et al., 2014), targeted cognitive-strategy and cognitive-process training regimens (Lustig et al., 2009; Park and Bischof, 2013), and general cognitive enrichment (Hertzog et al., 2008).

The use of an extreme-group cross-sectional approach in most aging studies precludes the possibility of determining the time of onset of changes within the adult auditory and cognitive systems. Several recent studies have used an intermediate age group to examine whether age-related deficits are already present in midlife (e.g., Ross et al., 2007; Schvartz et al., 2008; Humes et al., 2013a; Helfer and Freyman, 2014). More accurate estimates of when the first signs of aging in auditory and cognitive performance become apparent can be derived from cross-sectional studies sampling continuously across the entire adult life span (Bergman et al., 1976; Baltes and Lindenberger, 1997; Park et al., 2002; Salthouse, 2009; Füllgrabe, 2013) or longitudinal studies (Dubno et al., 2008; Payne et al., 2014). It is also those studies that will inform us about the shape of the trajectory of the decline throughout adulthood.

CONCLUSIONS

Taken together, the results show that, even in the absence of hearing loss as measured by the audiogram, SiN identification declines with age. Both consonant and sentence identification were poorer for the older participants, but possibly not for the same reasons. For both speech tasks, sensitivity to TFS information, which is thought to facilitate the parsing of auditory scenes into sound sources, was more important than sensitivity to TE information. When the target speech consisted of meaningful utterances presented against a background of interfering speech, identification performance was best predicted by cognitive abilities and, to a lesser extent, sensitivity to TFS information. Neither MMR nor SMR differed across age groups. These findings indicate a need for clinical tests in addition to the audiogram when assessing the hearing of older people, and confirm the need to take age into account in studies examining the effects of hearing loss.

ACKNOWLEDGMENTS

We thank Thomas Baer for his continuous assistance with various aspects of the study, Oliver Zobay for statistical advice, three reviewers for helpful comments, and the Cambridge University of the 3rd Age. During the project, CF was a visiting scholar in the Department of Psychology, University of Cambridge (UK). The Institute of Hearing Research is supported by the Medical Research Council (grant number U135097130). The project was partly funded by a Flexi-grant ("Effect of aging on auditory temporal processing") from Action on Hearing Loss (UK) to

Christian Füllgrabe and Brian C. J. Moore. Portions of this study were presented at the annual conference of the British Society of Audiology in 2011 (Füllgrabe et al., 2012).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnagi.2014.00347/abstract>

REFERENCES

- Agrawal, Y., Platz, E. A., and Niparko, J. K. (2008). Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the national health and nutrition examination survey, 1999–2004. *Arch. Intern. Med.* 168, 1522–1530. doi: 10.1001/archinte.168.14.1522
- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142
- Arbuthnott, K., and Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J. Clin. Exp. Neuropsychol.* 22, 518–528. doi: 10.1076/1380-3395(200008)22:4;1-0;FT518
- Arlinger, S. (2003). Negative consequences of uncorrected hearing loss—a review. *Int. J. Audiol.* 42(Suppl. 2), 2S17–2S20. doi: 10.3109/14992020309074639
- Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x
- Babkoff, H., Muchnik, C., Ben-David, N., Furst, M., Even-Zohar, S., and Hildesheimer, M. (2002). Mapping lateralization of click trains in younger and older populations. *Hear. Res.* 165, 117–127. doi: 10.1016/S0378-5955(02)00292-7
- Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359
- Baddeley, A., Logie, R., Nimmo-Smith, I., and Brereton, N. (1985). Components of fluent reading. *J. Mem. Lang.* 24, 119–131. doi: 10.1016/0749-596X(85)90019-1
- Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12
- Bate, A. J., Mathias, J. L., and Crawford, J. R. (2001). Performance on the test of everyday attention and standard tests of attention following severe traumatic brain injury. *Clin. Neuropsychol.* 15, 405–422. doi: 10.1076/clin.15.3.405.10279
- Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., and Margulies, M. K. (1976). Age-related decrement in hearing for speech. Sampling and longitudinal studies. *J. Gerontol.* 31, 533–538. doi: 10.1093/geronj/31.5.533
- Besser, J., Zekveld, A. A., Kramer, S. E., Rönnberg, J., and Festen, J. M. (2012). New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J. Speech Lang. Hear. Res.* 55, 194–209. doi: 10.1044/1092-4388(2011/11-0008)
- Bopp, K. L., and Verhaeghen, P. (2005). Aging and verbal memory span: a meta-analysis. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, P223–P233. doi: 10.1093/geronb/60.5.P223
- Bowie, C. R., and Harvey, P. D. (2006). Administration and interpretation of the trail making test. *Nat. Protoc.* 1, 2277–2281. doi: 10.1038/nprot.2006.390
- Broadway, J. M., and Engle, R. W. (2011). Individual differences in working memory capacity and temporal discrimination. *PLoS ONE* 6:e25422. doi: 10.1371/journal.pone.0025422
- BSA. (2004). *Pure Tone Air and Bone Conduction Threshold Audiometry With and Without Masking and Determination of Uncomfortable Loudness Levels*. Reading, UK: British Society of Audiology.
- Bunch, C. C. (1929). Age variations in auditory acuity. *Arch. Otolaryngol.* 9, 625–636. doi: 10.1001/archotol.1929.00620030657005
- Buss, E., Hall, J. W. III, and Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear Hear.* 25, 242–250. doi: 10.1097/01.AUD.0000130796.73809.09
- Cameron, S., Glyde, H., and Dillon, H. (2011). Listening in Spatialized Noise-Sentences test (LiSN-S): normative and retest reliability data for adolescents

- and adults up to 60 years of age. *J. Am. Acad. Audiol.* 22, 697–709. doi: 10.3766/jaaa.22.10.7
- CHABA. (1988). Speech understanding and aging. *J. Acoust. Soc. Am.* 83, 859–895. doi: 10.1121/1.395965
- Chan, R. C. (2000). Attentional deficits in patients with closed head injury: a further study to the discriminative validity of the test of everyday attention. *Brain Inj.* 14, 227–236. doi: 10.1080/026990500120709
- Christensen, K., Doblhammer, G., Rau, R., and Vaupel, J. W. (2009). Ageing populations: the challenges ahead. *Lancet* 374, 1196–1208. doi: 10.1016/S0140-6736(09)61460-4
- Clinard, C. G., Tremblay, K. L., and Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hear. Res.* 264, 48–55. doi: 10.1016/j.heares.2009.11.010
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/BF03196772
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *J. Acoust. Soc. Am.* 119, 1562–1573. doi: 10.1121/1.2166600
- Corso, J. F. (1963). Age and sex differences in pure-tone thresholds - Survey of hearing levels from 18 to 65 years. *Arch. Otolaryngol.* 77, 385–405. doi: 10.1001/archotol.1963.00750010399008
- Cox, R. M., and Alexander, G. C. (1995). The abbreviated profile of hearing-aid benefit. *Ear Hear.* 16, 176–186. doi: 10.1097/00003446-199504000-00005
- Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *J. Clin. Psychol.* 54, 585–591.
- Cruikshanks, K. J., Wiley, T. S., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., et al. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. *Am. J. Epidemiol.* 148, 879–886. doi: 10.1093/oxfordjournals.aje.a009713
- Crum, R. M., Anthony, J. C., Bassett, S. S., and Folstein, M. F. (1993). Population-based norms for the Mini-Mental-State-Examination by age and educational-level. *JAMA* 269, 2386–2391. doi: 10.1001/jama.1993.03500180078038
- Dalton, D. S., Cruikshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., and Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *Gerontologist* 43, 661–668. doi: 10.1093/geront/43.5.661
- Daneman, M., and Carpenter, P. A. (1980). Individual-differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Davis, A. (1995). *Hearing in Adults*. London: Whurr.
- Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., and Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish mental survey. *Intelligence* 28, 49–55. doi: 10.1016/S0160-2896(99)00031-8
- Delk, J. H., Glorig, A., Quiggle, R. R., and Summerfield, A. B. (1957). Predicting hearing loss for speech from pure tone audiograms. *Laryngoscope* 67, 1–15.
- Demeester, K. (2011). *The Phenotypic Variance of Presbycusis*. Ph.D. thesis, Universiteit Antwerpen, Antwerp.
- Drullman, R., Festen, J. M., and Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* 95, 2670–2680. doi: 10.1121/1.409836
- Drullman, R., Festen, J. M., and Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* 95, 1053–1064. doi: 10.1121/1.408467
- Dubno, J. R. (2013). Benefits of auditory training for aided listening by older adults. *Am. J. Audiol.* 22, 335–338. doi: 10.1044/1059-0889(2013)12-0080
- Dubno, J. R., and Ahlstrom, J. B. (1997). “Additivity of multiple maskers of speech,” in *Modeling Sensorineural Hearing Loss*, ed W. Jesteadt (Hillsdale NJ: Lawrence Erlbaum Associates), 253–272.
- Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (2000). Use of context by young and aged adults with normal hearing. *J. Acoust. Soc. Am.* 107, 538–546. doi: 10.1121/1.428322
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 111, 2897–2907. doi: 10.1121/1.1480421
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 113, 2084–2094. doi: 10.1121/1.1555611
- Dubno, J. R., Lee, F. S., Matthews, L. J., Ahlstrom, J. B., Horwitz, A. R., and Mills, J. H. (2008). Longitudinal changes in speech recognition in older persons. *J. Acoust. Soc. Am.* 123, 462–475. doi: 10.1121/1.2817362
- Dudley, H. (1939). Remaking speech. *J. Acoust. Soc. Am.* 11, 169–177. doi: 10.1121/1.1916020
- Durlach, N. I., Mason, C. R., Kidd, G. Jr., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. G. (2003). Note on informational masking. *J. Acoust. Soc. Am.* 113, 2984–2987. doi: 10.1121/1.1570435
- Eddins, D. A. (1993). Amplitude modulation detection of narrow-band noise: effects of absolute bandwidth and frequency region. *J. Acoust. Soc. Am.* 93, 470–479. doi: 10.1121/1.405627
- Ellis, R. J., and Munro, K. J. (2013). Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int. J. Audiol.* 52, 14–22. doi: 10.3109/14992027.2012.721013
- Eysenck, M. W., and Keane, M. T. (2000). *Cognitive Psychology: A Student's Handbook*. Hove: Psychology Press.
- Ferguson, M. A., Henshaw, H., Clark, D. P., and Moore, D. R. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74-year-olds with mild hearing loss. *Ear Hear.* 35, e110–e121. doi: 10.1097/AUD.0000000000000020
- Festen, J. M. (1993). Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice. *J. Acoust. Soc. Am.* 94, 1295–1300. doi: 10.1121/1.408156
- Foley, D. M., Frick, K. D., and Lin, F. R. (2014). Association between hearing loss and healthcare expenditures in older adults. *J. Am. Geriatr. Soc.* 62, 1188–1189. doi: 10.1111/jgs.12864
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *J. Acoust. Soc. Am.* 106, 3578–3588. doi: 10.1121/1.428211
- Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hear. Res.* 158, 1–27. doi: 10.1016/S0378-5955(01)00296-9
- Füllgrabe, C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 22, 313–315. doi: 10.1044/1059-0889(2013)12-0070
- Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.* 211, 74–84. doi: 10.1016/j.heares.2005.09.001
- Füllgrabe, C., and Lorenzi, C. (2003). The role of envelope beat cues in the detection and discrimination of second-order amplitude modulation. *J. Acoust. Soc. Am.* 113, 49–52. doi: 10.1121/1.1523383
- Füllgrabe, C., Meyer, B., and Lorenzi, C. (2003). Effect of cochlear damage on the detection of complex temporal envelopes. *Hear. Res.* 178, 35–43. doi: 10.1016/S0378-5955(03)00027-3
- Füllgrabe, C., and Moore, B. C. J. (2014). Effects of age and hearing loss on stream segregation based on interaural time differences. *J. Acoust. Soc. Am.* 136, EL185–EL191. doi: 10.1121/1.4890201
- Füllgrabe, C., Moore, B. C. J., Demany, L., Ewert, S. D., Sheft, S., and Lorenzi, C. (2005). Modulation masking produced by second-order modulators. *J. Acoust. Soc. Am.* 117, 2158–2168. doi: 10.1121/1.1861892
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2012). Speech-in-noise identification in elderly listeners with audiometrically normal hearing: contributions of auditory temporal processing and cognition. *Int. J. Audiol.* 51, 245.
- Füllgrabe, C., Stone, M. A., and Moore, B. C. J. (2009). Contribution of very low amplitude-modulation rates to intelligibility in a competing-speech task. *J. Acoust. Soc. Am.* 125, 1277–1280. doi: 10.1121/1.3075591
- Gatehouse, S., and Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int. J. Audiol.* 43, 85–99. doi: 10.1080/14992020400050014
- Gelfand, S. A., Piper, N., and Silman, S. (1985). Consonant recognition in quiet as a function of aging among normal hearing subjects. *J. Acoust. Soc. Am.* 78, 1198–1206. doi: 10.1121/1.392888
- Gelfand, S. A., Piper, N., and Silman, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *J. Acoust. Soc. Am.* 80, 1589–1598. doi: 10.1121/1.394323
- Gelfand, S. A., Ross, L., and Miller, S. (1988). Sentence reception in noise from one versus two sources: effects of aging and hearing loss. *J. Acoust. Soc. Am.* 83, 248–256. doi: 10.1121/1.396426

- George, E. L., Festen, J. M., and Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 120, 2295–2311. doi: 10.1121/1.2266530
- Gifford, R. H., and Bacon, S. P. (2005). Psychophysical estimates of nonlinear cochlear processing in younger and older listeners. *J. Acoust. Soc. Am.* 118, 3823–3833. doi: 10.1121/1.2126933
- Gifford, R. H., Bacon, S. P., and Williams, E. J. (2007). An examination of speech recognition in a modulated background and of forward masking in younger and older listeners. *J. Speech Lang. Hear. Res.* 50, 857–864. doi: 10.1044/1092-4388(2007/060)
- Glasberg, B. R., Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I. (1984). Dynamic range and asymmetry of the auditory filter. *J. Acoust. Soc. Am.* 76, 419–427. doi: 10.1121/1.391584
- Golomb, J. D., Peelle, J. E., and Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *J. Acoust. Soc. Am.* 121, 1701–1708. doi: 10.1121/1.2436635
- Gopinath, B., Wang, J. J., Schneider, J., Burlutsky, G., Snowden, J., McMahon, C. M., et al. (2009). Depressive symptoms in older adults with hearing impairments: the blue mountains study. *J. Am. Geriatr. Soc.* 57, 1306–1308. doi: 10.1111/j.1532-5415.2009.02317.x
- Gordon-Salant, S. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *J. Acoust. Soc. Am.* 80, 1599–1607. doi: 10.1121/1.394324
- Gordon-Salant, S., Frisina, R. D., Popper, A. N., and Fay, R. R. (eds.). (2010). *The Aging Auditory System*. Heidelberg: Springer. doi: 10.1007/978-1-4419-0993-0
- Gow, A. J., Johnson, W., Pattie, A., Brett, C. E., Roberts, B., Starr, J. M., et al. (2011). Stability and change in intelligence from age 11 to ages 70, 79, and 87: the Lothian Birth Cohorts of 1921 and 1936. *Psychol. Aging* 26, 232–240. doi: 10.1037/a0021072
- Grassi, M., and Borella, E. (2013). The role of auditory abilities in basic mechanisms of cognition in older adults. *Front. Aging Neurosci.* 5:59. doi: 10.3389/fnagi.2013.00059
- Grose, J. H., and Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear Hear.* 31, 755–760. doi: 10.1097/AUD.0b013e3181e627e7
- Grose, J. H., Mamo, S. K., and Hall, J. W. III. (2009). Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 30, 568–575. doi: 10.1097/AUD.0b013e3181ac128f
- Hall, J. W. III, and Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *J. Acoust. Soc. Am.* 96, 150–154. doi: 10.1121/1.410474
- Harris, J. D., Haines, H. L., and Myers, C. K. (1956). A new formula for using the audiogram to predict speech hearing loss. *Arch. Otolaryngol.* 63, 158–176. doi: 10.1001/archotol.1956.03830080044011
- He, N. J., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J. Acoust. Soc. Am.* 124, 3841–3849. doi: 10.1121/1.2998779
- Helfer, K. S., and Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *J. Acoust. Soc. Am.* 136, 748–759. doi: 10.1121/1.4887463
- Herman, G. E., Warren, L. R., and Wagener, J. W. (1977). Auditory lateralization: age differences in sensitivity to dichotic time and amplitude cues. *J. Gerontol.* 32, 187–191. doi: 10.1093/geronj/32.2.187
- Hertzog, C., Kramer, A. F., Wilson, R. S., and Lindenberger, U. (2008). Enrichment effects on adult cognitive development: can the functional capacity of older adults be preserved and enhanced? *Psychol. Sci. Public Interest* 9, 1–65. doi: 10.1111/j.1539-6053.2009.01034.x
- Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., and Moore, D. R. (2004). Auditory processing efficiency and temporal resolution in infants and adults. *J. Speech Lang. Hear. Res.* 47, 1022–1029. doi: 10.1044/1092-4388(2004/076)
- Hjalte, F., Brannstrom, J., and Gerdtham, U. G. (2012). Societal costs of hearing disorders: a systematic and critical review of literature. *Int. J. Audiol.* 51, 655–662. doi: 10.3109/14992027.2012.690077
- Hofer, S. M., Berg, S., and Era, P. (2003). Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychol. Aging* 18, 285–305. doi: 10.1037/0882-7974.18.2.285
- Hopkins, K., and Moore, B. C. J. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J. Acoust. Soc. Am.* 122, 1055–1068. doi: 10.1121/1.2749457
- Hopkins, K., and Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J. Acoust. Soc. Am.* 125, 442–446. doi: 10.1121/1.3037233
- Hopkins, K., and Moore, B. C. J. (2010). Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *Int. J. Audiol.* 49, 940–946. doi: 10.3109/14992027.2010.512613
- Hopkins, K., and Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J. Acoust. Soc. Am.* 130, 334–349. doi: 10.1121/1.3585848
- Howell, D. C. (2002). *Statistical Methods for Psychology, 4th Edn.* Belmont, CA: Duxbury.
- Huang, C. Q., Dong, B. R., Lu, Z. C., Yue, J. R., and Liu, Q. X. (2010). Chronic diseases and risk for depression in old age: a meta-analysis of published literature. *Ageing Res. Rev.* 9, 131–141. doi: 10.1016/j.arr.2009.05.005
- Humes, L. E. (1996). Speech understanding in the elderly. *J. Am. Acad. Audiol.* 7, 161–167.
- Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *J. Acoust. Soc. Am.* 112, 1112–1132. doi: 10.1121/1.1499132
- Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *J. Am. Acad. Audiol.* 18, 590–603. doi: 10.3766/jaaa.18.7.6
- Humes, L. E., Bussey, T. A., Craig, J., and Kewley-Port, D. (2013a). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten. Percept. Psychophys.* 75, 508–524. doi: 10.3758/s13414-012-0406-9
- Humes, L. E., and Dubno, J. R. (2010). “Factors affecting speech understanding in older adults,” in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (Heidelberg: Springer), 211–257.
- Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., et al. (2012). Central presbycusis: a review and evaluation of the evidence. *J. Am. Acad. Audiol.* 23, 635–666. doi: 10.3766/jaaa.23.8.5
- Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013b). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front. Syst. Neurosci.* 7:55. doi: 10.3389/fnsys.2013.00055
- Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., and Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *J. Speech Hear. Res.* 37, 465–474. doi: 10.1044/jshr.3702.465
- Ison, J. R., Tremblay, K. L., and Allen, P. D. (2010). “Closing the gap between neurobiology and human presbycusis: Behavioural and evoked potential studies of age-related hearing loss in animal models and in humans,” in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (Heidelberg: Springer), 75–110.
- Jerger, J., Jerger, S., and Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear Hear.* 12, 103–109. doi: 10.1097/00003446-199104000-00004
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577. doi: 10.1152/physrev.00029.2003
- Katz, J., Medwetsky, L., Burkard, R., and Hood, L. (eds.). (2009). *Handbook of Clinical Audiology*. Baltimore, MD: Lippincott Williams & Wilkins.
- Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., et al. (2003). Candidature for and delivery of audiological services: special needs of older people. *Int. J. Audiol.* 42(Suppl. 2), 2S92–2S101.
- Kilman, L., Zekveld, A., Hallgren, M., and Rönnerberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Front. Psychol.* 5:651. doi: 10.3389/fpsyg.2014.00651
- Kohlrusch, A., Fassel, R., and Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J. Acoust. Soc. Am.* 108, 723–734. doi: 10.1121/1.429605
- Kricos, P. B. (2006). Audiologic management of older adults with hearing loss and compromised cognitive/psychoacoustic auditory processing capabilities. *Trends Amplif.* 10, 1–28. doi: 10.1177/108471380601000102
- Kumar, A. U., and Sangamanatha, A. V. (2011). Temporal processing abilities across different age groups. *J. Am. Acad. Audiol.* 22, 5–12. doi: 10.3766/jaaa.22.1.2
- Lacher-Fougère, S., and Demany, L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *J. Acoust. Soc. Am.* 118, 2519–2526. doi: 10.1121/1.2032747

- Lamberty, G. J., Putnam, S. H., Chatel, D. M., Bieliauskas, L. A., and Adams, K. M. (1994). Derived Trail Making Test indexes - a preliminary-report. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 7, 230–234.
- Lemaire, P. (2010). Cognitive strategy variations during aging. *Curr. Dir. Psychol. Sci.* 19, 363–369. doi: 10.1177/0963721410390354
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375
- Li, L., Daneman, M., Qi, J. G., and Schneider, B. A. (2004). Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *J. Exp. Psychol. Hum. Percept. Perform.* 30, 1077–1091. doi: 10.1037/0096-1523.30.6.1077
- Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., and Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology* 25, 763–770. doi: 10.1037/a0024238
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868
- Lindenberger, U., Scherer, H., and Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: not due to sensory acuity reductions operating during cognitive assessment. *Psychol. Aging* 16, 196–205. doi: 10.1037/0882-7974.16.2.196
- Lorenzi, C., Dumont, A., and Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *J. Speech Lang. Hear. Res.* 43, 1367–1379. doi: 10.1044/jslhr.4306.1367
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18866–18869. doi: 10.1073/pnas.0607364103
- Lorenzi, C., and Moore, B. C. J. (2008). “Role of temporal envelope and fine structure cues in speech perception: A review,” in *Auditory Signal Processing in Hearing-Impaired Listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007)*, eds T. Dau, J. M. Buchholz, J. M. Harte, and T. U. Christiansen (Denmark: Centertryk A/S), 263–272.
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42(Suppl. 1), S49–S58. doi: 10.3109/14992020309074624
- Lustig, C., Shah, P., Seidler, R., and Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: a review and future directions. *Neuropsychol. Rev.* 19, 504–522. doi: 10.1007/s11065-009-9119-9
- Lutman, M. E., Gatehouse, S., and Worthington, A. G. (1991). Frequency resolution as a function of hearing threshold level and age. *J. Acoust. Soc. Am.* 89, 320–328. doi: 10.1121/1.400513
- MacLeod, A., and Summerfield, Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br. J. Audiol.* 24, 29–43. doi: 10.3109/03005369009077840
- Martin, D., Ellsworth, R., and Cranford, J. (1991). Limitations of analysis of covariance designs in aging research. *Ear Hear.* 12, 85–86. doi: 10.1097/00003446-199102000-00011
- Mattys, S. L., Brooks, J., and Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn. Psychol.* 59, 203–243. doi: 10.1016/j.cogpsych.2009.04.001
- McKellin, W. H., Shahin, K., Hodgson, M., Jamieson, J., and Pichora-Fuller, M. K. (2007). Pragmatics of conversation and communication in noisy settings. *J. Pragmat.* 39, 2159–2184. doi: 10.1016/j.pragma.2006.11.012
- Mohr, P. E., Feldman, J. J., Dunbar, J. L., McConkey-Robbins, A., Niparko, J. K., Rittenhouse, R. K., et al. (2000). The societal costs of severe to profound hearing loss in the United States. *Int. J. Technol. Assess. Health Care* 16, 1120–1135. doi: 10.1017/S0266462300103162
- Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. Chichester: John Wiley and Sons.
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Otolaryngol.* 9, 399–406. doi: 10.1007/s10162-008-0143-x
- Moore, B. C. J. (2012). “The importance of temporal fine structure for the intelligibility of speech in complex backgrounds,” in *Speech Perception and Auditory Disorders*, eds T. Dau, M. L. Jepsen, T. Poulsen, and J. C. Dalsgaard (Ballerup: The Danavox Jubilee Foundation), 21–32.
- Moore, B. C. J. (2014). *Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss*. Singapore: World Scientific. doi: 10.1142/9064
- Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2010). Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. *J. Acoust. Soc. Am.* 128, 360–371. doi: 10.1121/1.3436533
- Moore, B. C. J., and Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 110, 1067–1073. doi: 10.1121/1.1385177
- Moore, B. C. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., and Hopkins, K. (2012a). The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *J. Acoust. Soc. Am.* 131, 1003–1006. doi: 10.1121/1.3672808
- Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántara, J. I. (2000). A test for the diagnosis of dead regions in the cochlea. *Br. J. Audiol.* 34, 205–224. doi: 10.3109/03005364000000131
- Moore, B. C. J., and Ohgushi, K. (1993). Audibility of partials in inharmonic complex tones. *J. Acoust. Soc. Am.* 93, 452–461. doi: 10.1121/1.405625
- Moore, B. C. J., and Sek, A. (2009). Development of a fast method for determining sensitivity to temporal fine structure. *Int. J. Audiol.* 48, 161–171. doi: 10.1080/14992020802475235
- Moore, B. C. J., Stone, M. A., Füllgrabe, C., Glasberg, B. R., and Puria, S. (2008). Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss. *Ear Hear.* 29, 907–922. doi: 10.1097/AUD.0b013e31818246f6
- Moore, B. C. J., Vickers, D. A., and Mehta, A. (2012b). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *Int. J. Audiol.* 51, 715–721. doi: 10.3109/14992027.2012.690079
- Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., et al. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS ONE* 9:e107720. doi: 10.1371/journal.pone.0107720
- Moradi, S., Lidestam, B., Saremi, A., and Rönnerberg, J. (2014). Gated auditory speech perception: effects of listening conditions and cognitive capacity. *Front. Psychol.* 5:531. doi: 10.3389/fpsyg.2014.00531
- Neher, T., Laugesen, S., Jensen, N. S., and Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *J. Acoust. Soc. Am.* 130, 1542–1558. doi: 10.1121/1.3608122
- Olsen, W. O. (1998). Average speech levels and spectra in various speaking/listening conditions: a summary of the Pearson, Bennett, & Fidell (1977) report. *Am. J. Audiol.* 7, 21–25. doi: 10.1044/1059-0889(1998)012
- Park, D. C., and Bischof, G. N. (2013). The aging mind: neuroplasticity in response to cognitive training. *Dialogues Clin. Neurosci.* 15, 109–119.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., and Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* 17, 299–320. doi: 10.1037/0882-7974.17.2.299
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold. *J. Acoust. Soc. Am.* 72, 1788–1803. doi: 10.1121/1.388652
- Payne, B. R., Gross, A. L., Parisi, J. M., Sisco, S. M., Stine-Morrow, E. A., Marsiske, M., et al. (2014). Modelling longitudinal changes in older adults' memory for spoken discourse: findings from the ACTIVE cohort. *Memory* 22, 990–1001. doi: 10.1080/09658211.2013.861916
- Peters, R. W., and Moore, B. C. J. (1992). “Auditory filters and aging: filters when auditory thresholds are normal,” in *Auditory Physiology and Perception*, eds Y. Cazals, L. Demany, and K. Horner (Oxford: Pergamon), 179–185.
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47(Suppl. 2), S72–S82. doi: 10.1080/14992020802307404
- Pichora-Fuller, M. K., and MacDonald, E. (2008). “Auditory temporal processing deficits in older listeners: From a review to a future view of Presbycusis,” in *Auditory Signal Processing in Hearing-Impaired Listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007)*, eds T. Dau, J. M. Buchholz, J. M. Harte, and T. U. Christiansen (Denmark: Centertryk A/S), 291–300.
- Pichora-Fuller, M. K., and Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old subjects. *J. Acoust. Soc. Am.* 91, 2129–2135. doi: 10.1121/1.403673

- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Pickett, J. M. (1999). *The Acoustics of Speech Communication: The Fundamentals, Speech Perception Theory and Technology*. Boston, MA: Allyn & Bacon.
- Plomp, R. (1964). The ear as a frequency analyzer. *J. Acoust. Soc. Am.* 36, 1628–1636. doi: 10.1121/1.1919256
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Preacher, K. J., Rucker, D. D., MacCallum, R. C., and Nicewander, W. A. (2005). Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychol. Methods* 10, 178–192. doi: 10.1037/1082-989X.10.2.178
- Reitan, R. M. (1955). The relation of the trail making test to organic brain damage. *J. Consult. Psychol.* 19, 393–394. doi: 10.1037/h0044509
- Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Robertson, D., and Irvine, D. R. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471. doi: 10.1002/cne.902820311
- Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1994). *The Test of Everyday Attention: TEA*. Bury St. Edmunds, UK: Thames Valley Test Company.
- Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1996). The structure of normal human attention: the test of everyday attention. *J. Int. Neuropsychol. Soc.* 2, 525–534. doi: 10.1017/S1355617700001697
- Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnfors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Rönnberg, J., Rudner, M., and Lunner, T. (2011). Cognitive hearing science: the legacy of Stuart Gatehouse. *Trends Amplif.* 15, 140–148. doi: 10.1177/1084713811409762
- Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (2007). Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *J. Neurosci.* 27, 11172–11178. doi: 10.1523/JNEUROSCI.1813-07.2007
- Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4
- Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychol. Sci.* 2, 179–183. doi: 10.1111/j.1467-9280.1991.tb00127.x
- Salthouse, T. A. (2004). What and when of cognitive aging. *Curr. Dir. Psychol. Sci.* 13, 140–144. doi: 10.1111/j.0963-7214.2004.00293.x
- Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiol. Aging* 30, 507–514. doi: 10.1016/j.neurobiolaging.2008.09.023
- Salthouse, T. A. (2012). Are individual differences in rates of aging greater at older ages? *Neurobiol. Aging* 33, 2373–2381. doi: 10.1016/j.neurobiolaging.2011.10.018
- Sanchez-Cubillo, I., Perianez, J. A., Adrover-Roig, D., Rodriguez-Sanchez, J. M., Rios-Lago, M., Tirapu, J., et al. (2009). Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J. Int. Neuropsychol. Soc.* 15, 438–450. doi: 10.1017/S1355617709090626
- Santurette, S., and Dau, T. (2007). Binaural pitch perception in normal-hearing and hearing-impaired listeners. *Hear. Res.* 223, 29–47. doi: 10.1016/j.heares.2006.09.013
- Sayles, M., Füllgrabe, C., and Winter, I. M. (2013). Neurometric amplitude-modulation detection threshold in the guinea-pig ventral cochlear nucleus. *J. Physiol.* 591, 3401–3419. doi: 10.1113/jphysiol.2013.253062
- Schoof, T., and Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Front. Aging Neurosci.* 6:307. doi: 10.3389/fnagi.2014.00307
- Schvartz, K. C., Chatterjee, M., and Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *J. Acoust. Soc. Am.* 124, 3972–3988. doi: 10.1121/1.2997434
- Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *J. Neurosci.* 33, 13686–13694. doi: 10.1523/JNEUROSCI.1783-13.2013
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303
- Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). Effect of age, presentation method, and learning on identification of noise-vocoded words. *J. Acoust. Soc. Am.* 123, 476–488. doi: 10.1121/1.2805676
- Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *J. Acoust. Soc. Am.* 124, 1294–1305. doi: 10.1121/1.2949399
- Smoski, W. J., and Trahiotis, C. (1986). Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. *J. Acoust. Soc. Am.* 79, 1541–1547. doi: 10.1121/1.393680
- Sommers, M. S. (1997). Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment. *J. Acoust. Soc. Am.* 101, 2278–2288. doi: 10.1121/1.418208
- Sommers, M. S., and Humes, L. E. (1993). Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners. *J. Acoust. Soc. Am.* 93, 2903–2914. doi: 10.1121/1.405810
- Sörqvist, P., Ljungberg, J. K., and Ljung, R. (2010). A sub-process view of working memory capacity: evidence from effects of speech on prose memory. *Memory* 18, 310–326. doi: 10.1080/09658211003601530
- Sörqvist, P., and Rönnberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what is the role for working memory capacity? *J. Speech Lang. Hear. Res.* 55, 210–218. doi: 10.1044/1092-4388(2011/10-0353)
- Souza, P. E., and Boike, K. T. (2006). Combining temporal-envelope cues across channels: effects of age and hearing loss. *J. Speech Lang. Hear. Res.* 49, 138–149. doi: 10.1044/1092-4388(2006/011)
- Stone, M. A., Füllgrabe, C., Mackinnon, R. C., and Moore, B. C. J. (2011). The importance for speech intelligibility of random fluctuations in “steady” background noise. *J. Acoust. Soc. Am.* 130, 2874–2881. doi: 10.1121/1.3641371
- Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (2008). Benefit of high-rate envelope cues in vocoder processing: effect of number of channels and spectral region. *J. Acoust. Soc. Am.* 124, 2272–2282. doi: 10.1121/1.2968678
- Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317–326. doi: 10.1121/1.4725766
- Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., and Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: a longitudinal analysis. *Gerontologist* 40, 320–326. doi: 10.1093/geront/40.3.320
- Stuart, A., and Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing and presbycusis listeners. *Ear Hear.* 17, 478–489. doi: 10.1097/00003446-199612000-00004
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455
- Takahashi, G. A., and Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear. Res.* 35, 1410–1421. doi: 10.1044/jshr.3506.1410
- Troche, S. J., and Rammsayer, T. H. (2009). The influence of temporal resolution power and working memory capacity on psychometric intelligence. *Intelligence* 37, 479–486. doi: 10.1016/j.intell.2009.06.001
- Unsworth, N., and Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104–132. doi: 10.1037/0033-295X.114.1.104
- Van Rooij, J. C. G. M., and Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners.II: Multivariate analyses. *J. Acoust. Soc. Am.* 88, 2611–2624. doi: 10.1121/1.399981
- Van Rooij, J. C. G. M., and Plomp, R. (1992). Auditive and cognitive factors in speech perception by elderly listeners.III. Additional data and final discussion. *J. Acoust. Soc. Am.* 91, 1028–1033. doi: 10.1121/1.402628
- Van Rooij, J. C. G. M., Plomp, R., and Orlebeke, J. F. (1989). Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *J. Acoust. Soc. Am.* 86, 1294–1309. doi: 10.1121/1.398744

- Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *J. Acoust. Soc. Am.* 82, 1152–1161. doi: 10.1121/1.395251
- Verhaeghen, P., and De Meersman, L. (1998). Aging and the Stroop effect: a meta-analysis. *Psychol. Aging* 13, 120–126. doi: 10.1037/0882-7974.13.1.120
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 66, 1364–1380. doi: 10.1121/1.383531
- Walton, J. P., Frisina, R. D., and O'Neill, W. E. (1998). Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse. *J. Neurosci.* 18, 2764–2776.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – 3rd UK Edition (WAIS-IIIUK)*. Oxford: Harcourt Assessment.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: The Psychological Corporation.
- Weinstein, B. E., and Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *J. Speech Hear. Res.* 25, 593–599. doi: 10.1044/jshr.2504.593
- Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (2014). The perception of apparent auditory source width in hearing-impaired adults. *J. Acoust. Soc. Am.* 135, 3548–3559. doi: 10.1121/1.4875575
- Willott, J. F. (1996). Anatomic and physiologic aging: a behavioral neuroscience perspective. *J. Am. Acad. Audiol.* 7, 141–151.
- Wingfield, A. (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. *J. Am. Acad. Audiol.* 7, 175–182.
- Wingfield, A., and Grossman, M. (2006). Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. *J. Neurophysiol.* 96, 2830–2839. doi: 10.1152/jn.00628.2006
- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.* 17, 487–497. doi: 10.3766/jaaa.17.7.4
- Young, E. D., and Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibres. *J. Acoust. Soc. Am.* 66, 1381–1403. doi: 10.1121/1.383532
- Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., Van Beek, J. H., and Rönnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear.* 32, e16–e25. doi: 10.1097/AUD.0b013e318228036a
- Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., and Rönnberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Front. Neurosci.* 8:88. doi: 10.3389/fnins.2014.00088

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 November 2014; accepted: 23 December 2014; published online: 13 January 2015.

Citation: Füllgrabe C, Moore BCJ and Stone MA (2015) Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2015 Füllgrabe, Moore and Stone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The ups and downs of global motion perception: a paradoxical advantage for smaller stimuli in the aging visual system

Claire V. Hutchinson^{1*}, Tim Ledgeway² and Harriet A. Allen²

¹ College of Medicine, Biological Sciences and Psychology, School of Psychology, University of Leicester, Leicester, UK

² School of Psychology, University of Nottingham, Nottingham, UK

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Aurel Popa-Wagner, Clinic of

Psychiatry, Germany

John Andersen, University of

California, Riverside, USA

*Correspondence:

Claire V. Hutchinson, College of Medicine, Biological Sciences and Psychology, School of Psychology, University of Leicester, Leicester LE1 9HN, UK
e-mail: ch190@le.ac.uk

Recent evidence suggests that normal aging is typically accompanied by impairment in the ability to perceive the global (overall) motion of visual objects in the world. The purpose of this study was to examine the interplay between age-related changes in the ability to perceive translational global motion (up vs. down) and important factors such as the spatial extent (size) over which movement occurs and how cluttered the moving elements are (density). We used random dot kinematograms (RDKs) and measured motion coherence thresholds (% signal elements required to reliably discriminate global direction) for young and older adults. We did so as a function of the number and density of local signal elements, and the aperture area in which they were displayed. We found that older adults' performance was relatively unaffected by changes in aperture size, the number and density of local elements in the display. In young adults, performance was also insensitive to element number and density but was modulated markedly by display size, such that motion coherence thresholds decreased as aperture area increased (participants required fewer local elements to move coherently to determine the overall image direction). With the smallest apertures tested, young participants' motion coherence thresholds were considerably higher (~1.5 times worse) than those of their older counterparts. Therefore, when RDK size is relatively small, older participants were actually better than young participants at processing global motion. These findings suggest that the normal (disease-free) aging process does not lead to a general decline in perceptual ability and in some cases may be visually advantageous. The results have important implications for the understanding of the consequences of aging on visual function and a number of potential explanations are explored. These include age-related changes in spatial summation, reduced cortical inhibition, neural blur and attentional resource allocation.

Keywords: vision, global motion, age, aperture size, dot density

INTRODUCTION

The ability to effectively discriminate the movement of objects in the world around us is a fundamental part of visual perception. It requires us to be able to decode a constantly, and often rapidly, updated retinal image. These changing patterns of visual information that impinge on our retinas are generated, not only by multiple objects moving independently in the environment, often along different trajectories and at different velocities, but also by our own self-motion relative to stationary objects as we navigate our way through space.

Age-related declines in the ability to accurately decode the moving world have often been reported, particularly with respect to global motion patterns in which individual ("local") elements contribute to the overall pattern motion of a larger ("global") stimulus. The majority of studies that have investigated age-related deficits in global motion perception have employed random dot kinematograms (RDKs) in which local elements, or

dots, move along a translational trajectory (i.e., left or right; up or down). However, these studies have produced mixed results. Whilst some show a deleterious effect of age for translational global motion perception (e.g., Trick and Silverman, 1991; Wojciechowski et al., 1995; Tran et al., 1998; Billino et al., 2008), others have failed to find differences (e.g., Gilmore et al., 1992) between young and old participants' performance. Other studies still have failed to find age-related deficits under some conditions but found impaired performance for older participants under others (e.g., Atchley and Andersen, 1998; Snowden and Kavanagh, 2006; Allen et al., 2010). It is likely therefore that the degree of age-related deficits in global motion perception is heavily dependent on the precise parameters of the RDK patterns used. For example, we (Allen et al., 2010) have recently demonstrated that when high contrast (30% Michelson) dots move at a speed of 5.6 deg/s, older and younger adults' motion discrimination performance are equivalent. However when dots were lower contrast (but still

supra-threshold, $< \sim 4\%$) we found elevated motion coherence thresholds for older, compared to younger adults. These findings demonstrate the importance of contrast sensitivity in global motion perception.

The relative effects of other factors such as aperture area (overall stimulus size) and dot number/density on age-related global motion deficits are presently unclear, despite these being strong influences on performance in younger adults (see Hutchinson et al., 2012). As the size of a stimulus increases, detection and discrimination thresholds typically decrease (e.g., Lappin and Bell, 1976; Anderson and Burr, 1991; Watson and Turano, 1995; Hutchinson and Ledgeway, 2010). This is due to spatial summation whereby the larger stimuli contain more contrast energy and are thus more detectable. In older adults, however, there appears to be a further advantage for large, high contrast, moving gratings compared to younger adults. Betts et al. (2005) compared older and younger adults with high and low contrast sine wave gratings that subtended either 0.7 or 5 degrees. Whilst older adults required longer stimulus presentation times to discriminate motion direction than younger adults, with large high contrast stimuli they needed shorter durations than younger adults. In a similar task, with counter-phasing Gabor stimuli, matched for individual contrast sensitivity, younger adults showed less summation at higher contrasts than lower contrasts whereas older adults continued to show summation even at high contrasts. It has been suggested by some that these age-related differences may be due to differences in spatial suppression (Betts et al., 2005) in extrastriate cortical area V5/MT (Glasser and Tadin, 2010) or contrast sensitivity (Aaen-Stockdale et al., 2009). One might expect a similar advantage for large global motion RDK patterns, but given that the effects of aging on summation area are not necessarily consistent between different stimuli (e.g., Dannheim and Drance, 1971; Scheffrin et al., 1998) this is currently an unresolved issue.

Although studies suggest that increasing stimulus area may have differential effects in young and older adults for simple motion, none have investigated age-related differences in the context of global motion perception. In addition, the effects of age on performance for encoding other RDK parameters such as the number or density of individual elements (local dots) in the display are presently unknown. In the context of studies that have not specifically studied the effects of aging on motion vision, Barlow and Tripathy (1997) have shown, in a small sample of adults, that motion coherence thresholds decrease with increasing stimulus area. They also showed a nominal effect of local element density, although this only translated to an improvement of less than 20% over a 64-fold increase in local element density (1.7–111 dots/deg²). Others (e.g., Williams and Sekuler, 1984) have found no effect of local element density on the probability of perceiving unidirectional flow in RDKs, although Eagle and Rogers (1996) have shown that increasing local element density may lead to an increase in D_{\max} (the maximum displacement for reliable direction identification). Dakin et al. (2005) have suggested that the number of local elements in the display is the principal factor that determines performance for discriminating global motion. Using an equivalent noise paradigm, they showed that local and global limits on direction integration are determined by the number

of local elements in the display, irrespective of their density or indeed the size of the aperture in which they are displayed. Even if summation areas do not change with age, it is possible that the mechanism that serves to integrate over local motion elements is different, or at least subject to different constraints, in younger and older adults. Mapstone et al. (2008), for example, found no differences in older adults' heading discrimination with moving stimuli of different sizes unless a conflicting pattern was presented in the periphery.

To address these issues we conducted three experiments to investigate how varying the number and density of local signal elements, and the aperture area in which they are displayed affects performance for discriminating the global direction of translational RDKs in young and older adults.

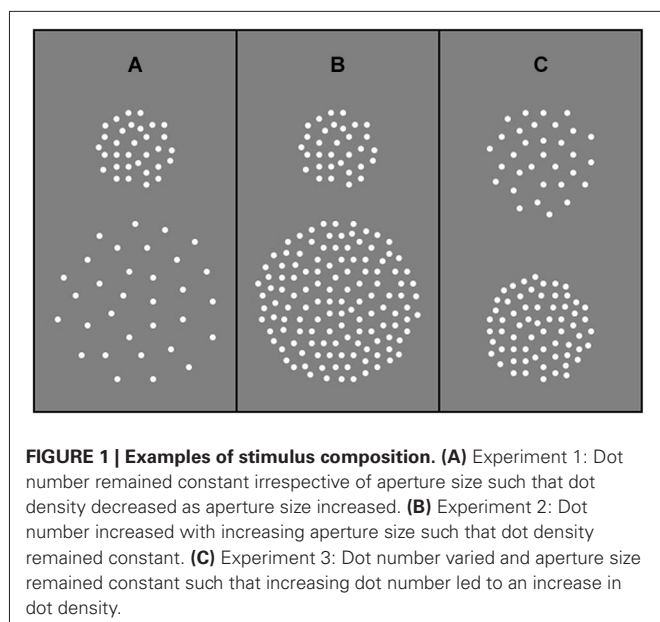
MATERIALS AND METHODS

PARTICIPANTS

Twelve young (mean age = 21.4 years, SD = 2.95) and 10 older (mean age = 72.9 years, SD = 2.33) participants took part in Experiment 1. 14 younger (mean age = 21.18 years, SD = 2.89) and 11 older (mean age = 72.82 years, SD = 2.23) participants took part in Experiment 2. 9 younger (mean age = 22.00 years, SD = 2.84) and 9 older (mean age = 71.33 years, SD = 4.21) participants took part in Experiment 3. All had normal or corrected-to-normal visual acuity and normal binocular vision. Different groups of participants took part in each experiment. Older participants were screened for major head injuries and dementia using the mini mental state examination (Folstein et al., 1983). All experimental methods adhered to the tenets of the Declaration of Helsinki and were approved by the relevant institutional ethics committees.

APPARATUS AND STIMULI

Stimuli were generated using a *Macintosh G4* and presented on a *P255f Professional* monitor (refresh rate 75 Hz) that was gamma-corrected with the aid of internal look-up tables. Stimuli were RDKs depicting translational (up vs. down) motion. Dots (0.47 deg diameter) were high contrast (30% Michelson) and were presented in a central aperture on a homogenous "gray" background (background luminance 64.72 cd/m²). Viewing distance was 92 cm. Each RDK was generated immediately prior to its presentation and was composed of a sequence of 8 images (each 53.3 ms), which when presented consecutively produced continuous motion lasting 426.7 ms. At the beginning of each motion sequence, the position of each dot was randomly assigned. On subsequent frames, each dot was shifted by 0.3 deg, resulting in a drift speed of 5.7 deg/s. When a dot dropped off the edge of the circular display window it was immediately re-plotted in a random spatial position within the window. In Experiment 1, aperture area varied in the range 28–227 deg² but dot number remained constant (64 dots), such that a two-fold increase in aperture area corresponded to a two-fold decrease in dot density (in the range 2.26–0.28 dots/deg²). In Experiment 2, aperture area varied from 14 to 227 deg² but dot density remained constant (1.13 dots/deg²) across experimental conditions, such that a two-fold increase in aperture area corresponded to an equivalent increase in dot number (in the range 16–256 dots).



In Experiment 3, aperture area remained constant at 113 deg² and two dot densities (0.44 and 1.13 dots/deg²) were presented. A stimulus schematic for each experiment is shown in **Figure 1**.

PROCEDURE

The global motion coherence level of the stimulus was manipulated by constraining a fixed proportion of “signal” dots on each image update to move coherently along a translational trajectory (either upwards or downwards on each trial with equal probability). The remainder (“noise” dots) moved in random directions. Global motion thresholds were measured monocularly using a single interval, forced choice, direction discrimination procedure. The participants’ task was to identify whether the global motion was upwards or downwards. The order of testing was randomized. Participants completed at least four 3-down, 1-up adaptive staircases (Edwards and Badcock, 1995) that varied the proportion of signal dots present on each trial, according to the observer’s recent response history. The staircase terminated after eight reversals and thresholds (79% correct performance) were taken as the mean of the last six reversals.

RESULTS

In Experiment 1, dot number remained constant at 64 dots such that dot density decreased as aperture size increased. Findings are presented in **Figure 2** which shows mean global motion coherence thresholds (% of signal dots required for 79% correct direction discrimination performance) for younger and older participants as a function of aperture size. Younger participants’ performance was more markedly affected by changes in aperture size/dot density than that of older participants. A 2 (age) × 4 (aperture size) mixed analysis of variance (ANOVA) showed that there was a significant interaction between age and aperture size [$F_{(1,535,30.709)} = 4.568$, $p < 0.05$]. Closer inspection of the data using one-way ANOVA, performed separately for each age group, confirmed

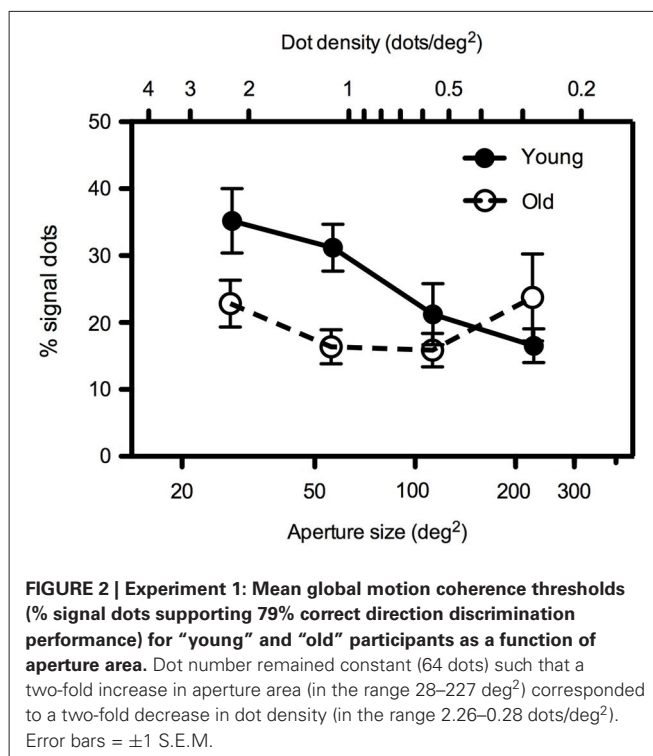


Table 1 | t-test results comparing motion coherence thresholds at each aperture size in younger and older participants in Experiment 1.

Aperture area (deg ²)	<i>t</i>	<i>df</i>	<i>p</i>
28.270	2.00	20	0.059
56.750	2.69	20	0.014
113.10	1.35	20	0.191
226.98	−1.16	20	0.261

that younger participants’ motion coherence thresholds decreased significantly as aperture size increased [$F_{(3,47)} = 5.401$, $p < 0.005$] whereas older participants’ performance remained relatively immune to changes in the spatial aperture [$F_{(3,39)} = 1.019$, ns]. Furthermore, for conditions in which the aperture size was relatively large ($\geq \sim 113$ deg²), there was no significant difference between younger and older participants’ performance. However at smaller aperture sizes ($\leq \sim 57$ deg²), younger participants exhibited higher motion coherence thresholds (performance was worse) than older participants (see **Table 1** for further details of *t*-test pairwise comparisons).

To separate the effects of aperture size and dot density, in Experiment 2, dot number increased with increasing aperture area such that dot density remained constant across conditions at 1.13 dots/deg². Mean global motion coherence for younger and older participants as a function of aperture size are shown in **Figure 3**. Even when dot density remained constant the findings were comparable to those in Experiment 1. A 2 (age) × 5 (aperture size) mixed ANOVA again showed a significant interaction between age and aperture size [$F_{(4,88)} = 7.945$, $p < 0.0001$]. For younger participants, motion coherence thresholds decreased

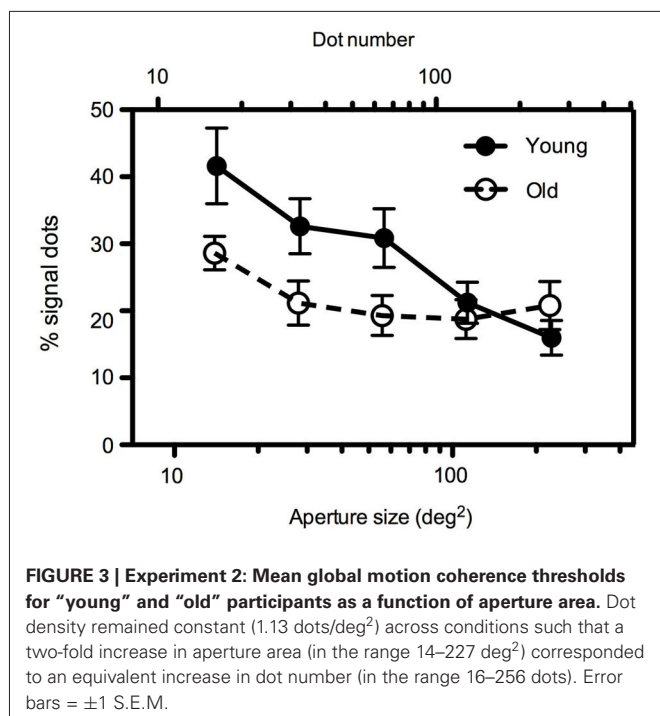
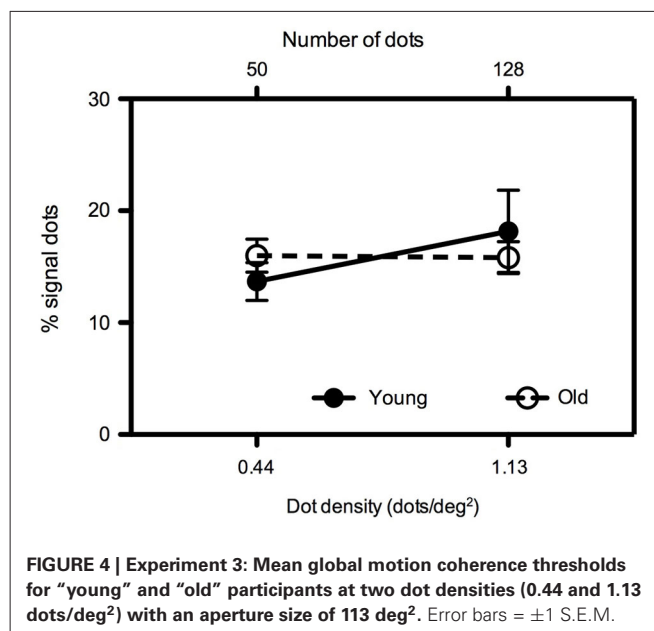


Table 2 | t-test results comparing motion coherence thresholds at each aperture size in younger and older participants in Experiment 2.

Aperture area (deg ²)	<i>t</i>	<i>df</i>	<i>p</i>
14.186	2.11	22	0.051
28.270	2.27	22	0.034
56.750	2.11	22	0.046
113.10	0.62	22	0.540
226.98	−1.11	22	0.279

significantly as aperture size increased [$F_{(4,64)} = 6.042$, $p < 0.0001$]. For older participants performance remained relatively consistent irrespective of the spatial extent of the image [$F_{(4,54)} = 1.956$, ns]. For large apertures (≥ 113 deg²), there was no significant difference between younger and older participants' performance, but for smaller aperture sizes younger participants motion coherence thresholds were higher than older participants (see Table 2 for further details).

To verify the robustness of our findings, in Experiment 3 mean global motion coherence thresholds were measured for younger and older participants for two dot densities (0.44 and 1.13 dots/deg²) with a constant aperture size of 113 deg² (Figure 4). We (Allen et al., 2010) have previously investigated the effects of age on translational motion perception using this particular aperture size and a dot density of 0.44 dots/deg² and found no difference between young and old participants when the dot contrast was relatively high (as it is in the present study). Consistent with our previous study, when the dot density was 0.44 dots/deg², younger and older participants' performance was equivalent ($t = -1.034$, $df = 16$, ns). This was also the case at the higher dot density of 1.13 dots/deg² ($t = 0.601$, $df = 16$, ns).



DISCUSSION

The findings of the present study have shown that changes in aperture size differentially affect global motion perception in young and old participants. At larger aperture sizes (≥ 113 deg²), performance for younger and older participants was equivalent. Indeed there was no advantage for older adults at larger stimulus sizes (as might be predicted by the findings of Betts et al., 2005) nor was there an overall impairment for global motion perception. Older participants' performance was relatively unaffected by changes in aperture size. As aperture size decreased, however, younger participants' coherence thresholds increased (performance became worse), the result being that older participants actually exhibited superior motion perception compared to younger participants with smaller apertures. Our finding of decreases in performance with stimulus size for young participants are in agreement with those found previously by Barlow and Tripathy (1997) and the overall pattern of results was similar irrespective of whether local element density increased (Experiment 1) as aperture area decreased or remained constant (Experiment 2). The findings of Experiment 3 confirmed that the age-related differences in performance observed with small apertures was driven by image size rather than local dot density or dot number. A number of factors may contribute to the pattern of findings presented here. Each potential explanation is addressed in turn below.

COMPARISON WITH OTHER KEY STUDIES

There were no deleterious effects of age in any of the experiments in the present study, in that even in the worst case older adults performed equivalently to younger adults with aperture areas of ≥ 113 deg² (12 deg diameter). These findings are in agreement with those of some other studies that have used high contrast dots displaced over similar spatial extents on each frame of the motion sequence (which was 0.3 deg in the present study) and presented within similar sized regions (apertures). Snowden and Kavanagh

(2006) for example found that motion coherence thresholds were unaffected by age when 400 dots were presented within a 9 deg square region (area 81 deg²) and displaced by 0.18 or 0.36 deg on each positional update. Arena et al. (2012) have shown that global motion perception for judging the direction of 100 dots within a 10 deg diameter aperture (area 79 deg²) is unaffected by age for dot displacements of 0.27 deg. Similarly, Allen et al. (2010) have shown that for high contrast dots, displaced by 0.3 deg on each frame, within a 12 deg diameter aperture (area 113 deg²) age does not significantly affect coherence thresholds for translational global motion.

Like the studies outlined above, Roudaia et al. (2010) have found that performance on global motion tasks is not invariably worse with aging, but rather depends on stimulus factors such as dot spatial displacement and inter-stimulus-interval. Moreover, the combination of these two factors appears to be important. They found that older adults' performance (% correct) for identifying the direction of two-frame apparent motion was markedly worse than their younger counterparts at relatively large spatial displacements and long inter-stimulus-intervals. However the two age groups performed at equivalent levels when the dots were displaced over 0.16 and 0.32 deg between frames, but only when the inter-stimulus-interval was relatively short (0.01–0.04 s for a spatial displacement of 0.16 deg and 0.01–0.02 s for a spatial displacement of 0.32 deg). In these studies, the size of the square region in which the 300 dots appeared was 6.4 deg (area 41 deg²). At a similar image size (area 57 deg²) in the present study, we found that older adults were better than younger adults. The differences between the findings of these two studies cannot be accounted for by differences in dot number/density as we discounted this in Experiment 3. As such, we speculate that these discrepant results could arise, in principle, due to other methodological differences between the two studies. For example the dots used by Roudaia et al. (2010) were approximately 8 times smaller in spatial extent than those used in the present experiments, perhaps making it especially difficult for the older participants to resolve the individual moving elements in their display. In a similar vein, the longer motion sequence duration in the current study (eight frames compared to only two in Roudaia et al., 2010) allowed older adults more integration time to determine the global motion direction and may have benefitted performance (e.g., Barlow and Tripathy, 1997). This highlights the importance of examining how changes in specific motion sequence parameters affect motion perception in older adults, by varying them in isolation and also in conjunction with other parameters, and warrants further study.

SPATIAL SUMMATION

The findings presented here could potentially, in part, reflect smaller spatial summation areas in older participants. Thresholds for older participants appeared to plateau at relatively small aperture sizes, whereas those for younger adults appeared to continue to decrease as image size increased. However, how these findings fit with the existing literature in the spatial domain is unclear. Larger, rather than smaller, summation areas for older compared to younger participants have been reported for static targets under scotopic viewing conditions (e.g., Scheffrin et al., 1998). However,

the majority of studies that have compared spatial summation for achromatic and chromatic static targets in young and old participants have, in the main, found no age-related differences in Ricco's area (Dannheim and Drance, 1971; Brown et al., 1989; Redmond et al., 2010). This also appears to be the case for temporally-modulated targets (Zeile et al., 2006). Drawing meaningful conclusions about the relative spatial summation areas in younger and older participants is problematic, however, both in the context of the present study and in the context of the existing literature, for a number of reasons: (1) there is no reason to assume that the findings in the spatial domain are applicable to moving stimuli; (2) measures of summation area are heavily influenced by experimental variables such as stimulus type, adaptation levels and the statistical techniques used to determine Ricco's area, of which there are many (Redmond et al., 2010); and (3) interpreting experimental findings is difficult because, even where the "young" visual system is concerned, the physiological mechanisms underlying Ricco's area are still unresolved (e.g., Hartline, 1940; Davila and Geisler, 1991; Swanson et al., 2004; Pan and Swanson, 2006). Finally, even if we accept that there are smaller summation areas for global motion in older adults, this does not address why their thresholds are so much lower than those of the younger participants at the smallest stimulus sizes tested.

CORTICAL INHIBITION

We are not the first to demonstrate enhanced motion perception in older adults, relative to their younger counterparts. Betts et al. (2005) for example have shown that older adults are better than younger adults at detecting moving high contrast sine-wave gratings when image size is relatively large. It has been suggested that these findings may reflect reduced cortical inhibition in older populations (e.g., Betts et al., 2005; Bennett et al., 2007), specifically, reduced center-surround antagonism in area V5/MT (e.g., Leventhal et al., 2003; Tadin et al., 2003). With high contrast stimuli, as image size increases, more of the stimulus falls into a putative detector's inhibitory region, reducing detection performance. For older adults, in their study, without this inhibitory region, performance continues to improve with increasing stimulus size.

This hypothesis can also be applied to our findings for global motion patterns as performance for encoding random dot motion has typically been assumed to reflect the underlying physiology of areas V1 and, particularly, area V5/MT. Direction-selective neurons in V1 encode the motion direction of local elements, the outputs of which project up-stream to area V5/MT (Movshon et al., 1985). The larger receptive fields in V5/MT integrate the local motion responses from V1 into a global representation (e.g., Livingstone et al., 2001). In our case, within this framework reduced inhibition may lead to better performance at relatively small image sizes simply due to the absence of an inhibitory surround even with smaller receptive field sizes. This assumption may not be unreasonable given that many foveal-centered receptive fields, in primate MT, sample a region of space comparable in spatial extent to the smallest aperture sizes used in the present study (e.g., Gattass and Gross, 1981; Albright and Desimone, 1987). One argument against the inhibitory surround explanation

of older adults improved sensitivity with larger images is that it is, in part, mediated by baseline differences in the measure of interest such as contrast sensitivity (Aaen-Stockdale et al., 2009). For our stimuli, however, at the contrasts used, there are no significant baseline differences in our measure of interest (see Allen et al., 2010). There is good support from elsewhere for the inhibitory hypothesis. Reduced cortical inhibition in the aging primate brain has been shown to produce improved cortical function and has been linked to age-related reductions in activity of the inhibitory neurotransmitter GABA (Leventhal et al., 2003). Superior global motion perception (relative to controls) has also been shown in certain patient groups such as those suffering from schizophrenia (Tadin et al., 2006) where inhibition is known to be weak. Most recently, Tadin et al. (2011) have shown that applying repetitive transcranial magnetic stimulation (TMS) over area V5/MT can improve motion discrimination for large stimuli, an effect that they attributed to a TMS-induced weakening of surround suppression strength. As such, our findings may, at least in part, reflect reduced center-surround antagonism and hence less spatial suppression in area V5/MT in older adults.

NEURAL BLUR

It is well known that old age leads to a contrast sensitivity loss at intermediate and high spatial frequencies, that is likely to be due mainly to neural rather than optical factors (e.g., Weale, 1975, 1986; Elliott et al., 1990). Within the context of our RDK stimulus, which necessarily contains a broad range of spatial frequencies, superior performance in older participants may reflect changes within the visual pathways that effectively result in lowpass spatial filtering of the image. The result of which may be to render the image “less dense” as higher frequency information is attenuated. There is evidence for example that the upper motion displacement threshold (D_{\max}) increases after lowpass spatial frequency filtering (Morgan, 1992). Furthermore, positive dioptric optical blur, which is known to degrade high spatial frequencies (Westheimer and McKee, 1980), can improve global motion perception in RDK displays (Barton et al., 1996). One notable difference between younger and older adults’ performance was that, whilst younger adults demonstrated the expected deterioration in performance at small image sizes, older adults’ performance was relatively unaffected by changes in image size. Younger adults exhibited poorest performance at the smallest image sizes. Performance improved markedly as image size initially increased and then became asymptotic at the largest aperture sizes tested. These findings may reflect the gradual encroachment of the RDK on the peripheral retina. Images in the visual periphery are effectively blurred relative to those in central vision due to high spatial frequency attenuation and hence poorer spatial acuity. Indeed, the blur-related advantage of high spatial frequency attenuation has been put forward to account for the finding that D_{\max} is greater in peripheral vision (Baker and Braddick, 1985; Cleary and Braddick, 1990). In the case of older participants, if the image is blurred in the fovea as well as the periphery, performance would be unlikely to deteriorate appreciably as the image size is reduced, given that high spatial frequencies are attenuated in both the periphery and the fovea.

VISUAL ATTENTION

Location-selective selective attention mechanisms operate early in visual processing (Cave and Bichot, 1999) and differences in attentional strategy and ability between older and younger adults are commonly found (e.g., Allen and Payne, 2012). Older adults are less able than their younger counterparts to extract information from a cluttered visual scene (Sekuler et al., 2000), show longer search times and larger set size effects than younger adults on visual search tasks (e.g., Madden, 2007). They are less able to divide their attention between salient environmental information and irrelevant distractors and to select a visual target whilst ignoring the distractors (e.g., Owsley et al., 1998). Older adults also show deficits on visual cueing tasks. They are slower to use cues, exhibit more errors than their younger counterparts and tend to have more difficulty distinguishing between valid and invalid cues (e.g., Hoyer and Familant, 1987; Lincourt et al., 1997).

Changes in how older adults employ selective attention might predict changes in global motion processing with age. Indeed older adults’ global motion performance appears to be predicted by their performance on tests of selective attention, whereas this does not appear to be the case for younger adults (Mapstone et al., 2008). Our findings for global motion may reflect a more narrow field of attentional focus in older, compared to younger adults. It may in fact be advantageous for older adults to restrict their field of view to a narrower region of visual space. Vision in the periphery in older adults is particularly poor (Haegerstrom-Portnoy et al., 1999). In real terms, this means that much of the information in the visual field may be invisible and, as such, effectively useless. As a result, older adults may attend to a narrow central window as a means of improving the quality of their vision. This notion fits with the zoom lens model of visual attention which predicts that reducing the window of attention improves resolution (Eriksen and Yeh, 1985). Superior global motion perception for small display sizes may therefore reflect perceptual self learning/training in older adults in response to poor peripheral vision. This notion is consistent with recent evidence for training-related changes in the structure and function in the older adult visual system (see Lustig et al., 2009).

CONCLUSIONS

We have shown that when the aperture within which a RDK stimulus is displayed is small, performance for determining the direction of translational global motion is markedly better in older, compared to younger, participants. These findings suggest that the normal aging process does not lead to a general decline in global motion. Indeed, in some cases, and in agreement with other studies, the normal (i.e., disease-free) aging process may be visually advantageous. As far as the neural underpinnings of our present findings are concerned, they are unlikely to be explainable simply in terms of gross differences in the spatial summation areas of younger and older participants. This is compounded by a lack of consensus in the existing literature as to what the effects of age on spatial summation actually are. In a similar vein, reduced GABA inhibition in older adults does not necessarily explain why their relative performance advantage is restricted to small image sizes. Our findings may reflect simple age-related reductions in

sensitivity to high spatial frequencies but it remains to be seen if these are sufficient to account for the magnitude of the differences found at the smallest stimulus sizes. Finally, although age-related narrowing of the attentional field of view may help to explain why the older adults' performance is largely invariant with increases in image size, it does not necessarily explain why they are so much better at the task than the younger participants with the smallest RDKs used. We are currently exploring some of these possibilities in greater detail in our laboratory.

In conclusion, this preliminary study represents somewhat surprising findings that will require further study. Indeed, in many respects they raise questions, rather than providing answers. We show that our findings cannot be explained by changes in dot density but future studies should investigate the potential interplay between image size and different RDK parameters such as signal dot trajectory, dot size (which in our study was relatively large), dot speed, spatial displacement and motion sequence duration. In doing so, to verify the robustness of the findings, testing would also ideally be conducted in a larger sample of participants.

ACKNOWLEDGMENTS

We thank all those who took part in this study. We are particularly indebted to the older adults for giving up their time to take part. Claire V. Hutchinson is grateful to the University of Leicester for a semester of study leave. This work was supported by a British Academy Grant (SG113210) to Claire V. Hutchinson.

REFERENCES

- Aaen-Stockdale, C. R., Thompson, B., Huang, P. C., and Hess, R. F. (2009). Low-level mechanisms may contribute to paradoxical motion percepts. *J. Vis.* 9:9. doi: 10.1167/9.5.9
- Albright, T. D., and Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Exp. Brain Res.* 65, 582–592. doi: 10.1007/bf00235981
- Allen, H. A., Hutchinson, C. V., Ledgeway, T., and Gayle, P. (2010). The role of contrast sensitivity in global motion processing deficits in the elderly. *J. Vis.* 10:15. doi: 10.1167/10.10.15
- Allen, H. A., and Payne, H. (2012). Similar behaviour, different brain patterns: age-related changes in neural signatures of ignoring. *Neuroimage* 59, 4113–4125. doi: 10.1016/j.neuroimage.2011.10.070
- Anderson, S. J., and Burr, D. C. (1991). Spatial summation of directionally selective mechanisms in human vision. *J. Opt. Soc. Am. A* 8, 1330–1339. doi: 10.1364/josaa.8.001330
- Arena, A., Hutchinson, C. V., and Shimozaki, S. S. (2012). The effects of age on the spatial and temporal integration of global motion. *Vision Res.* 58, 27–32. doi: 10.1016/j.visres.2012.02.004
- Atchley, P., and Andersen, G. J. (1998). The effect of age, retinal eccentricity and speed on the detection of optic flow components. *Psychol. Aging* 13, 297–308. doi: 10.1037//0882-7974.13.2.297
- Baker, C. L., and Braddick, O. J. (1985). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Res.* 25, 803–812. doi: 10.1016/0042-6989(85)90188-9
- Barlow, H., and Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *J. Neurosci.* 17, 7595–7966.
- Barton, J. S., Rizzo, M., Nawrot, N., and Simpson, T. (1996). Optical blur and the perception of global coherent motion in random dot cinematograms. *Vision Res.* 36, 3051–3059. doi: 10.1016/0042-6989(96)00063-6
- Bennett, P. J., Sekuler, R., and Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Res.* 47, 799–809. doi: 10.1016/j.visres.2007.01.001
- Betts, L. R., Taylor, C. P., Sekuler, A. B., and Bennett, P. J. (2005). Aging reduces center surround antagonism in visual motion processing. *Neuron* 45, 361–366. doi: 10.1016/j.neuron.2004.12.041
- Billino, J., Bremmer, F., and Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: evidence against general perceptual decline. *Vision Res.* 48, 1254–1261. doi: 10.1016/j.visres.2008.02.014
- Brown, B., Peterken, C., Bowman, K. J., and Crassini, B. (1989). Spatial summation in young and elderly observers. *Ophthalmic Physiol. Opt.* 9, 310–313. doi: 10.1111/j.1475-1313.1989.tb00913.x
- Cave, K. R., and Bichot, N. P. (1999). Visuospatial attention: beyond a spotlight model. *Psychon. Bull. Rev.* 6, 204–223. doi: 10.3758/bf03212327
- Cleary, R., and Braddick, O. J. (1990). Masking of low frequency information in short-range apparent motion. *Vision Res.* 30, 317–327. doi: 10.1016/0042-6989(90)90046-n
- Dakin, S. C., Mareschal, I., and Bex, P. J. (2005). Local and global limitations on direction integration assessed using equivalent noise. *Vision Res.* 45, 3027–3049. doi: 10.1016/j.visres.2005.07.037
- Dannheim, F., and Drance, S. M. (1971). Studies of spatial summation in central areas in normal people of all ages. *Can. J. Ophthalmol.* 6, 311–319.
- Davila, K. D., and Geisler, W. S. (1991). The relative contributions of pre-neural and neural factors to areal summation in the fovea. *Vision Res.* 31, 1369–1380. doi: 10.1016/0042-6989(91)90058-d
- Eagle, R. A., and Rogers, B. J. (1996). Motion detection is limited by element density not spatial frequency. *Vision Res.* 36, 545–558. doi: 10.1016/0042-6989(96)89252-2
- Edwards, M., and Badcock, D. (1995). Global motion perception: no interaction between the first- and second-order pathways. *Vision Res.* 35, 2589–2602. doi: 10.1016/0042-6989(95)00003-i
- Elliott, D. B., Whitaker, D., and MacVeigh, D. (1990). Neural contribution to spatiotemporal contrast sensitivity decline in healthy eyes. *Vision Res.* 30, 541–547. doi: 10.1016/0042-6989(90)90066-t
- Eriksen, C. W., and Yeh, Y. Y. (1985). Allocation of attention in the visual field. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 583–597. doi: 10.1037//0096-1523.11.5.583
- Folstein, M. F., Robins, L. N., and Helzer, J. E. (1983). The minimal state examination. *Arch. Gen. Psychiatry* 40, 812–812. doi: 10.1001/archpsyc.1983.01790060110016
- Gattass, R., and Gross, C. G. (1981). Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *J. Neurophysiol.* 46, 621–638.
- Gilmore, G. C., Wenk, H. E., Naylor, L. A., and Stuve, T. A. (1992). Motion perception and aging. *Psychol. Aging* 7, 654–660. doi: 10.1037/0882-7974.7.4.654
- Glasser, D. M., and Tadin, D. (2010). Low-level mechanisms do not explain paradoxical motion percepts. *J. Vis.* 10:20. doi: 10.1167/10.4.20
- Haegerstrom-Portnoy, G., Schneek, M. E., and Brabyn, J. A. (1999). Seeing into old age: vision function beyond acuity. *Optom. Vis. Sci.* 76, 141–158. doi: 10.1097/00006324-199903000-00014
- Hartline, H. K. (1940). The effects of spatial summation in the retina on excitation of the fibers of the optic nerve. *Am. J. Physiol.* 130, 700–711.
- Hoyer, W. J., and Familant, M. E. (1987). Adult age differences in the rate of processing expectancy information. *Cogn. Dev.* 2, 59–70. doi: 10.1016/s0885-2014(87)90035-9
- Hutchinson, C. V., Arena, A., Allen, H. A., and Ledgeway, T. (2012). Psychophysical correlates of global motion processing in the aging visual system: a critical review. *Neurosci. Biobehav. Rev.* 36, 1266–1272. doi: 10.1016/j.neubiorev.2012.02.009
- Hutchinson, C. V., and Ledgeway, T. (2010). Spatial summation of first-order and second-order motion in human vision. *Vision Res.* 50, 1766–1774. doi: 10.1016/j.visres.2010.05.032
- Lappin, J. S., and Bell, H. H. (1976). The detection of coherence in moving visual patterns. *Vision Res.* 16, 161–168. doi: 10.1016/0042-6989(76)90093-6
- Leventhal, A. G., Wang, Y. C., Pu, M. L., Zhou, Y. F., and Ma, Y. Y. (2003). GABA and its agonists improved visual cortical function in senescent monkeys. *Science* 300, 812–815. doi: 10.1126/science.1082874
- Lincoourt, A. E., Folk, C. L., and Hoyer, W. J. (1997). Effects of aging on voluntary and involuntary shifts of attention. *Aging Neuropsychol. Cogn.* 4, 290–303. doi: 10.1080/13825589708256654
- Livingstone, M. S., Pack, C. C., and Born, R. T. (2001). Two-dimensional substructure of MT receptive fields. *Neuron* 30, 781–793. doi: 10.1016/s0896-6273(01)00313-0

- Lustig, C., Shah, P., Seidler, R., and Reuter-Lorenz, P. A. (2009). Ageing, training and the brain: a review and future directions. *Neuropsychol. Rev.* 16, 504–522. doi: 10.1007/s11065-009-9119-9
- Madden, D. J. (2007). Aging and visual attention. *Curr. Dir. Psychol. Sci.* 16, 70–74. doi: 10.1111/j.1467-8721.2007.00478.x
- Mapstone, M., Dickerson, K., and Duffy, C. J. (2008). Distinct mechanisms of impairment in cognitive ageing and Alzheimer's disease. *Brain* 131, 1618–1629. doi: 10.1093/brain/awn064
- Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature* 355, 344–346. doi: 10.1038/355344a0
- Movshon, J. A., Adelson, E. H., Gizzi, M. S., and Newsome, W. T. (1985). "The analysis of moving visual patterns," in *Pattern Recognition Mechanisms*, ed R. G. A. C. G. C. Chagas (Rome: Vatican Press), 117–151.
- Owsley, C., Ball, K., McGwin, G., Sloane, M. E., Roenker, D. L., White, M. F., et al. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *JAMA* 279, 1083–1088. doi: 10.1001/jama.279.14.1083
- Pan, F., and Swanson, W. H. (2006). A cortical pooling model of spatial summation for perimetric stimuli. *J. Vis.* 6:2. doi: 10.1167/6.11.2
- Redmond, T., Zlatkova, M. B., Garway-Heath, D. F., and Anderson, R. S. (2010). The effect of age on the area of complete spatial summation for chromatic and achromatic stimuli. *Invest. Ophthalmol. Vis. Sci.* 51, 6533–6539. doi: 10.1167/iovs.10-5717
- Roudaia, E., Bennett, P. J., Sekuler, A. B., and Pilz, K. S. (2010). Spatiotemporal properties of apparant motion perception and aging. *J. Vis.* 10:5. doi: 10.1167/10.14.5
- Schefrin, B. E., Bieber, M. L., McLean, R., and Werner, J. S. (1998). The area of complete scotopic summation enlarges with age. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 15, 340–348. doi: 10.1364/josaa.15.000340
- Sekuler, A. B., Bennett, P. J., and Mamelak, M. (2000). Effects of aging on the useful field of view. *Exp. Aging Res.* 26, 103–120. doi: 10.1080/036107300243588
- Snowden, R. J., and Kavanagh, E. (2006). Motion perception and the ageing visual system: minimum motion, motion coherence and speed discrimination thresholds. *Perception* 35, 9–24. doi: 10.1068/p5399
- Swanson, W. H., Feliuss, J., and Pan, F. (2004). Perimetric defects and ganglion cell damage: interpreting linear relations using a two-stage neural model. *Invest. Ophthalmol. Vis. Sci.* 45, 466–472. doi: 10.1167/iovs.03-0374
- Tadin, D., Kim, J., Doop, M. L., Gibson, C., Lappin, J. S., Blake, R., et al. (2006). Weakened center-surround interactions in visual motion processing in schizophrenia. *J. Neurosci.* 2, 11403–11412. doi: 10.1523/jneurosci.2592-06.2006
- Tadin, D., Lappin, J. S., Gilroy, L. A., and Blake, R. (2003). Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature* 424, 312–315. doi: 10.1038/nature01800
- Tadin, D., Silvanto, J., Pascual-Leone, A., and Batelli, L. (2011). Improved motion perception and impaired spatial suppression following disruption of cortical area MT/V5. *J. Neurosci.* 31, 1279–1283. doi: 10.1523/JNEUROSCI.4121-10.2011
- Tran, D. B., Silverman, S. E., Zimmerman, K., and Feldon, S. E. (1998). Age-related deterioration of motion perception and detection. *Graefes Arch. Clin. Exp. Ophthalmol.* 236, 269–273. doi: 10.1007/s004170050076
- Trick, G. L., and Silverman, S. E. (1991). Visual sensitivity to motion: age-related changes and deficits in senile dementia of the Alzheimer type. *Neurology* 41, 1437–1440. doi: 10.1212/wnl.41.9.1437
- Watson, A. B., and Turano, K. (1995). The optimal motion stimulus. *Vision Res.* 35, 325–336. doi: 10.1016/0042-6989(94)00182-1
- Weale, R. A. (1975). Senile changes in visual acuity. *Trans. Ophthalmol. Soc. U K* 95, 36–38.
- Weale, R. A. (1986). Aging and vision. *Vision Res.* 26, 1507–1512. doi: 10.1016/0042-6989(86)90170-7
- Westheimer, G., and McKee, S. P. (1980). Stereoscopic acuity with defocused and spatially filtered retinal images. *J. Opt. Soc. Am.* 70, 772–778. doi: 10.1364/josa.70.000772
- Williams, D. W., and Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Res.* 24, 55–62. doi: 10.1016/0042-6989(84)90144-5
- Wojciechowski, R., Trick, G. L., and Steinman, S. B. (1995). Topography of the age-related decline in motion sensitivity. *Optom. Vis. Sci.* 72, 67–74. doi: 10.1097/00006324-199502000-00005
- Zeile, A. J., O'Loughlin, R. K., Guymer, R. H., and Vingrys, A. J. (2006). Disclosing disease mechanisms with a spatio-temporal summation paradigm. *Graefes Arch. Clin. Exp. Ophthalmol.* 244, 425–432. doi: 10.1007/s00417-005-0121-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 April 2014; accepted: 19 July 2014; published online: 08 August 2014.
 Citation: Hutchinson CV, Ledgeway T and Allen HA (2014) The ups and downs of global motion perception: a paradoxical advantage for smaller stimuli in the aging visual system. *Front. Aging Neurosci.* 6:199. doi: 10.3389/fnagi.2014.00199
 This article was submitted to the journal *Frontiers in Aging Neuroscience*.
 Copyright © 2014 Hutchinson, Ledgeway and Allen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Preserved fine-tuning of face perception and memory: evidence from the own-race bias in high- and low-performing older adults

Jessica Komes*, Stefan R. Schweinberger and Holger Wiese

DFG Research Unit Person Perception and Department of General Psychology and Cognitive Neuroscience, Friedrich Schiller University of Jena, Jena, Germany

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Guillaume A. Rousselet, University of Glasgow, UK

Andrea Hildebrandt, Ernst-Moritz-Arndt Universität Greifswald, Germany

*Correspondence:

Jessica Komes, DFG Research Unit Person Perception and Department of General Psychology and Cognitive Neuroscience, Friedrich Schiller University of Jena, Leutragraben 1, 07743 Jena, Germany
e-mail: jessica.komes@uni-jena.de

Previous research suggests specific deficits in face perception and memory in older adults, which could reflect a dedifferentiation in the context of a general broadening of cognitive architecture with advanced age. Such dedifferentiation could manifest in a less specialized face processing system. A promising tool to investigate the fine-tuning of face processing in older age is the own-race bias (ORB), a phenomenon reflecting more accurate memory for own-relative to other-race faces, which is related to an expertise-based specialization of early perceptual stages. To investigate whether poor face memory in older age is accompanied by reduced expertise-based specialization of face processing, we assessed event-related brain potential correlates of the ORB in high- vs. low-performing older adults (mean age = 69 years; $N = 24$ per group). Intriguingly, both older groups demonstrated an equivalent pattern of a behavioral ORB, and a parallel increase in N170 for other-race faces, reflecting less efficient early perceptual processing for this face category. Group differences only emerged independent of face ethnicity: whereas low-performers exhibited a right-lateralized N170, high-performers showed a more bilateral response. This finding may suggest a compensatory mechanism counteracting age-related decline in face perception enabling more efficient encoding into memory in high performers. Overall, our results demonstrate that even a less efficient face processing system in older adults can exhibit preserved expertise-related specialization toward own-race faces.

Keywords: face perception, face memory, cognitive aging, N170, own-race bias, expertise

INTRODUCTION

During aging, the human brain undergoes various changes in structure, function, and neural transmission (Sowell et al., 2003; Raz et al., 2005). These alterations most probably underlie, and certainly covary with cognitive functioning (Cabeza et al., 2004; Dennis and Cabeza, 2008; Nyberg et al., 2012). During childhood, the interaction of maturation, experience, and learning results in cortical differentiation (for review, see Johnson, 2001; Scherf et al., 2007), which is paralleled by more fine-grained cognitive abilities (Li et al., 2004; Werkle-Bergner et al., 2009). This pattern reverses during older age, when senescent changes include cognitive and sensorimotor dedifferentiation (Lindenberger and Baltes, 1994; Baltes and Lindenberger, 1997; Lindenberger et al., 2001) and diminished cortical specialization (Park et al., 2004, 2012).

Critically, age-related alterations and their putative consequences can be very diverse, and within-cohort differences can be similarly pronounced as between-cohort differences (Salthouse, 2013). More specifically, while some older adults show decreased performance in memory and executive processes (for overview, see Anderson and Craik, 2000), others maintain high levels of functioning, and match (see e.g., Cabeza et al., 2002; Duarte et al., 2006; Friedman et al., 2010) or even outperform their younger counterparts (Christensen et al., 1999; Henry et al., 2004; Federmeier et al., 2010). Results from functional brain imaging suggest that different processes in the aging brain could

mediate such performance differences: First, reduced activations of memory-related brain regions were found in older adults with poor performance (for review, see Grady, 2008). Second, increased activations in older participants were observed in prefrontal areas contralateral to those usually activated in younger adults (for reviews, see Cabeza, 2002; Park and Reuter-Lorenz, 2009; Reuter-Lorenz and Park, 2010; but see Nyberg et al., 2010, for potentially discrepant results from longitudinal data). Importantly, these latter findings were interpreted as indexing either dedifferentiation or compensation (for reviews, see Reuter-Lorenz and Cappell, 2008; Park and Reuter-Lorenz, 2009). While dedifferentiation connotes “negative plasticity” in terms of loss in specialization of neuro-cognitive structures, compensation connotes “positive plasticity,” allowing additional resource recruitment to optimize performance (Reuter-Lorenz and Park, 2010). This debate remains largely unresolved (Davis et al., 2012).

In line with the dedifferentiation view, several studies on object and face processing suggested the fading of distinctive neural representations with increasing age in ventral visual cortex. Of note, activation in fusiform gyrus, which is category-sensitive for faces in younger adults, was less selective in older adults (Park et al., 2004, 2012; Chee et al., 2006). Other studies used event-related potentials (ERPs) to examine the face-sensitive N170, a negative occipito-temporal peak at approximately 170 ms after stimulus onset (Bentin et al., 1996; Eimer, 2011) to elucidate

age-related changes in face processing. At potential variance with the above-described neuroimaging results, N170 category-selectivity (with larger N170 to faces than objects) was similarly found across age groups (Gao et al., 2009; Daniel and Bentin, 2010), suggesting preserved neural sensitivity for faces in older adults¹. However, the N170 is sometimes increased in older adults (Gao et al., 2009; Daniel and Bentin, 2010; Wiese et al., 2013a; see also Rousselet et al., 2009, who reported larger amplitudes in older participants slightly after the N170 peak), and is also larger for inverted relative to upright faces, which is commonly interpreted as reflecting disrupted configural face processing. Notably, smaller N170 inversion effects were also observed in older participants (Gao et al., 2009; Daniel and Bentin, 2010). Hence, a larger N170 for upright faces and a smaller inversion effect in older adults may indicate a general reduction in early stages of face perception with higher age, potentially affecting subsequent higher-order cognitive functions, such as memory (see also Chaby et al., 2011). Consequently, the degree of preserved vs. deficient face perception in older adults might be reflected in differential face memory performance. To our knowledge, this idea has not been addressed in previous work.

The ability to distinguish between exemplars within a face category and to discriminate individual faces develops early, with a specialization toward own species faces from around 6 months of age (Pascalis et al., 2002) and subsequent further tuning toward own-race faces during childhood (see e.g., Scherf and Scott, 2012). In adults, the own-race bias (ORB) reflects the consistent finding of better memory for own-relative to other-race faces (for overview, see Meissner and Brigham, 2001). However, whether an ORB can still be observed in older adults has been scarcely investigated (Wallis et al., 2012). Age-related dedifferentiation could result in a less specialized face processing system (particularly in those older adults with poor face memory), and therefore in a reduced ORB, given that this effect is commonly interpreted to reflect expertise-related fine tuning of face processing mechanisms (Valentine and Endo, 1992; Tanaka et al., 2013). Here, we examined memory for own- and other-race faces in two groups of high- and low-performing older adults to test whether poor face memory in low-performers reflects advancing dedifferentiation (as indexed by a reduced or absent ORB). In addition, we addressed the question, whether reduced memory performance is associated with deteriorated perceptual processing as indexed by the N170.

Several recent studies have identified neural correlates of the ORB using ERPs, and N170 seems to play a prominent role in the generation of this effect. Suggestive of reduced early perceptual processing, other-race faces were found to elicit a larger N170 than own-race faces (Balas and Nelson, 2010; Stahl et al., 2010; Caharel et al., 2011). In a recent study, we reported a larger N170

to other-race faces during the learning phases of a recognition memory paradigm, in both young Asian and Caucasian participants (Wiese et al., 2014). Importantly, this effect during encoding significantly correlated with the subsequent recognition advantage for own-race faces during test. Apart from the ORB, an own-age bias (OAB) in terms of better memory for own-age relative to other-age faces has also been consistently described (for reviews, see Rhodes and Anastasi, 2012; Wiese et al., 2013b). To avoid underestimating older adults' performance, we decided to examine their face recognition memory in a full-factorial design by using young and old own- and other-race faces (Wiese, 2012).

The aim of our investigation was two-fold: First, we tested whether low- relative to high-performing participants would exhibit a reduced (or even absent) behavioral ORB which would argue for a (incipiently) dedifferentiated face processing system in this participant group. Second, we examined whether N170 effects would co-vary with memory performance, indicating less efficient early face perception in those older participants with lower memory performance.

MATERIALS AND METHODS

PARTICIPANTS

Forty-eight older participants (27 females, mean age = 69.0, $SD = 4.7$), recruited in senior citizen groups and via a press release in a local newspaper, participated in the study and were reimbursed with 7.50 Euro per hour. All participants were Caucasian and reported to reside in independent living conditions with little or no contact to Asian people. Participants were right handed according to a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971). None reported psychiatric or neurological disorders or received central acting medication, and all participants reported normal or corrected-to-normal vision (visual acuity and contrast sensitivity were also formally tested, see below). Furthermore, all participants gave written informed consent and the study was approved by the local ethics committee.

The participant group was *post-hoc* subdivided via a median-split with respect to overall performance (mean d' across experimental conditions) in the main recognition experiment into 24 high-performing (14 females, mean age = 68.4, $SD = 5.2$) and 24 low-performing (13 females, mean age = 69.6, $SD = 4.4$) older adults. The groups did not differ with respect to age, $F < 1$, or education, Mann-Whitney- $U = 234.50$, $p_{(masymptotic)} = 0.240$.

STIMULI

Stimuli were identical to those used in Wiese (2012) and consisted of 480 gray-scale pictures showing 120 older Caucasian, 120 older eastern Asian, 120 young Caucasian, and 120 young eastern Asian faces (50% female, respectively), which were collected from diverse internet sources. Due to this stimulus selection procedure, the exact age of the persons depicted is unknown. However, stimuli have been rated for age by young participants in our previous study (rated age of young Caucasian faces = 28.61; $SD = 0.42$, young Asian faces = 29.28, $SD = 0.41$, older Caucasian faces = 66.16; $SD = 0.57$, older Asian faces = 72.29, $SD = 0.94$). All of the pictures displayed front views of neutral or moderately

¹Please note, that the above cited results obtained from fMRI and ERPs might not necessarily be discrepant, but may be reconciled, given that the N170 (unlike the N250 or N250r, see e.g., Schweinberger et al., 2004; Kaufmann et al., 2009) is most probably not generated in the fusiform gyrus. Please also note that these studies used peak measures to determine differences between age groups, which has been criticized (for alternative approaches, see Rousselet et al., 2009, 2011).

happy faces and were edited using Adobe Photoshop™ removing all information (clothing, background, etc.) apart from the face which was subsequently pasted in front of a black background. All stimuli were framed within an area of 170×216 pixels (6.0×7.6 cm), corresponding to a visual angle of $3.8^\circ \times 4.8^\circ$ at a viewing distance of 90 cm.

PROCEDURE

Prior to the recognition memory experiment, visual acuity and contrast sensitivity were measured for each participant via a computer-based test (FrACT, Version 3.5.5; Bach, 1996) at 90 cm viewing distance. Participants were asked to indicate the positions of Landolt's C gaps presented in different sizes (test for visual acuity) and gray-scales (contrast sensitivity). Visual acuity was determined by the logarithm of the minimum angle of resolution (logMAR). Contrast sensitivity was measured by Michelson Contrast scores, which refer to the difference between highest and lowest luminance values divided by the sum of the two values.

The procedure of the main experiment was identical to Wiese (2012). Participants were seated in a dimly lit, electrically shielded, and sound-attenuated chamber (400A-CT_Special, Industrial Acoustics, Niederkrüchten, Germany) with their heads in a chin rest. Approximate distance between eyes and computer screen was 90 cm. Each experimental session began with a series of practice trials on different stimuli, which were excluded from data analysis. On each trial, a face stimulus was presented for various durations (depending on study vs. test phases, see below), preceded by a fixation cross for 500 ms and followed by a blank screen for 500 ms indicating the end of a trial.

The main experiment consisted of 12 blocks, each divided into a study and a test phase. In each study phase 10 young and 10 older faces, 50% Caucasian and 50% Asian, respectively, were presented for 5000 ms each. Half of the participants were asked to categorize the face on the screen as fast and correctly as possible according to age (elderly vs. young), whereas the other half was asked to categorize the face on the screen as fast and correctly as possible according to ethnicity (Asian vs. Caucasian). Furthermore, participants were instructed to memorize the faces. Between learning and test phases a fixed break of 30 s duration was inserted. During each test phase all of the 20 faces from the directly preceding study phase and 20 new faces (50% older, 50% Asian) were presented for 2000 ms each. Participants were instructed to indicate as fast as possible and without compromising accuracy whether the faces have been encountered in the preceding study phase. Between each block, participants were allowed a self-timed period of rest. During study and test, stimuli were presented in a randomized order, and key assignment and allocation of stimuli to learned and non-learned conditions were counterbalanced across participants. During study phases, mean reaction time (RT, correct responses only) and accuracy was analyzed. Data from the test phases were sorted into hits (correctly identified studied faces), misses (studied faces incorrectly classified as new), correct rejections (CR, new faces correctly classified as new), and false alarms (FA, new faces incorrectly classified as studied). Measures of sensitivity (d') and response bias (C) were calculated (Green and Swets, 1966).

ERP RECORDING AND ANALYSIS

Thirty-two-channel EEG was recorded using a BioSemi Active II system (BioSemi, Amsterdam, Netherlands). The active sintered Ag/Ag-Cl-electrodes were mounted in an elastic cap. EEG was recorded continuously from Fz, Cz, Pz, Iz, FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, F9, F10, FT9, FT10, TP9, TP10, P9, P10, PO9, PO10, I1, I2, with a 512-Hz sample rate from DC to 155 Hz. Please note that BioSemi systems work with a “zero-Ref” set-up with ground and reference electrodes

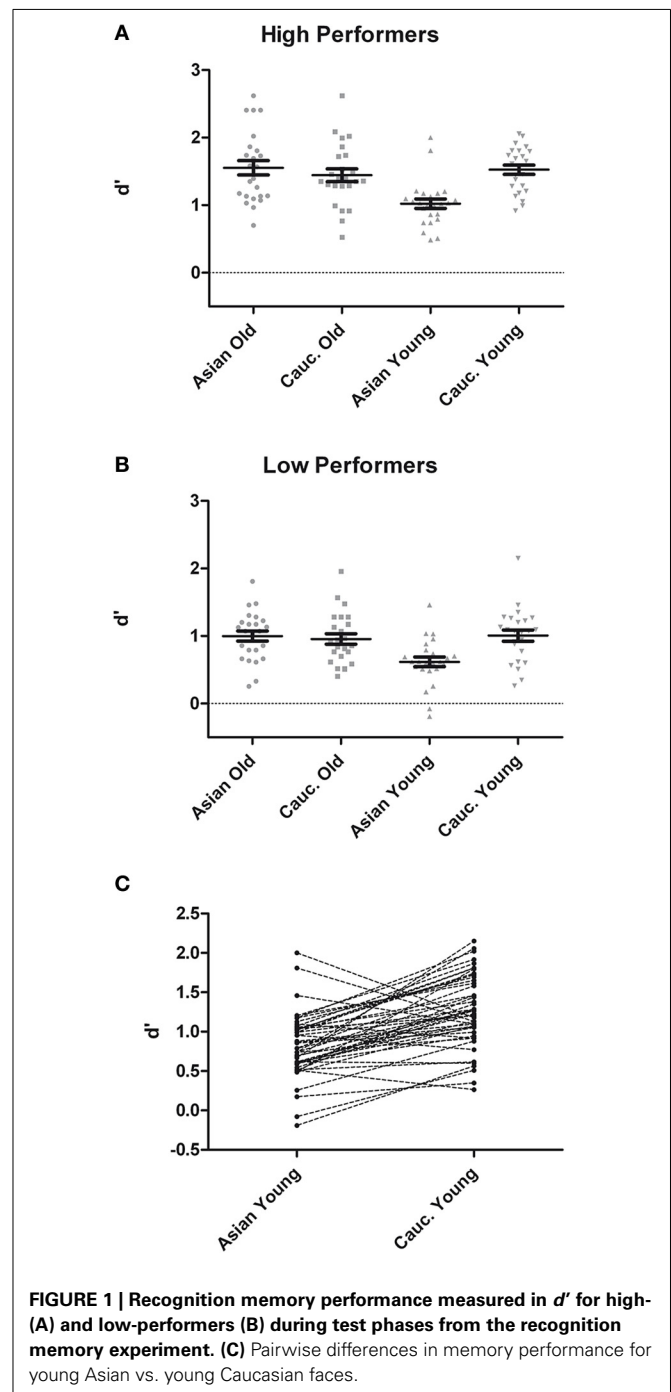


FIGURE 1 | Recognition memory performance measured in d' for high- (A) and low-performers (B) during test phases from the recognition memory experiment. (C) Pairwise differences in memory performance for young Asian vs. young Caucasian faces.

replaced by a CMS/DRL circuit (for further information, see www.biosemi.com/faq/cms&drl.htm).

Contributions of blink artifacts were corrected using the algorithm implemented in BESA 5.1 (MEGIS Software GmbH, Graefelfing, Germany). EEG was segmented from -200 until 1000 ms relative to stimulus onset, with the first 200 ms as baseline. Trials contaminated by non-ocular artifacts and saccades were rejected from further analysis. Artifact rejection was carried out using the BESA 5.1 tool, with an amplitude threshold of $100 \mu\text{V}$, as well as a gradient criterion of $75 \mu\text{V}$. Remaining trials were recalculated to average reference, digitally low-pass filtered at 40 Hz (12 dB/oct, zero phase shift), and averaged according to the following four experimental conditions during learning for the first six study blocks (see below): young Asian, young Caucasian, elderly Asian, elderly Caucasian. The mean number of trials contributing to an individual averaged ERP for these conditions was 27, 27, 26, and 27, respectively². The minimum number of trials contributing to an individual waveform was 16.

In the resulting waveforms, mean amplitudes for P1 were determined at O1/O2 and between 110 and 160 ms, for N170 between 180 and 220 ms, and for a subsequent time window from 220 to 260 ms at TP9/P9/PO9 and TP10/P10/PO10. Extensive analyses were also performed throughout the entire epoch, which can be found in the Supplementary Material. Statistical analyses

²Please note that only trials from the first half of the experiment were analyzed, see Results section.

were performed by calculating mixed-model analyses of variance (ANOVA), with degrees of freedom corrected according to the Greenhouse-Geisser procedure where appropriate.

RESULTS

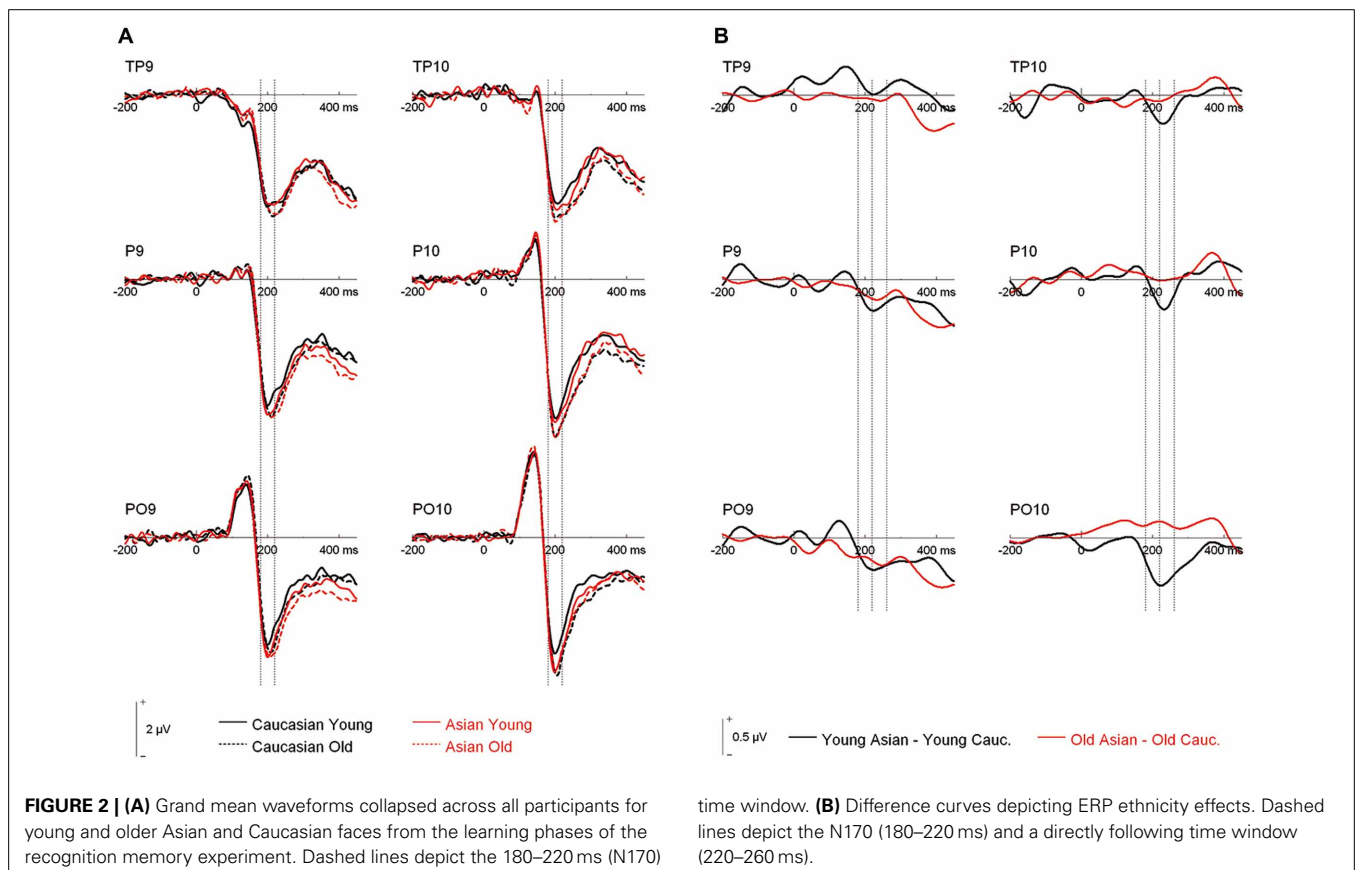
VISUAL ACUITY/CONTRAST VISION

An ANOVA on the logMAR measure of visual acuity with the between-subjects factor group (high- vs. low-performers) revealed no significant difference, $F_{(1, 46)} = 1.73$, $p = 0.195$, $\eta_p^2 = 0.04$. Similarly, an ANOVA on the Michelson Contrast indicated no group difference, $F < 1$.

PERFORMANCE

Given the length of the experiment (>70 min plus preparation times for EEG) and frequent reports of exhaustion toward the end of the session, we conducted an initial ANOVA on d' with the factor "block" ($1-6$ vs. $7-12$) to test for potential effects of fatigue. A significant block effect, $F = 6.85$, $p = 0.012$; $\eta_p^2 = 0.13$, indicated decreased performance in the second half of the experiment. As overall performance was a critical aspect, we decided to analyze data in the first six blocks only, to avoid contamination of any effects by fatigue.

Analysis of d' at test revealed main effects of group, $F_{(1, 46)} = 55.42$, $p < 0.001$, $\eta_p^2 = 0.55$, face ethnicity, $F_{(1, 46)} = 11.06$, $p = 0.002$, $\eta_p^2 = 0.19$, face age, $F_{(1, 46)} = 11.62$, $p = 0.001$, $\eta_p^2 = 0.20$, and a significant interaction of Face Ethnicity \times Face Age, $F_{(1, 46)} = 19.85$, $p < 0.001$, $\eta_p^2 = 0.30$. *Post-hoc* tests indicated



better recognition for young Caucasian vs. young Asian faces, $t_{(48)} = 6.86$, $p < 0.001$, $d = 1.05$, but no difference between older Caucasian and Asian faces, $t_{(48)} = 1.18$, $p = 0.245$, $d = 0.20$. Furthermore, recognition was more accurate for older Asian as compared to young Asian faces, $t_{(48)} = 5.66$, $p < 0.001$, $d = 0.98$, whereas there was no difference between older Caucasian and young Caucasian faces, $t_{(48)} = 1.20$, $p = 0.237$, $d = 0.19$. Importantly, these effects were not modulated by group, $F < 1$ (see **Figure 1**). As the division of participants into high and low performing groups may have led to a loss of statistical power due to dichotomization of the continuous memory score (see e.g., Cohen, 1983), we additionally calculated a correlation between overall d' and the ORB for young faces, which revealed no significant result, $r = 0.07$, $p = 0.62$.

For the sake of completeness, we also analyzed response bias (C) but consider the finding as Supplementary Material.

EVENT-RELATED POTENTIALS

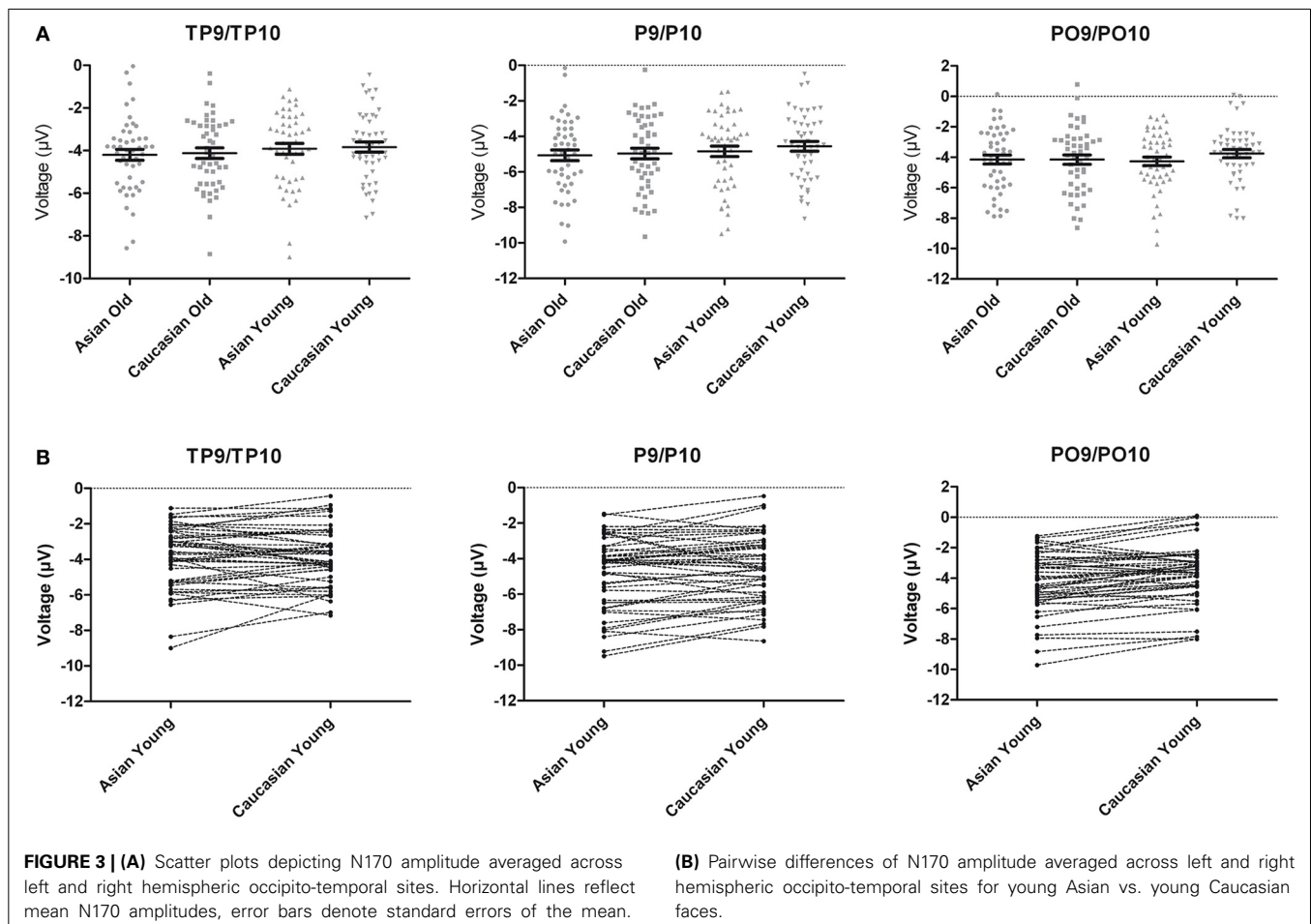
We analyzed the study phase ERPs, as previous research suggested the study phase N170 as a neural correlate of the ORB (Wiese et al., 2014). Note that effects involving topographic factors (hemisphere, site) are only reported when interacting with experimental factors. In accordance with our behavioral data

analysis strategy, we restricted our ERP analyses to the first six blocks of the experiment.

A mixed-model ANOVA on P1 with the within-subject factors hemisphere (left, right), face ethnicity, and face age, and the between-subjects factor group resulted in no significant effects, $F \leq 3.62$, $p \geq 0.064$, $\eta_p^2 \leq 0.07$.

An ANOVA on N170 with the within-subject factors hemisphere, site (TP/P/PO), face ethnicity and face age, and the between-subjects factor group resulted in effects of face ethnicity, $F_{(1, 46)} = 5.77$, $p = 0.020$, $\eta_p^2 = 0.11$, with larger N170 for Asian faces, and face age, $F_{(1, 46)} = 8.88$, $p = 0.005$, $\eta_p^2 = 0.16$, with larger N170 for old faces, as well as in an interaction of Face ethnicity \times Face age \times Site, $F_{(1.59, 73.01)} = 3.61$, $p = 0.042$, $\eta_p^2 = 0.073$. Separate *post-hoc* analyses for each site and ethnicity yielded larger amplitudes for young Asian as compared to young Caucasian faces at P9/P10, $F_{(1, 47)} = 4.46$, $p = 0.040$, $\eta_p^2 = 0.09$ and PO9/PO10, $F_{(1, 47)} = 15.45$, $p < 0.001$, $\eta_p^2 = 0.25$, but not at TP9/TP10, $F < 1$. No corresponding effects were detected for older faces, all $F < 1$ (see **Figures 2, 3**).

Additionally, the three-way interaction of Site \times Hemisphere \times Group, $F_{(1.44, 66.12)} = 3.56$, $p = 0.048$, $\eta_p^2 = 0.072$, was significant (see **Figures 4, 5**). Separate ANOVAs at each site and for high- and low-performers



were suggestive of a right-lateralized N170 in the low-performers at TP9/TP10, $F_{(1, 23)} = 3.27$, $p = 0.084$, $\eta_p^2 = 0.13$, and at P9/P10, $F_{(1, 23)} = 2.86$, $p = 0.104$, $\eta_p^2 = 0.11$. By contrast, there was no evidence of right lateralization in high-performers, $F_{(1, 23)} = 2.05$, $p = 0.166$, $\eta_p^2 = 0.08$, who even exhibited numerically larger N170 over the left hemisphere at the more anterior electrode sites (e.g., for TP9/TP10; see **Figure 4**).

An additional analysis on N170 amplitude was carried out for the second half of the experiment to test whether group differences in N170 lateralization would hold with increasing fatigue. While the general pattern was similar to the first half (with larger left hemispheric amplitudes at more anterior sites for high relative to low performers), the respective mixed-model ANOVA revealed no significant interaction of Hemisphere \times Site \times Group, $F_{(1.47, 67.43)} = 2.19$, $p = 0.134$, $\eta_p^2 = 0.045$.

Moreover, given the continuous nature of the memory scores, we performed additional correlation analyses to establish a potential relationship between memory performance and ERP effects. First, to test whether the N170 ethnicity effect for young faces was related to overall performance, we calculated difference scores for N170 amplitude (Young Asian – Young Caucasian) at those electrode sites with significant effects in ANOVA reported above (P and PO). This measure was not correlated with overall performance, all $r_{(46)} < 0.23$, all $p > 0.14$, confirming the results

from the median split analysis. Second, N170 lateralization scores (right hemispheric – left hemispheric amplitudes across experimental conditions) were calculated for TP, P, and PO sites separately. This measure correlated significantly with overall memory scores at TP, $r_{(46)} = 0.29$, $p = 0.046^3$, but not at P, $r_{(46)} = 0.20$, $p = 0.179$, and PO sites, $r_{(46)} = 0.08$, $p = 0.586$ (see **Figure 5B**), again confirming the results from the median split analysis.

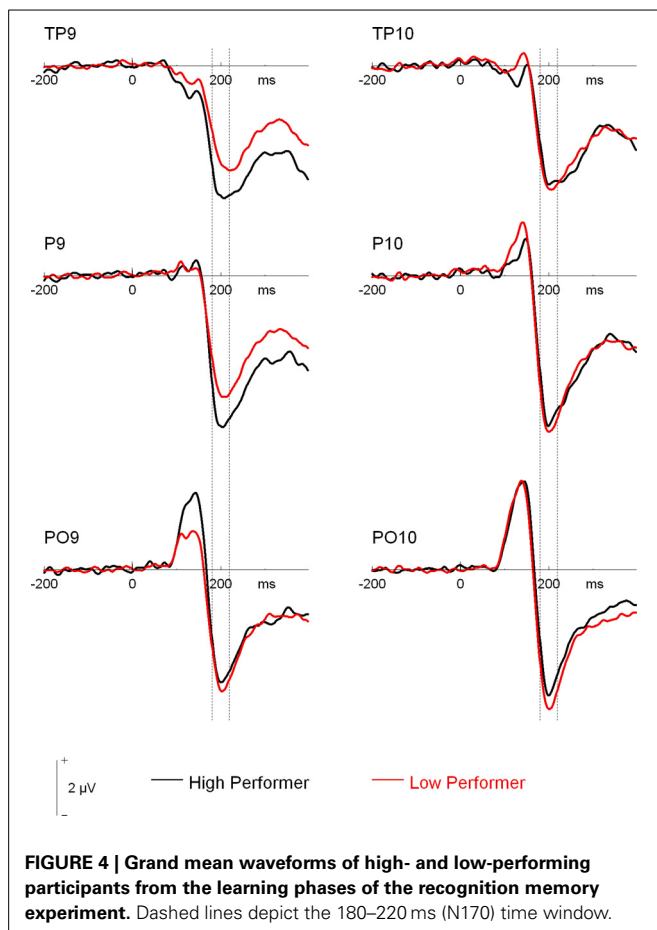
Following inspection of the ERP difference curves for the ethnicity effects (see **Figure 2B**), we performed an additional analysis using a time window from 220 to 260 ms. This analysis revealed main effects of face age, $F_{(1, 46)} = 44.74$, $p < 0.001$, $\eta_p^2 = 0.49$, and ethnicity, $F_{(1, 46)} = 8.01$, $p = 0.007$, $\eta_p^2 = 0.15$, which were qualified by a significant interaction of hemisphere \times age \times ethnicity, $F_{(1, 46)} = 4.74$, $p = 0.035$, $\eta_p^2 = 0.09$. Follow-up tests revealed significantly more negative amplitudes for young Asian relative to young Caucasian faces over the right, $F_{(1, 47)} = 9.79$, $p = 0.003$, $\eta_p^2 = 0.17$, but not over the left hemisphere, $F_{(1, 47)} = 1.65$, $p = 0.206$, $\eta_p^2 = 0.03$ (see **Figure 6**). No significant ethnicity effects were detected for old faces, neither over the left, $F_{(1, 47)} = 3.35$, $p = 0.074$, $\eta_p^2 = 0.07$, nor over the right hemisphere, $F < 1$. No significant effects involving the group factor were detected, all $F < 2.03$, all $p > 0.15$. Moreover, further analyses revealed no significant correlations between the lateralization scores and overall d' , TP9/TP10: $r = 0.27$, $p = 0.062$, P9/P10: $r = 0.23$, $p = 0.12$, PO9/PO10: $r = 0.11$, $p = 0.47$. Similarly, the ethnicity effect for young faces over the right hemisphere did not correlate with overall d' , $r = -0.06$, $p = 0.68$, supporting the above finding of no relation between ERP ethnicity effects and overall memory performance.

We also conducted extensive data analyses on the later P2 (260–400 ms), N2 (400–600 ms), and a late slow-wave (600–1000 ms) at occipito-temporal electrode sites. These analyses essentially replicated previous findings, with a prominent left lateralized ethnicity effect starting in the P2-time range, and a right lateralized age effect starting in the N2-time range (see Wiese, 2012). As these findings were not of primary interest for the present study, we refrain from a detailed report in the main paper (please see Supplementary Material).

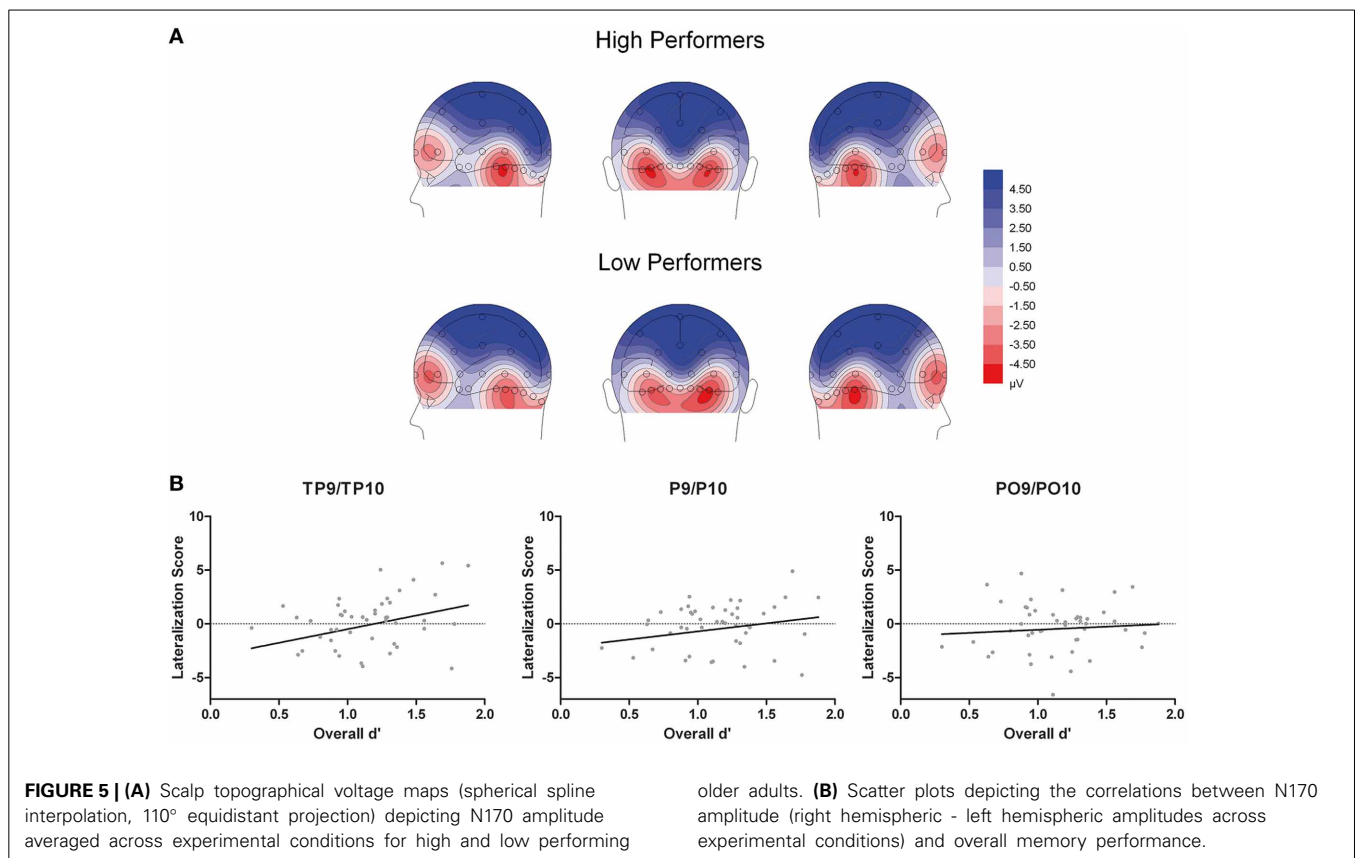
DISCUSSION

The present study examined the ORB in face recognition memory in high- and low-performing older adults. Despite prominent overall performance differences, both groups demonstrated an equivalent magnitude of the ORB. Paralleling these behavioral results, ERP effects of face ethnicity first showed up in a larger N170 to young Asian vs. Caucasian faces in both groups alike. Additionally, performance-related ERP differences were observed between groups: While low-performing older adults showed a tendency for a right-lateralization of N170, this was not the case in high-performers. These findings are discussed in more detail below.

Our first aim was to test whether poor overall performance would be accompanied by dedifferentiation of processing within



³Please note that this was also significant when calculating a Spearman correlation, $Rho = 0.29$, $p = 0.048$, which is less sensitive to outliers as it uses ranked data.



the category of faces, in terms of a reduced ORB in low-performers. Such a finding would complement previous reports of attenuated between-category (faces vs. objects) neural distinctiveness in older adults (Park et al., 2004, 2012), by indicating a degree of loss in the fine-tuning of the face processing system. We observed a clear ORB for young faces, which was virtually identical in both high- and low-performing participants. We therefore conclude that the fine-tuning of the face processing system of low-performers was not compromised when compared to high-performers, despite their reduced overall face memory. This conclusion is also in line with our finding of similar ethnicity effects in the N170 and the directly following time window in the two older groups (convergent with the additional result of no significant correlation of the ERP ethnicity effects with overall memory performance in both time segments), which further suggests similar encoding of facial race in the older groups.

Our observation of a preserved fine-tuning of the face processing system in older adults is generally in line with findings from an individual difference perspective. Hildebrandt et al. (2011) observed no dedifferentiation of individual differences in the latent constructs of face cognition as compared to general cognitive functioning. This finding was further substantiated in a subsequent study comparing object and face processing, in which only a slight dedifferentiation was observed, which was not specific for faces but operated on a more general cognitive level (Hildebrandt et al., 2013). In the present study, the behavioral ORB was equivalent between groups, although overall memory was reduced in low performers. Hence, we suggest that low

performers did not exhibit poor face memory due to less specialized face processing, but rather due to more general cognitive decline.

Of note, although a previous study from our group found a significant correlation between the N170 ethnicity effect and the behavioral ORB (Wiese et al., 2014), two studies using a latent variable approach reported no relationship between the amplitude of the N170 and face cognition (Herzmann et al., 2010; Kaltwasser et al., in press). Yet, the latent variable approach produces combined measures of a number of tests, which may well tap into different processes than those reflected by N170 amplitude. More precisely, N170 is typically thought to represent very early perceptual processing stages (such as structural encoding or the detection of a face-like pattern, see e.g., Amihai et al., 2011; Eimer, 2011). Thus, less efficient processing of other-race faces at this early processing stage appears to result in a cascade of less efficient subsequent processes, which may ultimately result in less accurate recognition memory. Importantly, none of the tests used in the studies by Herzmann et al. (2010) or Kaltwasser et al. (in press) examined the processing of other-race faces. Early perceptual processing stages may have been highly efficient in all tasks used by Herzmann and others (Herzmann et al., 2010; Kaltwasser et al., in press), but not in the other-race conditions of the present and our previous study.

At a more specific level it is notable that ethnicity effects were prominent for young but not old faces, both in the N170 component and the subsequent time window, as well as in the behavioral ORB. The absence of an ORB for old faces in memory replicates

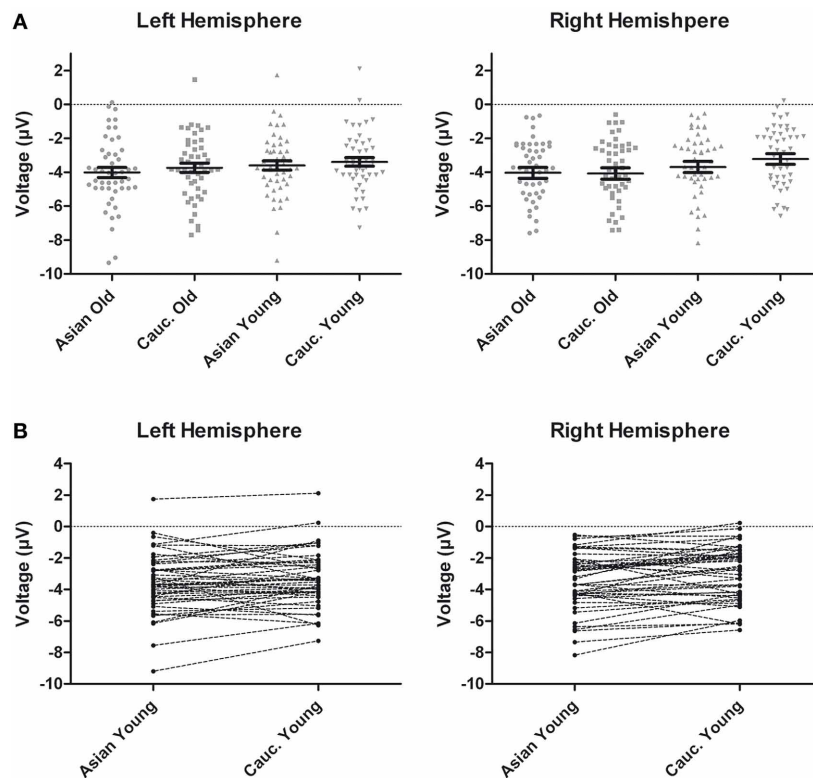


FIGURE 6 | (A) Scatter plots depicting ERP amplitudes in the 220–260 ms time window averaged across left and right hemispheric occipito-temporal sites. Horizontal lines reflect mean amplitudes, error bars denote standard

errors of the mean. **(B)** Pairwise differences of ERP amplitudes in the 220–260 ms time window averaged across left and right hemispheric occipito-temporal sites for young Asian vs. young Caucasian faces.

two recent studies (Wallis et al., 2012; Wiese, 2012). We attribute this consistent absence of an ORB for old faces to decreased perceptual salience of ethnicity information in these faces, along with an increased salience of general age-related changes in facial shape and skin texture. The second aim was to examine whether the previously observed larger N170 amplitude in older as compared to young adults (e.g., Wiese et al., 2013a) was related to a decrease in early perceptual processing, which would in turn result in reduced face memory in older adults. This idea was based on the N170 inversion effect, reflecting an increased N170 amplitude for inverted, less configurally processed faces (e.g., Jacques et al., 2007), an effect which is reduced in older participants (Gao et al., 2009; Daniel and Bentin, 2010; Saavedra et al., 2012). In the present study, a decrease in face memory was not associated with larger N170 amplitudes, arguing against the idea of less efficient early face perception as the basis for reduced memory. If anything, N170 was larger in high- relative to low-performing participants, which, however, was apparent only over the left hemisphere.

Related to this latter finding, we detected a significant interaction involving the factors hemisphere and group, which suggested a more right-lateralized N170 in low-performers and a more bilateral pattern in high-performers. This finding was additionally confirmed in a correlation analysis, in which a more bilateral N170 was associated with higher overall performance. N170 is typically lateralized to the right hemisphere in young participants (e.g., Bentin et al., 1996; Amihai et al., 2011; Eimer, 2011),

whereas a reduced lateralization of N170 amplitude in older adults has been previously described (Gao et al., 2009; Daniel and Bentin, 2010). Based on previous neuroimaging results (e.g., Cabeza, 2002; Cabeza et al., 2002) this finding has been interpreted as reflecting compensation of age-related decrements in face perception (Gao et al., 2009; Daniel and Bentin, 2010). The present study adds to this idea, by revealing that a reduced laterality of N170 can indeed relate to performance levels. Our finding is generally in line with the idea of reduced hemispheric asymmetry as a function of compensation rather than dedifferentiation (for reviews, see Cabeza, 2002; Reuter-Lorenz and Cappell, 2008). This may not only occur for higher-order cognitive processes, but also for (early) perceptual processes (De Sanctis et al., 2008) and the present findings extend this idea to the domain of face processing. However, although a similar pattern compared to the first half was observed, the interaction of hemisphere by group was no longer significant in the second half of the experiment, in which overall performance was clearly reduced in both groups. This finding may suggest that compensatory neural activity requires effortful processing and that fatigue hampers such activity in high performing older adults.

Although the more bilateral N170 in high-performers argues for a compensatory mechanism, the specific interpretation of their enhanced left-hemispheric N170 is subject to debate. One possibility considers that the left hemisphere may be more involved in feature-based than configural or holistic face

processing (Rossion et al., 2000; Scott and Nelson, 2006). Accordingly, high-performers may engage in more feature-based processing during encoding, enabling them to exhibit better memory performance than low-performers at test. While such a strategy could provide more effective encoding fostering clearer subsequent representations, it is noteworthy that even high-performers in the present study perform at levels clearly below those of young participants in an identical experiment (see Wiese, 2012). Accordingly, even high-performers may not be able to fully compensate for age-related decline in face memory.

The present study is, to our knowledge, the first to examine ERP correlates of the ORB in older adult participants. The finding of an ORB independent of overall performance indicates that the fine-tuning of the face processing system toward faces of particular expertise is preserved in older adults. In line with this interpretation, the specific pattern of N170 ethnicity effects was found to parallel the behavioral ORB. In addition, a more bilateral N170 response in high-performing older adults suggests a partial compensation for general age-related decline in face perception by recruiting additional neural resources in the left hemisphere. In conclusion, the present results indicate that older adults' face processing system, even when working less efficiently, may still exhibit preserved expertise-related specialization toward own-race faces.

ACKNOWLEDGMENTS

This work was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) to Holger Wiese (Wi 3219/4-2). The authors gratefully acknowledge help during data collection by Kathrin Rauscher, Kristin Oehler, Franziska Krahmer, and Julia Festini.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnagi.2014.00060/abstract>

REFERENCES

- Amihai, I., Deouell, L. Y., and Bentin, S. (2011). Neural adaptation is related to face repetition irrespective of identity: a reappraisal of the N170 effect. *Exp. Brain Res.* 209, 193–204. doi: 10.1007/s00221-011-2546-x
- Anderson, N. D., and Craik, F. I. M. (2000). "Memory in the aging brain," in *The Oxford Handbook of Memory*, eds E. Tulving and F. I. M. Craik (Oxford: Oxford University Press, Inc.), 411–425.
- Bach, M. (1996). The Freiburg Visual Acuity test-automatic measurement of visual acuity. *Optom. Vis. Sci.* 73, 49. doi: 10.1097/00006324-199601000-00008
- Balas, B., and Nelson, C. A. (2010). The role of face shape and pigmentation in other-race face perception: an electrophysiological study. *Neuropsychologia* 48, 498–506. doi: 10.1016/j.neuropsychologia.2009.10.007
- Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12
- Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565. doi: 10.1162/jocn.1996.8.6.551
- Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol. Aging* 17, 85–100. doi: 10.1037/0882-7974.17.1.85
- Cabeza, R., Anderson, N. D., Locantore, J. K., and McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Cabeza, R., Nyberg, L., and Park, D. C. (2004). "Cognitive neuroscience of aging: the birth of a new discipline," in *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging*, eds R. Cabeza, L. Nyberg, and D. C. Park (New York: Oxford University Press), 3–15.
- Caharel, S., Montalan, B., Fromager, E., Bernard, C., Lalonde, R., and Mohamed, R. (2011). Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: an event-related brain potential study. *Int. J. Psychophysiol.* 79, 266–271. doi: 10.1016/j.ijpsycho.2010.10.018
- Chaby, L., Narme, P., and George, N. (2011). Older adults' configural processing of faces: role of second-order information. *Psychol. Aging* 26, 71–79. doi: 10.1037/a0020873
- Chee, M. W. L., Goh, J. O. S., Venkatraman, V., Tan, J. C., Gutchess, A., Sutton, B., et al. (2006). Age-related changes in object processing and contextual binding revealed using fMRI adaptation. *J. Cogn. Neurosci.* 18, 495–507. doi: 10.1162/jocn.2006.18.4.495
- Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P., et al. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: individual differences in change scores as a function of age. *Psychol. Aging* 14, 365–379. doi: 10.1037/0882-7974.14.3.365
- Cohen, J. (1983). The cost of dichotomization. *Appl. Psychol. Meas.* 7, 249–253. doi: 10.1177/014662168300700301
- Daniel, S., and Bentin, S. (2010). Age-related changes in processing faces from detection to identification: ERP evidence. *Neurobiol. Aging* 33, 1–28. doi: 10.1016/j.neurobiolaging.2010.09.001
- Davis, S., Kragel, J., Madden, D., and Cabeza, R. (2012). The architecture of cross-hemispheric communication in the aging brain: linking behavior to functional and structural connectivity. *Cereb. Cortex* 22, 232–242. doi: 10.1093/cercor/bhr123
- Dennis, N., and Cabeza, R. (2008). "Neuroimaging of healthy cognitive aging," in *The Handbook of Aging and Cognition*, Vol. 3, eds F. I. M. Craik and T. A. Salthouse (New York, NY: Psychology Press), 1–54.
- De Sanctis, P., Katz, R., Wylie, G. R., Sehatpour, P., Alexopoulos, G. S., and Foxe, J. J. (2008). Enhanced and bilateralized visual sensory processing in the ventral stream may be a feature of normal aging. *Neurobiol. Aging* 29, 1576–1586. doi: 10.1016/j.neurobiolaging.2007.03.021
- Duarte, A., Ranganath, C., Trujillo, C., and Knight, R. T. (2006). Intact recollection memory in high-performing older adults: ERP and behavioral evidence. *J. Cogn. Neurosci.* 18, 33–47. doi: 10.1162/089892906775249988
- Eimer, M. (2011). "The face-sensitive N170 component of the event-related brain potential," in *The Oxford Handbook of Face Perception*, eds A. J. Calder, G. Rhodes, M. H. Johnson, and J. V. Haxby (Oxford: Oxford University Press), 339–344.
- Federmeier, K. D., Kutas, M., and Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain Lang.* 115, 149–161. doi: 10.1016/j.bandl.2010.07.006
- Friedman, D., de Chastelaine, M., Nessler, D., and Malcolm, B. (2010). Changes in familiarity and recollection across the lifespan: an ERP perspective. *Brain Res.* 1310, 124–141. doi: 10.1016/j.brainres.2009.11.016
- Gao, L., Xu, J., Zhang, B., Zhao, L., Harel, A., and Bentin, S. (2009). Aging effects on early-stage face perception: an ERP study. *Psychophysiology* 46, 970–983. doi: 10.1111/j.1469-8986.2009.00853.x
- Grady, C. L. (2008). "Cognitive neuroscience of aging," in *Year in Cognitive Neuroscience 2008*, Vol. 1124, eds A. Kingstone and M. B. Miller (Oxford, UK: Blackwell Publishing), 127–144.
- Green, D., and Swets, J. (1966). *Signal Detection Theory and Psychophysics* (A reprint, with corrections of the original 1966 ed.). Huntington, NY: Robert E. Krieger Publishing Co.
- Henry, J. D., MacLeod, M. S., Phillips, L. H., and Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychol. Aging* 19, 27–39. doi: 10.1037/0882-7974.19.1.27
- Herzmann, G., Kunina, O., Sommer, W., and Wilhelm, O. (2010). Individual differences in face cognition: brain-behavior relationships. *J. Cogn. Neurosci.* 22, 571–589. doi: 10.1162/jocn.2009.21249
- Hildebrandt, A., Wilhelm, O., Herzmann, G., and Sommer, W. (2013). Face and object cognition across adult age. *Psychol. Aging* 28, 243–248. doi: 10.1037/a0031490
- Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G., and Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychol. Aging* 26, 701–715. doi: 10.1037/a0023056

- Jacques, C., d'Arripe, O., and Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *J. Vis.* 7, 1–9. doi: 10.1167/7.8.3
- Johnson, M. H. (2001). Functional brain development in humans. *Nat. Rev. Neurosci.* 2, 475–483. doi: 10.1038/35081509
- Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O., and Sommer, W. (in press). Neurocognitive mechanisms of individual differences in face cognition: a replication and extension. *Cogn. Affect. Behav. Neurosci.* doi: 10.3758/s13415-013-0234-y
- Kaufmann, J. M., Schweinberger, S. R., and Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *J. Cogn. Neurosci.* 21, 625–641. doi: 10.1162/jocn.2009.21080
- Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., and Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychol. Sci.* 15, 155–163. doi: 10.1111/j.0956-7976.2004.01503003.x
- Lindenberger, U., and Baltes, P. B. (1994). Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339–355. doi: 10.1037/0882-7974.9.3.339
- Lindenberger, U., Scher, H., and Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: not due to sensory acuity reductions operating during cognitive assessment. *Psychol. Aging* 16, 196–205. doi: 10.1037//0882-7974.16.2.196
- Meissner, C. A., and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces - A meta-analytic review. *Psychol. Public Policy Law* 7, 3–35. doi: 10.1037/1076-8971.7.1.3
- Nyberg, L., Lovden, M., Riklund, K., Lindenberger, U., and Backman, L. (2012). Memory aging and brain maintenance. *Trends Cogn. Sci.* 16, 292–305. doi: 10.1016/j.tics.2012.04.005
- Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., et al. (2010). Longitudinal evidence for diminished frontal cortex function in aging. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22682–22686. doi: 10.1073/pnas.1012651108
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Park, D. C., Polk, T. A., Park, R., Minear, M., Savage, A., and Smith, M. R. (2004). Aging reduces neural specialization in ventral visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13091–13095. doi: 10.1073/pnas.0405148101
- Park, D. C., and Reuter-Lorenz, P. A. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* 60, 173–196. doi: 10.1146/annurev.psych.59.103006.093656
- Park, J., Carp, J., Kennedy, K. M., Rodrigue, K. M., Bischof, G. N., Huang, C.-M., et al. (2012). Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the face network in a large lifespan sample. *J. Neurosci.* 32, 2154–2158. doi: 10.1523/jneurosci.4494-11.2012
- Pascalis, O., de Haan, M., and Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science* 296, 1321–1323. doi: 10.1126/science.1070223
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb. Cortex* 15, 1676–1689. doi: 10.1093/cercor/bhi044
- Reuter-Lorenz, P., and Cappell, K. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Reuter-Lorenz, P., and Park, D. C. (2010). Human neuroscience and the aging mind: at old problems a new look. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 65, 405–415. doi: 10.1093/geronb/gbq035
- Rhodes, M. G., and Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychol. Bull.* 138, 146–174. doi: 10.1037/a0025750
- Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., de Gelder, B., et al. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *J. Cogn. Neurosci.* 12, 793–802. doi: 10.1162/089892900562606
- Rousset, G. A., Gaspar, C. M., Wiczorek, K. P., and Pernet, C. R. (2011). Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). *Front. Psychol.* 2:137. doi: 10.3389/fpsyg.2011.00137
- Rousset, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC Neuroscience* 10:114. doi: 10.1186/1471-2202-10-114
- Saavedra, C., Olivares, E. I., and Iglesias, J. (2012). Cognitive decline effects at an early stage: evidence from N170 and VPP. *Neurosci. Lett.* 518, 149–153. doi: 10.1016/j.neulet.2012.04.075
- Salthouse, T. A. (2013). Within-cohort age-related differences in cognitive functioning. *Psychol. Sci.* 24, 123–130. doi: 10.1177/0956797612450893
- Scherf, K. S., Behrmann, M., Humphreys, K., and Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* 10, F15–F30. doi: 10.1111/j.1467-7687.2007.00595.x
- Scherf, K. S., and Scott, L. S. (2012). Connecting developmental trajectories: biases in face processing from infancy to adulthood. *Dev. Psychobiol.* 54, 643–663. doi: 10.1002/dev.21013
- Schweinberger, S. R., Huddy, V., and Burton, A. M. (2004). N250r: a face-selective brain response to stimulus repetitions. *Neuroreport* 15, 1501–1505. doi: 10.1097/01.wnr.0000131675.00319.42
- Scott, L. S., and Nelson, C. A. (2006). Featural and configural face processing in adults and infants: a behavioral and electrophysiological investigation. *Perception* 35, 1107–1128. doi: 10.1068/p5493
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., and Toga, A. W. (2003). Mapping cortical change across the human life span. *Nat. Neurosci.* 6, 309–315. doi: 10.1038/nn1008
- Stahl, J., Wiese, H., and Schweinberger, S. R. (2010). Learning task affects ERP-correlates of the own-race bias, but not recognition memory performance. *Neuropsychologia* 48, 2027–2040. doi: 10.1016/j.neuropsychologia.2010.03.024
- Tanaka, J., Heptonstall, B., and Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. *Vis. Cogn.* 21, 1183–1201. doi: 10.1080/13506285.2013.826315
- Valentine, T., and Endo, M. (1992). Towards an exemplar model of face processing: the effects of race and distinctiveness. *Q. J. Exp. Psychol. A* 44, 671–703. doi: 10.1080/14640749208401305
- Wallis, J., Lipp, O. V., and Vanman, E. J. (2012). Face age and sex modulate the other-race effect in face recognition. *Atten. Percept. Psychophys.* 74, 1712–1721. doi: 10.3758/s13414-012-0359-z
- Werkle-Bergner, M., Shing, Y. L., Mueller, V., Li, S.-C., and Lindenberger, U. (2009). EEG gamma-band synchronization in visual coding from childhood to old age: evidence from evoked power and inter-trial phase locking. *Clin. Neurophysiol.* 120, 1291–1302. doi: 10.1016/j.clinph.2009.04.012
- Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biol. Psychol.* 89, 137–147. doi: 10.1016/j.biopsycho.2011.10.002
- Wiese, H., Kachel, U., and Schweinberger, S. (2013a). Holistic face processing of own- and other-age faces in young and older adults: ERP evidence from the composite face task. *Neuroimage* 74, 306–317. doi: 10.1016/j.neuroimage.2013.02.051
- Wiese, H., Kaufmann, J., and Schweinberger, S. R. (2014). The neural signature of the own-race bias: evidence from event-related potentials. *Cereb. Cortex* 24, 826–835. doi: 10.1093/cercor/bhs369
- Wiese, H., Komes, J., and Schweinberger, S. R. (2013b). Aging faces in aging minds: a review on the own-age bias in face recognition. *Vis. Cogn.* 21, 1337–1363. doi: 10.1080/13506285.2013.823139

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 January 2014; accepted: 18 March 2014; published online: 04 April 2014.
 Citation: Komes J, Schweinberger SR and Wiese H (2014) Preserved fine-tuning of face perception and memory: evidence from the own-race bias in high- and low-performing older adults. *Front. Aging Neurosci.* 6:60. doi: 10.3389/fnagi.2014.00060
 This article was submitted to the journal *Frontiers in Aging Neuroscience*.
 Copyright © 2014 Komes, Schweinberger and Wiese. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Reduced audiovisual recalibration in the elderly

Yu Man Chan, Michael J. Pianta and Allison M. McKendrick*

Department of Optometry and Vision Sciences, University of Melbourne, Parkville, VIC, Australia

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Eugenie Roudaia, Trinity College Dublin, Ireland

Denton DeLoss, University of California, Riverside, USA

***Correspondence:**

Allison M. McKendrick, Department of Optometry and Vision Sciences, University of Melbourne, Parkville, VIC 3010, Australia
e-mail: allisonm@unimelb.edu.au

Perceived synchrony of visual and auditory signals can be altered by exposure to a stream of temporally offset stimulus pairs. Previous literature suggests that adapting to audiovisual temporal offsets is an important recalibration to correctly combine audiovisual stimuli into a single percept across a range of source distances. Healthy aging results in synchrony perception over a wider range of temporally offset visual and auditory signals, independent of age-related unisensory declines in vision and hearing sensitivities. However, the impact of aging on audiovisual recalibration is unknown. Audiovisual synchrony perception for sound-lead and sound-lag stimuli was measured for 15 younger (22–32 years old) and 15 older (64–74 years old) healthy adults using a method-of-constant-stimuli, after adapting to a stream of visual and auditory pairs. The adaptation pairs were either synchronous or asynchronous (sound-lag of 230 ms). The adaptation effect for each observer was computed as the shift in the mean of the individually fitted psychometric functions after adapting to asynchrony. Post-adaptation to synchrony, the younger and older observers had average window widths (\pm standard deviation) of 326 (\pm 80) and 448 (\pm 105) ms, respectively. There was no adaptation effect for sound-lead pairs. Both the younger and older observers, however, perceived more sound-lag pairs as synchronous. The magnitude of the adaptation effect in the older observers was not correlated with how often they saw the adapting sound-lag stimuli as asynchronous. Our finding demonstrates that audiovisual synchrony perception adapts less with advancing age.

Keywords: audiovisual, multisensory, synchrony judgement, aging, adaptation

INTRODUCTION

It is important to correctly combine visual and auditory signals to obtain a coherent percept of events occurring in our surrounds. However, this is a non-trivial task due to the relative difference in the transmission speed of visual and auditory signals in air, and in the nervous system (King and Palmer, 1985). At a distance of around ten meters, the slower speed of sound in air is compensated by the faster speed of sound processing in the neural system, thus visual and auditory signals arrive at a common brain area at the same time. For source distances within ten meters, there is an increase in auditory-lead with decreasing source distance. Beyond ten meters, the amount of auditory-lag within an audiovisual signal increases with increasing source distance (e.g., lightning and thunder). Audiovisual processing needs to be adaptable to accommodate the different arrival times at different viewing/hearing distances. A common real world example is spectator sports (for example, tennis) where, when watching from the top of the stands, there is an asynchrony between the visual image of the racquet hitting the ball and the sound of the contact. This perceived audiovisual asynchrony is typically only noticeable for a brief period, and is no longer noticed as the game proceeds. The ability to adapt to crossmodal asynchrony is important for correctly relating events across different distances (Heron et al., 2007; Parsons et al., 2013).

Previous work using a range of different stimulus types has demonstrated such a shift in audiovisual synchrony perception after adapting to audiovisual asynchrony (Fujisaki et al., 2004;

Vroomen et al., 2004; van Eijk et al., 2008; Vatakis et al., 2008). A classic study by Fujisaki et al. (2004) presented a continuous stream of asynchronous (auditory-lead or -lag) flash-click stimuli to young healthy participants for 3 min. The participants were then asked to judge if a subsequent audiovisual stimulus pair was synchronous or asynchronous. By measuring synchrony judgements across a range of stimulus onset asynchronies before and after the adaptation, the authors reported a shift in perceived synchrony in the direction of the adapted asynchrony. In other words, some stimuli that were perceived as asynchronous before adaptation were perceived as synchronous after the short-term adaptation. A similar shift in synchrony perception occurs for more complex and natural stimuli (Fujisaki et al., 2004; van Eijk et al., 2008; Vatakis et al., 2008; Asakawa et al., 2009, 2012; Tanaka et al., 2009, 2011).

Older people are more likely to perceive synchrony, or are more likely to have trouble separating temporally offset visual and auditory signals that are not relevant to each other (Hay-McCutcheon et al., 2009; DeLoss et al., 2013; Chan et al., 2014). We have recently shown that this observation cannot be entirely accounted for by an age-related reduction in unisensory detection thresholds (Chan et al., 2014). We scaled stimulus visual Gabor contrast and auditory sound pip intensity to individual detection thresholds, yet the older adults still had wider audiovisual synchrony windows (average width of 224 ms) compared to the younger group (average width of 166 ms). Our findings indicate that age-related differences in the ability to separate auditory and visual signals in time are not due to peripheral

visual or hearing decline. A decrease in the ability to perceive asynchrony may predict a reduction in audiovisual asynchrony adaptation.

Besides synchrony judgements, other methods used to assess audiovisual temporal perception include the sound induced flash illusion, as well as temporal order judgements. For the former, audiovisual interaction is quantified as the susceptibility to the illusion. For the latter, participants are required to judge whether the visual or the auditory signal is presented first within an audiovisual pair. Previous studies have shown that both younger and older people are equally susceptible to the sound induced flash illusion when the flash and sound are presented 70 ms apart. However, the older group experienced the sound induced flash illusion more often than the younger group when the flash and sound signals are separated by 270 ms (Setti et al., 2011a,b). This finding has been interpreted to indicate an increased audiovisual interaction resulting in difficulty in separating temporally offset visual and auditory signals with age, consistent with reports for audiovisual synchrony judgements (Hay-McCutcheon et al., 2009; Chan et al., 2014). However, in a temporal order judgment task, Fiacconi et al. (2013) failed to find the same age effect. Love et al. (2013) and van Eijk et al. (2008) have compared the results from audiovisual synchrony judgment and temporal order judgment tasks and suggest that the two tasks tap into different underlying neural mechanisms for temporal perception (van Eijk et al., 2008; Love et al., 2013). Audiovisual synchrony judgment gives a more accurate measure of the perception of subjective simultaneity, whereas temporal order judgment provides a better measure of the smallest audiovisual asynchrony detectable by the perceptual system (van Eijk et al., 2008).

Our study was designed to test whether healthy older individual exhibit altered adaptation to audiovisual asynchrony. After adapting older and younger observers to sound-lag asynchrony, both groups showed an expansion of their synchrony window in the direction of the adapted asynchrony, but the degree of expansion was smaller for the older group. However, in contrast to predictions, the reduced expansion in the older group could not be accounted for by the perceived synchrony of the adapting stimuli.

MATERIALS AND METHODS

PARTICIPANTS

Fifteen younger (22–32 years old) and 15 older (64–74 years old) adults participated in the experiment. Younger participants were recruited from the University of Melbourne, and older adults were recruited from the university and the community via advertisements posted in community newspapers. Inclusion criteria included having normal or corrected-to-normal vision of 6/7.5 or better, and normal hearing for age. Hearing was assessed in a quiet laboratory space using an audiometer with headphones (Garson Stadler GSI 18 audiometer, Eden Prairie, MN, USA). Normal hearing was defined as having audiometric thresholds less than 35 decibels hearing level (dB HL) at 4 kHz, and less than 25 dB HL at all other tested frequencies (0.25, 0.5, 1, and 2 kHz), according to the International Organization for Standardization (ISO) standard on hearing by age and gender (ISO 7029:2000 Acoustics). The study was approved by the Human Research Ethics

Committee of University of Melbourne and informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

EQUIPMENT

The experiment was controlled by software written in MATLAB 7.6.0 (R2008a; Mathworks, Boston, MA, USA) and run on a personal computer (Dell Precision T3500, Round Rock, TX, USA). The visual stimulus was presented using a ViSaGe (Cambridge Research Systems, Cambridge, UK) to drive a cathode ray tube monitor (Sony Trinitron Multiscan G520 – mean luminance: 100 cd/m², frame rate: 100 Hz, 1024 × 768 pixels, Tokyo, Japan) that was gamma corrected on a weekly basis. Responses were collected using a CB6 response box. The ViSaGe also initiated sound presentation through a set of headphones (Sennheiser HD 205, Wedemark, Germany), by triggering a multifunction processor [Tucker-Davis Technologies (TDT) RX6, Alachua, FL, USA] that drove a programmable attenuator (TDT PA5) and a headphone driver (TDT HB7). Timing precision was verified prior to starting on the main experiment using an oscilloscope. Participants stabilized their head position by resting on a chin rest positioned 100 cm from the monitor.

STIMULI

The visual stimulus was a vertically striped Gabor of 3 c/deg (85% contrast), with the standard deviation of the Gaussian envelope defined as the reciprocal of the spatial frequency ($= 0.33^\circ$), that was presented for one monitor frame (frame rate: 100 Hz) (Figure 1A). The auditory stimulus was a pure tone pip (20 dB, 10 ms duration; 2.5 ms onset and offset ramp) of 500 Hz presented binaurally through headphones over a pure tone mask (75 dB, 1.5 s duration, 100 ms onset and offset ramp) of the same frequency (Figure 1B). The onset of the tone pip was jittered between 200 and 300 ms from the auditory mask onset. Stimulus onset asynchrony was defined as the time difference between the onset timings of the tone pip and the Gabor. Adaptation pairs were either synchronous or asynchronous with a fixed sound-lag asynchrony of 230 ms. By comparing the shift in synchrony perception to each observer's perception after adapting to synchrony, we reduced the amount of inter-subject variability which can result from individual differences in prior experience (Navarra et al., 2010; Alm and Behne, 2013).

PROCEDURE

Each test run consisted of an initial adaptation phase, followed by a repeated looping of a test phase and a top-up adaptation phase until all test trials were completed. In the initial adaptation phase, participants were exposed to 120 adaptation pairs that were either synchronous or asynchronous (Figure 1C). The adaptation pairs were separated by a duration randomly chosen from a uniform distribution from 1000 to 1100 ms. In order to maintain attention during the adaptation phase, 20 randomly occurring catch trials were interleaved where the visual Gabor contained horizontal instead of vertical stripes. Participants were instructed to press on a button if the orientation of the Gabor was horizontal. All participants responded correctly to all of the catch trials. The adaptation phase was approximately 3 min in duration.

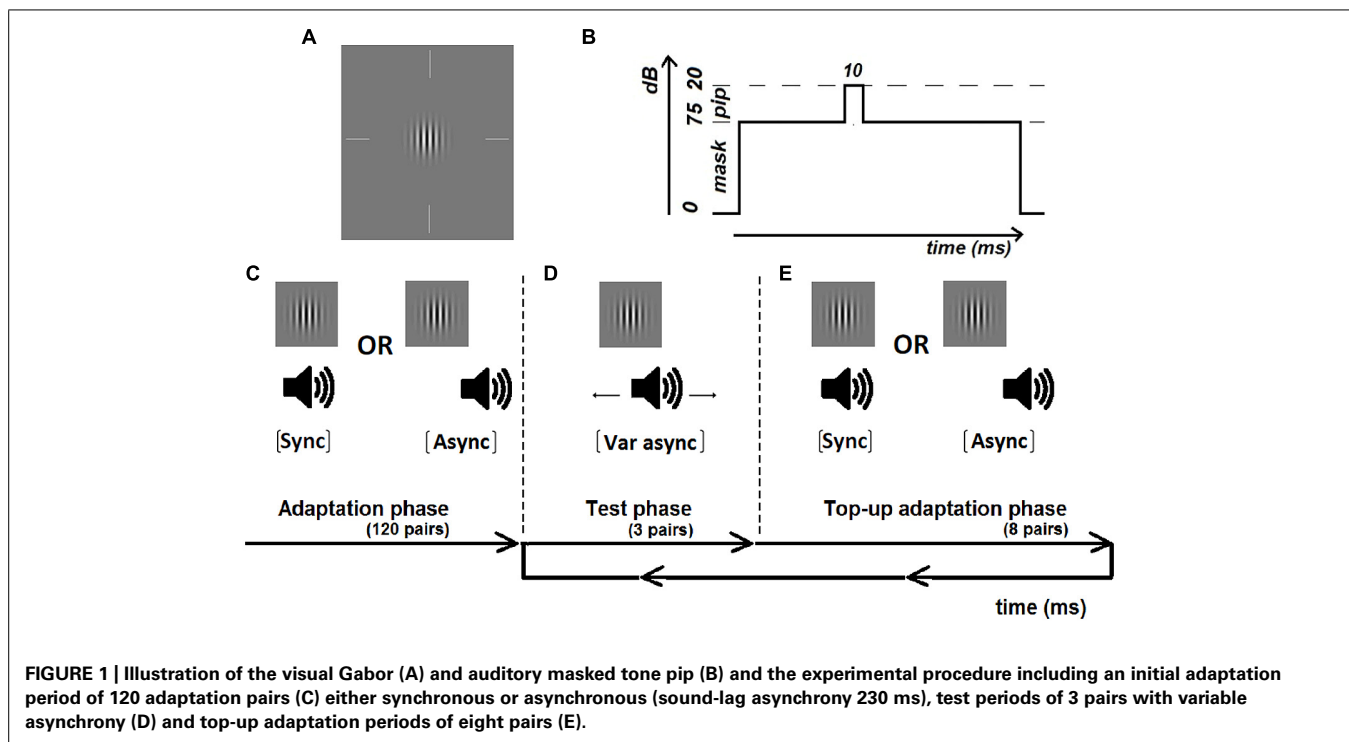


FIGURE 1 | Illustration of the visual Gabor (A) and auditory masked tone pip (B) and the experimental procedure including an initial adaptation period of 120 adaptation pairs (C) either synchronous or asynchronous (sound-lag asynchrony 230 ms), test periods of 3 pairs with variable asynchrony (D) and top-up adaptation periods of eight pairs (E).

At the end of the adaptation phase, participants judged the synchrony for three test pairs (Figure 1D) before being re-exposed to eight top-up adaptation pairs (Figure 1E). Participants indicated whether the test stimuli appeared synchronous or asynchronous via a button press. No feedback was given to the participants regarding the likelihood of encountering synchronous or asynchronous pairs. Responses were self-paced, with the next test-pair presented 500 ms after the button press. Individual synchrony windows were measured across eleven asynchronies using a method-of-constant-stimuli (MOCS). These asynchronies were manually adjusted by the researcher (YMC) for each individual to span from approximately 100% synchronous response to 0% synchronous response. Based on previous work in the lab, a test range of ± 450 and ± 550 ms was sufficient to reach asymptotic response (i.e., 100 and 0% proportion of synchronous responses) for younger and older participants, respectively. Therefore each younger participant began with a practice run with MOCS steps of ± 450 , ± 330 , ± 190 , ± 100 , ± 50 , and 0 ms. If asymptotic response was not achieved with the test range used in the practice run, the test range for the actual test runs was extended to ± 500 , ± 400 , ± 300 , ± 200 , ± 100 , 0 ms. Each older participant began with a practice run with MOCS steps of ± 550 , ± 450 , ± 350 , ± 250 , ± 100 , and 0 ms. If asymptotic response was not achieved with this test range, the MOCS steps used in the actual test runs were changed to ± 600 , ± 450 , ± 350 , ± 250 , ± 100 , and 0 ms. Asymptotic response was obtained in all of the actual test runs for all participants. Within each run, test pairs were tested in sets of three until each of the eleven asynchrony steps were tested for four repeats. The initial adaptation pairs and top up adaptation pairs in each test run were always identical, either synchronous or asynchronous. Each adaptation

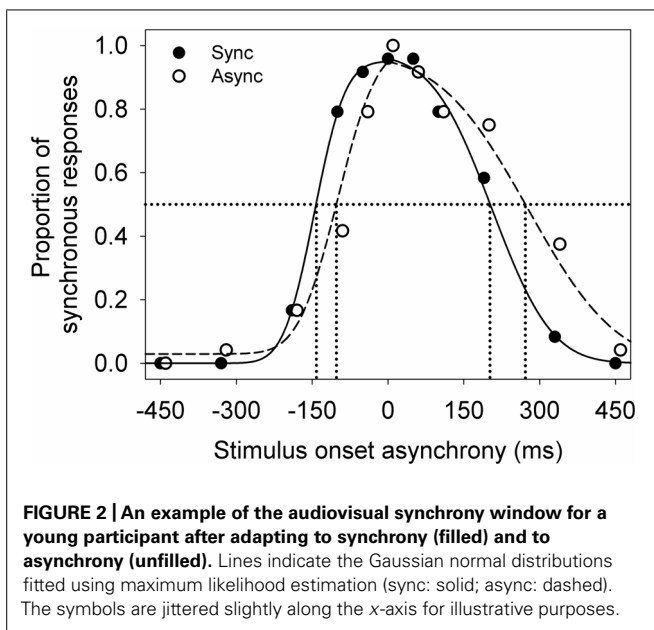
condition was tested for four repeats (total of 16 repeats at each asynchrony step). The order of the eight test runs was randomized within and between participants to avoid order effects. Each run, including the adaptation phase and test and top-up adaptation phases, lasted for no more than 10 min. Total test time to complete the two adaptation conditions lasted around 80 min. Participants were given breaks (no fixed duration) whenever required. Consequently each observer participated in ~ 180 min in total of testing including initial screening, practice runs and breaks.

ANALYSIS

Each participant's data was averaged across the four test repeats for each adapting condition. Then two independent cumulative Gaussian distributions were fitted to the averaged data using least sum of squares (Figure 2). One cumulative Gaussian distribution was fit to the data from earliest sound-lead asynchrony tested to 0 ms, and the second cumulative Gaussian was fit to the data from 0 ms to the latest sound-lag asynchrony tested. The cumulative Gaussian distribution was defined by

$$f(t) = FP + (1 - FP - FN) \times [G(t, \mu, \sigma)] \quad (1)$$

where $G(t, \mu, \sigma)$ was the cumulative Gaussian distribution with mean (μ) and standard deviation (σ) for stimulus asynchrony value t . FP and FN represented the proportions of false positive and false negative responses, respectively, (i.e., the asymptotic error values). The means (μ) of the fitted distributions gave sound-lead and sound-lag synchrony thresholds (i.e., the sound-lead asynchrony that was perceived as synchronous 50% of the time, and the sound-lag that was perceived as synchronous 50% of the time). The standard deviations (σ) defined the participant's asynchrony



discrimination for sound-lead and sound-lag pairs. The width of the audiovisual synchrony window was calculated as the difference between the sound-lead and sound-lag thresholds. The magnitude of the adaptation effect was quantified as the difference between these parameters (sound-lead threshold, sound-lag threshold, asynchrony discrimination sensitivities and window width) for the two adapting conditions. The curve fit aided in the data analysis but was not intended to have any physiological meaning. The root mean squared error of the individually fitted psychometric functions fell below 0.183 for the younger cohort and below 0.179 for the older group. We used a repeated-measures analysis of variance (RM-ANOVA) to test for effects of age or adapted condition on the sound-lead synchrony threshold, sound-lag synchrony threshold, and window width estimates.

RESULTS

Figure 2 shows an example of the audiovisual synchrony windows obtained from a young observer. There is a high proportion of synchronous response when the visual and auditory signals were presented at synchrony (0 ms on the x-axis) and at a small stimulus onset asynchrony. The proportion of synchronous responses decreases with increasing asynchrony (moving leftward and rightward along the x-axis). This observer showed a smaller shift in sound-lead threshold than sound-lag threshold after adapting to asynchrony (unfilled circles, dashed line).

Figure 3 illustrates the synchrony thresholds for sound-lead and sound-lag stimuli for the two adapting conditions. In a mixed design repeated measures ANOVA that compared between the two adapting conditions, two threshold types and two age groups, there was a main effect of threshold type [$F(1,28) = 31.82$, $p < 0.001$], adapting condition [$F(1,28) = 12.43$, $p = 0.001$] and age [$F(1,28) = 4.76$, $p = 0.04$] on the synchrony thresholds. The main effect of threshold type was not dependent on the age group (no significant interaction between threshold

type and age group: $F(1,28) = 0.04$, $p = 0.84$). The significantly different synchrony thresholds between the adapting conditions was, however, dependent on age group (significant interaction between adapting condition and age group: $F(1,28) = 11.74$, $p = 0.002$) and on threshold type (significant interaction between adapting condition, threshold type and age group: $F(1,28) = 4.53$, $p = 0.04$). No other statistics were significant.

SYNCHRONY THRESHOLDS FOR SOUND-LEAD STIMULI

Previous reports showed that adaptation to sound-lag asynchrony expanded the audiovisual synchrony window asymmetrically toward greater sound-lag asynchrony (Fujisaki et al., 2004). We examined the effect of adaptation on the sound-lead and sound-lag thresholds independently in two separate ANOVAs. The left panels in **Figure 3** plot the sound-lead thresholds for the two adapting conditions. There was a main effect of age on sound-lead threshold [$F(1,28) = 4.27$, $p = 0.048$] where older observers on average required greater sound-lead asynchrony to perceive asynchrony (adapted to synchrony: -206 , adapted to asynchrony: -187 ms) than the average younger group (adapted to synchrony: -157 , adapted to asynchrony: -156 ms). There was no effect of adaptation on sound-lead thresholds for both age groups [no main effect of adapting condition: $F(1,28) = 1.05$, $p = 0.31$]. No other statistics were significant.

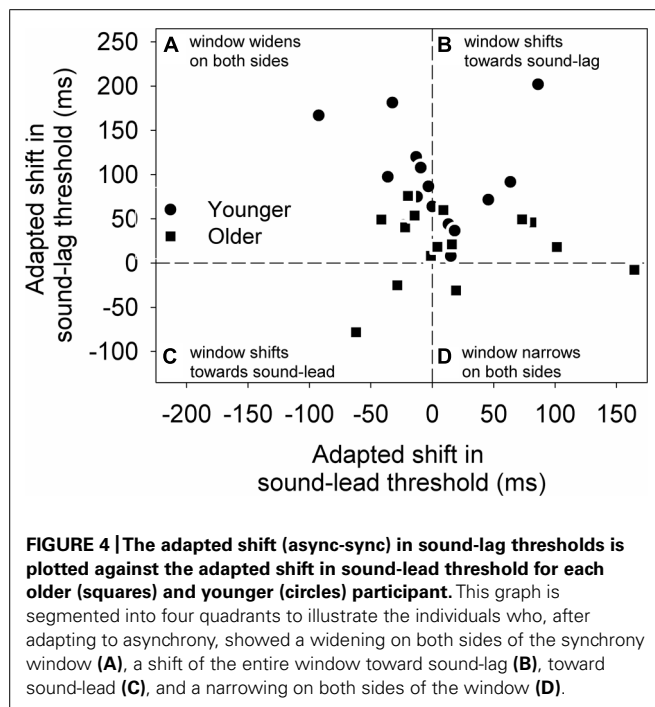
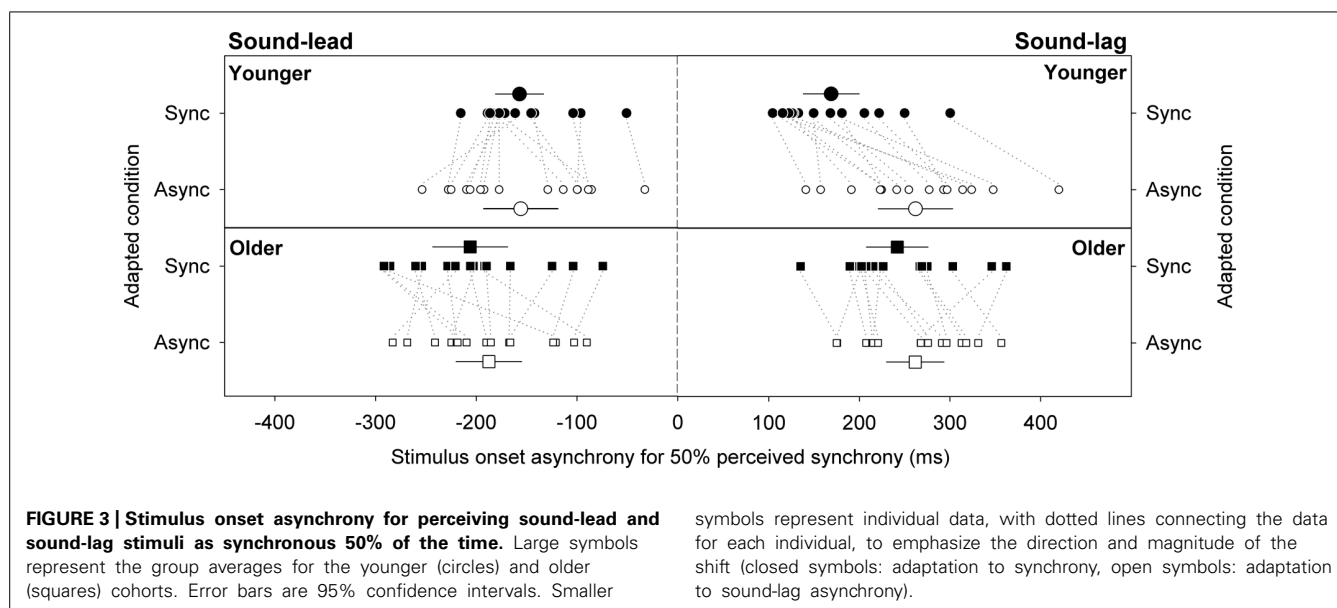
SYNCHRONY THRESHOLDS FOR SOUND-LAG STIMULI

The right panels of **Figure 3** show the sound-lag thresholds. Adaptation to sound-lag shifts mean sound-lag thresholds toward greater asynchronies (i.e., in the direction of the adapted asynchrony) for both age groups [main effect of adapting condition: $F(1,28) = 39.96$, $p < 0.001$], however, the group average shift in the young observers (adapted to synchrony: 169 ms, adapted to asynchrony: 262 ms; shift: +93 ms) is greater than the group average shift in the older cohort [adapted to synchrony: 242 ms, adapted to asynchrony: 262 ms; shift: +20 ms; significant interaction between age group and adapting condition: $F(1,28) = 16.78$, $p < 0.001$]. No other statistics were significant.

To better illustrate the change in the overall synchrony window post-adapting to asynchrony, we calculated the difference between the two adapted conditions (adapted shift in threshold) for each individual's sound-lead and sound-lag thresholds (**Figure 4**). For both age groups, most data points appear above zero on the y-axis, meaning the majority of our participants either showed a widening on both sides of their synchrony window (**Figure 4A**) or a shift of the entire window toward sound-lag (**Figure 4B**).

AUDIOVISUAL SYNCHRONY WINDOW WIDTHS

For both adapting conditions, synchrony windows were on average wider in the older (448, 449 ms after adaptation to synchrony and asynchrony, respectively) than the younger (326, 418 ms) group. After adapting to asynchrony, the synchrony window widened for the younger group (an increase of 92 ms), but the change was small for the older group (an increase of 1 ms).



AUDIOVISUAL ASYNCHRONY DISCRIMINATION SENSITIVITY

The standard deviation of the fitted cumulative Gaussians provided an estimate of each participant's asynchrony discrimination sensitivity for sound-lead and sound-lag pairs. After adapting to synchrony, the group averaged standard deviation for the younger group was 36 ms (95% confidence interval: ± 10 ms) for sound-lead pairs and was 48 ms (± 11 ms) for sound-lag pairs. The group averaged standard deviation for the older adults was 41 ms (± 13 ms) for sound-lead pairs, and 57 ms (± 20 ms) for sound-lag pairs. After adapting to asynchrony, the group averaged standard deviations for the younger cohort were 37 ms (± 16 ms) and 57 ms

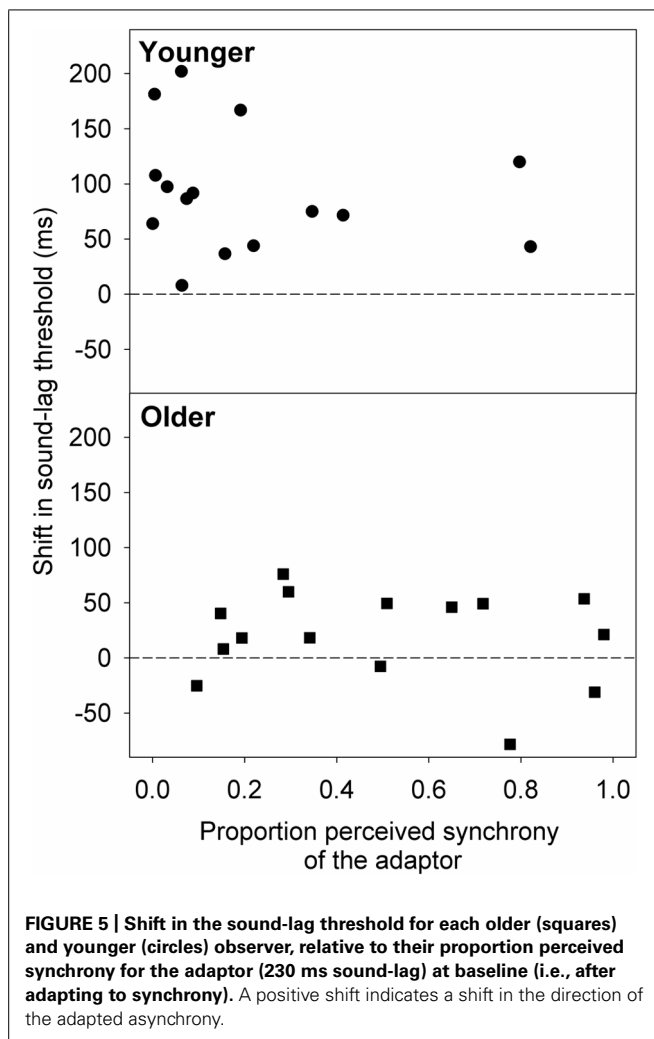
(± 21 ms) for sound-lead and sound-lag pairs, respectively, and in the older cohort were 61 ms (± 19 ms) and 65 ms (± 18 ms). These estimates were analyzed in a mixed design ANOVA that compared between the two adapted conditions, two age groups and two synchrony threshold types ($2 \times 2 \times 2$). There was no main effect of age [$F(1,28) = 2.04$, $p = 0.16$], nor of adaptation condition [$F(1,28) = 2.70$, $p = 0.11$], and no significant interaction effects (all $p > 0.05$).

IS THE ADAPTATION EFFECT RELATED TO HOW THE ADAPTOR WAS PERCEIVED?

We extended our analysis to see if the smaller magnitude of adaptation in the older group was due to the adaptor appearing synchronous more often to the older than the younger observers. Therefore, we used the best-fit psychometric functions obtained after adaptation to synchrony to estimate the proportion of perceived synchrony for the adaptor asynchrony (sound-lag asynchrony of 230 ms). We then conducted linear regression analyses on the relationship between the shift in sound-lag threshold and the proportion of perceived synchrony for the adaptor asynchrony (Figure 5). No statistically significant linear dependence of the adaptation shift magnitude on perceived synchrony of the adaptor was detected for either age group [younger: slope ($\pm 95\%$ CI) = -38.26 (-159.51 , 82.99); $t(13) = -0.68$, $p = 0.51$; older: slope = -20.73 (-98.51 , 56.70); $t(13) = -0.58$, $p = 0.57$].

IS THE ADAPTATION EFFECT RELATED TO INDIVIDUAL SYNCHRONY WINDOW WIDTHS?

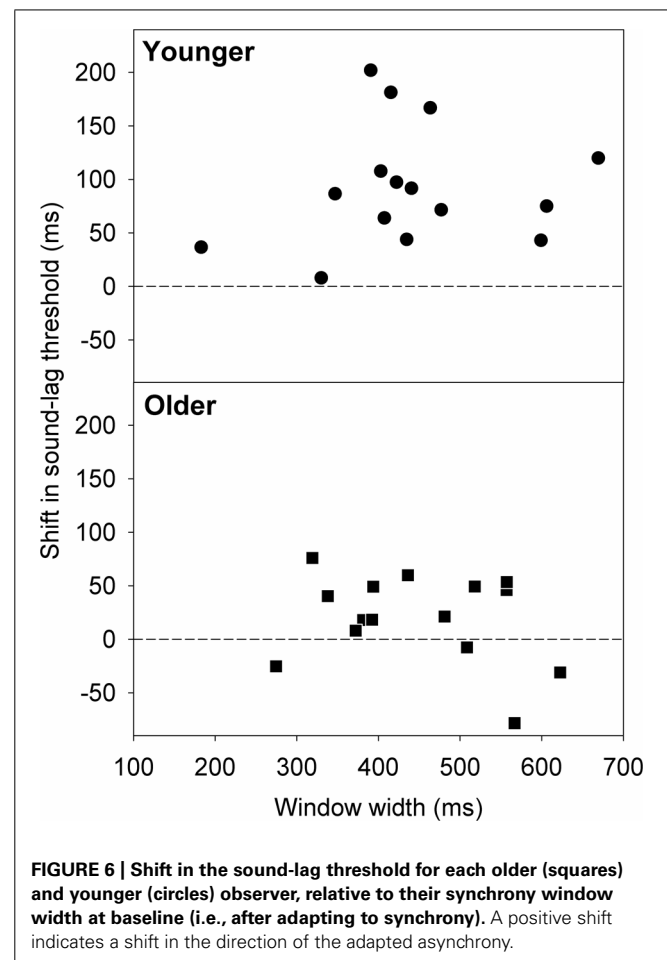
Van der Burg et al. (2013) found that rapid asynchrony adaptation to the audiovisual pair presented immediately before the test pair was dependent on individual synchrony window width. Participants who had wider synchrony windows exhibited a larger magnitude of adaptation effect (Van der Burg et al., 2013). We investigated if this trend applies similarly to our data. Figure 6 plots the shift in the sound-lag threshold as a function of window width. A simple linear regression analysis showed no statistically



significant linear dependence of the adaptation shift magnitude on synchrony window width [younger: slope = 0.10; 95% confidence interval: $-0.31, 0.51$; $t(13) = 0.54$, $p = 0.60$; older: slope = -0.11 ; 95% confidence interval: $-0.33, 0.12$; $t(13) = -1.00$, $p = 0.33$].

DISCUSSION

Our results show an age-related widening of the audiovisual synchrony time window, consistent with recent reports on aging in synchrony perception (Hay-McCutcheon et al., 2009; Chan et al., 2014) and in the audiovisual sound-induced illusion (DeLoss et al., 2013). We also demonstrate that, with healthy ageing, elderly observers recalibrate their sound-lag threshold to a lesser extent when they are exposed to the same asynchrony adaptation as younger adults. The degree to which the adaptor is perceived as asynchronous does not influence the size of the shift in perceived sound-lag synchrony in either age group, so the smaller adaptation effect seen in the elderly observers is unlikely to be due to differences in how the adaptor was perceived by the young and older cohorts. Although rapid adaptation to asynchronous pairs was dependent on synchrony window width (Van der Burg et al., 2013), our study on short term adaptation showed no



significant relationship between adaptation effect and window width.

We did not explicitly measure cognitive performance of our participant, however, it is unlikely that differences in cognitive capacity or differential task difficulty could have resulted in the age effect found. Our older participants were recruited from the community, were fit and active, and passed general screening of vision and hearing to ensure no significant age-related sensory organ damage. There were four current or retired university staff. The rest were active elderly citizens who were still involved in casual paid work and volunteer work in the community. Recent aging study on the audiovisual sound-induced illusion also showed an unlikely role of selective attention on the increased audiovisual integration (increased illusion) with age (DeLoss et al., 2013). Our older participants were active elderly citizens (64–74 years) recruited from the university and the community. All passed the inclusion criteria of having healthy vision and hearing that are normal for their age. As expected for individuals of this age, hearing thresholds were less than 35 decibels hearing level (dB HL) at 4 kHz, and less than 25 dB HL at all other tested frequencies (0.25, 0.5, 1, and 2 kHz; ISO 7029:2000 Acoustics). Data collected during the testing was clean, as demonstrated by all participants reaching 100% synchronous responses when the audiovisual stimulus was at physical synchrony, and

0% synchronous responses when the visual and auditory signals were separated by at most 600 ms. There was also no difference in the spread of the psychometric functions between groups hence the ability to make asynchrony discrimination judgements was similar between the two groups. Our data provides no evidence for either differential levels of task learning, fatigue, or attention. Selective attention has been previously shown to be an unlikely explanation for age-related increase in the audiovisual sound-induced flash illusion (more likely to combine the beeps with the flash; DeLoss et al., 2013). All participants were trained with practice trials to ensure they were confident with the task and were performing it correctly before proceeding onto the main experiment. It is also worth noting that previous studies have shown that procedural and perceptual learning in older adults is similar to that of younger adults for visual tasks (McKendrick and Battista, 2013), and is not simply an effect of task practice but rather a change in the underlying neural process (Andersen et al., 2010).

After adaptation to synchronous audiovisual stimuli, elderly observers have a wider synchrony window compared to young observers. This is consistent with previous studies on the effect of aging on audiovisual synchrony judgment without adaptation (Hay-McCutcheon et al., 2009; Chan et al., 2014). By measuring asynchrony-synchrony judgements for simple flash-pip stimuli in a two-interval-forced-choice design, previous data from our laboratory measured wider synchrony windows in the older observers (465 ms; 61–72 years, mean age of 66 years; seven males) than in younger observers (319 ms; 21–32 years, mean age of 25 years; five males). This widening of the window was independent of response criteria bias and age-related decline in visual contrast sensitivity and hearing thresholds, since data were collected using stimuli scaled to visual and auditory sensitivity for each individual. The width estimates in our younger (326 ms) and older (448 ms) cohorts in the current study are comparable to those we have measured previously.

Adaptation to sound-lag produces a change in the synchrony window of young observers. Consistent with previous data (Fujisaki et al., 2004; Vroomen et al., 2004; Navarra et al., 2009, 2012), there was no significant change to sound-lead thresholds, whereas sound-lag thresholds increased, thus resulting in an asymmetric widening of the window. The elderly observers showed a similar pattern of results, but the magnitude of the adaptation-induced shift was reduced. This reduced recalibration in the elderly observers was not related to how frequently the adaptation stimuli were perceived as synchronous or asynchronous (Figure 5) or synchrony window width (Figure 6). It is, however, possible that the trend for a greater variance in perceived synchrony of the adaptor in the older group could have contributed to the lack of a significant relationship between the perceived synchrony of the adaptor and adaptation effect. On the other hand, Van der Burg et al. (2013) reported a direct relationship between rapid asynchrony adaptation effect and synchrony window width in their young adults. The absence of this relationship in our data on short term adaptation can be argued as the recruitment of different neural mechanism for rapid and short term audiovisual asynchrony adaptation. Rapid adaptation is more likely to be an early sensory effect, whereas

the short term adaptation in our study alters later higher level neural processes (Van der Burg et al., 2013). It is, however, worth noting that such rapid recalibration may possibly influence our findings as the data was collected in test triplets. There is insufficient data to analyse test order effects within the triplets directly.

After adapting to asynchrony, sound-lag thresholds shifted to an average of 262 ms for both age groups. One possible explanation for our results is that 262 ms may be the average optimal position for the sound-lag threshold when the audiovisual system is exposed to an adaptor with a sound-lag asynchrony of 230 ms. Consequently, the reduced response in the elderly may simply reflect the fact that the average sound-lag end of their synchrony window is closer to this limit prior to adaptation. Such an explanation would predict that those individuals closer to this optimal point would demonstrate less adaptation than those further away. However, the individual data, as plotted in Figure 3, are not readily consistent with this suggestion.

The smaller adaptation with age could be explained by the need for longer adaptation duration in the older group to reach the same amount of adaptation effect as the younger group. In a visuomotor experiment that compared the motor response before and after visual prism adaptation, older people required longer adaptation duration to the prism before they were able to correctly point to the visual target (Fernández-Ruiz et al., 2000). However, this is a purely vision-based adaptation that is possibly confined to the neural processes responsible for visual processing only. We do not know if the same explanation of longer adaptation duration with age can be directly applied to the multisensory context of our study.

The computational approach and neural basis by which the brain encodes temporal information regarding auditory and visual stimuli is poorly understood. One proposed mechanism is that unisensory neural processing speed is altered by the adaptation process (Di Luca et al., 2009; Navarra et al., 2009). It has been argued that such a model predicts a uniform recalibration, independent of the specific test stimulus asynchronies that may be presented post-adaptation (Di Luca et al., 2009; Navarra et al., 2009). Roach et al. (2011) demonstrate that the magnitude of recalibration is non-uniform, but instead varies according to the timing offsets between the auditory and visual pairs. The pattern of human behavior that they observed was fit by a computational population-coding model of audiovisual timing tuned neurons. The observed variable bias in recalibration as a function of audiovisual timing offset was well captured by a model that incorporated reduced response gain in the neurons tuned to the adapted asynchrony (Roach et al., 2011).

The neural and anatomical locus of such a population of audiovisual timing tuned neurons, and whether they are functionally or structurally affected by aging is not known. Neurons with such temporal specificity may be located in multisensory brain areas like the superior colliculus, as shown in single cell recording in cats (Meredith et al., 1987). In aged primates, there is evidence for the tuning properties of visual neurons to become less selective with advancing age, for example broader direction tuning and orientation tuning in the primary visual cortex of cats (Hua et al., 2006) and primates (Schmolesky et al., 2000).

A proposed mechanism is a reduction of inhibition which is supported by the presence of fewer GABAergic neurons in the cat visual striate cortex (Hua et al., 2008), and by experiments demonstrating that orientation tuning can be regained by administration of inhibitory GABA agonists (Leventhal et al., 2003). To our knowledge, similar neurophysiological data is not available for multisensory areas. Broadening of neuronal tuning properties may be a generalized feature of aging in sensory cortices, and it is also possible that the number of neurons contributing to a population of cells encoding audiovisual timing might also reduce with age. Future planned experiments should be able to collect data suitable for specific application of a population-coding model similar to that described in Roach et al. (2011). Comparison of age groups using such strategies may enable insight regarding whether differences in the patterns of behavior between older and younger adults are consistent with either (or both) a broadening of neuronal tuning or drop-out of cells. Electrophysiological experiments comparing evoked potentials between age groups before and after adaptation to asynchrony might also provide evidence for or against the alternate suggestion of an altered neural latency mechanism (Di Luca et al., 2009; Navarra et al., 2009).

In conclusion, our findings demonstrate that the recalibration response in older people differs from that of younger adults. For the stimulus conditions used in our experiments, older adults demonstrated reduced recalibration to the prevailing audiovisual timing environment, which was not related to their baseline percept of the adapting asynchrony. The specific neural basis for this difference, and how it impacts on sensory performance in more naturalistic environments, warrants further study.

ACKNOWLEDGMENT

This research was funded by the Australian Research Council FT0990930 to author Allison M. McKendrick.

REFERENCES

- Alm, M., and Behne, D. (2013). Audio-visual speech experience with age influences perceived audio-visual asynchrony in speech. *J. Acoust. Soc. Am.* 134, 3001–3010. doi: 10.1121/1.4820798
- Andersen, G. J., Ni, R., Bower, J. D., and Watanabe, T. (2010). Perceptual learning, aging, and improved visual performance in early stages of visual processing. *J. Vis.* 10, 4. doi: 10.1167/10.13.4
- Asakawa, K., Tanaka, A., and Imai, H. (2009). Temporal recalibration in audio-visual speech integration using a simultaneity judgment task and the McGurk identification task. *Paper Presented at the 31st Annual Meeting of the Cognitive Science Society*, Amsterdam.
- Asakawa, K., Tanaka, A., and Imai, H. (2012). Audiovisual temporal recalibration for speech in synchrony perception and speech identification. *Kansei Eng. Int. J.* 111, 35–40. doi: 10.5057/kei.11.35
- Chan, Y. M., Pianta, M. J., and McKendrick, A. M. (2014). Older age results in difficulties separating auditory and visual signals in time. *J. Vis.* (in press).
- DeLoss, D. J., Pierce, R. S., and Andersen, G. J. (2013). Multisensory integration, aging, and the sound-induced flash illusion. *Psychol. Aging* 283, 802–812. doi: 10.1037/a0033289
- Di Luca, M., Machulla, T.-K., and Ernst, M. O. (2009). Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *J. Vis.* 9, 7. doi: 10.1167/9.12.7
- Fernández-Ruiz, J., Hall, C., Vergara, P., and Díaz, R. (2000). Prism adaptation in normal aging: slower adaptation rate and larger aftereffect. *Cogn. Brain Res.* 93, 223–226. doi: 10.1016/S0926-6410(99)00057-9
- Fiacconi, C. M., Harvey, E. C., Sekuler, A. B., and Bennett, P. J. (2013). The influence of aging on audiovisual temporal order judgments. *Exp. Aging Res.* 39, 179–193. doi: 10.1080/0361073X.2013.761896
- Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778. doi: 10.1038/nn1268
- Hay-McCutcheon, M. J., Pisoni, D. B., and Hunt, K. K. (2009). Audiovisual asynchrony detection and speech perception in hearing-impaired listeners with cochlear implants: a preliminary analysis. *Int. J. Audiol.* 48, 321–333. doi: 10.1080/14992020802644871
- Heron, J., Whitaker, D., McGraw, P. V., and Horoshenkov, K. V. (2007). Adaptation minimizes distance-related audiovisual delays. *J. Vis.* 7, 5. doi: 10.1167/7.13.5
- Hua, T., Kao, C., Sun, Q., Li, X., and Zhou, Y. (2008). Decreased proportion of GABA neurons accompanies age-related degradation of neuronal function in cat striate cortex. *Brain Res. Bull.* 751, 119–125. doi: 10.1016/j.brainresbull.2007.08.001
- Hua, T., Li, X., He, L., Zhou, Y., Wang, Y., and Leventhal, A. G. (2006). Functional degradation of visual cortical cells in old cats. *Neurobiol. Aging* 27, 155–162. doi: 10.1016/j.neurobiolaging.2004.11.012
- King, A., and Palmer, A. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500. doi: 10.1007/BF00236934
- Leventhal, A. G., Wang, Y., Pu, M., Zhou, Y., and Ma, Y. (2003). GABA and its agonists improved visual cortical function in senescent monkeys. *Science* 300, 812–815. doi: 10.1126/science.1082874
- Love, S. A., Petrini, K., Cheng, A., and Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS ONE* 81:e54798. doi: 10.1371/journal.pone.0054798
- McKendrick, A. M., and Battista, J. (2013). Perceptual learning of contour integration is not compromised in the elderly. *J. Vis.* 13, 5. doi: 10.1167/13.1.5
- Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci.* 7, 3215–3229.
- Navarra, J., Alsius, A., Velasco, I., Soto-Faraco, S., and Spence, C. (2010). Perception of audiovisual speech synchrony for native and non-native language. *Brain Res.* 1323, 84–93. doi: 10.1016/j.brainres.2010.01.059
- Navarra, J., García-Morera, J., and Spence, C. (2012). Temporal adaptation to audio-visual asynchrony generalizes across different sound frequencies. *Front. Psychol.* 3:152. doi: 10.3389/fpsyg.2012.00152
- Navarra, J., Hartcher-O'Brien, J., Piazza, E., and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9169–9173. doi: 10.1073/pnas.0810486106
- Parsons, B. D., Novich, S. D., and Eagleman, D. M. (2013). Motor-sensory recalibration modulates perceived simultaneity of cross-modal events at different distances. *Front. Psychol.* 4:46. doi: 10.3389/fpsyg.2013.00046
- Roach, N. W., Heron, J., Whitaker, D., and McGraw, P. V. (2011). Asynchrony adaptation reveals neural population code for audio-visual timing. *Proc. Biol. Sci.* 278, 1314–1322. doi: 10.1098/rspb.2010.1737
- Schmolesky, M. T., Wang, Y., Pu, M., and Leventhal, A. G. (2000). Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nat. Neurosci.* 3, 384–390. doi: 10.1038/73957
- Setti, A., Burke, K. E., Kenny, R. A., and Newell, F. N. (2011a). Is inefficient multisensory processing associated with falls in older people? *Exp. Brain Res.* 209, 375–384. doi: 10.1007/s00221-011-2560-z
- Setti, A., Finnigan, S., Sobolewski, R., McLaren, L., Robertson, I. H., Reilly, R. B., et al. (2011b). Audiovisual temporal discrimination is less efficient with aging: an event-related potential study. *Neuroreport* 22, 554–558. doi: 10.1097/WNR.0b013e328348c731
- Tanaka, A., Asakawa, K., and Imai, H. (2009). Recalibration of audiovisual simultaneity in speech. *Paper Presented at the International Conference on Audio-Visual Speech Processing 2009*, University of East Anglia, Norwich.
- Tanaka, A., Asakawa, K., and Imai, H. (2011). The change in perceptual synchrony between auditory and visual speech after exposure to asynchronous speech. *Neuroreport* 22, 684–688. doi: 10.1097/WNR.0b013e32834a2724
- Van der Burg, E., Alais, D., and Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *J. Neurosci.* 33, 14633–14637. doi: 10.1523/JNEUROSCI.1182-13.2013
- van Eijk, R. L., Kohlrausch, A., Juola, J. F., and van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept. Psychophys.* 70, 955–968. doi: 10.3758/PP.70.6.955

- Vatakis, A., Navarra, J., Soto-Faraco, S., and Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. *Exp. Brain Res.* 185, 521–529. doi: 10.1007/s00221-007-1168-9
- Vroomen, J., Keetels, M., de Gelder, B., and Bertelson, P., (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn. Brain Res.* 22, 32–35. doi: 10.1016/j.cogbrainres.2004.07.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 June 2014; accepted: 08 August 2014; published online: 27 August 2014.
Citation: Chan YM, Pianta MJ and McKendrick AM (2014) Reduced audiovisual recalibration in the elderly. *Front. Aging Neurosci.* 6:226. doi: 10.3389/fnagi.2014.00226
This article was submitted to the journal *Frontiers in Aging Neuroscience*.
Copyright © 2014 Chan, Pianta and McKendrick. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



How does aging affect the types of error made in a visual short-term memory 'object-recall' task?

Raju P. Sapkota^{1*}, Ian van der Linde^{1,2} and Shahina Pardhan¹

¹ Vision and Eye Research Unit, Postgraduate Medical Institute, Anglia Ruskin University, Cambridge, UK

² Department of Computing and Technology, Anglia Ruskin University, Cambridge, UK

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Kevin Dent, University of Essex, UK
Summer Sheremata, George Washington University, USA

*Correspondence:

Raju P. Sapkota, Vision and Eye Research Unit, Postgraduate Medical Institute, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK
e-mail: raju.sapkota@anglia.ac.uk

This study examines how normal aging affects the occurrence of different types of incorrect responses in a visual short-term memory (VSTM) object-recall task. Seventeen young (Mean = 23.3 years, SD = 3.76), and 17 normally aging older (Mean = 66.5 years, SD = 6.30) adults participated. Memory stimuli comprised two or four real world objects (the memory load) presented sequentially, each for 650 ms, at random locations on a computer screen. After a 1000 ms retention interval, a test display was presented, comprising an empty box at one of the previously presented two or four memory stimulus locations. Participants were asked to report the name of the object presented at the cued location. Errors rates wherein participants reported the names of objects that had been presented in the memory display but not at the cued location (*non-target* errors) vs. objects that had not been presented at all in the memory display (*non-memory* errors) were compared. Significant effects of aging, memory load and target recency on error type and absolute error rates were found. Non-target error rate was higher than non-memory error rate in both age groups, indicating that VSTM may have been more often than not populated with partial traces of previously presented items. At high memory load, non-memory error rate was higher in young participants (compared to older participants) when the memory target had been presented at the earliest temporal position. However, non-target error rates exhibited a reversed trend, i.e., greater error rates were found in older participants when the memory target had been presented at the two most recent temporal positions. Data are interpreted in terms of proactive interference (earlier examined non-target items interfering with more recent items), false memories (non-memory items which have a categorical relationship to presented items, interfering with memory targets), slot and flexible resource models, and spatial coding deficits.

Keywords: age differences, object-recall, memory objects, memory load, recency

INTRODUCTION

An important component of working memory model that enables recent visual events to be remembered is visual short-term memory (VSTM; Baddeley and Hitch, 1974; Baddeley, 1986, 2003). VSTM was originally proposed as the means by which recently acquired visual information is transferred into longer-term storage (Phillips, 1974). However, more recent views emphasize the importance of VSTM in the online cognition that underpins everyday tasks, such as noticing change (change detection), seeking objects (visual search), and more generally in the perception of complex scenes (Hollingworth et al., 2008).

Measurement of VSTM performance in the laboratory usually entails the brief presentation of a finite set of 'to-be-remembered' visual items, known as memory stimuli. After a brief retention interval, during which a blank field is typically presented, participants' memory for the previously presented items is tested using paradigms such as change detection, yes–no recognition, or whole/partial recall (Pashler, 1988; Irwin, 1991; Irwin and Andrews, 1996; Luck and Vogel, 1997; Irwin and Zelinsky, 2002). These studies typically report that VSTM can store around 3–4 multifaceted items at any time. This apparent ceiling on the

storage capacity of VSTM implies that, where more than four items are viewed, competition for storage occurs. One popular model (the *slot model*) proposes that VSTM has a fixed number of discrete *all-or-none* storage compartments (or slots); when the number of to-be-remembered items is less than the capacity of VSTM, items are remembered without a significant loss in their visual detail (Luck and Vogel, 1997; Zhang and Luck, 2008). In this model, it is believed that competition does not occur between memory items unless all slots are occupied. A more recent alternative, the *resource model*, proposes that VSTM resources are shared between the items featuring in a memory display in a more continuous fashion, such that when the number of items to be remembered exceeds the capacity of VSTM, some (correspondingly impoverished) information that pertains to a larger number of items is retained (Wilken and Ma, 2004; Bays and Husain, 2008).

Competition between items held in VSTM has been found to be influenced by several factors, including, but not limited to the items' visual, spatial, and temporal properties. For example, items that are more familiar, visually salient and that were seen more recently in time are retained with greater accuracy than

items that are less familiar, less salient (with respect to other presented items), and that were seen earlier in time (Phillips, 1983; Alvarez and Cavanagh, 2004; Zelinsky and Loschky, 2005; Hollingworth, 2007).

Another important phenomenon known to influence object representation in VSTM is our ability to inhibit information that is irrelevant (or no longer relevant) to the current task. This has been studied typically in terms of *interference* (or *intrusion*) effects (Zelinsky and Loschky, 2005; Makovski and Jiang, 2008; Fiore et al., 2012). These studies relate this interference (or intrusion) to the memory-diminishing effects arising from the items other than the memory target(s) that were also presented during the memory display (non-target memory items). In addition, interference may also originate from items retrieved from long-term memory that were not present in the memory display. These may be guesses elicited in the absence of any available memory trace, or may bear a semantic/categorical, structural, positional, or other relationship to the forgotten items that were presented in the memory display.

Error rates that arise from participant's reporting non-target memory items for a memory target during a VSTM task have been studied extensively in young healthy participants (Irwin and Zelinsky, 2002; Zelinsky and Loschky, 2005; Makovski and Jiang, 2008; Sapkota et al., 2011). As a result of the aging population, and the commensurate increase in the prevalence of neurodegenerative conditions that affect memory performance in the elderly, research that examines age-related changes in VSTM function is becoming increasingly important. A number of studies have improved our understanding of the differences between clinically significant changes in memory function and healthy aging (De Beni and Palladino, 2004; Borella et al., 2008; Fiore et al., 2012). These studies have shown that older participants are less adept at suppressing non-target memory items compared to young participants during memory retrieval, and consequently experience greater memory distraction. The findings are compatible with the proposal that memory performance decreases with age as a result of a general deterioration in the ability of older participants to inhibit irrelevant visual information (Hasher and Zacks, 1988; Zacks and Hasher, 1994). However, previous studies that have investigated the effect of aging on VSTM performance have largely overlooked the possibility (especially in object-recall tasks) that incorrect responses may also arise due to interference from non-memory items (i.e., novel items that were not presented in the memory display). The degree to which aging may affect our ability to overcome distraction from these two types of irrelevant items during an object-recall task has not been directly compared, despite that such an investigation could have important implications for understanding the mechanisms that underpin age-related VSTM decline.

A non-target memory error can occur during an object-recall task when the binding between an item and its location has been lost, and when one of the non-target items has instead bound with the location of the memory target item. A non-memory error can occur when a memory target has been forgotten, such that interference from items that had not been presented in the memory display is greater than the interference from the items presented in the memory display. It is also possible that a well-remembered

non-target memory item can be successfully excluded, leaving the participant with no choice but to guess an item that had not been presented when the memory target item has been forgotten. This may occur (for example) due to confusion between a memory target and a previously unexamined item that belongs to the same object category as the memory target (rather than report a previously examined non-target memory item, as may happen when spatiotemporal confusion between items occurs).

In this study we compared error rates in VSTM in which non-target memory items vs. non-memory items were reported during a location cued object-recall task in healthy young and normally aging older participants. In contrast to previous studies that have measured VSTM performance (e.g., capacity, longevity) wherein errors are only considered in terms of their effect on overall performance rate, a detailed analysis of the nature of error responses in this study will enable us to acquire a greater understanding of the effect of aging on the occurrence of different types of memory interference, such as proactive and retroactive interference (memory-diminishing effects arising, respectively, from stimuli examined *before* and *after* a memory target), interference from an item belonging to the same object category, and interference from other spatially nearby items (spatial proximity). To our knowledge, the influence of aging on VSTM performance during a location cued object-recall task in terms of pro/retroactive interference, spatial proximity, object category, error type, *viz.*, the reporting of non-target vs. non-memory items, and its relationship with memory load and stimulus recency has not been examined before.

We predict a significant effect of age group (young or older) on the occurrence of different object-recall error types (non-target vs. non-memory), as our ability to suppress distraction arising from irrelevant (or no longer relevant) items decreases with the age (Hasher and Zacks, 1988). Furthermore, we hypothesize that older participants will exhibit greater confusion due to interference from items that are (spatially) nearby to the memory target. Also, higher error rates are expected for older participants in the high memory load condition, and for earlier presented items.

MATERIALS AND METHODS

PARTICIPANTS

Seventeen young (Mean = 23.3 years, SD = 3.76) and 17 normally aging older adults (Mean = 66.5 years, SD = 6.30) participated, all with normal/corrected-to-normal vision. All participants had a mini-mental state examination (MMSE) score ≥ 27 , confirming normal cognitive function. Young and older participants were matched for gender (9-female, 8-male), and the minimum number of years of formal education (13 years). All participants were naïve to the purpose of the study, and were paid for their participation. Ethical clearance was obtained from Anglia Ruskin University's Faculty Research Ethics Panel before data collection commenced. Participants were treated in accordance with applicable ethical guidelines that followed the tenets of the Helsinki Declaration.

APPARATUS

Stimuli were displayed on a 17" LCD screen set at a spatial resolution of 1024 × 768 pixels and a refresh rate of 75 Hz. The

screen was positioned at 57 cm from observers (such that the spatial extent of the display was $\sim 34^\circ \times 27^\circ$). A chin/forehead rest was used to stabilize viewing position and distance. Ambient light was held constant across trials and between participants.

STIMULI

Stimuli comprised 170 line drawings of real world objects (Snodgrass and Vanderwart, 1980), each centered within an invisible square subtending 2.5° of visual angle at a testing distance of 57 cm. Stimuli belonged to one of 14 semantic categories (four-footed animals, birds, kitchen utensils, etc.). Example stimuli are shown in **Figure 1**. Stimulus presentation was controlled by MATLAB (Mathworks, Natick, MA, USA) with the Psych-Toolbox/VideoToolbox extensions (Brainard, 1997; Pelli, 1997). Stimulus background was set to mid-gray.

PROCEDURE

Experimental procedures were preceded by a stimulus learning routine, during which all 170 stimuli were displayed sequentially in random order; participants were asked to name each stimulus as it appeared. When participants could not name/recognize a stimulus, the experimenter familiarized them with it by speaking its name aloud (i.e., a verbal prompt). Next, stimuli that participants could not name originally were re-presented one at a time, and participants were asked again to name them as they appeared. All participants were able to name all the stimuli correctly on their second attempt. Of the 13 (from 34) participants who were given verbal prompts (seven older, six young), 11 required a prompt for only one stimulus (from a total of 170), and two required a prompt for just two stimuli, indicating stimuli were nearly universally recognized, and that any priming effects arising due to stimuli being presented more than once during the learning routine were negligible. A practice block of 20 trials followed that used stimuli that were not part of the main experiment. The procedure in the practice trials was identical to that used in the main experiment (see below).

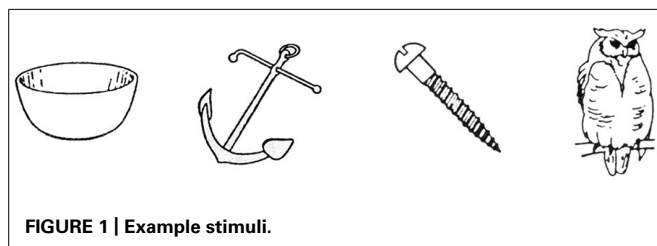
A schematic representation of the experimental procedure is provided in **Figure 2**, in which an example trial at sequence length 4 (see below) is shown. Each trial began with a 2.5° fixation cross displayed for 800 ms at the screen center (Frame 1). This ensured that all participants fixated upon a common screen position prior to the memory display. Next, a two-digit number was shown at the display screen center for 800 ms (Frame 2), followed by the presentation of a memory display (Frame 3–6), in which a sequence of either two or four stimuli [hereafter referred to as sequence length (SL)] were shown, each for 650 ms. This display duration was sufficiently long to enable stimuli to be encoded in VSTM (Vogel et al.,

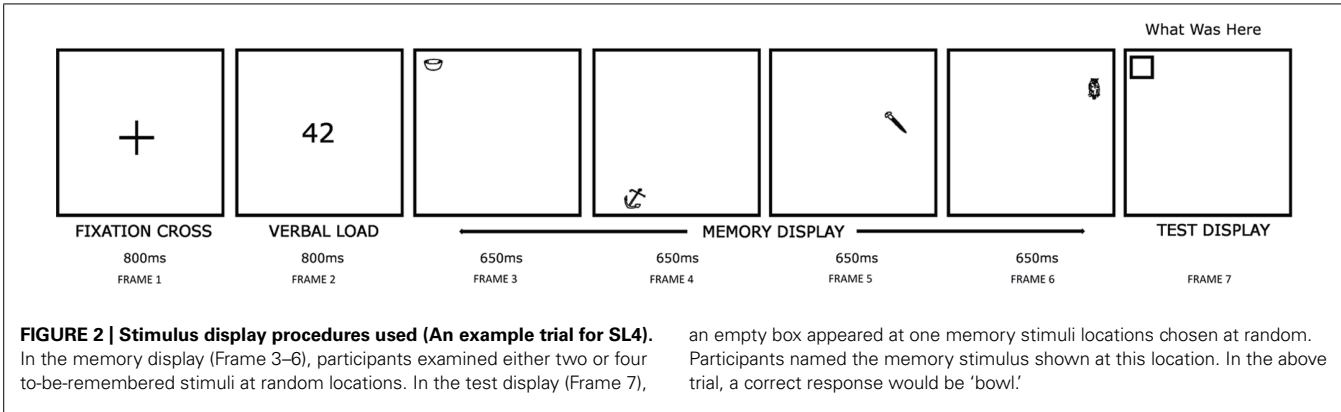
2006). Participants were asked to remember the object-location pairing of each stimulus (i.e., the location in the memory display at which a given stimulus was presented). Participants read aloud the two-digit number (above) whilst examining each memory stimulus to discourage verbal encoding, i.e., a *verbal suppression task* (Baddeley, 1986; Todd and Marois, 2004). Suppression tasks like this have been used previously by researchers investigating age-related changes in VSTM (Brockmole et al., 2008). Participants who were found not to be complying with the procedure were cautioned immediately by the experimenter, and the trial continued; such trials were noted down by the experimenter.

In any given trial, two or four stimuli were used in order to test memory performance at the higher and the lower end of the commonly cited 3–4 item VSTM capacity (Pashler, 1988; Luck and Vogel, 1997; Irwin and Zelinsky, 2002). Stimuli were chosen from different object categories to avoid any processing competition that may arise if the same category stimuli were used, owing to a greater number of shared properties (Bright et al., 2005). Across trials, stimuli were repeated occasionally, but with the constraint that items from different object categories were tested equally often in order to minimize ‘pop out’ effects (Jonides and Yan-tis, 1988). The use of a sequential (rather than simultaneous) stimulus presentation method ensured that configurational cues that could be produced by the relative spatial locations of stimuli were avoided (Jiang et al., 2000; Cowan et al., 2006; Blalock and Clegg, 2010). This isolated observed effects to VSTM specifically, and ensured that each stimulus had been fixated. In addition, a sequential display procedure ensured that any possible spatial crowding effects that may be produced by a simultaneous display were avoided, and enabled us to study the effect of stimulus recency. Each memory stimulus appeared at a new, unique screen location (chosen randomly from one of the 64 imaginary positions of an 8×8 square, each 2.5°); this square window, which covered the central $20^\circ \times 20^\circ$ display screen area, represented the total possible area within which the memory stimuli could be displayed in any trial. Stimulus positions were never repeated within a single trial, and were at least 2.5° apart from one another. There was no delay between successive stimuli, enabling eye movements to be executed immediately after the preceding stimulus was offset (i.e., rather than drifting randomly). Moreover, the execution of eye movements between successive stimuli disrupted iconic memory that could otherwise have supported temporal integration (Eriksen and Hoffman, 1972). After a blank interval of 1000 ms, a test display was presented. This comprised a written command ‘what was here?’ and an empty square box (2.5°) at one of the randomly chosen locations used to present memory stimuli (Frame 7); participants were required to verbally report the name of the stimulus presented at that location. Where participants were entirely unsure, they were asked to report ‘I do not know.’

Participant’s (verbal) responses were recorded manually by the experimenter. The next trial started when the participant clicked a computer mouse button. The importance of accuracy (rather than speed) was emphasized to participants.

Each participant completed two blocks of 56 trials (i.e., 112 trials in total), distributed equally between sequence lengths, i.e., for each participant, for SL2, there were 56 trials in total or 28





trials per block (14 trials per temporal position per block). Similarly for SL4, there were 56 trials in total or 28 trials per block (seven trials per temporal position per block). Approximately 1 h of data capture was required per participant. Rest breaks were given between blocks; data were captured within a single session for each participant.

Performance was measured in terms of error rate, i.e., the proportion of trials in which participants incorrectly reported the name of one of the previously presented (non-target) memory items or an entirely new (non-memory) item. When non-target items were reported, the minimum 2-D Euclidean distances between each non-target item selected and the target item were calculated, and then ranked. For sequence length 4, error rates were grouped as spatial offset rank 1 (wherein the incorrectly selected non-target item was *closest* to the location of the target item), spatial offset rank 2 (where the incorrectly selected non-target item was second closest to the target item), and spatial offset rank 3 (where the incorrectly selected non-target item was *furthest* from the target item). For sequence length 2, only spatial offset rank 1 was possible.

Data were analyzed using ANOVA and *t*-tests as appropriate (see later). Where the assumption of sphericity was violated (identified using Mauchly's test), degrees of freedom were adjusted using the Greenhouse-Geisser procedure.

RESULTS

In less than 5% of trials, participants could not report any item at test; these trials were not subject to further analyses (except for identifying whether older participants used a more stringent criterion to choose a non-memory item, see later). In those incorrect trials in which an object name was reported at test, non-target items were reported in 56.38% of trials, and non-memory items in 43.62% of trials. **Figure 3** shows the mean error rate (pooled across sequence length) for non-target items and non-memory items for both age groups. To examine whether these pooled error rates differed significantly between age groups, a mixed ANOVA with object-recall error type (non-target and non-memory) as a within-subjects factor, and age group (young and older) as a between-subjects factor was performed. Significant main effects of age group [$F(1,32) = 9.01, p = 0.005$] and object-recall error type [$F(1,32) = 7.82, p = 0.009$] were found. A significant interaction between age group and object-recall error type was also

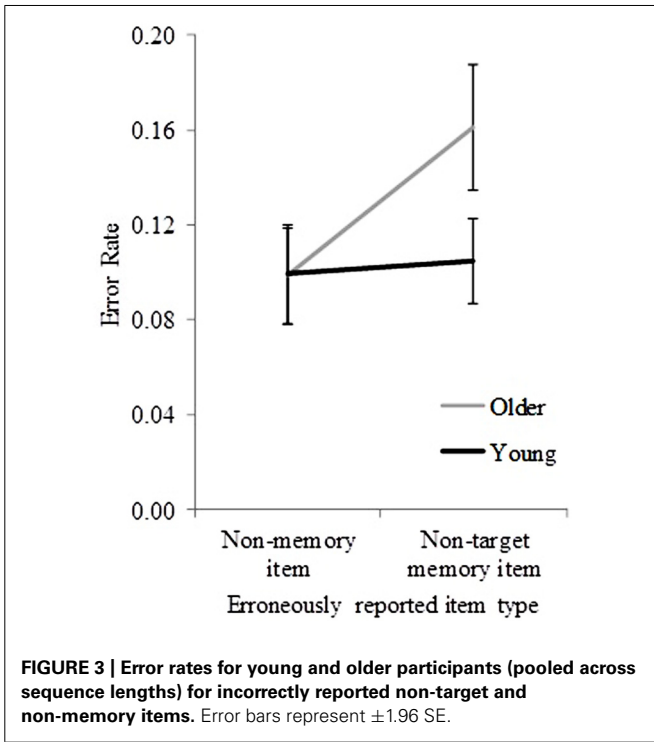


Table 1 | Mean error rates for non-target memory items and non-memory items for individual sequence length for young and older participants.

SL	Object-recall error type	Older age group	Young age group
2	Non-target	0.05 (SD = 0.04)	0.01 (SD = 0.01)
	Non-memory	0.07 (SD = 0.04)	0.06 (SD = 0.05)
4	Non-target	0.28 (SD = 0.09)	0.20 (SD = 0.07)
	Non-memory	0.13 (SD = 0.06)	0.14 (SD = 0.06)

found [$F(1,32) = 5.37, p = 0.03$]. The results suggest a significant effect of age group on the occurrence of different object-recall error types (non-target vs. non-memory).

Table 1 shows mean error rates for non-target memory items and non-memory items for individual sequence length for young and older participants.

To identify how non-target and non-memory error rates differed between age groups across memory target temporal positions for high (SL4) and low (SL2) memory loads, a mixed ANOVA was performed (separately for non-target and non-memory items), in which temporal position (two levels for SL2, four levels for SL4) served as a within-subjects factor and age group (two levels, young and older) served as a between-subjects factor. Non-target error rates were significantly greater for older participants than young participants for both sequence lengths (**Table 2A**).

A significant interaction between age group and memory target temporal position was found for SL4 but not SL2 (**Table 2A**). This suggests that interference from non-target memory items at high memory load affected each age group differently, depending upon whether the memory target was probed at earlier or more recent temporal positions. Non-memory error rates, however, did not differ significantly between age groups for either sequence length, although a significant interaction between age group and memory target temporal position was found for SL4 (**Table 2A**), suggesting that interference from non-memory items at high memory load also affected each age group differently, depending upon whether the memory target was probed at earlier or more recent temporal positions. Furthermore, these results suggest that older participants incurred greater intrusion (compared to young participants)

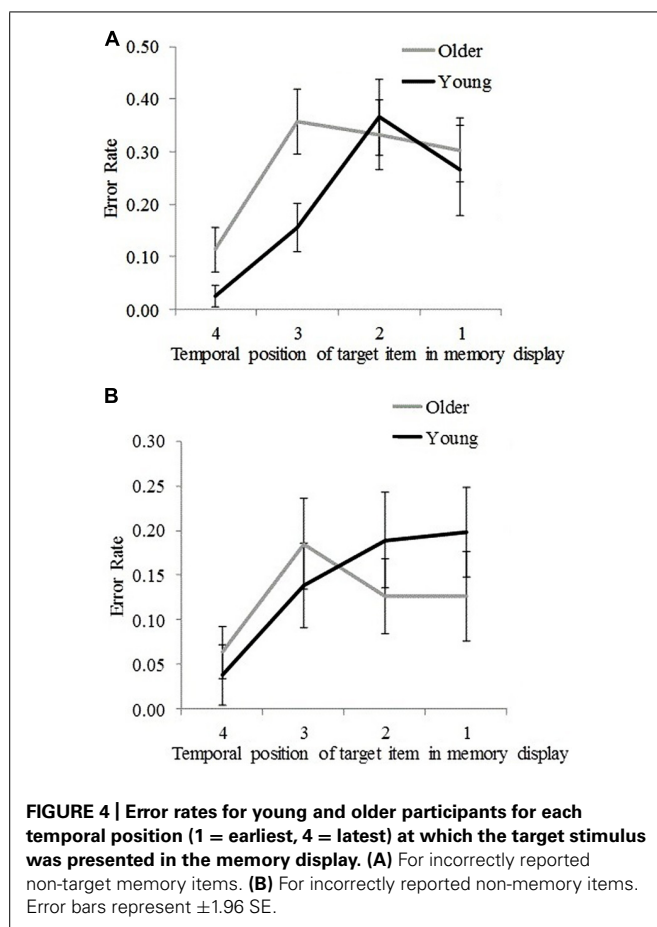
from non-target memory (but not from non-memory) items at both high and low memory loads.

Plotting error rate as a function of the memory target temporal position enables us to visualize the significant interaction effects (reported above) between age group and memory target temporal position for non-target and non-memory items for SL4 (**Figures 4A,B**).

To determine at which temporal positions non-target error rates differed significantly between the two age groups, an independent samples *t*-test was performed. Results are presented in **Table 2B**. Significantly greater non-target error rates were found for older participants compared to young participants for the third and the fourth temporal positions but not for the first and the second temporal positions (**Table 2B**). This suggests that older participants incurred greater proactive interference compared to young participants. An inverse trend was observed for non-memory error rates (*cf.* **Figures 4A,B**), in which younger participants showed greater (nearing the significance threshold of 0.05) error rates compared to older participants when the memory target was presented at the earliest two temporal positions (**Table 2B**). This effect may have arisen because older participants used a more stringent criterion for a non-memory item response when the memory target was presented at the first and second temporal positions. This assertion is supported by an analysis of the frequency of those trials in which participants could not report any item at test (although these occurred in less than 5% of all trials), which were more common

Table 2 | (A) ANOVA results for comparison of error rates between young and older participants for incorrectly reported non-target memory items and non-memory items for SL2 and SL4. (B) *t*-test results for comparison of non-target error rates and non-memory error rates between young and older participants at each temporal position of target item presentation for SL4. (C) ANOVA results for comparison of error rates between incorrectly reported non-target memory items and non-memory items within each age group.

(A)					
SL	Object-recall error type	Main effect of age group	Main effect of temporal position	Interaction between age group and temporal position	
2	Non-target	$F(1,32) = 13, p = 0.001$	$F(1,32) = 0.90, p = 0.35$	$F(1,32) = 2.71, p = 0.11$	
	Non-memory	$F(1,32) = 0.9, p = 0.35$	$F(1,32) = 7.18, p = 0.01$	$F(1,32) = 1.0, p = 0.32$	
4	Non-target	$F(1,32) = 6.62, p = 0.01$	$F(2.53,81.03) = 40.36, p < 0.001$	$F(2.53,81.03) = 6.91, p = 0.001$	
	Non-memory	$F(1,32) = 0.64, p = 0.43$	$F(3,96) = 13.05, p < 0.001$	$F(3,96) = 3.94, p = 0.01$	
(B)					
SL	Object-recall error type	Temporal position 1 (young vs. older)	Temporal position 2 (young vs. older)	Temporal position 3 (young vs. older)	Temporal position 4 (young vs. older)
4	Non-target	$t(32) = -0.73, p = 0.47$	$t(32) = 0.61, p = 0.51$	$t(32) = -5.13, p < 0.001$	$t(32) = -3.57, p = 0.001$
4	Non-memory	$t(32) = 1.92, p = 0.06$	$t(32) = 1.83, p = 0.08$	$t(32) = -1.30, p = 0.20$	$t(32) = -1.10, p = 0.28$
(C)					
SL	Age group	Main effect of error type	Main effect of temporal position	Interaction between error type and temporal position	
2	Older	$F(1,16) = 2.94, p = 0.11$	$F(1,16) = 0.12, p = 0.73$	$F(1,16) = 3.39, p = 0.84$	
	Young	$F(1,16) = 15.22, p = 0.001$	$F(1,16) = 5.95, p = 0.03$	$F(1,16) = 2.48, p = 0.13$	
4	Older	$F(1,16) = 26.02, p < 0.001$	$F(3,48) = 25.75, p < 0.001$	$F(3,48) = 5.14, p = 0.004$	
	Young	$F(1,16) = 5.11, p = 0.04$	$F(3,48) = 50.04, p < 0.001$	$F(3,48) = 3.72, p = 0.02$	



for older participants than young participants, $t(32) = 2.76$, $p = 0.009$.

To examine temporal position effects in SL4 trials, a one-way repeated measures ANOVA was performed on combined error rate data (non-target and non-memory), using memory target temporal position (four levels for SL4) as a within-subjects factor. A significant temporal position effect was found for both age groups [older participants, $F(3,48) = 26.13$, $p < 0.001$; young participants, $F(3,48) = 50.36$, $p < 0.001$], suggesting that more recent items were remembered with greater accuracy (i.e., global recency effects).

To examine how object-recall error types differed from each other *within* each age group across memory target temporal positions, separate analyses were performed for SL2 and SL4 as follows: 2 (error type, non-memory and non-target) \times 2 (temporal position) repeated measures ANOVA for SL2; and 2 (error type, non-memory and non-target) \times 4 (temporal position) repeated measures ANOVA for SL4. Results are presented in **Table 2C**. At high memory load (SL4), non-target error rates were significantly greater than non-memory error rates for both age groups. A significant interaction between object-recall error type and memory target temporal position was found. This suggests that, for both age groups, at high memory load, object-recall error types varied differently as a function of memory target recency. At low memory load (SL2), a reverse trend (non-memory error rates greater

than non-target error rates) was observed, although the difference attained statistical significance only for young participants (**Table 2C**). No significant interaction between the object-recall error type (non-target and non-memory) and memory target temporal position was found for SL2. These results suggest that, as memory load increases from 2 to 4 items, greater interference occurs from non-target memory items when compared to non-memory items in both age groups.

To examine whether the incorrectly reported non-target memory items were more likely to be spatially proximal to the memory target (spatial proximity), a one-way repeated measures ANOVA was performed, separately for each age group, on non-target error rates for SL4 using spatial offset (of the non-target item reported relative to the memory target) ranks 1–3 as a within-subjects factor. A significant main effect of spatial offset rank was found in both young [$F(2,32) = 4.30$, $p = 0.02$] and older [$F(2,32) = 16.63$, $p < 0.001$] participants. To establish, specifically, at which spatial offset ranks non-target error rates differed significantly between age groups, an independent samples t -test was performed on error rates at each spatial offset rank of 1–3 between young and older participants. Older participants were found to exhibit greater error rates at spatial offset rank 1 [$t(32) = 2.60$, $p = 0.01$], and rank 2 [$t(32) = 3.76$, $p = 0.01$] compared to younger participants. No significant difference was found at spatial offset rank 3 [$t(32) = 1.20$, $p = 0.24$] between the two age groups. This suggests that age-related differences in VSTM error rate are influenced significantly by non-target items presented at nearby locations (albeit sequentially) compared to those that are more distant. The average calculated distance between a memory target and the non-target item reported was found to be less than the average distance between all non-target items presented in SL4 trials [older participants, 10.99 vs. 12.06°; young participant, 11.55 vs. 11.95°]. To identify whether these distances differed significantly from one another within each age group, a paired samples t -test was performed separately for young and older participants. A significant difference in older [$t(16) = 2.56$, $p = 0.02$], but not in young participants [$t(16) = 1.29$, $p = 0.21$] was found, suggesting that older participants are more likely to select non-target items that are closer to the memory target. This is further supported by a mixed ANOVA performed on absolute distances (i.e., between the memory target and erroneously reported non-target items) using spatial offset ranks (1–3) as a within-subjects factor, and age group (young and older) as a between-subjects factor, in which a significant main effect of age group was also found [$F(1,32) = 4.49$, $p = 0.04$].

One may suggest that non-memory items presented in earlier trials at the location where a memory target is presented in subsequent trial may intrude significantly upon the memory target item. However, errors for non-memory items that had been presented during earlier trials at the location where a memory target was subsequently presented occurred in less than 1% of trials in both age groups, suggesting that non-significant intrusion occurred from the items presented in earlier trials at the location of a memory target item.

Table 3 shows the distribution of errors depending upon whether incorrectly reported non-memory items belonged to the

Table 3 | Non-memory errors (pooled across sequence lengths) for young and older participants depending upon whether the incorrectly recalled item belonged to the same or different object category to the memory target.

Incorrectly recalled object category	Older participants	Young participants
Different to memory target	75% (SD = 13%)	72% (SD = 9%)
Same to memory target	25% (SD = 13%)	28% (SD = 9%)

same or a different object category relative to the memory target. In 25% of trials for older participants, and 28% of trials for young participants, an item that was not examined in the memory display, but belonged to the same object category as the memory target was reported, suggesting that category information was retained across trials even if detailed item information may have been lost. However, there was no significant difference between the two age groups [$t(32) = 0.84, p = 0.40$], demonstrating that VSTM performance does not differ significantly between age groups as a consequence of competition from unseen items belonging to the same object category as the memory target.

DISCUSSION

In this study, we investigated how different types of VSTM errors (i.e., the incorrect reporting of non-target memory and non-memory items) in a location-cued object-recall task differed within and across age groups. We examined how these differences varied with memory load (the number of items to be remembered) and target stimulus recency. Overall, a greater error rate for non-target memory items compared to non-memory items was found in both young and older participants; however, older participants performed less well overall. At the high memory load (SL4), non-target error rates occurred more often than non-memory error rate in both age groups. However, at the low memory load (SL2), non-memory error rates occurred more often than non-target error rates, although the difference attained statistical significance only in young participants. Furthermore, at the high memory load, young participants exhibited greater error rates (compared to older participants) for non-memory items when the memory target was presented at the earliest two temporal positions. This is possibly due to older participants using a more stringent criterion for producing a non-memory item response. However, a reverse trend was observed for non-target error rates, which were significantly greater for older participants compared to young participants when the memory target was presented at the two most recent temporal positions (Figures 4A,B), suggesting that older participants incurred more proactive interference compared to young participants (Bowles and Salthouse, 2006). Furthermore, non-target error rates (in both age groups) were found to asymptote at the earliest (first and second) temporal positions, countering the predictions of the slot-based model of working memory. According to this model, a fixed number of slots are assumed to exist in VSTM, such that with each newly examined item the probability of forgetting an earlier examined item increases for high memory load. If it was the case that lower memory performance (specifically, greater non-target error rate)

for older participants was a consequence of their having fewer slots, the probability that earlier items would be forgotten would be greater (compared to young participants) in SL4 trials. This was not found with our data. The finding that young participants exhibited reduced non-target error rates compared to older participants when the memory target item was presented at the two most recent temporal positions rather suggests that the memory resources allocated for more recently examined items are prioritized better by young participants above less recently examined items, potentially as a result of greater encoding fidelity, attention, resource redeployment to favor more recent items, or superior executive control. This assumption is compatible with the flexible resource model (Wilken and Ma, 2004; Bays and Husain, 2008).

The finding of a greater error rate overall for older participants may be explained by inhibitory deficits, i.e., the ability to inhibit non-target memory items decreases with age (De Beni and Palladino, 2004; Bowles and Salthouse, 2006; Fiore et al., 2012). In addition, our findings provide evidence for the following factors that influence object representation in VSTM: (i) at high memory loads we incur greater interference from non-target memory items, and at low memory loads we incur greater interference from non-memory items; (ii) we are less adept at remembering less recent non-target memory items, as a consequence of interference incurred from subsequent items (retroactive interference).

One may argue that our findings could have been confounded by systematic differences between young and older participants' gaze control (i.e., shifting the gaze from one stimulus in the sequence to the next). It should be noted that our stimuli were presented one at a time with a display duration that was sufficiently long (at 650 ms) for each object to be captured by the visual system, and gaze to be shifted to the subsequent object. Although we did not record the eye movements directly, we noted those trials in which participants either forgot to read aloud the verbal load (which they were supposed to read aloud every time a memory stimulus was presented), or read it aloud an incorrect number of times (i.e., differing from the number of stimuli presented in the memory display). Such errors were found to occur in less than 2% of trials. Furthermore, no significant differences were found in these trials between young and older participants, suggesting that our results were not confounded by systematic age-related differences in eye-gaze control.

A number of hypotheses are offered as to why VSTM performance may decline with advancing age. Salthouse (1990, 1994) suggests a generalized slowing of overall cognitive processes, while Craik and Byrd (1982) argue for a progressive deterioration in available attentional resources. Others propose impoverished memory representations owing to lower-fidelity sensory inputs (Rabbitt, 1991; Lindenberger and Baltes, 1994; Baltes and Lindenberger, 1997; Murphy et al., 2000), or a general deterioration in our ability to inhibit visual information belonging to objects that are irrelevant to our current goals (Hasher and Zacks, 1988; Zacks and Hasher, 1994). Yet another hypothesis proposes that VSTM performance declines with advancing age due to a decreased ability of older participants to encode and retrieve associations between constituent object features stored in VSTM (Chalfonte

and Johnson, 1996; Mitchell et al., 2000; Naveh-Benjamin, 2000; Cowan et al., 2006). While the aim of this research was not to test one or the other of these hypotheses, our finding of significant age-related differences in proactive interference is better explained by inhibitory deficit hypothesis (Hasher and Zacks, 1988; Zacks and Hasher, 1994). Our findings add to the inhibitory deficit hypothesis by proposing that age-related VSTM decline is influenced by interference originating from items examined *earlier* than a memory target (proactive interference), but not from items examined *after* a memory target (retroactive interference). Similarly, age-related decline in VSTM may also occur due to interference from non-target items that are examined spatially nearby to the memory target, suggesting that spatial coding deficits become more pronounced in VSTM with advancing age.

To summarize, our data demonstrate that normal aging affects VSTM performance in an object-recall task generally, and more specifically that these aging effects were modulated by memory load and target stimulus recency. Overall error rates differed significantly between age groups when the incorrectly reported item was one of the previously presented non-target items, but not when an entirely new item (non-memory item) was reported. For both young and older participants, at the high memory load, error rates for non-target memory items were greater, whereas at low memory load, error rates for non-memory items were greater. Older participants showed higher non-target error rates when the memory target was presented at the two most recent temporal positions (but not at earlier temporal positions). This suggests that proactive interference was greater for older participants compared to young participants. Similarly, greater interference from non-target items that were spatially nearer to the memory target was found for older participants compared to young participants, suggesting impaired spatial coding for older participants. Future studies might consider the influence of changes in other variables, such as stimulus duration, retention interval, and response time, in order to compare age-related differences in the memory decline associated with non-target memory items and non-memory items at each stage of VSTM processing, *viz.* encoding, maintenance, and retrieval.

Our findings have important implications for understanding the mechanisms that underpin age-related VSTM performance decline, and suggest that a less cluttered visual environment may be particularly beneficial to the elderly by reducing the number of irrelevant visual items (that they are less adept at inhibiting), which may improve their performance in everyday visual tasks requiring VSTM. Furthermore, our results are relevant to ongoing debate concerning the most appropriate working memory model, such as the resource (Wilken and Ma, 2004; Bays and Husain, 2008) and slot models (Zhang and Luck, 2008). A lower non-target error rate for young participants compared to older participants was found when the memory target was presented at the two most recent temporal positions in SL4. This may suggest that young participants are more efficient at redeploying memory resources allocated to earlier presented items for more recent items, exhibit greater attention or executive control, or dynamically prioritize new stimuli over earlier stimuli, an assumption that is compatible with the flexible resource model (Wilken and Ma, 2004; Bays and Husain,

2008). Furthermore, our results suggest that VSTM resources are not only shared between recently examined items, but also with items that were examined in the distant past (i.e., from our prior visual experiences), which can produce false memories.

ACKNOWLEDGMENTS

This research was supported by Patrick Geoghegan Health & Well-being Academy, Anglia Ruskin University (in collaboration with South Essex Partnership University NHS Foundation Trust).

REFERENCES

- Alvarez, G. A., and Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15, 106–111. doi: 10.1111/j.0963-7214.2004.01502006.x
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201
- Baddeley, A. D., and Hitch, G. J. (1974). "Working memory," in *Recent Advances in Learning and Motivation*, Vol. 8, ed. G. A. Bower (New York: Academic Press), 47–90.
- Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12
- Bays, P. M., and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854. doi: 10.1126/science.1158023
- Blalock, L. D., and Clegg, B. A. (2010). Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. *Q. J. Exp. Psychol.* 63, 856–862. doi: 10.1080/17470211003690680
- Borella, E., Carretti, B., and De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychol. (Amst.)* 128, 33–44. doi: 10.1016/j.actpsy.2007.09.008
- Bowles, R. P., and Salthouse, T. A. (2006). Assessing the age-related effects of proactive interference on working memory tasks using the rasch model. *Psychol. Aging* 18, 608–615. doi: 10.1037/0882-7974.18.3.608
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Bright, P., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2005). The anatomy of object processing: the role of antero-medial temporal cortex. *Q. J. Exp. Psychol.* 58, 361–377. doi: 10.1080/02724990544000013
- Brockmole, J. R., Parra, M. A., Sala, S. D., and Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychon. Bull. Rev.* 15, 543–547. doi: 10.3758/PBR.15.3.543
- Chalfonte, B. L., and Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Mem. Cogn.* 24, 403–416. doi: 10.3758/BF03200930
- Cowan, N., Naveh-Benjamin, M., Kilb, A., and Saults, J. S. (2006). Life span development of visual working memory: when is feature binding difficult? *Dev. Psychol.* 42, 1089–1102. doi: 10.1037/0012-1649.42.6.1089
- Craik, F. I. M., and Byrd, M. (1982). "Aging and cognitive deficits: the role of attentional resources," in *Aging and Cognitive Processes*, eds F. I. M. Craik and S. Trehub (New York: Plenum Press), 191–211.
- De Beni, R., and Palladino, P. (2004). Decline in working memory updating through ageing: intrusion error analyses. *Memory* 12, 75–89. doi: 10.1080/09658210244000568
- Eriksen, C. W., and Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Percept. Psychophys.* 12, 201–204. doi: 10.3758/BF03212870
- Fiore, F., Borella, E., Mammarella, I. C., and De Beni, R. (2012). Age differences in verbal and visuo-spatial working memory updating: evidence from analysis of serial position curves. *Memory* 20, 14–27. doi: 10.1080/09658211.2011.628320
- Hasher, L., and Zacks, R. T. (1988). Working memory, comprehension, and aging. A review and a new view. *Psychol. Learn. Motiv.* 22, 193–225. doi: 10.1016/S0079-7421(08)60041-9
- Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 31–47. doi: 10.1037/0096-1523.33.1.31

- Hollingworth, A., Richard, A. M., and Luck, S. J. (2008). Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *J. Exp. Psychol. Gen.* 137, 163–181. doi: 10.1037/0096-3445.137.1.163
- Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cogn. Psychol.* 23, 420–456. doi: 10.1016/0010-0285(91)90015-G
- Irwin, D. E., and Andrews, R. V. (1996). "Integration and accumulation of information across saccadic eye movements," in *Attention and Performance XVI: Information Integration in Perception and Communication* eds T. Inui and J. L. McClelland (Cambridge, MA: MIT Press), 125–155.
- Irwin, D. E., and Zelinsky, G. (2002). Eye movements and scene perception: memory for things observed. *Percept. Psychophys.* 64, 882–895. doi: 10.3758/BF03196793
- Jiang, Y., Olson, I. R., and Chun, M. M. (2000). Organization of visual short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 683–702. doi: 10.1037/0278-7393.26.3.683
- Jonides, J., and Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Percept. Psychophys.* 43, 346–354. doi: 10.3758/BF03208805
- Lindenberger, U., and Baltes, P. B. (1994). Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339–355. doi: 10.1037/0882-7974.9.3.339
- Luck, S. J., and Vogel, E. K. (1997). Capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846
- Makovski, T., and Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Mem. Cogn.* 36, 43–52. doi: 10.3758/MC.36.1.43
- Mitchell, K. J., Johnson, M. K., D'Episito, M., Raye, C. L., and Mather, M. (2000). Aging and reflective processes of working memory: binding and test load deficits. *Psychol. Aging* 15, 527–541. doi: 10.1037/0882-7974.15.3.527
- Murphy, D. R., Craik, F. I. M., and Li, K. Z. H. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: tests of an associative deficit hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1170–1187. doi: 10.1037/0278-7393.26.5.1170
- Pashler, H. (1988). Familiarity and visual change detection. *Percept. Psychophys.* 44, 369–378. doi: 10.3758/BF03210419
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* 16, 283–290. doi: 10.3758/BF03203943
- Phillips, W. A. (1983). Short-term visual memory. *Philos. Trans. R. Soc. B Biol. Sci.* 302, 295–309. doi: 10.1098/rstb.1983.0056
- Rabbitt, P. M. A. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol.* 476, 167–176. doi: 10.3109/00016489109127274
- Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. *Dev. Rev.* 10, 101–124. doi: 10.1016/0273-2297(90)90006-P
- Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology*, 8, 535–543. doi: 10.1037/0894-4105.8.4.535
- Sapkota, R., Pardhan, S., and van der Linde, I. (2011). Object-position binding in visual short-term memory for sequentially presented unfamiliar stimuli. *Perception* 40, 538–548. doi: 10.1068/p6899
- Snodgrass, J. G., and Vanderwart, M. (1980). A standardised set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215. doi: 10.1037/0278-7393.6.2.174
- Todd, J. J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754. doi: 10.1038/nature02466
- Vogel, E. K., Woodman, G. F., and Luck, S. J. (2006). The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1436–1451. doi: 10.1037/0096-1523.32.6.1436
- Wilken, P., and Ma, W. J. (2004). A detection theory account of change detection. *J. Vis.* 4, 1120–1135. doi: 10.1167/4.12.11
- Zacks, R. T., and Hasher, L. (1994). "Directed ignoring: inhibitory regulation of working memory," in *Inhibitory Processes in Attention, Memory, and Language* eds D. Dagenbach and T. H. Carr (San Diego, CA: Academic press), 241–264.
- Zelinsky, G. J., and Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Percept. Psychophys.* 67, 676–690. doi: 10.3758/BF03193524
- Zhang, W. W., and Luck, S. S. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233–235. doi: 10.1038/nature06860

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 June 2014; accepted: 22 December 2014; published online: 20 January 2015.

Citation: Sapkota RP, Van Der Linde I and Pardhan S (2015) How does aging affect the types of error made in a visual short-term memory 'object-recall' task? *Front. Aging Neurosci.* 6:346. doi: 10.3389/fnagi.2014.00346

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2015 Sapkota, Van Der Linde and Pardhan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration

David P. McGovern^{1*}, Eugenie Roudaia¹, John Stapleton¹, T. Martin McGinnity² and Fiona N. Newell¹

¹ Trinity College Institute of Neuroscience, Trinity College Dublin, College Green, Dublin, Ireland

² Intelligent Systems Research Centre, University of Ulster, Londonderry, UK

Edited by:

Katherine Roberts, University of Warwick, UK

Reviewed by:

Mark T. Wallace, Vanderbilt University, USA
Adele Diederich, Jacobs University Bremen, Germany

*Correspondence:

David P. McGovern, Trinity College Institute of Neuroscience, Lloyd Building, Trinity College Dublin, College Green, Dublin 2, Ireland
e-mail: mcgoved1@tcd.ie

While aging can lead to significant declines in perceptual and cognitive function, the effects of age on multisensory integration, the process in which the brain combines information across the senses, are less clear. Recent reports suggest that older adults are susceptible to the sound-induced flash illusion (Shams et al., 2000) across a much wider range of temporal asynchronies than younger adults (Setti et al., 2011). To assess whether this cost for multisensory integration is a general phenomenon of combining asynchronous audiovisual input, we compared the time courses of two variants of the sound-induced flash illusion in young and older adults: the *fission* illusion, where one flash accompanied by two beeps appears as two flashes, and the *fusion* illusion, where two flashes accompanied by one beep appear as one flash. Twenty-five younger (18–30 years) and older (65+ years) adults were required to report whether they perceived one or two flashes, whilst ignoring irrelevant auditory beeps, in bimodal trials where auditory and visual stimuli were separated by one of six stimulus onset asynchronies (SOAs). There was a marked difference in the pattern of results for the two variants of the illusion. In conditions known to produce the fission illusion, older adults were significantly more susceptible to the illusion at longer SOAs compared to younger participants. In contrast, the performance of the younger and older groups was almost identical in conditions known to produce the fusion illusion. This surprising difference between sound-induced fission and fusion in older adults suggests dissociable age-related effects in multisensory integration, consistent with the idea that these illusions are mediated by distinct neural mechanisms.

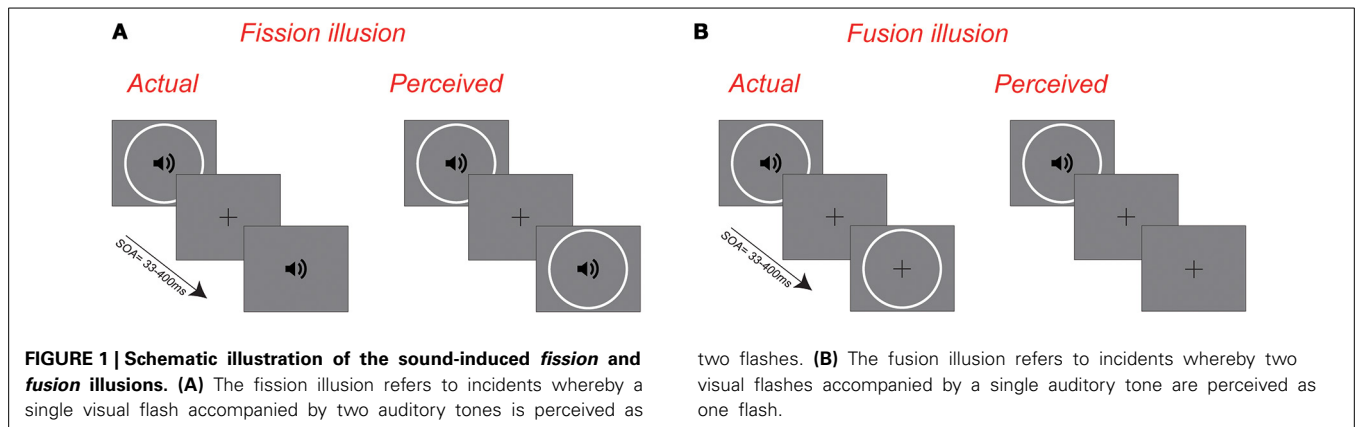
Keywords: multisensory integration, sound-induced flash illusion, perception, time window of integration, aging

INTRODUCTION

The aging process is accompanied by a gradual decline in many aspects of perceptual function. Perhaps the most salient examples of sensory decline are found in vision and audition, where visual acuity and auditory sensitivity decrease in an age-dependent manner (Pitts, 1982; Liu and Yan, 2007). More recent work has examined whether aging affects the way in which information is combined across the senses, and whether this multisensory integration could help to compensate for the reduced sensitivity to unisensory information. So far, however, the evidence for an age-related benefit of multisensory integration remains surprisingly equivocal with examples of multisensory enhancement (Laurienti et al., 2006; Peiffer et al., 2007; Diederich et al., 2008) and impairment (Setti et al., 2011; Stapleton et al., 2014), as well as instances where integration appears to be reduced in older adults (Stephen et al., 2010; Roudaia et al., 2013). As such, it remains an open question as to whether multisensory integration generally compensates for age-related unisensory deficits or whether specific aspects of this integration process also decline with age.

Audiovisual illusions provide a useful means of assaying multisensory integration in human observers. The sound-induced double-flash illusion (Shams et al., 2000), for example, refers

to instances whereby a single visual flash accompanied by two auditory tones is erroneously perceived as two flashes (see **Figure 1A**). Whereas younger adults perceive this *fission* illusion only when the time interval between tones is relatively short (Shams et al., 2002), older adults are susceptible to this illusion across a much wider range of temporal asynchronies (Setti et al., 2011; Stapleton et al., 2014), presumably owing to an enlarged temporal window of integration (e.g., Diederich et al., 2008). This finding suggests that, under certain conditions, the integration of incongruous audiovisual signals leads to an age-related cost in perception. We wondered whether this cost of integration was due to the specific conditions of this illusion or represented a more general phenomenon associated with combining audiovisual input. For instance, it may be that older adults are generally more prone to perceiving multisensory illusions across a wider range of temporal asynchronies, indicative of a general cost associated with multisensory integration. On the other hand, the extent of the temporal window of integration is known to vary with different stimuli and task demands (e.g., Vatakis and Spence, 2006; Stevenson and Wallace, 2013) and it may be that older adults display an enhanced susceptibility to some multisensory illusions, but not others.



A lesser-known variant of the sound-induced flash illusion exists, in which two visual flashes accompanied by a single auditory beep are perceived as a single flash (Andersen et al., 2004). This fusion effect (see **Figure 1B**) is assessed in an identical manner to the fission illusion and is observed across a similar range of stimulus onset asynchronies in younger participants (Apthorp et al., 2013). As such, it provides an excellent means for assessing whether the age-dependent cost for multisensory integration extends to other conditions that involve the combination of audiovisual signals. To test this possibility, 25 younger and older participants performed an audiovisual task in which they had to indicate whether they perceived one or two flashes, whilst ignoring irrelevant auditory tones. Conditions known to produce fission and fusion were randomly interleaved with unisensory and multisensory control conditions to allow us to determine measures of response bias and participant lapses.

To preview our results, we replicated the original finding by Setti et al. (2011) that older adults are susceptible to the sound-induced fission illusion over a much wider range of cross-modal stimulus onset asynchronies than younger adults. However, this was not the case for trials that induced the fusion illusion where the performance of older adults was very similar to the younger group. Specifically, both groups were susceptible to the illusion when stimuli were separated by a short interval, but responded accurately in conditions with moderate to long stimulus onset asynchronies. These results point to dissociable age-related effects of multisensory integration and suggest that caution is required in interpreting multisensory behavioral effects in older adults.

METHODS

PARTICIPANTS

Twenty five younger (10 male, age range: 18–30 years, mean age: 24 years old) and 25 older participants (9 male, age range: 65–88 years, mean age: 71 years old) volunteered to take part in the study. The younger participants were recruited from the student population of Trinity College Dublin and were compensated with research credits for their time. Older volunteers were community-living adults recruited through advertisements in local newspapers and community groups and were compensated for their travel expenses. All participants were naive to the purposes of the study and provided written consent to participate.

We assessed all older adults across a range of sensory functions and also on cognitive ability in order to screen for cognitive impairment. Visual acuity in near and far ranges was measured in older participants using the SLOAN Two-Sided ETDRS Near Vision and the 4 m 200 Series Revised ETDRS charts (Precision Vision, La Salle, and Illinois, USA), respectively. Contrast sensitivity was estimated using the Pelli-Robson Contrast Sensitivity Test. All older participants showed normal or corrected-to-normal acuity and contrast sensitivity for their age. Hearing ability was assessed using a modified version of the Hughson-Westlake method via a Kamplex BA 25 screening audiometer. All participants included in the study displayed thresholds within the normal limits for their age. The Montreal Cognitive Assessment was also administered to all the older participants to screen for cognitive impairment. Participants who scored below a score of 24/30 were excluded from the study. The Trinity College School of Psychology ethics board approved all recruitment and experimental procedures, and the experiment protocol was conducted in accordance with the principles of the Declaration of Helsinki.

APPARATUS AND STIMULI

Stimulus generation and presentation were controlled by an Apple Mac Pro computer on a HP L1710 monitor at a refresh rate of 60 Hz and a spatial resolution of 1280 × 1024 pixels. Participants were positioned at a distance of 57 cm from the screen, with head position supported by a chin rest. Experimental testing was conducted in a darkened, windowless room.

Stimuli were created and displayed in Matlab version 7.14 (R2012a) using Psychtoolbox (Brainard, 1997; Pelli, 1997). The visual stimulus was a hard-edged annulus presented at maximum luminance and displayed for 17 ms. The inner and outer edges of the annulus stimulus extended 8.5 and 10° from the center of the screen, respectively. The auditory stimulus was a brief auditory tone with a frequency of 3.5 KHz, which was presented for 10 ms via Sennheiser HD 202 headphones at a sound pressure level of 65 dB.

PROCEDURE

On each trial, participants were presented with one or two visual flashes, accompanied by one, two or no auditory beeps. Thus, there were six conditions in total, representing all possible combinations of flashes and beeps. For convenience, we will

subsequently refer to these conditions by an abbreviation, which relates to their veridical percept. For example, trials described as 2F1B refer to those where two flashes were accompanied by one beep. For all conditions, participants were required to report how many flashes they perceived and were instructed to ignore the auditory beeps, which were irrelevant to the task. Participants indicated their response with a key press. Conditions known to produce the fission (1F2B) and fusion (2F1B) illusions were randomly interleaved with unisensory (1F and 2F) and multisensory (1F1B and 2F2B) control trials, such that there were six different conditions comprising an equal number of trials.

In conditions containing two flashes or two beeps, auditory and visual stimuli were separated by different stimulus onset asynchronies (SOAs). Seventeen younger and 25 older participants completed the experiment with 6 SOAs ranging between 33 and 400 ms (33, 50, 100, 150, 200, 400 ms), while 8 younger participants completed the experiment with 4 SOAs (50–200 ms). Participants collected a minimum of 10 trials per SOA for each condition, leading to total number of 360 trials (6 SOAs \times 6 conditions \times 10 repeats). At regular intervals over the course of the experiment, participants were prompted to take a self-timed break to avoid fatigue. The experiment lasted approximately 25–30 min.

DATA ANALYSIS

To examine the temporal bounds of susceptibility to the sound-induced flash illusion, we first examined the proportion of incorrect responses across all SOAs in the 1F2B and 2F1B conditions, known to produce the fission and fusion illusions, respectively. In a second analysis, we used signal detection theory to determine whether the change in the proportion of illusory reports resulted from changes in perceptual sensitivity, response bias, or both. This approach was based on that of Rosenthal et al. (2009) and further details of the analysis can be found there. Briefly, for each participant, we calculated the sensitivity (d') and response bias (c) for discriminating between one and two flashes at each SOA and for each auditory beep condition separately. Perceptual sensitivity was calculated with the following equation:

$$d' = z(H) - z(FA) \quad (1)$$

where H denotes the proportion of correctly reported multiple flashes (i.e., hits), FA denotes the proportion of incorrectly reported multiple flashes, and $z(p)$ represents the inverse of the cumulative normal distribution. Using these same definitions, response bias was calculated as:

$$c = 0.5 * (z(H) + z(FA)) \quad (2)$$

In cases where $p = 0$ or $p = 1$ (i.e., where participants reported all hits or no false alarms), these variables were approximated by $1/n$ and $1-1/n$, respectively (where n is the total number of trials used to calculate H and FA).

Statistical analyses were conducted using mixed-model ANOVAs to analyse the effects of age group and SOA on the proportion of illusory responses, d' and c . When appropriate, the Greenhouse-Geisser correction was applied to adjust the degrees

of freedom of within-subject tests to correct for violations of the sphericity assumption and, in these cases, the adjusted p -value is reported. When multiple one-sample t -tests or pairwise comparisons were performed, the Bonferroni correction was used to maintain a family-wise Type I error rate at 0.05 and the adjusted p -value is reported.

RESULTS

We first compared the performance between younger and older groups in the conditions known to produce the fission illusion (1F2B). **Figure 2** shows the group-averaged proportion of illusory responses to the 1F2B trials for both groups as a function of SOA. Young participants experienced the sound-induced fission illusion when the auditory beep stimuli were separated by short intervals, but their performance improved with increasing SOA, consistent with previous reports (Shams et al., 2002; Setti et al., 2011; Aphorpe et al., 2013). In contrast, whereas older adults were as susceptible to the illusion as younger adults at shorter SOAs (33–50 ms), they remained susceptible to the illusion even for the longest SOA presented between the auditory beeps. A mixed-model 2 (age group) \times 6 (SOA) ANOVA on performance to the 1F2B trials revealed significant main effects of SOA [$F_{(5, 200)} = 8.92$, $p_{\text{adj}} < 0.001$, GGeps = 0.48] and age group [$F_{(1, 40)} = 8.09$, $p = 0.007$], as well as a significant age group \times SOA interaction [$F_{(5, 200)} = 5.19$, $p_{\text{adj}} = 0.005$, GGeps = 0.48], indicating that the temporal limits of the fission illusion differed in younger and older groups. To determine the range of SOAs producing the illusion in each age group, we compared the proportion of illusory responses at each SOA with veridical performance (i.e., error rate = 0) using one-sample t -tests. The younger group showed a significant fission effect for SOAs ranging between 33 and 150 ms [$t_{(24)} > 4.04$, $p_{\text{adj}} < 0.002$] only, and not for longer SOAs [200 ms: $t_{(24)} = 2.83$, $p_{\text{adj}} = 0.06$; 400 ms:

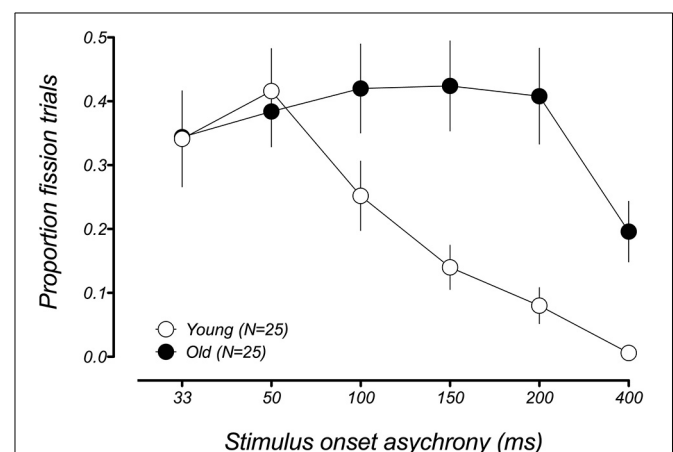


FIGURE 2 | Mean proportion of illusory responses in the fission condition (1F2B) as a function of stimulus onset asynchrony (SOA) for younger (white data points) and older (black data points) groups. Both age groups experience the fission illusion on a similar amount of trials for short SOAs. However, whereas the illusion becomes less frequent for younger adults at longer SOAs, older adults still report the illusion on a significant proportion of trials for the longest SOA. Error bars represent ± 1 SE of the mean.

$t_{(16)} = 1.0$, $p_{\text{adj}} = 1.0$]. In contrast, the older group showed a significant fission effect at all six SOAs [33–400 ms, $t_{(24)} > 4.13$, $p_{\text{adj}} < 0.002$].

This result replicates a previous finding from Setti et al. (2011), who reported that older adults were more susceptible to the sound-induced fission illusion across a wide range of SOAs compared to younger adults. One difference in the current study was the inclusion of the 400 ms condition. This was in part motivated by the fact that it was not clear from the Setti et al. study what interval between the auditory stimuli would be required to facilitate a return to veridical performance in older adults. Although there was a significant fission effect at 400 ms in older adults tested in the current study, it is clear that the illusion occurs less frequently than at shorter SOAs and a return to veridical performance at a longer duration appears likely. Coupled with the similarities in performance to the younger group at short SOAs, this improved performance at longer SOAs suggests that these effects may arise from an enlarged temporal window of integration of multisensory inputs in older adults (e.g., Diederich et al., 2008), a point we return to in the discussion.

A very different pattern of results was observed in the 2F1B condition, which is known to produce the fusion illusion (Figure 3A). Similar to the fission illusion, the performance of the younger participants suggested a large fusion effect at short SOAs, while incidents of the illusion were relatively rare at longer intervals. While older adults appear to be more susceptible to the illusion at shorter SOAs, the temporal constraints of the effect were very similar to their younger counterparts. The higher proportion of illusion responses in the older group most likely reflects group differences in the unisensory conditions containing two flashes. Indeed, a paired sample t -test revealed that older participants were significantly less accurate than younger adults in the 2F condition at short SOAs [at 50 ms the proportion correct in younger and older groups was 0.91 and 0.64, respectively, $t_{(24)} = 3.5527$, $p = 0.002$]. This result is consistent with declines in temporal acuity with age (e.g., Misiak, 1951). Since the primary interest of the current study was in multisensory interactions, each participant's data were normalized by their accuracy level in the 2F condition at each SOA, to better reflect the proportion of fusion reports that occur as a result of the auditory tone rather than poor visual temporal resolution. Figure 3B shows the data replotted from Figure 3A, following this baseline correction. Represented this way, the data from the younger and older groups are almost identical and there was no significant effect of age group [$F_{(1, 40)} = 0.048$, $p = 0.83$], and no interaction between age group and SOA [$F_{(5, 200)} = 1.26$, $p_{\text{adj}} = 0.29$, GGeps = 0.57], indicating that the magnitude and temporal limits of the illusion were very similar across younger and older adults. Moreover, comparing the normalized proportions of illusory responses with veridical performance at each SOA revealed that both younger and older participants experience the illusion for the same range of SOAs, namely 33–100 ms [younger: $t_{(24)}s > 4.08$, $p_{\text{adj}} < 0.003$; older $t_{(24)}s > 4.33$, $p_{\text{adj}} < 0.002$], while performance did not differ from veridical performance for SOAs of 150 ms and longer [younger 150 ms: $t_{(24)} = 2.53$, $p_{\text{adj}} = 0.11$; older 150 ms: $t_{(24)} = 2.65$, $p_{\text{adj}} = 0.08$].

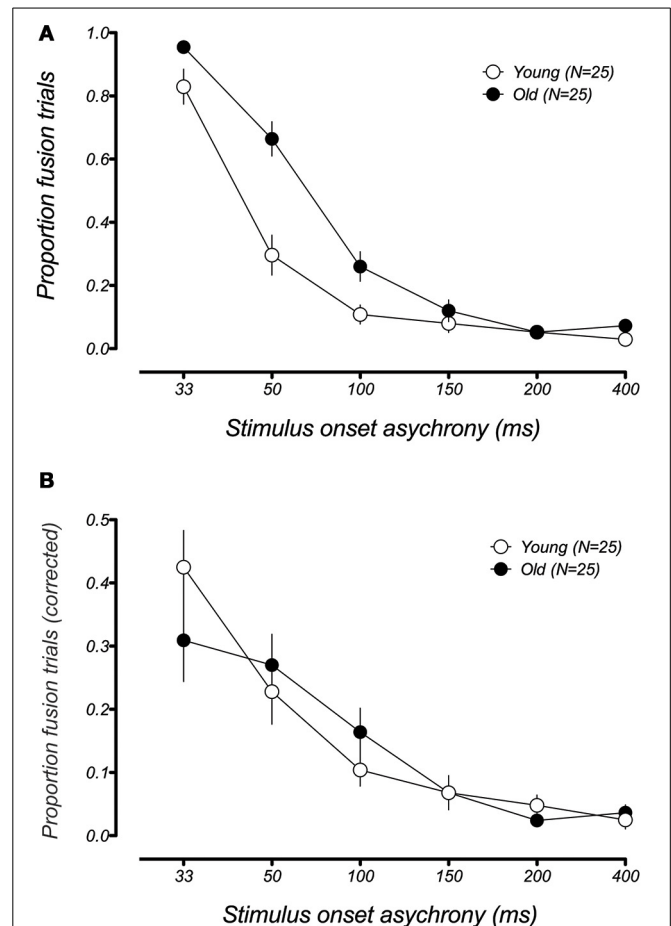
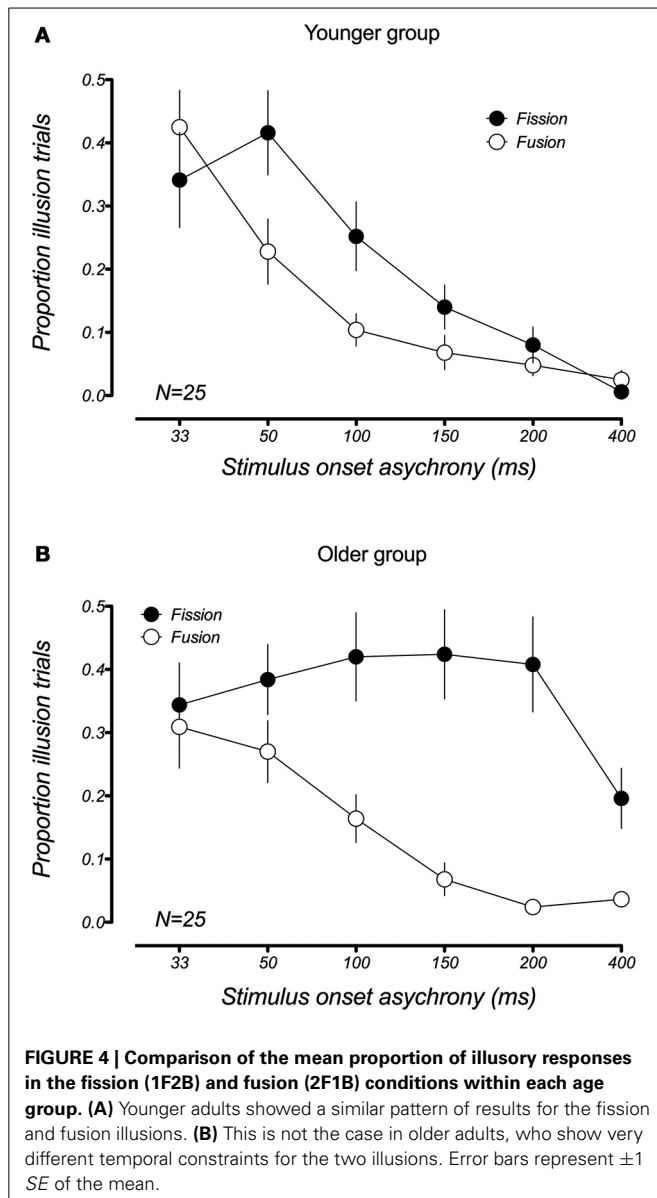


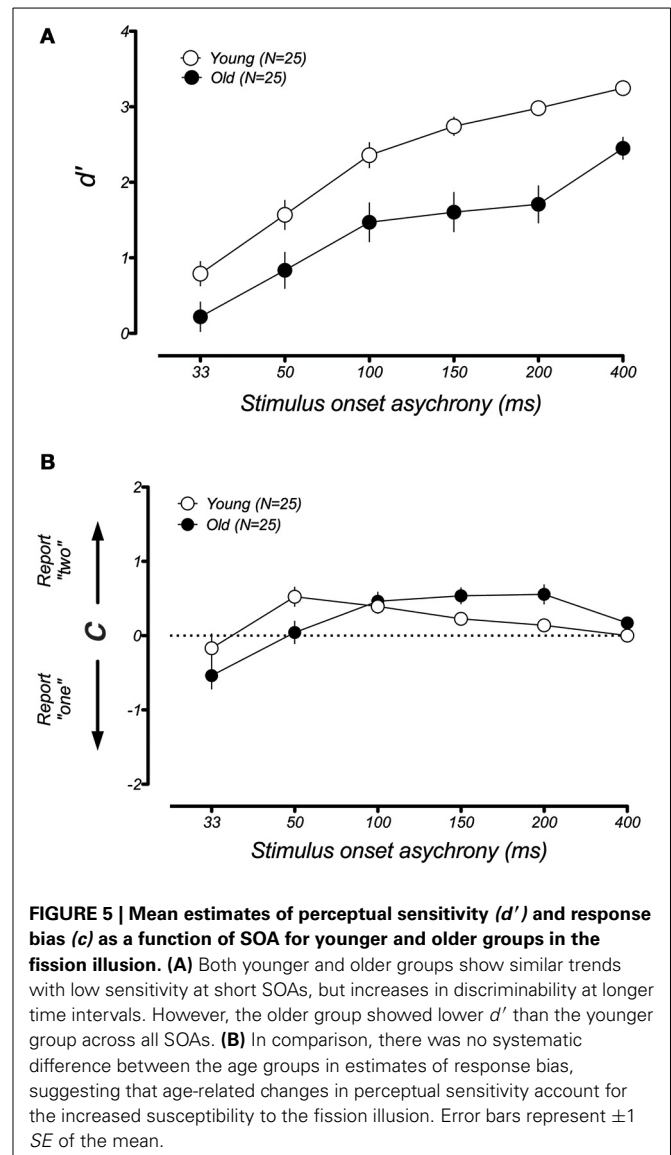
FIGURE 3 | Mean proportion of illusory responses in the fusion condition (2F1B) as a function of SOA for younger (white data points) and older (black data points) groups. (A) Younger and older groups display similar performance on fusion trials, with both groups showing susceptibility to the illusion at short SOAs and a rapid decline in the illusion as the interval between flashes is increased. **(B)** Mean proportion of illusory fusion responses from (A) after being normalized by individual performance in the unisensory 2F condition. When individual and group differences in performance in the unisensory condition are taken into account in this way, the curves for both groups overlap each other, demonstrating that the temporal bounds of the fusion illusion does not differ across age groups. Error bars represent ± 1 SE of the mean.

Although the fission and fusion variants of the sound-induced flash illusion appear similar from a behavioral perspective, there is debate as to whether the two illusions are driven by the same or different neural processes (Mishra et al., 2007, 2008; Aphorpe et al., 2013). Aphorpe et al. suggested that two illusions stemmed from a common mechanism based on the high degree of similarity between the time courses of the fission and fusion illusion in younger participants. Consistent with this finding, our data also show this similarity in the temporal bounds of the illusions in younger participants (Figure 4A) and a 2 (illusion type) \times 6 (SOA) repeated measures ANOVA confirmed that there was no significant main effect of illusion type [$F_{(1, 16)} = 0.07$, $p = 0.79$], and no significant illusion \times SOA interaction [$F_{(5, 75)} = 1.84$,

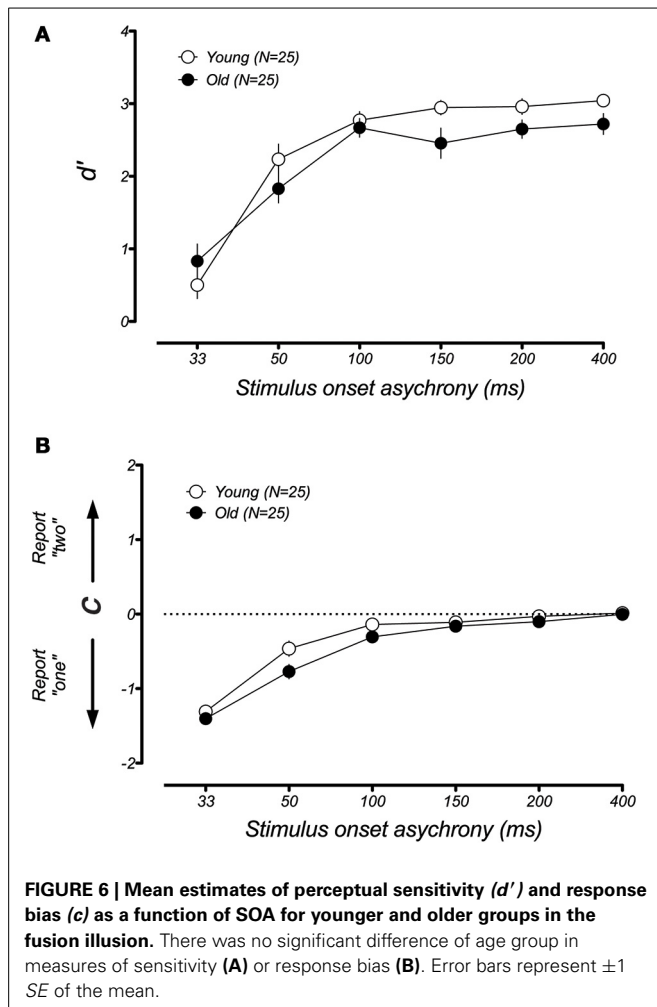


$p_{\text{adj}} = 0.16$, GGeps = 0.53]. However, this was clearly not the case for older participants (Figure 4B), where the fission and fusion effects had very different temporal constraints. For the older group, the ANOVA revealed a significant main effect of illusion type [$F_{(1, 24)} = 19.5$, $p < 0.001$], as well as a significant illusion \times SOA interaction [$F_{(5, 120)} = 6.16$, $p_{\text{adj}} = 0.001$, GGeps = 0.54]. This difference in the temporal constraints of the two illusions in older adults supports the hypothesis that the two illusions result from distinct neural mechanisms.

In younger participants, the sound-induced fission illusion has been shown to result from both a decrease in visual sensitivity and a shift in criterion (McCormick and Mamassian, 2008). We wondered which of these changes could explain the large performance difference between the younger and older groups in the fission condition. For instance, it could be that aging causes a genuine change in the perception of the visual flash during the trials

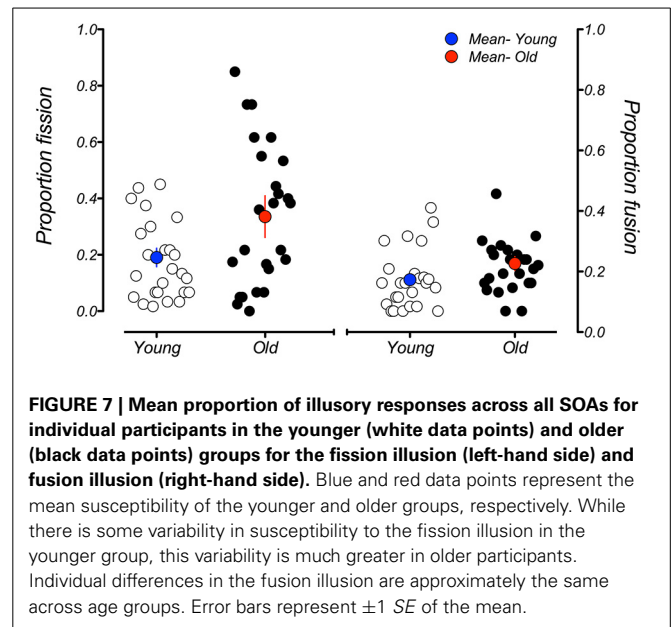


that lead to the illusion. Alternatively, older participants might simply become confused or distracted by the presence of the auditory tone (e.g., Andres et al., 2006), leading to a larger response bias. To address this issue, we used a signal detection analysis to separate the changes in participant sensitivity from general shifts in response bias (see Methods and Rosenthal et al., 2009). Figure 5 plots the changes in d' and response bias as a function of SOA for the fission illusion. For both groups, discrimination was poor at short SOAs and improved when the time interval between stimuli was made longer. However, a lower d' was found for the older group than the younger group across all SOAs and there was a significant difference between the age groups [$F_{(1, 40)} = 19.49$, $p < 0.001$]. In contrast, estimates of response bias were similar across both young and old groups and there was no significant group effect [$F_{(1, 40)} = 0.34$, $p = 0.53$]. Thus, it appears that the age-related difference in performance in the sound-induced fission illusion resulted from a reduction in perceptual sensitivity of



older adults. In the case of the fusion illusion (Figure 6), there was no significant differences between the age groups for measures of d' [$F_{(1, 40)} = 0.97$, $p = 0.33$] or response bias [$F_{(1, 40)} = 2.41$, $p = 0.13$].

A commonly reported observation regarding the sound-induced fission illusion is its high degree of between-subject variability (e.g., Mishra et al., 2007; Stevenson et al., 2011; de Haas et al., 2012). We also observed individual differences of this nature and were interested in how much it could account for the age-related differences we see in the group-averaged data. Figure 7 plots each individual's performance, averaged across all SOAs, for the fission and fusion illusions. Plotted this way it is clear that, while individual differences do exist for the fission illusion in younger adults, the degree of variability is much higher in older adults. The data also suggest that there may exist two distinct groups within the current older cohort: those whose performance is located below the group mean data point, who experienced the fission illusion to a similar extent to the younger group, and those whose performance is above the mean data point, who were much more susceptible to the fission illusion. Importantly, this difference cannot be explained by an effect of aging within our older sample. When the magnitude of individual fission effects in older



participants were plotted as function of participant age, there was a modest positive slope in the linear regression line fitted through the data (0.011) indicating some relationship between these variables. However, the slope did not significantly differ from zero ($p = 0.192$), suggesting that this split in the older group is not solely due to age of the participants. The reasons for the difference between these individuals are not clear, but ongoing work in our laboratory is attempting to understand the factors that lead to this variability in performance across participants. In contrast, the variation in performance in the fusion conditions was much smaller and there were no obvious differences between the young and old groups.

DISCUSSION

The aging process leads to significant changes in all sensory systems and a variety of cognitive functions. Multisensory integration plays a key role in bridging the gap between these sensory functions and higher-order cognitive processing, yet research into the effects of aging on this process has been equivocal (Laurienti et al., 2006; Poliakoff et al., 2006; Peiffer et al., 2007; Setti et al., 2011, for a review see Mozolic et al., 2012). In the current study, we replicated previous findings showing that older adults are susceptible to the sound-induced fission illusion across a wider range of SOAs than younger adults (e.g., Setti et al., 2011). We extended this line of research by showing that this enhanced susceptibility to the illusion results from changes in perceptual sensitivity, rather than changes in response bias, and that the older group was significantly more variable in their susceptibility to the fission illusion than their younger counterparts. Surprisingly, however, we did not observe equivalent age-related changes in susceptibility to the sound-induced fusion illusion, with older adults performing on a par with the younger group. In the following section, we discuss the role of cognitive factors on our results and suggest potential mechanisms to explain the discrepancy in the susceptibility of older adults to the fission and fusion illusions.

Multisensory integration plays an important intermediary role between perception and cognition, where the brain must merge bottom-up, stimulus-driven input from primary sensory areas with top-down guidance from a range of cognitive processes. For example, integration of perceptual signals across the senses can capture attention through bottom-up processes, while top-down selective attention can facilitate the integration of multisensory inputs or lead to a spread of attention across the senses depending on the particular task demands (e.g., Talsma and Woldorff, 2005; Talsma et al., 2010). As such, the role of cognitive processes, such as top-down attention, must be considered in explaining any multisensory effect. This task becomes considerably more difficult when assessing multisensory perception in older adults, given the systematic and age-dependent decreases in unisensory function, as well as changes in selective attention that are known to influence both unisensory (e.g., Hasher et al., 1991; Alain and Woods, 1999) and cross-modal perception (e.g., Andres et al., 2006; Poliakoff et al., 2006). For example, some studies show that older adults are more likely to become distracted by irrelevant auditory information when performing an auditory-visual oddball task (Andres et al., 2006), suggesting an impaired ability to filter out task-irrelevant auditory noise. However, other evidence suggests that older adults are not impaired in their ability to selectively attend to the visual modality (Hugenschmidt et al., 2009).

Could age differences in attentional control explain the increased susceptibility to the sound-induced fission illusion experienced by older adults? We suspect not for two reasons. First, one might expect that any increase in distractibility would lead to an increase in response bias, rather than affecting perceptual sensitivity, as the participant would be tempted to respond in line with the number of auditory stimuli presented. However, our signal detection analysis revealed that age-related increases in susceptibility to the illusion were primarily due to decreases in sensitivity (d'), indicating that the auditory beeps generated illusory second flashes (see also McCormick and Mamassian, 2008). Second, if a deficit in selective attention underpinned the increased susceptibility to the fission illusion, we would expect to observe a similar pattern of results for the fusion illusion, given the similarities in the task structure and experimental conditions. This was not the case, however, with the younger and older groups showing similar performance in fusion conditions. Thus, it is unlikely that the current results can be explained by age-related deficits in suppressing irrelevant auditory stimuli. This is not to fully rule out the influence of cognitive factors on our results, however, which undoubtedly play a role. In particular it seems likely that differences in cognitive function might help to explain the increased inter-subject variability in the older group performance on the fission illusion. This is a hypothesis we are currently pursuing.

The current data show that while older adults are susceptible to the fission illusion across a wider range of SOAs than their younger counterparts, their performance does appear to recover if the interval between the auditory tones is long enough. Similar to findings in other studies of multisensory processing in older adults (Laurienti et al., 2006; Peiffer et al., 2007; Diederich et al., 2008), this pattern of results is consistent with the notion of an extended temporal window of integration in older

adults. This broader temporal window is believed to arise from slowing of peripheral sensory processing, rather than general cognitive decline (Diederich et al., 2008), and has been proposed to help compensate for unisensory deficits in certain conditions (Laurienti et al., 2006; Peiffer et al., 2007). Based on these findings, we expected older adults would also be susceptible to the fusion illusion across a larger range of SOAs than younger participants. This did not prove to be the case, however, with older adults displaying a similar level of performance to the younger group.

What could cause these large differences in age-related susceptibility to the fission and fusion illusions? One possible explanation is that the fusion illusion does not constitute a genuine example of multisensory integration. This explanation seems unlikely, however. Both EEG and fMRI studies (Watkins et al., 2007; Mishra et al., 2008) investigating the neural correlates of the fusion illusion have demonstrated a causal role between activity in the superior temporal sulcus, a brain region believed to play a pivotal role in multisensory integration (Beauchamp et al., 2004), and the subjective experience of the illusion. For instance, Watkins et al. (2007) compared trials in which participants experienced the fusion illusion with those where they did not. In trials where participants reported the illusion, there was increased BOLD activity in the right superior temporal sulcus and decreased activity in primary visual cortex and this was not the case when participants responded veridically. This pattern of results is consistent with the idea that the subjective experience of the illusion is mediated by feedback connections from polymodal areas to primary sensory cortices (Mishra et al., 2008) and suggests that the sound-induced fusion illusion is a *bona fide* example of multisensory integration.

A more likely explanation for this discrepancy between age-related fission and fusion effects is that each variant of the illusion has its own temporal integration window derived from distinct networks of activation in the brain (Mishra et al., 2007, 2008). This explanation is consistent with previous research showing large differences in the size of temporal integration windows depending on the type of stimulus to be integrated and the task to be performed (Vatakis and Spence, 2006; Stevenson and Wallace, 2013). It is also consistent with EEG studies that show significant differences in the timing and localization of the major ERP components associated with each illusion (Mishra et al., 2007, 2008). In these studies, Mishra et al. combined an ERP difference analysis with source localization techniques to identify the different patterns of cortical activity underlying each variant of the illusion. A trial-by-trial analysis of ERPs suggested that the latencies of the major components underlying the fission illusion occurred at 110, 120, and 130 ms, reflecting activity in auditory, visual and superior temporal cortices, respectively (Mishra et al., 2007). On the other hand, the major components for the fusion illusion were observed at much later latencies (180 and 240 ms), but again involved feedback from superior temporal sulcus to visual cortex (Mishra et al., 2008). Interestingly, the overall pattern of cortical activity for each illusion differed markedly from that of the congruent audiovisual input (i.e., 1F1B or 2F2B). From these findings the authors concluded that, despite appearing to be reciprocal perceptual phenomena, very different neural circuits underlie the fission and fusion illusions.

These findings may help to explain the dissociation between the older group's performance on the fission and fusion illusions in the current study. For instance, it is conceivable that the cortical network underlying the fusion illusion could be preserved in older adults, while one or more areas within the fission network may experience a decrease in processing speed. Indeed, Peiffer et al. (2009) found that the pattern of cross-modal deactivation of the visual cortex by auditory stimuli differed between younger and older adults, indicating age-related changes in cross-modal interactions between sensory cortices. Importantly, the current results, together with those from Mishra et al. (2007, 2008), suggest that a degree of caution is required in interpreting how multisensory integration is affected by age on the basis of a single behavioral effect. Rather, these findings suggest that a more prudent approach would be to treat each study independently, considering factors such as the particular task demands, the quality and type of sensory inputs involved, as well as the underlying neural mechanisms that give rise to the effect in question.

ACKNOWLEDGMENTS

This work was funded by the Government of Ireland Postdoctoral Research Fellowship awarded to David P. McGovern by the Irish Research Council, and by the DELNI Cross-Border Research and Development Funding Programme "Strengthening the All-Island Research Base," awarded to the Intelligent Systems Research Centre, University of Ulster and Institute of Neuroscience, Trinity College Dublin.

REFERENCES

- Alain, C., and Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. *Psychol. Aging* 14, 507–519. doi: 10.1037/0882-7974.14.3.507
- Andersen, T. S., Tiippana, K., and Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cogn. Brain Res.* 21, 301–308. doi: 10.1016/j.cogbrainres.2004.06.004
- Andres, P., Parmentier, F. B., and 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, 2564–2568. doi: 10.1016/j.neuropsychologia.2006.05.005
- Apthorp, D., Alais, D., and Boenke, L. L. (2013). Flash illusions induced by visual, auditory and audiovisual stimuli. *J. Vis.* 13, 1–15. doi: 10.1167/13.5.3
- Beauchamp, M. S., Argall, B. D., Bordurka, J., Duyn, J. H., and Martin, A. (2004). Unravelling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* 7, 1190–1192. doi: 10.1038/nn1333
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- de Haas, B., Kanai, R., Jalkanen, L., and Ress, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proc. Biol. Sci.* 279, 4955–4961. doi: 10.1098/rspb.2012.2132
- Diederich, A., Colonius, H., and Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia* 46, 2556–2562. doi: 10.1016/j.neuropsychologia.2008.03.026
- Hasher, L., Stolz, E. R., Zacks, R., and Rypma, B. (1991). Age and inhibition. *J. Exp. Psychol. Learn.* 17, 163–169. doi: 10.1037/0278-7393.17.1.163
- Hugenschmidt, C. E., Mozolic, J. L., and Laurienti, P. J. (2009). Suppression of multisensory integration by modality-specific attention in aging. *Neuroreport* 20, 349–353. doi: 10.1097/WNR.0b013e328323ab07
- Laurienti, P. J., Burdette, J. H., Maldjian, J. A., and Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol. Aging* 27, 1155–1163. doi: 10.1016/j.neurobiolaging.2005.05.024
- Liu, X., and Yan, D. (2007). Ageing and hearing loss. *J. Pathol.* 211, 188–197. doi: 10.1002/path.2102
- McCormick, D., and Mamassian, M. (2008). What does the illusory-flash look like? *Vision Res.* 48, 63–69. doi: 10.1016/j.visres.2007.10.010
- Mishra, J., Martinez, A., and Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Res.* 1242, 102–115. doi: 10.1016/j.brainres.2008.05.023
- Mishra, J., Martinez, A., Sejnowski, T. J., and Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J. Neurosci.* 27, 4120–4131. doi: 10.1523/JNEUROSCI.4912-06.2007
- Misiak, H. (1951). The decrease of critical flicker frequency with age. *Science* 113, 551–551. doi: 10.1126/science.113.2941.551
- Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., and Laurienti, P. J. (2012). "Multisensory integration and aging," in *The Neural Bases of Multisensory Processes*, eds M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press), 381–395.
- Peiffer, A. M., Hugenschmidt, C. E., Maldjian, J. A., Casanova, E., Srikanth, R., Hayasaka, S., et al. (2009). Aging and the interaction of sensory cortical function and structure. *Hum. Brain Mapp.* 30, 228–240. doi: 10.1002/hbm.20497
- Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., and Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual task. *Neuroreport* 18, 1077–1081. doi: 10.1097/WNR.0b013e3281e72ae7
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Pitts, D. (1982). "The effects of aging on selected visual functions: dark adaptation, visual acuity, stereopsis," in *Aging and Human Visual Function*, eds R. Sekuler, D. Kline, and K. Dismukes (New York, NY: Alan R. Liss), 131–159.
- Poliakoff, E., Ashworth, S., Lowe, C., and Spence, C. (2006). Vision and touch in ageing: crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia* 44, 507–517. doi: 10.1016/j.neuropsychologia.2005.07.004
- Rosenthal, R., Shimojo, S., and Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topogr.* 21, 185–192. doi: 10.1007/s10548-009-0090-9
- Roudaia, E., Sekuler, A. B., Bennett, P. J., and Sekuler, R. (2013). Aging and audio-visual and multi-cue integration in motion. *Front. Psychol.* 4:267. doi: 10.3389/fpsyg.2013.00267
- Setti, A., Burke, K. E., Kenny, R. A., and Newell, F. N. (2011). Is inefficient multisensory processing associated with falls in older people? *Exp. Brain Res.* 209, 375–384. doi: 10.1007/s00221-011-2560-z
- Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature* 408, 788–788. doi: 10.1038/35048669
- Shams, L., Kamitani, Y., and Shimojo, S. (2002). A visual illusion induced by sound. *Cogn. Brain Res.* 14, 147–152. doi: 10.1016/S0926-6410(02)00069-1
- Stapleton, J., Setti, A., Doherty, E. P., Kenny, R. A., and Newell, F. N. (2014). A standing posture is associated with increased susceptibility to the sound-induced flash illusion in fall-prone older adults. *Exp. Brain Res.* 232, 423–434. doi: 10.1007/s00221-013-3750-7
- Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., and Aine, C. J. (2010). Ageing-related changes in auditory and visual integration measures with MEG. *Neurosci. Lett.* 484, 76–80. doi: 10.1016/j.neulet.2010.08.023
- Stevenson, R. A., and Wallace, M. T. (2013). Multisensory integration: task and stimulus dependencies. *Exp. Brain Res.* 227, 249–261. doi: 10.1007/s00221-013-3507-3
- Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. (2011). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1517–1529. doi: 10.1037/a0027339
- Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410. doi: 10.1016/j.tics.2010.06.008
- Talsma, D., and Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* 17, 1098–1114. doi: 10.1162/0898929054475172
- Vatakis, A., and Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain Res.* 1111, 134–142. doi: 10.1016/j.brainres.2006.05.078

Watkins, S., Shams, L., Josephs, O., and Rees, G. (2007). Activity in human V1 follows multisensory perception. *Neuroimage* 37, 572–578. doi: 10.1016/j.neuroimage.2007.05.027

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 June 2014; accepted: 04 September 2014; published online: 24 September 2014.

Citation: McGovern DP, Roudaia E, Stapleton J, McGinnity TM and Newell FN (2014) The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Front. Aging Neurosci.* 6:250. doi: 10.3389/fnagi.2014.00250

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 McGovern, Roudaia, Stapleton, McGinnity and Newell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A rightward shift in the visuospatial attention vector with healthy aging

Christopher S.Y. Benwell^{1,2}, Gregor Thut¹, Ashley Grant² and Monika Harvey^{2*}

¹ Centre for Cognitive Neuroimaging, School of Psychology, University of Glasgow, Glasgow, UK

² School of Psychology, University of Glasgow, Glasgow, UK

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Kelly M. Goedert, Seton Hall University, USA

Mark E. McCourt, North Dakota State University, USA

*Correspondence:

Monika Harvey, School of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, UK
e-mail: Monika.Harvey@glasgow.ac.uk

The study of lateralized visuospatial attention bias in non-clinical samples has revealed a systematic group-level leftward bias (pseudoneglect), possibly as a consequence of right hemisphere (RH) dominance for visuospatial attention. Pseudoneglect appears to be modulated by age, with a reduced or even reversed bias typically present in elderly participants. It has been suggested that this shift in bias may arise due to disproportionate aging of the RH and/or an increase in complementary functional recruitment of the left hemisphere (LH) for visuospatial processing. In this study, we report rightward shifts in subjective midpoint judgment relative to healthy young participants whilst elderly participants performed a computerized version of the landmark task (in which they had to judge whether a transection mark appeared closer to the right or left end of a line) on three different line lengths. This manipulation of stimulus properties led to a similar behavioral pattern in both the young and the elderly: a rightward shift in subjective midpoint with decreasing line length, which even resulted in a systematic rightward bias in elderly participants for the shortest line length (1.98° of visual angle, VA). Overall performance precision for the task was lower in the elderly participants regardless of line length, suggesting reduced landmark task discrimination sensitivity with healthy aging. This rightward shift in the attentional vector with healthy aging is likely to result from a reduction in RH resources/dominance for attentional processing in elderly participants. The significant rightward bias in the elderly for short lines may even suggest a reversal of hemisphere dominance in favor of the LH/right visual field under specific conditions.

Keywords: visuospatial attention, aging, landmark task, line bisection, hemispatial neglect, lateralization, spatial bias, pseudoneglect

INTRODUCTION

Studies of lateralized visuospatial attention in non-clinical samples have consistently revealed a slight but systematic group-level bias favoring the left visual field in young adults, a phenomenon termed “pseudoneglect” (see Voyer et al., 2012; Brooks et al., 2014 and Jewell and McCourt, 2000 for reviews). This behavioral bias is deemed to arise due to a right hemisphere (RH) dominance for visuospatial attention processing (Reuter-Lorenz et al., 1990; Fink et al., 2000a,b, 2001; Fierro et al., 2001; Foxe et al., 2003; Thiebaut de Schotten et al., 2005, 2011; Bultitude and Aimola-Davies, 2006; Ghacibeh et al., 2007; Waberski et al., 2008; Çiçek et al., 2009; Cavérian et al., 2012; Cai et al., 2013; Benwell et al., 2014) and RH dominance also seems to underlie the tendency for visuospatial neglect symptoms to occur more frequently and severely after right as compared to left hemisphere (LH) stroke (Driver and Mattingley, 1998; Vallar, 1998; Halligan et al., 2003; Harvey and Rossit, 2012). The degree of lateralized visuospatial attention bias is often assessed using variants of the horizontal line bisection task, in both clinical (Milner and Harvey, 1995; Urbanski and Bartolomeo, 2008) and non-clinical samples (Bowers and Heilman, 1980; Milner et al., 1992; Jewell and McCourt, 2000).

Though bisection performance has proven to be less consistent in older healthy adults, the systematic leftward bias appears to be attenuated, eliminated, or even reversed with age (Fukatsu et al., 1990; Stam and Bakker, 1990; Fujii et al., 1995; Jewell and McCourt, 2000; Failla et al., 2003; Goedert et al., 2010; Nagamatsu et al., 2011; Hatin et al., 2012; Loureiro et al., 2013; Brooks et al., 2014; Veronelli et al., 2014). Additionally, recent evidence suggests potential sex-differences in age-related changes in manual line bisection performance, with aging effects being strongest in males vs. relatively intact performance with aging in females (Varnava and Halligan, 2007; Barrett and Craver-Lemley, 2008; Chen et al., 2011; however see Beste et al., 2006 for discrepant results). In order to minimize the influence of motor factors on bisection decisions, Schmitz and Peigneux (2011) recently employed the Landmark Task (a non-manual, perceptual variant of line bisection) to investigate age-related changes in pseudoneglect. In this task participants are asked to estimate which of two segments of a pre-bisected line is shortest or longest (Milner et al., 1992, 1993; Harvey et al., 1995; Milner and Harvey, 1995). They found that young participants perceived the left side of equally bisected lines to be longer than the right side (typical of pseudoneglect),

whereas elderly participants presented the opposite pattern, and were more accurate when unevenly bisected lines were divided on the left side. Overall, a rightward shift in the performance of older participants was found as compared to young participants, in line with previous studies (Sex of the participants was not distinguished in the study, Schmitz and Peigneux, 2011).

Several candidate models may account for the observed change in pseudoneglect with aging. One is that of Hemispheric Asymmetry Reduction in Older Adults (i.e., the HAROLD model, Cabeza, 2002). The HAROLD model suggests that functional recruitment of the non-dominant hemisphere for a given task helps to compensate for age-related unilateral working efficiency decline, resulting in reduced asymmetry in processing for the task at hand (Cabeza, 2002; Reuter-Lorenz and Cappell, 2008; Li et al., 2009). The HAROLD model has largely been investigated in the context of memory tasks and its predictions have often been supported (Bäckman et al., 1997; Grady et al., 2002; Logan et al., 2002; Cabeza et al., 2004; Rossi et al., 2004; Solé-Padullés et al., 2006; Schmitz et al., 2013). Using positron emission tomography (PET), Reuter-Lorenz et al. (2000) found prefrontal cortex (PFC) activity to be lateralized to the respective dominant hemisphere for a given stimulus in young participants. However, in elderly participants the activity was bilateral for all stimulus types. Although mainly observed in the PFC, the HAROLD model may also apply to other regions and tasks (Collins and Mohr, 2013). Nielson et al. (2002) found that during an inhibition task, parietal activity was right lateralized in young participants yet bilateral in older participants. Thus in the context of visuospatial attention biases, when performing the landmark task, elderly participants may recruit supplementary contralateral (left) brain areas in a compensatory manner, resulting in the observed absence or reversal of pseudoneglect.

Another model emphasizes accelerated aging in the right relative to the LH (Brown and Jaffe, 1975; Goldstein and Shelly, 1981), which may in turn reduce the functional dominance of visuospatial attention processing in the RH. Using a test battery designed to diagnose lateralized brain injury, it has previously been found that the performance of elderly participants is analogous to that of RH damaged patients (Klisz, 1978) and more recently specific RH impairment in elderly participants has been found during performance of a variety of psychophysical tasks (Jenkins et al., 2000; Lux et al., 2008; Nagamatsu et al., 2011; Chokron et al., 2013). The absence or reversal of pseudoneglect presented by elderly participants may therefore reflect general RH decline. However, evidence supporting greater aging of the RH in comparison to the left has been mixed (Dolcos et al., 2002; Sowell et al., 2003; Raz et al., 2004).

Additionally, rightward spatial biases are often associated with states of both tonic and chronic reduced arousal (Bellgrove et al., 2004; Manly et al., 2005; Fimm et al., 2006; Dufour et al., 2007; Dodds et al., 2008; Heber et al., 2008; Matthias et al., 2010; Benwell et al., 2013a,b; Newman et al., 2013). It is possible that a reduction in general alertness/vigilance over the lifespan (Robinson and Kertzman, 1990; Buysse et al., 2005; Nebes et al., 2009; Goedert et al., 2010) may also contribute to the chronic attenuation of pseudoneglect in the elderly.

Interestingly, the degree of visuospatial bias displayed during landmark task performance is modulated within participants by stimulus properties such as line length. Recent studies employing the landmark task in healthy young participants have shown that while long lines (subtending $>6^\circ$ horizontal visual angle (VA) in length) induce a systematic (usually left) bias, short lines (subtending $<2^\circ$ VA) induce either no bias or a right bias (McCourt and Jewell, 1999; Rueckert et al., 2002; Rueckert and McFadden, 2004; Heber et al., 2010; Thomas et al., 2012; Benwell et al., 2013a, 2014). The line length effect appears to arise due to asymmetrical hemispheric contributions (in favor of the RH) to the perceived salience of line stimuli that is more pronounced for long than short lines and hence a left bias arises more prominently for long lines (Anderson, 1996; Benwell et al., 2014). In a recent study, we manipulated both time-on-task/vigilance and line length in a sample of healthy young participants (Benwell et al., 2013a). We found the rightward shifting effects of time-on-task and line length to be additive: at baseline the common group-level leftward bias was observed in long lines whereas no systematic bias was observed in short lines (group average not significantly different from veridical center). After 1 h of prolonged performance of the landmark task with long lines, both long and short line performance were tested again. A rightward shift in bias was evident in that the left bias was now absent in long lines, and intriguingly the rightward shift also transferred to the un-practiced short lines which now evidenced a right bias significantly different from veridical center. The additive effects of reduced line length and increased time-on-task suggest that both manipulations may result in down-regulation of RH attention network engagement and hence the observed rightward shifts in spatial bias. Additionally, an overall task performance decrement (as indexed by the curve width of the fitted psychometric function) was observed with prolonged time-on-task, further suggesting a degradation of attentional resources.

Elucidating how the established bias modulators of age and line length interact to influence lateralized visuospatial bias as displayed during landmark task performance will allow for a refinement of models of visual attention processing changes with healthy aging. To investigate this, we compared landmark task performance on three different line lengths (short, medium and long) between young and elderly healthy participants. In line with previous studies, we predicted a systematic leftward bias for long lines in young participants that would be attenuated with reducing line length. If hemispheric asymmetry reduction alone accounts for the attenuation of pseudoneglect with aging then we would expect to see no systematic bias for any line length in the elderly and also relatively preserved overall performance on the task. Alternatively, if reduced RH function and/or chronic reduced arousal play a role in the attenuation of bias then we would expect to see a pattern of performance in the elderly analogous to that previously observed in young participants following prolonged time-on-task: namely no bias in long lines and a systematic rightward bias for short lines along with an overall task performance decrement (Benwell et al., 2013a).

MATERIAL AND METHODS

PARTICIPANTS

Twenty right-handed young (12 males, mean age = 23.25 years; SD = 2.83, max = 31, min = 18) and 20 right-handed elderly participants (11 males, mean age = 68.45 years; SD = 4.95, max = 77, min = 60) took part in the experiment. Written informed consent was obtained from each participant. All participants were volunteers naive to the experimental hypothesis being tested. All participants had normal or corrected-to-normal vision and reported no history of neurological disorder. The experiment was carried out within the Institute of Neuroscience and Psychology at the University of Glasgow and was approved by the local ethics committee.

INSTRUMENTATION AND STIMULI

Stimuli were presented using the E-Prime software package (Schneider et al., 2002) on a CRT monitor with a 1280 × 1024 pixel resolution and 85 Hz refresh rate. Adapted from experiment 2 of Benwell et al. (2013b), the paradigm represented a computerized version of the landmark task (Milner et al., 1992; McCourt and Olafson, 1997; Olk and Harvey, 2002). Lines of 100% Michelson contrast were presented on a gray background (luminance = 179, hue = 179). **Figure 1** shows examples of line stimuli used in the experiment. Three different lengths of line were presented. “Long” lines measured 24.3 cm in length by 0.5 cm in height and at a viewing distance of 70 cm subtended 19.67° (width) by 0.4° (height) of VA. At the same viewing distance, “medium” lines measuring 12.15 cm × 0.5 cm subtended 9.92° × 0.4° of VA and “short” lines measuring 2.43 cm × 0.5 cm subtended 1.98° by 0.4° of VA.

All three line lengths were transected at 1 of 13 points ranging from $\pm 7.5\%$ (distance between transector locations = 1.25%) of absolute line length to veridical center. In long lines, this represented a range of -1.48° to 1.48° of VA with a distance between transector locations of 0.25° of VA. In medium lines a range of -0.74° to 0.74° of VA with a distance between transector locations of 0.12° of VA was presented and in short lines, a range of -0.15° to 0.15° of VA with a distance between transector locations of 0.02° of VA was presented. All lines were displayed with the transector location centered on the vertical midline of the display (i.e., aligned to a central fixation cross which preceded the presentation of the lines, see below).

PROCEDURE

Participants were seated with their midsagittal plane aligned with the display monitor. Viewing distance (70 cm) was kept constant using a chin rest. Each trial began with presentation of a fixation cross (0.4° (height) × 0.4° (width) of VA) for 1 s followed by presentation of the transected line (150 ms). The transection mark was always aligned with the fixation cross (i.e., the eccentricity of the line endpoints varied across trials while the transection point always appeared at the same central position), therefore preventing use of the fixation cross as a reference point for bisection judgments. The fixation cross then reappeared for the duration of the response period during which participants indicated which end of the line had appeared longest/shortest to them by pressing



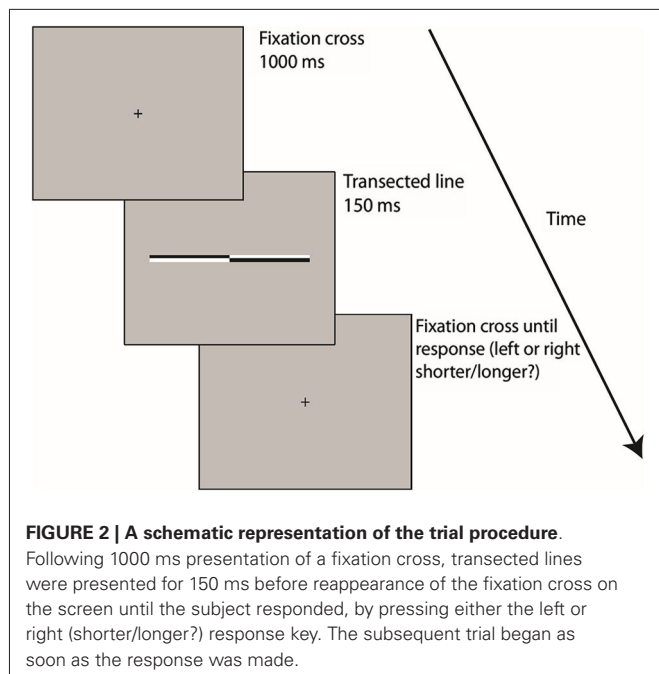
FIGURE 1 | Examples of line stimuli used in the experiment. Lines were transected at 1 of 13 locations ranging symmetrically from $\pm 7.5\%$ (distance between transector locations = 1.25%) of absolute line length relative to and including veridical center. In long lines, this represented a range of -1.48° to 1.48° of visual angle (VA) with a distance between transector locations of 0.25° of VA. In medium lines a range of -0.74° to 0.74° of VA with a distance between transector locations of 0.12° of VA was presented and in short lines, a range of -0.15° to 0.15° of VA with a distance between transector locations of 0.02° of VA was presented. All lines were displayed with the transector location centered on the vertical midline of the display (i.e., aligned to a central fixation cross which preceded the presentation of the lines). Lines A, C and E are transected to the left of veridical center whereas lines B, D and F are transected to the right of veridical center. Lines of varying contrast polarity appeared with equal frequency and the order of appearance was randomized.

either the left or right response key. Half of the participants were asked to judge which end of each line was longest and the other half were asked to judge which end was shortest, in order to prevent any possible group-level response bias (increased likelihood of pressing either the left or right response key regardless of the visual percept, especially in cases of uncertainty (see Morgan et al., 2012; García-Pérez and Alcalá-Quintana, 2013)) from contaminating the perceptual midpoint analysis.

Participants always responded using their dominant right hand (right index and middle finger respectively) and were instructed to hold their gaze on the center of the screen throughout each trial. The subsequent trial began as soon as the response was made. Trials lasted approximately 2 s. Trial type (location of transector in line) was selected at random. Each participant completed 91 trials of each line length (Overall = 273 trials, 7 judgments at each of the 13 transector locations) split into 7 short blocks (lasting approximately 2–3 min). Participants were allowed to take as long a break as they wished between blocks. A block of 20 practice trials was performed immediately prior to the beginning of the experimental blocks. The entire experiment lasted approximately 20–25 min (see **Figure 2** for schematic representation of the trial procedure).

ANALYSIS

In order to obtain an objective measure of perceived line midpoint for each line length in each subject, psychometric functions (PFs) were derived using the method of constant stimuli. The dependent measure was the proportion of trials on which participants indicated that the transector had appeared closer to the left end of the line. Non-linear least-squares regression was used to fit a cumulative logistic function to the data for each line length in each subject. The cumulative logistic function is described by



the equation:

$$f(\mu, x, s) = 1/(1 + \exp((x - \mu)/s))$$

where x are the tested transector locations, μ corresponds to the x -axis location with a 50% “left” and 50% “right” response rate and s indexes the width of the nonasymptotic region of the fitted curve. The 50% location is known as the point of subjective equality point of subjective equality (PSE) and represents an objective measure of perceived line midpoint. The width of the PF provides a measure of the precision of participants’ line midpoint judgments per line length. A low width value indicates that the PF is steep and that the observer can discriminate differences between transector locations relatively easily, whereas a high width value indicates that the PF is shallow and that the observer can only discriminate relatively coarse differences (Fründ et al., 2011). Inferential statistical analyses were performed on the individually fitted PF PSE and width estimates.

RESULTS

SUBJECTIVE MIDPOINT (P.S.E) ANALYSIS

Figures 3A–C present group-averaged PFs for both experimental groups at each line length. For each line length, black filled circle symbols (young participants) and gray open diamond symbols (elderly participants) plot mean percentage left response as a function of transector location. The black (young) and gray (elderly) smooth curves represent the best-fitting least-squares cumulative logistic PFs (95% confidence interval represented by black (young) and gray (elderly) dotted lines). Where black (young) and gray (elderly) vertical dashed lines cross the black horizontal dashed lines indicate the transector locations corresponding to the 50% response rates (PSEs).

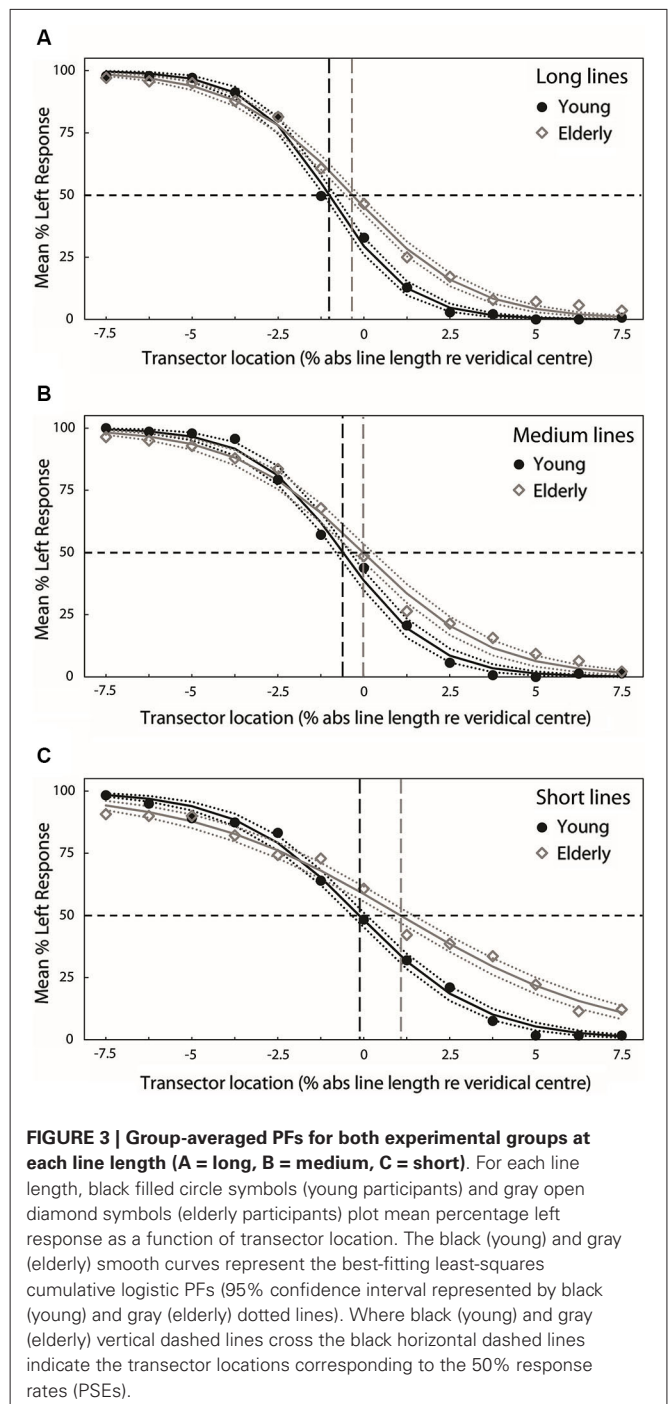
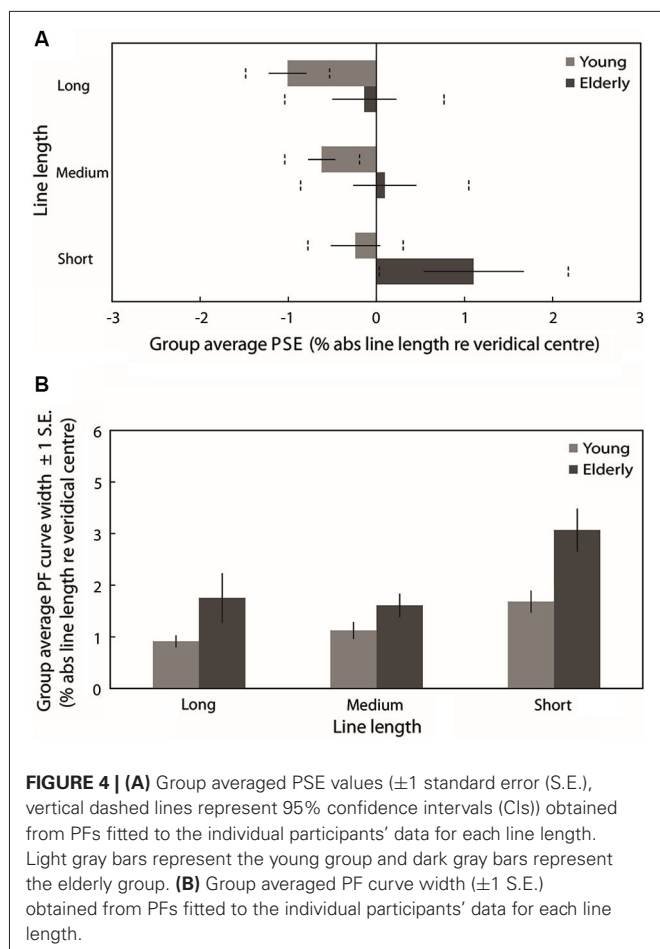


Figure 4A plots the group mean PSEs (± 1 standard error (S.E.), vertical dashed lines represent 95% confidence intervals (CIs)) obtained from PFs fitted to the individual participants’ data for each line length. These are in close agreement with the group averaged PF PSEs. In line with previous studies of pseudoneglect, mean long line PSE in the young group was displaced to the left of veridical center by -1% of absolute line length and this leftward bias was significantly different from veridical center (95% CI does not include 0) whereas



in the elderly group the mean PSE was slightly to the left (-0.14%) but not significantly different from veridical center (95% includes 0). Mean medium line PSE in the young group was displaced to the left of veridical center by -0.62% and this leftward bias was also significantly different from veridical center (95% CI does not include 0). In contrast, the medium line elderly PSE was very slightly to the right of center by 0.1% but again not significantly different from veridical center (95% CI includes 0). In the short lines, mean PSE in the young group was -0.24% to the left of veridical center but the difference from veridical center was not significant (95% CI includes 0) whereas mean PSE in the elderly group was significantly displaced to the right of veridical center by 1.1% (95% CI does not include 0).

A 2 (Age group: young vs. elderly) \times 3 (Line length: long vs. medium vs. short) ANOVA on individually fitted PF PSEs revealed a significant main effect of age group ($F_{(1,38)} = 5.830$, $p = 0.021$, $\eta_p^2 = 0.133$), a significant main effect of line length ($F_{(2,76)} = 6.509$, $p = 0.002$, $\eta_p^2 = 0.146$) but no significant age group \times line length interaction ($F_{(2,76)} = 0.524$, $p = 0.524$, $\eta_p^2 = 0.017$). The overall subjective midpoint was significantly more to the left in the young participants than in the elderly (as indexed by the PSEs), indicating a group level rightward shift in the attentional vector with age (as is clearly displayed

in **Figure 4A**). Pairwise comparisons (Bonferroni-corrected) to analyze the simple effects of line length revealed no statistically significant difference in subjective midpoint between either long and medium lines ($t_{(39)} = -1.846$, $p = 0.226$, Cohen's $d = -0.292$) or medium and short lines ($t_{(39)} = -2.163$, $p = 0.111$, Cohen's $d = -0.345$) but a significant rightward shift in subjective midpoint from long to short lines ($t_{(39)} = -3.022$, $p = 0.014$, Cohen's $d = -0.482$) regardless of age (again displayed in **Figure 4A**). Additionally, a within-subjects linear contrast analysis revealed a significant linear shift in bias with line length ($F_{(1,38)} = 9.017$, $p = 0.005$, $\eta_p^2 = 0.192$).

PSYCHOMETRIC FUNCTION CURVE WIDTH ANALYSIS

Figure 4B plots the mean PF curve width (± 1 S.E.) obtained from PFs fitted to the individual participants' data for each line length. A 2 (Age group: young vs. elderly) \times 3 (Line length: long vs. medium vs. short) ANOVA revealed a significant main effect of age group ($F_{(1,38)} = 8.674$, $p = 0.005$, $\eta_p^2 = 0.186$), a significant main effect of line length ($F_{(2,76)} = 11.637$, $p < 0.001$, $\eta_p^2 = 0.234$) and no significant age group \times line length interaction ($F_{(2,76)} = 1.706$, $p = 0.188$, $\eta_p^2 = 0.043$). PF curve widths were significantly shallower in elderly participants than in young participants, indicating reduced discrimination sensitivity with age. Pairwise comparisons (Bonferroni-corrected) to analyze the simple effects of line length revealed no statistically significant difference in PF width between long and medium lines ($t_{(39)} = -0.155$, $p = 1$, Cohen's $d = -0.033$) but a significant increase in width from both long to short lines ($t_{(39)} = -3.409$, $p = 0.005$, Cohen's $d = -0.542$) and from medium to short lines ($t_{(39)} = -4.845$, $p < 0.001$, Cohen's $d = -0.881$). A within-subjects linear contrast analysis revealed a significant linear shift in curve width with line length ($F_{(1,38)} = 11.56$, $p = 0.002$, $\eta_p^2 = 0.233$). Discrimination sensitivity for the task was significantly lower for short lines than for long and medium lines regardless of age (as displayed in **Figure 4B**).

ADDITIONAL GENDER ANALYSIS

Recent evidence from studies employing manual line bisection has suggested potential sex-differences in age-related changes in bisection performance, with aging effects being strongest in males vs. relatively intact performance with aging in females (Varnava and Halligan, 2007; Barrett and Craver-Lemley, 2008; Chen et al., 2011; however see Beste et al., 2006 for discrepant results). In order to test for any such gender effects in age-related changes in landmark task performance, we re-analyzed (*post hoc*) the PSE and width values with an additional between-subjects factor of gender (female, male) included in the ANOVAs. The PSE re-analysis revealed no additional main effect of gender ($F_{(1,36)} = 0.019$, $p = 0.892$, $\eta_p^2 = 0.001$) and no significant interaction between either age group \times gender ($F_{(1,36)} = 0.411$, $p = 0.525$, $\eta_p^2 = 0.011$), length \times gender ($F_{(2,72)} = 0.337$, $p = 0.715$, $\eta_p^2 = 0.009$) nor age group \times length \times gender ($F_{(2,72)} = 0.608$, $p = 0.547$, $\eta_p^2 = 0.017$).

The width re-analysis also revealed no main effect of gender ($F_{(1,36)} = 0.970$, $p = 0.331$, $\eta_p^2 = 0.026$), no significant interaction between either age group \times gender ($F_{(1,36)} = 0.299$, $p = 0.588$, $\eta_p^2 =$

0.008), length \times gender ($F_{(2,72)} = 0.615$, $p = 0.543$, $\eta_p^2 = 0.017$) nor age group \times length \times gender ($F_{(2,72)} = 0.958$, $p = 0.388$, $\eta_p^2 = 0.026$).

DISCUSSION

Recent studies have shown age-related changes in the expression of visual pseudoneglect (Fukatsu et al., 1990; Stam and Bakker, 1990; Fujii et al., 1995; Jewell and McCourt, 2000; Failla et al., 2003; Barrett and Craver-Lemley, 2008; Goedert et al., 2010; Nagamatsu et al., 2011; Schmitz and Peigneux, 2011; Hatin et al., 2012; Loureiro et al., 2013; Veronelli et al., 2014). We aimed to investigate, for the first time, how the established line bisection bias modulator of line length interacts with healthy aging to influence lateralized visuospatial bias as displayed during landmark task performance. For this purpose, we compared landmark task performance on three different line lengths (short, medium and long) between young (18–31 years old) and elderly (60–77) healthy participants.

As expected, young participants displayed a group-level systematic leftward bias (pseudoneglect) during long line landmark task performance. This leftward bias was reduced for the medium length lines and no systematic bias was observed for performance of the task with short lines, confirming the previously reported line-length effect (McCourt and Jewell, 1999; Rueckert et al., 2002; Rueckert and McFadden, 2004; Heber et al., 2010; Thomas et al., 2012; Benwell et al., 2013a, 2014). Moreover, the results revealed a group-level rightward shift in the visuospatial attention vector in the elderly as compared to the young participants, in line with previous findings of an attenuation or reversal of pseudoneglect with healthy aging (Fukatsu et al., 1990; Stam and Bakker, 1990; Fujii et al., 1995; Jewell and McCourt, 2000; Failla et al., 2003; Barrett and Craver-Lemley, 2008; Goedert et al., 2010; Nagamatsu et al., 2011; Schmitz and Peigneux, 2011; Hatin et al., 2012; Loureiro et al., 2013; Veronelli et al., 2014). Importantly, no interaction was observed between age group and line length suggesting that the elderly participants were subject to the line length effect in a similar manner to the young (i.e., a rightward shift in subjective midpoint with reduced line length). We found no effect of gender on landmark task performance in either the young or the elderly.

Our results replicate and extend those of Schmitz and Peigneux (2011) who found suppression, and near reversal, of the leftward pseudoneglect bias in their elderly sample during long line landmark performance. In their study, the line stimuli remained onscreen until the participant responded (free-viewing). The authors note that this absence of control of ocular scanning in their study precluded them from dissociating a true perceptual bias shift with aging from a failure of inhibition of return (IOR). IOR represents a mechanism by which the viewer disengages from previously processed aspects of a stimulus in order to facilitate perception of its entirety (Posner and Cohen, 1984). Using a stimulus duration of 150 ms only (and thus preventing eye movements), we here confirm that the observed rightward shift in the attention vector with healthy aging is unlikely to occur as a result of a failure of IOR.

POTENTIAL NEURAL MECHANISMS OF THE RIGHTWARD PERCEPTUAL SHIFT WITH AGING

Accelerated right hemisphere aging/HAROLD model

Previous studies exploring age-related variability in neurocognitive function have posited a decline in hemispheric specialization of task-related neural activity to represent a form of compensation for age-related deficits that supports task performance (Reuter-Lorenz and Lustig, 2005; Reuter-Lorenz and Cappell, 2008; Angel et al., 2011). However, the functional significance of the observed neural activation of regions not primarily associated with task performance in young participants, and whether such “recruitment” is restricted to elderly participants, remains unclear (Reuter-Lorenz and Park, 2010; Friedman, 2013).

Though the rightward shift in the visual attention vector with age observed in the current study would support an increased involvement of the LH in task processing in the elderly compared to the young participants (Cabeza, 2002; Reuter-Lorenz and Cappell, 2008; Li et al., 2009), the HAROLD model alone appears to be inconsistent with the findings of a significant rightward bias for short lines in the elderly in the current study along with previous reports of group-level rightward bisection biases in elderly samples (Stam and Bakker, 1990; Fujii et al., 1995). The HAROLD model would predict symmetrical bisection behavior in elderly participants but it would not predict systematic right biases beyond the veridical midline (Brooks et al., 2014). Additionally, we found overall performance precision (as indexed by the curve width of the fitted psychometric functions) to be lower in elderly participants suggesting reduced discrimination sensitivity with aging. Although the influence of low level visual deficits (such as reduced visual resolution) cannot be ruled out, elderly participants were less able to successfully discriminate between the different transector locations (for all three line lengths) and so “compensatory” recruitment of the LH for landmark task processing does not equate to preserved task performance ability equivalent to that of young participants.

Moreover, increased LH involvement could occur as a result of reduced inhibitory influence of the RH, in line with an interhemispheric competition account of spatial attention control (Kinsbourne, 1977; Duecker et al., 2013; Szczepanski and Kastner, 2013) in combination with accelerated RH aging (Brown and Jaffe, 1975; Goldstein and Shelly, 1981; Nagamatsu et al., 2011) and/or a decline in corpus callosum integrity with age (Hausmann et al., 2003; Sullivan and Pfefferbaum, 2006; Koch et al., 2007, 2011).

Potential role of arousal and/or perceptual load

Rightward spatial biases are often associated with states of both tonic and chronic reduced arousal (Bellgrove et al., 2004; Manly et al., 2005; Fimm et al., 2006; Dufour et al., 2007; Dodds et al., 2008; Heber et al., 2008; Matthias et al., 2010; Newman et al., 2013; Benwell et al., 2013a,b). In fact, after 1 h of landmark task performance with long lines, a rightward shift in the attentional vector was displayed by the young participants in our previous study (including a rightward bias for short lines that was significantly different from veridical center) (Benwell et al.,

2013a). This pattern of bisection behavior was remarkably similar to that displayed at baseline by the elderly sample in the current study. It is possible that a reduction in general alertness over the lifespan (Robinson and Kertzman, 1990; Goedert et al., 2010; Buysse et al., 2005; Nebes et al., 2009), and/or a reduction in functional interaction between RH ventral and dorsal networks subserving visuospatial attention (see Thiebaut de Schotten et al., 2011 and the discussion of Benwell et al., 2013b), may contribute to a chronic attenuation of pseudoneglect in aged individuals. Additionally, the increased difficulty of performing the task with short lines (as indexed by the shallower PF curve width values) may further hinder RH contribution to the task in states of sub-optimal function (such as with aging (Brown and Jaffe, 1975; Goldstein and Shelly, 1981; Nagamatsu et al., 2011) or reduced vigilance/increased time-on-task (Fimm et al., 2006; Benwell et al., 2013a,b)) and hence bring about the observed rightward biases.

LINE LENGTH EFFECT AND AGING

Potential neural mechanisms

The current results show for the first time that, despite an overall rightward shift in midpoint judgments in the elderly, reducing line length results in the same pattern of behavior in the elderly as in the young (i.e., a rightward shift in subjective midpoint) during landmark task performance. The rightward shifting effects of age and line length on midpoint judgment appear to be additive. In a mathematical model of bisection behavior, the line length effect was posited to arise due to asymmetrical hemispheric contributions (in favor of the RH) to the perceived salience of line stimuli that is more pronounced for long than short lines (Anderson, 1996). We have recently investigated the neural correlates of the line length effect in neurologically normal young participants during performance of the landmark task (Benwell et al., 2014). Our EEG results showed that increased engagement of regions of the right lateralized ventral attention network in long relative to short lines contributes to the genesis of the spatial bias: we found an ERP response which showed higher amplitude to long as compared to short lines, corresponded in its timing to the N1-component and was right lateralized to areas of the temporo-parietal junction (TPJ; Benwell et al., 2014). Furthermore, the difference in peak N1-amplitude between long and short line processing correlated with the difference in line bisection bias between long and short lines across participants, thereby providing empirical support for Anderson's (1996) model. The TPJ represents a key node in the ventral frontoparietal attention network implicated in both the orienting of visuospatial attention and the maintenance of arousal (Corbetta and Shulman, 2002, 2011). De-regulation of RH TPJ activity is thought in turn to reduce activation of the bihemispheric dorsal frontoparietal network (implicated in the distribution of visuospatial attention across the visual field) and has been linked to rightward shifts in visuospatial bias in healthy participants (O'Connell et al., 2011; Newman et al., 2013; Benwell et al., 2013b). We posit that these neural correlates may also underlie the length effect observed here in the elderly, over and above any age-related changes in task processing.

No evidence for gender specific effects

Varnava and Halligan (2007) employed manual line bisection to investigate the effects of age and gender on bisection performance in healthy participants on three different line lengths comparable to those used in the current study. In their study, only males showed a rightward shift in bisection bias with age and only for long line performance. This effect of gender on manual line bisection performance with aging has been supported by subsequent studies, with the effect of aging appearing to be strongest for males (Barrett and Craver-Lemley, 2008; Chen et al., 2011). A possible explanation for the discrepant finding of no sex difference in the current study could be the use of the landmark task instead of manual line bisection (Varnava and Halligan, 2007; Barrett and Craver-Lemley, 2008; Chen et al., 2011). In general, differences in experimental procedure (such as the viewing distance employed (see McCourt and Garlinghouse, 2000; Varnava et al., 2002; Longo and Lourenco, 2006)), sample demographics and analysis techniques across studies may contribute to the differential findings. Treating age as a continuous variable in a sample of participants largely over 40 years old (mean age = 58.7, only 5 out of 44 participants <40), Chen et al. (2011) dissociated "where" perceptual errors from "aiming" motor errors during line bisection and found a rightward shift in perceptual midpoint with aging in men only. Thus, further research should aim to explore, ideally in larger samples and utilizing the deployment of multiple visuospatial tasks and analysis techniques, the reasons underlying these discrepancies in gender- and age-related effects on visuospatial bias. Although the current experiment was not explicitly set up to investigate gender differences, we would propose that non-perceptual factors may contribute to the previously observed gender specific aging effects in pseudoneglect, and that both sexes appear to experience a rightward perceptual shift in the visuospatial attention vector with healthy aging.

Comparison to neglect

The pattern of the line length effect displayed by our elderly sample is in the opposite direction to that often observed in unilateral neglect patients. In these patients, a reduction in line length generally results in a systematic reduction of the severe rightward bias typically exhibited on long lines, with a leftward bias sometimes being displayed on very short lines (the "crossover" effect Halligan and Marshall, 1988; Marshall and Halligan, 1989; Harvey et al., 1995; Anderson, 1996, 1997; Monaghan and Shillcock, 1998, 2004; Ricci and Chatterjee, 2001; Mennemeier et al., 2005; Veronelli et al., 2014). We therefore think it unlikely that the performance of elderly participants can be seen as a mild version of spatial neglect. What seems to be the case is that the elderly participants show an overall rightward shift in the attentional vector, that is most pronounced for the short lines. However, the comparison of findings from healthy participants with those in neglect patients and the "crossover" literature is complicated by the large variance of line bisection performance patterns both within and across patients (Halligan et al., 1990) and common concurrent primary visual and motor deficits post-stroke (Doricchi et al., 2005; Binetti et al., 2011; Kerkhoff and Schenk, 2011). The 150 ms landmark task presentation duration employed here minimizes the influence of non-perceptual motor

components such as hand use and visual scanning on bisection decisions (Milner et al., 1992; Luh, 1995; Bisiach et al., 1998; Toraldo et al., 2004). Employing the paradigm from the current study in RH stroke neglect patients both with and without concomitant primary visual deficits would be highly informative in terms of elucidating further purely perceptual contributions to the line length effect in neglect and the potential role played by primary visual deficits in the commonly observed “crossover” effect (Doricchi et al., 2005; Binetti et al., 2011).

FUTURE DIRECTIONS

The neural origin(s) of the additive effects of aging and line length remain unclear. It is possible that two independent processes influencing spatial bias are at play, one affected by aging (leading to a rightward shift) and the other unaffected (preserving the line length effect in healthy aging). The introduction of neuroimaging techniques is likely to represent an important step with regard to answering this and many more of the open questions pertaining to visuospatial processing in the elderly. To our knowledge, neuroimaging studies of bisection task performance to date have been restricted to young healthy participants, revealing strong RH dominance for task processing (Fink et al., 2000a,b, 2001; Foxe et al., 2003; Waberski et al., 2008; Çiçek et al., 2009; Thiebaut de Schotten et al., 2011; Cavézian et al., 2012; Benwell et al., 2014). Using EEG and a passive viewing task, De Sanctis et al. (2008) showed reduced hemispheric asymmetry of early-visual processing in elderly compared to young participants. As mentioned, we have linked the genesis of the landmark task bias to the RH amplitude of an early component (N1) of the visual evoked potential (Benwell et al., 2014). In addition, the magnitude and direction of bias have also been linked to the relative anatomical hemispheric lateralization of a parieto-frontal white matter pathway (Thiebaut de Schotten et al., 2011). Investigation of these neural modulators of visuospatial bias in the elderly represents a natural and potentially illuminating next step.

AUTHOR CONTRIBUTIONS

Christopher S.Y. Benwell conceived the experiment, analyzed the data and co-wrote the manuscript. Monika Harvey and Gregor Thut supervised the entire work and co-wrote the manuscript. Ashley Grant collected and analyzed the data.

ACKNOWLEDGMENTS

We wish to thank Maria Smithers for help with data collection and all those who participated in the study. This work was supported by the Economic and Social Research Council [grant number ES/I02395X/1].

REFERENCES

- Anderson, B. (1996). A mathematical model of line bisection behaviour in neglect. *Brain* 119, 841–850. doi: 10.1093/brain/119.3.841
- Anderson, B. (1997). Pieces of the true crossover effect in neglect. *Neurology* 49, 809–812. doi: 10.1212/wnl.49.3.809
- Angel, L., Fay, S., Bouazzaoui, B., and Isingrini, M. (2011). Two hemispheres for better memory in old age: role of executive functioning. *J. Cogn. Neurosci.* 23, 3767–3777. doi: 10.1162/jocn_a_00104
- Bäckman, L., Almkvist, O., Andersson, J., Nordberg, A., Winblad, B., Reineck, R., et al. (1997). Brain activation in young and older adults during implicit and explicit retrieval. *J. Cogn. Neurosci.* 9, 378–391. doi: 10.1162/jocn.1997.9.3.378
- Barrett, A. M., and Craver-Lemley, C. E. (2008). Is it what you see, or how you say it? Spatial bias in young and aged subjects. *J. Int. Neuropsychol. Soc.* 14, 562–570. doi: 10.1017/S1355617708080764
- Bellgrove, M. A., Dockree, P. M., Aimola, L., and Robertson, I. H. (2004). Attenuation of spatial attentional asymmetries with poor sustained attention. *Neuroreport* 15, 1065–1069. doi: 10.1097/00001756-200404290-00027
- Benwell, C. S. Y., Harvey, M., and Thut, G. (2014). On the neural origin of pseudoneglect: EEG-correlates of shifts in line bisection performance with manipulation of line length. *Neuroimage* 86, 370–380. doi: 10.1016/j.neuroimage.2013.10.014
- Benwell, C. S. Y., Harvey, M., Gardner, S., and Thut, G. (2013a). Stimulus- and state-dependence of systematic bias in spatial attention: additive effects of stimulus-size and time-on-task. *Cortex* 49, 827–836. doi: 10.1016/j.cortex.2011.12.007
- Benwell, C. S. Y., Thut, G., Learmonth, G., and Harvey, M. (2013b). Spatial attention: differential shifts in pseudoneglect direction with time-on-task and initial bias support the idea of observer subtypes. *Neuropsychologia* 51, 2747–2756. doi: 10.1016/j.neuropsychologia.2013.09.030
- Beste, C., Hamm, J. P., and Hausmann, M. (2006). Developmental changes in visual line bisection in women throughout adulthood. *Dev. Neuropsychol.* 30, 753–767. doi: 10.1207/s15326942dn3002_6
- Binetti, N., Aiello, M., Merola, S., Bruschini, M., Lecce, F., Macci, E., et al. (2011). Positive correlation in the bisection of long and short horizontal Oppel-Kundt illusory gradients: implications for the interpretation of the “cross-over” effect in spatial neglect. *Cortex* 47, 608–616. doi: 10.1016/j.cortex.2010.08.003
- Bisiach, E., Ricci, R., Lualdi, M., and Colombo, M. R. (1998). Perceptual and response bias in unilateral neglect: two modified versions of the milner landmark task. *Brain Cogn.* 37, 369–386. doi: 10.1006/brcg.1998.1003
- Bowers, D., and Heilman, K. M. (1980). Pseudoneglect: effects of hemispace on a tactile line bisection task. *Neuropsychologia* 18, 491–498. doi: 10.1016/0028-3932(80)90151-7
- Brooks, J. L., Della Sala, S., and Darling, S. (2014). Representational pseudoneglect: a review. *Neuropsychol. Rev.* 24, 148–165. doi: 10.1007/s11065-013-9245-2
- Brown, J. W., and Jaffe, J. (1975). Hypothesis on cerebral dominance. *Neuropsychologia* 13, 107–110. doi: 10.1016/0028-3932(75)90054-8
- Bultitude, J., and Aimola-Davies, A. M. (2006). Putting attention on the line: investigating the activation-orientation hypothesis of pseudoneglect. *Neuropsychologia* 44, 1849–1858. doi: 10.1016/j.neuropsychologia.2006.03.001
- Buyse, D. J., Monk, T. H., Carrier, J., and Begley, A. (2005). Circadian patterns of sleep, sleepiness and performance in older and younger adults. *Sleep* 28, 1365–1376.
- Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol. Aging* 17, 85–100. doi: 10.1037/0882-7974.17.1.85
- Cabeza, R., Daselaar, S. M., Dolcos, F., Prince, S. E., Budde, M., and Nyberg, L. (2004). Task-independent and task-specific age effects on brain activity during working memory, visual attention and episodic retrieval. *Cereb. Cortex* 14, 364–375. doi: 10.1093/cercor/bhg133
- Cai, Q., Van der Haegen, L., and Brysbaert, M. (2013). Complementary hemispheric specialization for language production and visuospatial attention. *Proc. Nat. Acad. Sci. U S A* 110, E322–E330. doi: 10.1073/pnas.1212956110
- Cavézian, C., Valadao, D., Hurwitz, M., Saoud, M., and Danckert, J. (2012). Finding centre: ocular and fMRI investigations of bisection and landmark task performance. *Brain Res.* 1437, 89–103. doi: 10.1016/j.brainres.2011.12.002
- Chen, P., Goedert, K. M., Murray, E., Kelly, K., Ahmeti, S., and Barrett, A. M. (2011). Spatial bias and right hemisphere function: sex-specific changes with aging. *J. Int. Neuropsychol. Soc.* 17, 455–462. doi: 10.1017/S135561771100004X
- Chokron, S., Helft, G., and Perez, C. (2013). Effects of age and cardiovascular disease on selective attention. *Cardiovasc. Psychiatry Neurol.* 2013:185385. doi: 10.1155/2013/185385
- Çiçek, M., Deouell, L. Y., and Knight, R. T. (2009). Brain activity during landmark and line bisection tasks. *Front. Human. Neurosci.* 3:7. doi: 10.3389/neuro.09.007.2009
- Collins, K., and Mohr, C. (2013). Performance of younger and older adults in lateralised right and left hemisphere asymmetry tasks supports the HAROLD model. *Laterality* 18, 491–512. doi: 10.1080/1357650X.2012.724072
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755

- Corbetta, M., and Shulman, G. L. (2011). Spatial neglect and attention networks. *Annu. Rev. Neurosci.* 34, 569–599. doi: 10.1146/annurev-neuro-061010-113731
- De Sanctis, P., Katz, R., Wylie, G. R., Sehatpour, P., Alexopoulos, G. S., and Foxe, J. J. (2008). Enhanced and bilateralized visual sensory processing in the ventral stream may be a feature of normal aging. *Neurobiol. Aging* 29, 1576–1586. doi: 10.1016/j.neurobiolaging.2007.03.021
- Dodds, C. M., van Belle, J., Peers, P. V., Dove, A., Cusack, R., Duncan, J., et al. (2008). The effects of time-on-task and concurrent cognitive load on normal visuospatial bias. *Neuropsychology* 22, 545–552. doi: 10.1037/0894-4105.22.4.545
- Dolcos, F., Rice, H. J., and Cabeza, R. (2002). Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction. *Neurosci. Biobehav. Rev.* 26, 819–825. doi: 10.1016/s0149-7634(02)00068-4
- Doricchi, F., Guariglia, P., Figliozzi, F., Silvetti, M., Bruno, G., and Gasparini, M. (2005). Causes of cross-over in unilateral neglect: between-group comparisons, within-patient dissociations and eye movements. *Brain* 128, 1386–1406. doi: 10.1093/brain/awh461
- Driver, J., and Mattingley, J. B. (1998). Parietal neglect and visual awareness. *Nat. Neurosci.* 1, 17–22. doi: 10.1038/217
- Duecker, F., Formisano, E., and Sack, A. T. (2013). Hemispheric differences in the voluntary control of spatial attention: direct evidence for a right-hemispheric dominance within frontal cortex. *J. Cogn. Neurosci.* 25, 1332–1342. doi: 10.1162/jocn_a_00402
- Dufour, A., Touzalin, P., and Candas, V. (2007). Time-on-task effect in pseudoneglect. *Exp. Brain Res.* 176, 532–537. doi: 10.1007/s00221-006-0810-2
- Failla, C. V., Sheppard, D. M., and Bradshaw, J. L. (2003). Age and responding-hand related changes in performance of neurologically normal subjects on the line-bisection and chimeric-faces tasks. *Brain Cogn.* 52, 353–363. doi: 10.1016/s0278-2626(03)00181-7
- Fierro, B., Brighina, F., Piazza, A., Oliveri, M., and Bisiach, E. (2001). Timing of right parietal and frontal cortex activity in visuo-spatial perception: a TMS study in normal individuals. *Neuroreport* 12, 2605–2607. doi: 10.1097/00001756-200108080-00062
- Fimm, B., Willmes, K., and Spijkers, W. (2006). The effect of low arousal on visuo-spatial attention. *Neuropsychologia* 44, 1261–1268. doi: 10.1016/j.neuropsychologia.2006.01.027
- Fink, G. R., Marshall, J. C., Shah, N. J., Weiss, P. H., Halligan, P. W., Grosse-Ruyken, M., et al. (2000a). Line bisection judgments implicate right parietal cortex and cerebellum as assessed by fMRI. *Neurology* 54, 1324–1331. doi: 10.1212/wnl.54.6.1324
- Fink, G. R., Marshall, J. C., Weiss, P. H., and Zilles, K. (2001). The neural basis of vertical and horizontal line bisection judgments: an fMRI study of normal volunteers. *Neuroimage* 14, S59–S67. doi: 10.1006/nimg.2001.0819
- Fink, G. R., Marshall, J. C., Weiss, P. H., Shah, N. J., Toni, I., Halligan, P. W., et al. (2000b). 'Where' depends on 'what': a differential functional anatomy for position discrimination in one- versus two-dimensions. *Neuropsychologia* 38, 1741–1748. doi: 10.1016/s0028-3932(00)00078-6
- Foxe, J. J., McCourt, M. E., and Javitt, D. C. (2003). Right hemisphere control of visuospatial attention: line-bisection judgments evaluated with high-density electrical mapping and source analysis. *Neuroimage* 19, 710–726. doi: 10.1016/s1053-8119(03)00057-0
- Friedman, D. (2013). The cognitive aging of episodic memory: a view based on the event-related brain potential. *Front. Behav. Neurosci.* 7:111. doi: 10.3389/fnbeh.2013.00111
- Fründ, I., Valentin Haenel, N., and Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behaviour. *J. Vis.* 11:16. doi: 10.1167/11.6.16
- Fujii, T., Fukatsu, R., Yamadori, A., and Kimura, I. (1995). Effect of age on the line bisection test. *J. Clin. Exp. Neuropsychol.* 17, 941–944. doi: 10.1080/01688639508402443
- Fukatsu, R., Fujii, T., Kimura, I., Saso, S., and Kogure, K. (1990). Effects of hand and spatial conditions on visual line bisection. *Tohoku J. Exp. Med.* 161, 329–333. doi: 10.1620/tjem.161.329
- García-Pérez, M. A., and Alcalá-Quintana, R. (2013). Shifts of the psychometric function: distinguishing bias from perceptual effects. *Q. J. Exp. Psychol. (Hove)* 66, 319–337. doi: 10.1080/17470218.2012.708761
- Ghacibeh, G. A., Shenker, J. I., Winter, K. H., Triggs, W. J., and Heilman, K. M. (2007). Dissociation of neglect subtypes with transcranial magnetic stimulation. *Neurology* 69, 1122–1127. doi: 10.1212/01.wnl.0000276950.77470.50
- Goedert, K. M., LeBlanc, A., Tsai, S. W., and Barrett, A. M. (2010). Asymmetrical effects of adaptation to left- and right-shifting prisms depends on pre-existing attentional biases. *J. Int. Neuropsychol. Soc.* 16, 795–804. doi: 10.1017/S1355617710000597
- Goldstein, G., and Shelly, C. (1981). Does the right hemisphere age more rapidly than the left? *J. Clin. Neuropsychol.* 3, 65–78. doi: 10.1080/01688638108403114
- Grady, C. L., Bernstein, L. J., Beig, S., and Siegenthaler, A. L. (2002). The effects of encoding task on age-related differences in the functional neuroanatomy of face memory. *Psychol. Aging* 17, 7–23. doi: 10.1037//0882-7974.17.1.7
- Halligan, P. W., Fink, G. R., Marshall, J. C., and Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends Cogn. Sci.* 7, 125–133. doi: 10.1016/s1364-6613(03)00032-9
- Halligan, P. W., and Marshall, J. C. (1988). How long is a piece of string? A study of line bisection in a case of visual neglect. *Cortex* 24, 321–328. doi: 10.1016/s0010-9452(88)80040-6
- Halligan, P. W., Manning, L., and Marshall, J. C. (1990). Individual variation in line bisection: a study of four patients with right hemisphere damage and normal controls. *Neuropsychologia* 28, 1043–1051. doi: 10.1016/0028-3932(90)90139-f
- Harvey, M., Milner, A. D., and Roberts, R. C. (1995). Differential effects of line length on bisection judgements in hemispatial neglect. *Cortex* 31, 711–722. doi: 10.1016/s0010-9452(13)80022-6
- Harvey, M., and Rossit, S. (2012). Visuospatial neglect in action. *Neuropsychologia* 50, 1018–1028. doi: 10.1016/j.neuropsychologia.2011.09.030
- Hatin, B., Tottenham, L. S., and Oriet, C. (2012). The relationship between collisions and pseudoneglect: is it right? *Cortex* 48, 997–1008. doi: 10.1016/j.cortex.2011.05.015
- Hausmann, M., Waldie, K. E., and Corballis, M. C. (2003). Developmental changes in line bisection: a result of callosal maturation? *Neuropsychology* 17, 155–160. doi: 10.1037//0894-4105.17.1.155
- Heber, I., Siebertz, S., Wolter, M., Kühlen, T., and Fimm, B. (2010). Horizontal and vertical pseudoneglect in peri- and extrapersonal space. *Brain Cogn.* 73, 160–166. doi: 10.1016/j.bandc.2010.04.006
- Heber, I. A., Valvoda, J. T., Kühlen, T., and Fimm, B. (2008). Low arousal modulates visuospatial attention in three-dimensional virtual space. *J. Int. Neuropsychol. Soc.* 14, 309–317. doi: 10.1017/S135561770808034X
- Jenkins, L., Myerson, J., Joerding, J. A., and Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychol. Aging* 15, 157–175. doi: 10.1037//0882-7974.15.1.157
- Jewell, G., and McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia* 38, 93–110. doi: 10.1016/s0028-3932(99)00045-7
- Kerkhoff, G., and Schenk, T. (2011). Line bisection in homonymous visual field defects—Recent findings and future directions. *Cortex* 47, 53–58. doi: 10.1016/j.cortex.2010.06.014
- Kinsbourne, M. (1977). Hemineglect and hemisphere rivalry. *Adv. Neurol.* 18, 41–49.
- Klisz, D. (1978). "Neuropsychological evaluation in older persons," in *The Clinical Psychology of Aging*, eds M. Storandt, I. C. Siegler and M. F. Elias (New York: Plenum Publishing), 71–95.
- Koch, G., Cercignani, M., Bonni, S., Giacobbe, V., Bucchi, G., Versace, V., et al. (2011). Asymmetry of parietal interhemispheric connections in humans. *J. Neurosci.* 31, 8967–8975. doi: 10.1523/JNEUROSCI.6567-10.2011
- Koch, G., Fernandez Del Olmo, M., Cheeran, B., Ruge, D., Schippling, S., Caltagirone, C., et al. (2007). Focal stimulation of the posterior parietal cortex increases the excitability of the ipsilateral motor cortex. *J. Neurosci.* 27, 6815–6822. doi: 10.1523/jneurosci.0598-07.2007
- Li, Z., Moore, A. B., Tyner, C., and Hu, X. (2009). Asymmetric connectivity reduction and its relationship to "HAROLD" in aging brain. *Brain Res.* 1295, 149–158. doi: 10.1016/j.brainres.2009.08.004
- Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., and Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: dissociable neural mechanisms associated with aging. *Neuron* 33, 827–840. doi: 10.1016/s0896-6273(02)00612-8
- Longo, M. R., and Lourenco, S. F. (2006). On the nature of near space: effects of tool use and the transition to far space. *Neuropsychologia* 44, 977–981. doi: 10.1016/j.neuropsychologia.2005.09.003

- Loureiro, M., d'Almeida, O. C., Mateus, C., Oliveiros, B., and Castelo-Branco, M. (2013). The effect of normal development and aging on low-level visual field asymmetries. *Perception* 42, 138. doi: 10.1068/v130385
- Luh, K. E. (1995). Line bisection and perceptual asymmetries in normal individuals: what you see is not what you get. *Neuropsychology* 9, 435–448. doi: 10.1037//0894-4105.9.4.435
- Lux, S., Marshall, J. C., Thimm, M., and Fink, G. R. (2008). Differential processing of hierarchical visual stimuli in young and older healthy adults: implications for pathology. *Cortex* 44, 21–28. doi: 10.1016/j.cortex.2005.08.001
- Manly, T., Dobler, V. B., Dodds, C. M., and George, M. A. (2005). Rightward shift in spatial awareness with declining alertness. *Neuropsychologia* 43, 1721–1728. doi: 10.1016/j.neuropsychologia.2005.02.009
- Marshall, J. C., and Halligan, P. W. (1989). When right goes left: an investigation of line bisection in a case of visual neglect. *Cortex* 25, 503–515. doi: 10.1016/s0010-9452(89)80065-6
- Matthias, E., Bublak, P., Muller, H. J., Schneider, W. X., Krummenacher, J., and Finke, K. (2010). The influence of alertness on spatial and nonspatial components of visual attention. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 38–56. doi: 10.1037/a0017602
- McCourt, M. E., and Garlinghouse, M. (2000). Asymmetries of visuospatial attention are modulated by viewing distance and visual field elevation: pseudoneglect in peripersonal and extrapersonal space. *Cortex* 36, 715–731. doi: 10.1016/s0010-9452(08)70548-3
- McCourt, M. E., and Jewell, G. (1999). Visuospatial attention in line bisection: stimulus modulation of pseudoneglect. *Neuropsychologia* 37, 843–855. doi: 10.1016/s0028-3932(98)00140-7
- McCourt, M. E., and Olafson, C. (1997). Cognitive and perceptual influences on visual line bisection: psychophysical and chronometric analyses of pseudoneglect. *Neuropsychologia* 35, 369–380. doi: 10.1016/s0028-3932(96)00143-1
- Mennemeier, M., Pierce, C. A., Chatterjee, A., Anderson, B., Jewell, G., Woods, A. J., et al. (2005). Biases in attentional orientation and magnitude estimation explain crossover: neglect is a disorder of both. *J. Cogn. Neurosci.* 17, 1194–1211. doi: 10.1162/0898929055002454
- Milner, A. D., Brechmann, M., and Pagliarini, L. (1992). To halve and to halve not: an analysis of line bisection judgements in normal subjects. *Neuropsychologia* 30, 515–526. doi: 10.1016/0028-3932(92)90055-q
- Milner, A. D., and Harvey, M. (1995). Distortion of size perception in visuospatial neglect. *Curr. Biol.* 5, 85–89. doi: 10.1016/s0960-9822(95)00020-0
- Milner, A. D., Harvey, M., Roberts, R. C., and Forster, S. V. (1993). Line bisection errors in visual neglect: misguided action or size distortion? *Neuropsychologia* 31, 39–49. doi: 10.1016/0028-3932(93)90079-f
- Monaghan, P., and Shillcock, R. (1998). The cross-over effect in unilateral neglect. Modelling detailed data in the line-bisection task. *Brain* 121, 907–921. doi: 10.1093/brain/121.5.907
- Monaghan, P., and Shillcock, R. (2004). Hemispheric asymmetries in cognitive modelling: connectionist modelling of unilateral visual neglect. *Psychol. Rev.* 111, 283–308. doi: 10.1037/0033-295x.111.2.283
- Morgan, M., Dillenburger, B., Raphael, S., and Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity. *Atten. Percept. Psychophys.* 74, 185–193. doi: 10.3758/s13414-011-0222-7
- Nagamatsu, L. S., Carolan, P., Liu-Ambrose, T. Y. L., and Handy, T. C. (2011). Age-related changes in the attentional control of visual cortex: a selective problem in the left visual hemifield. *Neuropsychologia* 49, 1670–1678. doi: 10.1016/j.neuropsychologia.2011.02.040
- Nebes, R. D., Buysse, D. J., Halligan, E. M., Houck, P. R., and Monk, T. H. (2009). Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 64B, 180–187. doi: 10.1093/geronb/gbn037
- Newman, D. P., O'Connell, R. G., and Bellgrove, M. A. (2013). Linking time-on-task, spatial bias and hemispheric activation asymmetry: a neural correlate of rightward attention drift. *Neuropsychologia* 51, 1215–1223. doi: 10.1016/j.neuropsychologia.2013.03.027
- Nielson, K. A., Langenecker, S. A., and Garavan, H. (2002). Differences in the functional neuroanatomy of inhibitory control across the adult lifespan. *Psychol. Aging* 17, 56–71. doi: 10.1037//0882-7974.17.1.56
- O'Connell, R. G., Schneider, D., Hester, R., Mattingley, J. B., and Bellgrove, M. A. (2011). Attentional load asymmetrically affects early electrophysiological indices of visual orienting. *Cereb. Cortex* 21, 1056–1065. doi: 10.1093/cercor/bhq148
- Olk, B., and Harvey, M. (2002). Effects of visible and invisible cueing on line bisection and landmark performance in hemispatial neglect. *Neuropsychologia* 40, 282–290. doi: 10.1016/s0028-3932(01)00095-1
- Posner, M. I., and Cohen, Y. (1984). Components of visual orienting. *Atten. Perform.* 32, 531–556.
- Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A., and Acker, J. D. (2004). Aging, sexual dimorphism and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol. Aging* 25, 377–396. doi: 10.1016/S0197-4580(03)00118-0
- Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Reuter-Lorenz, P. A., and Lustig, C. (2005). Brain aging: reorganizing discoveries about the aging mind. *Curr. Opin. Neurobiol.* 15, 245–251. doi: 10.1016/j.conb.2005.03.016
- Reuter-Lorenz, P. A., and Park, D. C. (2010). Human neuroscience and the aging mind: a new look at old problems. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 65, 405–415. doi: 10.1093/geronb/gbq035
- Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marchuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12, 174–187. doi: 10.1162/089892900561814
- Reuter-Lorenz, P. A., Kinsbourne, M., and Moscovitch, M. (1990). Hemispheric control of spatial attention. *Brain Cogn.* 12, 240–266. doi: 10.1016/0278-2626(90)90018-j
- Ricci, R., and Chatterjee, A. (2001). Context and crossover in unilateral neglect. *Neuropsychologia* 39, 1138–1143. doi: 10.1016/s0028-3932(01)00054-9
- Robinson, D. L., and Kertzman, C. (1990). Visuospatial attention: effects of age, gender and spatial reference. *Neuropsychologia* 28, 291–301. doi: 10.1016/0028-3932(90)90022-g
- Rossi, S., Miniussi, C., Pasqualetti, P., Babiloni, C., Rossini, P. M., and Cappa, S. F. (2004). Age-related functional changes of prefrontal cortex in long-term memory: a repetitive transcranial magnetic stimulation study. *J. Neurosci.* 24, 7939–7944. doi: 10.1523/JNEUROSCI.0703-04.2004
- Rueckert, L., Deravanesian, A., Baboorian, D., Lacalmita, A., and Repplinger, M. (2002). Pseudoneglect and the cross-over effect. *Neuropsychologia* 40, 162–173. doi: 10.1016/s0028-3932(01)00082-3
- Rueckert, L., and McFadden, H. G. (2004). Context effects in pseudo-neglect measured with a free vision landmark task. *Laterality* 9, 421–432. doi: 10.1080/13576500342000257
- Schmitz, R., and Peigneux, P. (2011). Age-related changes in visual pseudoneglect. *Brain Cogn.* 76, 382–389. doi: 10.1016/j.bandc.2011.04.002
- Schmitz, R., Dehon, H., and Peigneux, P. (2013). Lateralized processing of false memories and pseudoneglect in aging. *Cortex* 49, 1314–1324. doi: 10.1016/j.cortex.2012.06.005
- Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools.
- Solé-Padullés, C., Bartrés-Faz, D., Junqué, C., Clemente, I. C., Molinuevo, J. L., Bargalló, N., et al. (2006). Repetitive transcranial magnetic stimulation effects on brain function and cognition among elders with memory dysfunction. A randomized sham-controlled study. *Cereb. Cortex* 16, 1487–1493. doi: 10.1093/cercor/bhj083
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., and Toga, A. W. (2003). Mapping cortical change across the human life span. *Nat. Neurosci.* 6, 309–315. doi: 10.1038/nn1008
- Stam, C. J., and Bakker, M. (1990). The prevalence of neglect: superiority of neuropsychological over clinical methods of estimation. *Clin. Neurol. Neurosurg.* 92, 229–235. doi: 10.1016/0303-8467(90)90025-z
- Sullivan, E. V., and Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neurosci. Biobehav. Rev.* 30, 749–761. doi: 10.1016/j.neubiorev.2006.06.002
- Szczepanski, S. M., and Kastner, S. (2013). Shifting attentional priorities: control of spatial attention through hemispheric competition. *J. Neurosci.* 33, 5411–5421. doi: 10.1523/JNEUROSCI.4089-12.2013
- Thiebaut de Schotten, M., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G., et al. (2011). A lateralized network for visuospatial attention. *Nat. Neurosci.* 14, 1245–1246. doi: 10.1038/nm.2905
- Thiebaut de Schotten, M., Urbanski, M., Duffau, H., Volle, E., Levy, R., Dubois, B., et al. (2005). Direct evidence for a parietal-frontal pathway subserving spatial awareness in humans. *Science* 309, 2226–2228. doi: 10.1126/science.1116251

- Thomas, N. A., Loetscher, T., and Nicholls, M. E. R. (2012). Central fixations with rightward deviations: saccadic eye movements on the landmark task. *Exp. Brain Res.* 220, 29–39. doi: 10.1007/s00221-012-3113-9
- Torraldo, A., McIntosh, R. D., Dijkerman, C. H., and Milner, D. H. (2004). A revised method for analysing neglect using the landmark task. *Cortex* 40, 415–431. doi: 10.1016/s0010-9452(08)70136-9
- Urbanski, M., and Bartolomeo, P. (2008). Line bisection in left neglect: the importance of starting right. *Cortex* 44, 782–793. doi: 10.1016/j.cortex.2007.04.001
- Vallar, G. (1998). Spatial hemineglect in humans. *Trends Cogn. Sci.* 2, 87–97. doi: 10.1016/s1364-6613(98)01145-0
- Varnava, A., and Halligan, P. W. (2007). Influence of age and sex on line bisection: a study of normal performance with implications for visuospatial neglect. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 14, 571–585. doi: 10.1080/13825580600826454
- Varnava, A., McCarthy, M., and Beaumont, J. G. (2002). Line bisection in normal adults: direction of attentional bias for near and far space. *Neuropsychologia* 40, 1372–1378. doi: 10.1016/s0028-3932(01)00204-4
- Veronelli, L., Vallar, G., Marinelli, C. V., Primativo, S., and Arduino, L. S. (2014). Line and word bisection in right-brain-damaged patients with left spatial neglect. *Exp. Brain Res.* 232, 133–146. doi: 10.1007/s00221-013-3726-7
- Voyer, D., Voyer, S. D., and Tramonte, L. (2012). Free-viewing laterality tasks: a multilevel meta-analysis. *Neuropsychology* 26, 551–567. doi: 10.1037/a0028631
- Waberski, T. D., Gobbele, R., Lamberty, K., Buchner, H., Marshall, J. C., and Fink, G. R. (2008). Timing of visuo-spatial information processing: electrical source imaging related to line bisection judgments. *Neuropsychologia* 46, 1201–1210. doi: 10.1016/j.neuropsychologia.2007.10.024

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 February 2014; accepted: 20 May 2014; published online: 10 June 2014.

Citation: Benwell CSY, Thut G, Grant A and Harvey M (2014) A rightward shift in the visuospatial attention vector with healthy aging. *Front. Aging Neurosci.* 6:113. doi: 10.3389/fnagi.2014.00113

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Benwell, Thut, Grant and Harvey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The drive-wise project: driving simulator training increases real driving performance in healthy older drivers

Gianclaudio Casutt^{1,2,3*}, Nathan Theill⁴, Mike Martin^{5,6}, Martin Keller^{1,7} and Lutz Jäncke^{1,3,5,8*}

¹ Division of Neuropsychology, Department of Psychology, University of Zurich, Zurich, Switzerland

² Division of Gerontopsychology, Department of Psychology, University of Zurich, Zurich, Switzerland

³ International Normal Aging and Plasticity Research Centre Zurich, Zurich, Switzerland

⁴ Division of Psychiatry Research, University of Zurich, Zurich, Switzerland

⁵ University Research Priority Program "Dynamics of Healthy Aging," University of Zurich, Zurich, Switzerland

⁶ Center for Gerontology, University of Zurich, Zurich, Switzerland

⁷ Department of Neurology, Rehabilitation Centre Valens, Valens, Switzerland

⁸ Department of Special Education, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Asgar Zaheer, University of Iowa

Hospitals and Clinics, USA

David Crundall, Nottingham Trent University, UK

*Correspondence:

Gianclaudio Casutt and Lutz Jäncke,
Division of Neuropsychology,
Department of Psychology,
University of Zurich,
Binzmühlestrasse 14/25, Zurich
CH-8050, Switzerland
e-mail: g.casutt@psychologie.uzh.ch;
l.jaencke@psychologie.uzh.ch

Background: Age-related cognitive decline is often associated with unsafe driving behavior. We hypothesized that 10 active training sessions in a driving simulator increase cognitive and on-road driving performance. In addition, driving simulator training should outperform cognitive training.

Methods: Ninety-one healthy active drivers (62–87 years) were randomly assigned to one of three groups: (1) a driving simulator training group, (2) an attention training group (vigilance and selective attention), or (3) a control group. The main outcome variables were on-road driving and cognitive performance. Seventy-seven participants (85%) completed the training and were included in the analyses. Training gains were analyzed using a multiple regression analysis with planned orthogonal comparisons.

Results: The driving simulator-training group showed an improvement in on-road driving performance compared to the attention-training group. In addition, both training groups increased cognitive performance compared to the control group.

Conclusion: Driving simulator training offers the potential to enhance driving skills in older drivers. Compared to the attention training, the simulator training seems to be a more powerful program for increasing older drivers' safety on the road.

Keywords: cognitive training, training effects, driving simulator, on-road driving performance, cognitive performance

INTRODUCTION

Due to the changing age structure in industrial countries more and more older drivers (>65 years) will drive a car on public roads either for the sake of mobility, leisure activities, or business reasons. However, there is ample evidence that on average driving performance declines and crash risks increase with increasing age (Lyman et al., 2002; Casutt et al., 2013). This decline in driving performance is also associated with a decline in perception (sensory functions), cognition (perceptual speed, higher order cognitive functions) and physiological functions as well as medical conditions (Anstey et al., 2005). Many driving errors result as a consequence of a reduction in cognitive performance, which however should be improved by training and practice (Anstey and Wood, 2011). Thus, there is growing interest in many countries to cope with increasing crash risks and decreasing driving performance in older drivers (OECD, 2001). Many strategies have been proposed so far to reduce age-related crash risks comprising specific educational programs (Stalvey and Owsley, 2003; Owsley et al., 2004; Baldock et al., 2008), withdrawal of the driving license at a particular age (Langford et al., 2004), or training cognitive functions supposed to underlie driving performance (Roegner et al., 2003; Edwards et al., 2009a,b; Ball et al., 2011).

Cognitive training regimes in older adults have consistently demonstrated improvements in the trained cognitive tasks (e.g., Karbach and Kray, 2009; von Bastian et al., 2013). However, most of these studies demonstrated transfer effects only to very similar tasks (near transfer) (Lustig et al., 2009), and virtually no transfer to other domains (far transfer; Lustig et al., 2009; Zelinski, 2009). However, the complexity of cognitive training approaches seems to be an important variable influencing far transfer of cognitive training. In fact, several studies showed that the complexity of the cognitive training increases far transfer to other cognitive domains (Basak et al., 2008; Karbach and Kray, 2009; Marmeleira et al., 2009) most likely because several cognitive functions are simultaneously trained.

In line with these findings and focusing on driving problems in older subjects there is evidence that training of particular cognitive functions can exert beneficial effects on the driving behavior. Cassavaugh and Kramer (2009) found in their driving simulator study that cognitive performance was associated with driving simulator performance. Furthermore, practicing several cognitive functions (including sensorimotor control, selective attention, working memory, and dual tasking) for eight sessions across several days resulted in improved

driving performance (lane change, distance from the vehicle ahead, less driving errors, shorter reaction time). A further set of studies explored the effects of a “speed of processing training” on driving performance and identified improved performance in driving-related functions like UFOV (useful field of view), driving safety (Roenker et al., 2003), or reduced number of crashes (Ball et al., 2011) as well as number of self-reported driving difficulties (Edwards et al., 2009a,b). Interestingly, in some studies the cognitive training regimes resulted in long-lasting beneficial influences on driving behavior. For example, in the Roenker et al. (2003) study positive effects have been identified 18 months after the cognitive training. Ball et al. (2011) even reported reduced number of crashes in an observation period of 5 years.

A further strategy to improve driving performance in older adults is to practice *active* driving on a driving simulator. Simulators are frequently and intensively used in the context of various transportation situations (rail, aviation, maritime transport, space travel) especially where vehicles are very expensive in relation to a simulator. Lees et al. (2010) postulated in their review that driving simulators offer important opportunities for an efficient and valid training (interactivity, complexity, simultaneous use of different domains) not only for novice drivers, but also for older drivers.

As described above cognitive training (e.g., speed of processing) positively influences driving related variables like reduction of dangerous maneuvers, driving cessation, and driving errors (Roenker et al., 2003; Edwards et al., 2009a,b; Ball et al., 2011). Additionally, several studies have been published so far using driving simulator training approaches to improve specific and accident related driving behaviors in older adults. In older drivers reduced hazard perception was associated with reduced UFOV performance (Horswill et al., 2008). Hazard perception training in a driving simulator resulted in faster anticipation of hazardous traffic situations (Horswill et al., 2010). Other driving simulator studies investigated different aspects of problematic driving behavior (e.g., visual scanning at intersections, use of mirror while overtaking). Romoser and Fisher (2009) trained older drivers visual scanning at intersections with a driving simulator. After simulator training visual scanning (secondary looks) was improved during driving in the simulator and on the road. Furthermore, after training they examined an increase in *Rey-Osterreith Complex Figure* test (ROCFT) performance, which is associated with cognitive functions like attention, planning, and working memory (executive functions). The improved performance for “secondary looks” was observed in a follow-up 2 years later (Romoser, 2013). In another driving simulator study the use of side and rear mirrors while overtaking were trained in a sample of older drivers. Following the training, frequency of blind spot inspection increased in comparison to a training group who received no feedback (Lavallière et al., 2012). Taken together different aspects of the driving performance in older drivers (e.g., visual scanning in intersection, hazard perception, use of mirror during lane change, visuo-spatial memory) can be improved with appropriate driving simulator training (Romoser and Fisher, 2009; Horswill et al., 2010; Lavallière et al., 2012). These studies have also shown that driving simulator

training can positively influence very specific aspects of cognition (e.g., executive functions) (e.g., Romoser and Fisher, 2009).

While these studies have shown that driving simulator and cognitive training regimes both do have the potential to change very specific aspects of driving (and cognition) we are more interested to examine whether general driving performance differentially benefits from a driving simulator or cognitive training. The cognitive training was designed to practice cognitive functions, which have been shown to be essential for effective driving (e.g., vigilance and selective attention) (Anstey et al., 2005; Selander et al., 2011; Casutt et al., 2014). Different to the aforementioned studies we were interested to examine whether our driving simulator training improves real on-road driving in general and not behavior in specific driving situations (e.g., visual scanning in intersection, use of mirrors during lane change). Our driving simulator training approach (practicing driving through towns, on highways, rural roads with changing traffic situations etc.) was based on a practical everyday behavior. Therefore, the used scenarios were comparable to on-road driving, which is a complex behavior and needs several psychological functions (Hakamies-Blomqvist, 1994). Our training approach is similar to multi- or dual-task training approaches, which have been shown to be more effective than single-task training (Basak et al., 2008; Marmeleira et al., 2009; Anguera et al., 2013). Real driving is a highly demanding task requiring the orchestration of many psychological functions to process many information simultaneously (traffic observation, speed control, scanning for hazard events, traffic rules, car handling). If demands increase, also the likelihood of driving errors increase (Holm et al., 2009). The relation between reduced multitasking ability and unsafe driving in older drivers and the use of compensatory strategies is well known (Sheridan, 2004; Cantin et al., 2009). Therefore, our training approach for the driving simulator training was to increase the multitasking demands in a realistic way.

Since on-road driving is difficult to assess and strongly depends on local aspects (e.g., traffic density, specific population, and specific traffic rules) we used a new on-road driving test specifically designed for a major European city (Zurich in Switzerland) with dense traffic to test whether intensive driving simulator training improves real on-road driving. In addition, we were also interested to examine whether an intensive attention training of psychological functions known to be involved in controlling driving might influence real on-road driving performance. In this context we also paid attention to examine whether our driving simulator and cognitive training exert different effects on the on-road driving performance.

Based on the results of the afore-mentioned studies we hypothesize that our driving simulator training will induce stronger improvements in on-road driving than attention training since the driving simulator training needs stronger multitasking skills and seems more attractive than attention training. In addition, we hypothesize that both training regimes (driving simulator and attention training) will improve cognitive performance **and** on-road driving compared to a no-training control group.

MATERIALS AND METHOD

PARTICIPANTS

Participants were recruited via a newspaper articles and a newspaper advertisement about the *Drive-Wise* project. A total of 244 participants indicated interest in study participation. All of them received detailed study information and a short medical condition questionnaire (driving relevant illness, e.g., all kinds of neurological and psychiatric disorders, orthopedic problems of the upper and lower extremities etc.), medication influencing driving (e.g., drug intake influencing the central nervous system), sensory impairment (e.g., visual field < 140°). In addition, the active driving status (annual driving distance, years of possession of driving license, driving context) was assessed with a questionnaire. Participants who did not drive in all common driving contexts (urban, rural, motorway) were excluded. Ninety-one participants agreed to participate in the study and fulfilled all inclusion criteria. It is worth to mention that in Switzerland, drivers older than 70 years must undergo a screening test every 2 years (medical and cognitive screening) for renewal of their drivers' license. All participants had an original valid driver license.

Participants were not financially compensated for their travel expenses or participation, but received a detailed written feedback about performance in cognition and driving after finishing their participation. Before data collection, participants were randomly allocated either to a simulator training condition ($n = 39$), a cognitive training condition ($n = 26$), or a control group ($n = 26$). However, 14 participants dropped out during data collection. Seventy-seven individuals (55 men, 71.4%) with a mean age of 72.36 ± 5.61 (range 62–87) completed the study (Table 1). The three groups were influenced differently by the dropouts. In the simulator training group seven participants (six female) dropped out due to simulator sickness (SS) and one participant stopped due to excessive experimental demands. In the cognitive training group as well as in the control group, three participants (two female per group) finished participation due to time constraints (the entire study lasted for approximately 2 years). Study information for all groups was identical except for the particular information to run the driving simulator and cognitive training. The training setting was not explained in detail. Participants in the control group were offered the simulator training sessions (according to time of training for the two training groups) after finishing their study participation (two assessment evaluations with a 5-week waiting period in between).

This project (*Drive-Wise*) was approved by the Cantonal Ethic-Commission of Zurich, University Hospital of Zurich (KEK-ZH-NR: 2010-0090/0). Furthermore, traffic and police

departments have granted permission to conduct the on-road test assessment. The private car of participants was labeled during the on-road test. According to information by the ethics committee, participants were informed that participation would not impinge on their driving license and that they had permission to terminate the study at any time without any negative consequences.

EXPERIMENTAL APPARATUS

The on-road test drive was conducted in the participants' private car. Start and end point was always the department of psychology, Zurich. All tests for the cognitive test battery were conducted on a Windows Computer with a 15" screen (resolution 1280 × 1024), distance of approximately 40 cm to the participant. Response panel and other hardware were products of Schuhfried GmbH (Schuhfried, n.d.). The training sessions of participants in the attention training group were conducted on that system as well (Phasic and tonic alertness and vigilance; CogniPlus Software from Schuhfried GmbH). Participants in the simulator training group conducted their sessions on a driving simulator type "Trainer F12PT-1L40," software version 12 of Dr. Foerst GmbH, 32," Samsung LCD-screen (resolution 1920 × 1080), distance of approximately 70 cm to the steering wheel (Jäncke and Klimmt, 2011). Participants sat in a driver's seat of a *Ford Focus*® equipped with a steering wheel, a starter lock, a tachometer, signalers for light and blinker, wiper control switch, clutch, breaking, and throttle pedals as well as gearshift (Figure 1). The software recorded participants driving behavior. The software automatically produced traffic scenarios on a Windows 7 operating system. Moreover there were two operator screens in the same room for controlling the training sessions and giving feedback after training.

ON-ROAD TEST DRIVING ASSESSMENT

The on-road test drive was conducted on public roads including district and urban streets, suburb and rural roads and a motorway passage with a total length of approximately 25 km. The test track was used as a regular basis for official on-road test exercises. A licensed driving instructor (DI) blinded to condition, sat in the front passenger seat and rated participants driving behavior directly after finishing the driving session. During the ride the instructors made notes about the driving performance, which they used for the final evaluation. The evaluation sheet (*Zurich On-road test Assessment, ZOA*), is a modified version of the formal evaluation sheet used by the DI. The DI was instructed to evaluate only cognitive aspects of driving behavior but not car handling. Seven different dimensions (Table 2) were implemented in the

Table 1 | Demographic characteristics.

Variable	N	Age (years)	Gender		Years of license	km/year	Gearbox		
		M (SD)	Male	Female	M (SD)	M (SD)	A	M	A/M
Simulator training	31	71.74 (5.18)	22	9	49.77 (5.16)	11909 (7338)	14	11	6
Cognitive training	23	72.30 (6.46)	15	8	50.21 (5.95)	8973 (6067)	11	10	2
No training	23	73.26 (5.38)	18	5	51.34 (6.85)	10934 (5079)	8	10	5

A, automatic; M, manual.

ZOA using six to eight items on a 5-point-scale (poor = 1, slightly insufficient = 2, sufficient = 3, good = 4, excellent = 5). An overall on-road driving performance was calculated as the mean of all on-road driving measures. This composite measure was used as dependent variable for the evaluation of on-road driving performance (Cronbach's $\alpha = 0.95$). The internal reliability coefficients for each dimension ranges between 0.62–0.83.



FIGURE 1 | Still photo of the used driving simulator.

COGNITIVE TEST BATTERY

This test battery is a well-established and standardized computer-based version of the *Expert System Traffic XPSV* (Schuhfried, 2005) often used as a standard test for evaluating driving-related cognitive performance in European countries (Sommer et al., 2008, 2009, 2010). A recent paper has shown that this test battery explains 50% of the on-road test variance (Risser et al., 2008). **Table 3** gives an overview of all subtests.

The *Reaction Test* (RT) is a simple choice reaction time task. From three different stimuli (yellow or red circle and an acoustic signal), participants have to discriminate the simultaneous presentation of the yellow circle and the pitched sound by pushing a corresponding target button with the right index finger as fast as possible. In all other conditions, single yellow or red circle, single pitched sound, combination of red circle and pitched sound, participants have to suppress a movement. Decision speed (DS) is measured in milliseconds by the latency from stimulus onset until lifting off the start button while the physical motor speed (MS) in milliseconds is defined as the movement time from the start button to the target button.

The *Cognitrone Test* (COG) measures selective attention. During test administration, different geometrical figures are presented block-wise. Each block comprises 60 trials. During each trial two different stimulus types are presented: four reference stimuli and one test stimulus. The four reference stimuli are presented as an array above the test stimulus. The subject's task is to decide whether the test stimulus is identical to one of the reference

Table 2 | ZOA (Zurich On-road Assessment).

Traffic observation (0.62)*	Speed at intersection (0.78)*	Gaze behavior (0.83)*	Change of direction (0.67)*	District dependent behavior (0.82)*	Use of different speed limits (0.73)*	Lane behavior (0.81)*
Motor vehicle	Intersection	Interior mirror	Turn left/right	District area	Urban streets	Straight ahead
Bicyclist	Lane change	Side view mirror	Lane change	Suburb area	Rural roads	In curves
Pedestrian	Entrance roundabouts	While start-up	Overtaking	Urban area	Knowing of speed limit	Turn left
Anticipation	Exit roundabouts	While stopping	While move along	Town area	Driving on speed limit	Turn right
Ready to slow	Entrance motorway	During lane change	Roundabouts	Rural area	Distance to cars urban streets	On light signals
Right-of-way	Exit Motorway	During overtaking	Entrance motorway	Motorway area	Distance to cars on rural road	Into account of bikes
	Acceleration	While move along	Exit motorway		Distance to cars on motorway	

*, Cronbach's α .

Table 3 | Expert System Traffic XPSV Schuhfried.

Tests	Cognitive paradigm	Output variables	Reliability	Literature
Reaction Test (RT)	Simple choice reaction	Reaction time decision (DS) and motor speed (MS)	DS = 0.94 MS = 0.98	Schuhfried and Prieler, 1997
Cognitrone (COG)	Selective attention	Correct response, required processing time (CIAn)	0.95	Wagner and Karner, 2001
Determination Test (DT)	Complex choice reaction	Correct reactions (CR)	0.98	Schuhfried, 1998
Peripheral Perception (PP)	Field of vision, divided attention	Angular dimension (FV), tracking deviation (DA)	FV = 0.96 DA = 0.98	Schuhfried et al., 2002
Tachistoscopic Traffic Perception Test (ATAVT)	Perceptual speed	Correct response (PS)	0.80	Sommer et al., 2008, 2009
Matrices Test (AMT)	Fluid intelligence	Correct answer (FI)	0.70	Hornke et al., 2003

stimuli by pressing one of two corresponding buttons (identical vs. different). For task completion there is no time limit. Mean reaction time of correct and incongruent responses are calculated and used as a measure of selective attention (CIAn).

The *Vienna Determination Test* (DT) is used to measure reactive stress tolerance and the related reaction speed. In principle the DT requires to discriminate colors and acoustic signals, to memorize the relevant characteristics of stimulus configurations and response buttons as well as the assignment rules. In addition it is necessary to select the relevant reactions according to the assignment rules laid down in the instructions and / or learned during the course of the test. The difficulty of the DT-Test lies in the production of continuous, sustained rapid and varied reactions to rapidly changing stimuli¹. During the 4-min test administration each subject works at the limit of his performance ability. The number of correct responses (CR) is the main variable and represents reactive stress.

The *Peripheral Perception Test* (PP) utilizes a field of vision (FV) and divided attention (DA) paradigm. Participants are sitting in front of a computer screen and perform a primary task. Beside the computer vertical panels with diodes are placed on the right and left side. The participants have to keep track of the changing diode in the periphery (secondary task) while performing the primary task. As primary task the participants have to move a cross-wire on a computer screen in order to minimize the position difference of the cross-wire with a computer-controlled moving red ball. Participants have to work simultaneously on the primary and secondary task. Every time when vertical lines appeared in the periphery they are instructed to pressing a foot pedal as fast as possible. DA is measured as the performance in the primary task (tracking deviation). FV is measured as the widest field angle at which the vertical lines of diodes are detected (during the secondary task) with respect to the distance of the screen.

The *Adaptive Tachistoscopic Traffic Perception Test* (ATAVT) is an object perception task. Photographs of traffic situations with different complexity (defined as the number of objects depicted on the photo) are presented for a short time (700–1300 ms). Participants have to decide what types of objects were presented: (1) vehicles, (2) bicycles, (3) pedestrians, (4) road signs, or (5) traffic lights. These objects are presented alone or in groups of up to five objects. The test is administered as a computerized adaptive test (CAT). The number of correctly identified objects weighted by complexity of the photograph is the dependent measure for perceptual speed (PS).

The *Adaptive Matrices Test* (AMT) is a fluid intelligence (FI) test. The stimuli are comparable to classical matrices (e.g., the Raven test). Participants have to identify the figural pattern among eight alternatives.

In previous research an overall index as a composite measure representing cognition performance has been computed on the basis of multivariate classification algorithms (artificial neural networks; NN) (Risser et al., 2008). Based on empirical evidence of Austrian and German practical on-road tests, Schuhfried GmbH categorized this composite measure into five categories.

NN measures ≥ 4 indicate an insufficient driving behavior and also indicating that the participant would fail an on-road test. NN measures ≤ 3 indicate that the participant would pass the on-road test (Risser et al., 2008; Sommer et al., 2008, 2009, 2010). The NN validly estimates the composite score as demonstrated by a good jack-knife validity coefficient of $R = 0.77$. For a better data overview in the present study the scores were changed in their direction comparable to the on-road assessment measures. Therefore, scores 1 and 2 indicate that participant fail in an on-road test. Participants with a score 3 or greater would successfully complete an on-road test. This composite variable was used as dependent variable for the evaluation of cognitive performance.

DRIVING SIMULATOR TRAINING

The goal of this training approach was to increase the mental workload of correct driving in a realistic multitasking driving setting. Therefore, complexity and difficulty were gradually increased from session to session. A training session took 40 min active driving and a short verbal feedback (feedback on reaction time, number of errors). Participants were instructed to drive with adequate speed and follow the instructions of the “simulated trainer.” The “simulated trainer” was a computer-based program: a male voice giving information about the direction of travel. These instructions were delivered according to Swiss traffic rules. The first training session included four scenarios (interurban, suburb, town, and motorway) without other vehicles in order to familiarize with the simulator. In the remaining nine training sessions six different traffic scenarios (interurban, suburb, town, motorway, overtaking and traffic rules scenario) with three different levels of difficulty were presented. Four or five scenarios were conducted in each training session (time duration of each scenario depends on the participants driving speed, no longer than 15 min for one scenario). Levels of difficulty were defined in each scenario with an increasing traffic frequency, increase of virtual drivers ignoring traffic rules (e.g., right of way rule) and an increase of hazardous traffic situations (e.g., child runs into the street). Additionally the complexity of traffic situation increased from interurban to suburban with highest complexity in the town scenario (see **Figure 2**). Furthermore, weather conditions were varied: in the third, sixth, and ninth training session it was raining infrequently or it was foggy. In training sessions four, seven, and ten participants had to drive in nighttime conditions. This training plan was fixed and participants had no possibility to adapt their subjective condition. As described above, one participant stopped due to excessive private demands.

The training progress was evaluated in four scenarios (three rural, one urban), which were not included in the training. In these scenarios, driving performance was measured using *driving errors* (accidents, traffic rule violation, leaving the lane, no use of direction indicator etc.), *top speed*, *mean speed*, *lane accuracy*, *lane variability*, and *reaction time* to unexpected stimuli (hazardous events). The driving simulator software automatically recorded the six variables for each of the four scenarios. These scenarios were conducted after the second, sixth, and tenth training session. The variables from the three rural scenarios were averaged.

¹http://www.lafayetteevaluation.com/product_detail.asp?ItemID=353

Training progress was analyzed with an ANOVA for repeated measures and showed a significant and positive effects, expect for the variable *lane accuracy* ($p > 0.05$). The detailed results are listed in **Table 4**.

ATTENTION TRAINING

The goal of this training approach was to increase specific driving-relevant cognitive functions. To prevent multitasking each scenario (intrinsic, phasic alertness, vigilance) was trained consecutively. A training session took 40 min active training and a short feedback (feedback on reaction time, number of errors and achieved level). The training regimes contained three different approaches. Each of the 10 training sessions was composed of 10 min intrinsic alertness training, followed by 10 min of

phasic alertness training (Hauke et al., 2011) and 20 min of vigilance training, conducted by a software of Schuhfried GmbH (Schuhfried, n.d.). In all training sessions participants were seated on a chair.

In the two alertness training sessions, they saw a motorcycle from a driver's viewpoint, in motion. The motorcycle drove automatically a predefined circuit in a realistic driving scene. Speed and steering was controlled automatically by the software. Participants were instructed to react as fast as possible to objects and situations, which appeared during the ride. Objects were falling trees or rocks, cars, or animals crossing the street, and traffic lights changing to red. The visualized objects only require a reaction by pushing a corresponding button if they block the road. If participants reacted more than eight times fast enough (regularly stop to prevent a crash with an object) and/or made no further errors (e.g., anticipation), the software automatically increased the level of difficulty (e.g., increasing driving speed). In case of poorer performance during a training session level decreased.

In the vigilance training a virtual driver cabin's viewpoint was presented. The car drove automatically straight ahead with constant speed. Infrequently the car was overtaken. If the brake lights of the car now being in front light up, participants had to push a corresponding button as fast as possible. If participants did not push the button after 3 s, the brake light started to flash before an error was registered. The level of difficulty was controlled automatically by the software and increased after 15 CR with a reduction in overtaking maneuvers and reduction in surrounding visual stimulation (e.g., buildings, trees).

The training progress was evaluated for each session (reached training level) and was analyzed with an ANOVA for repeated measures. Participants in this training group showed a significant and positive training progress. The detailed results are listed in **Table 5**.

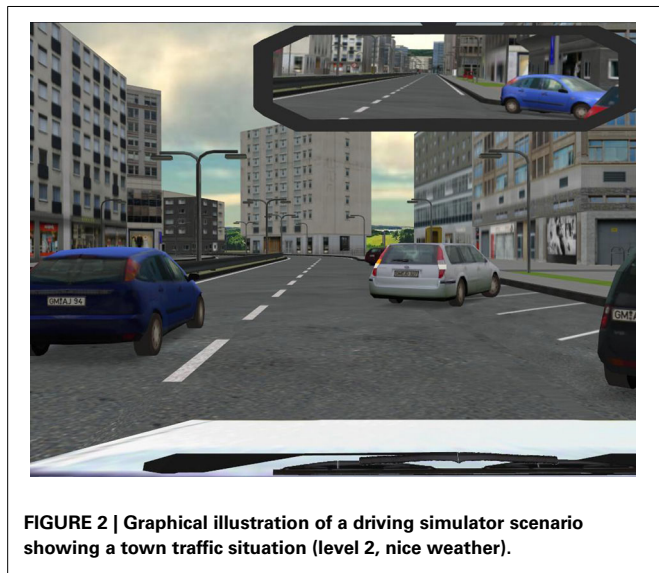


FIGURE 2 | Graphical illustration of a driving simulator scenario showing a town traffic situation (level 2, nice weather).

Table 4 | Training progress: simulator training.

Variable	Second training session ($N = 31$)	Sixth training session ($N = 31$)	Tenth training session ($N = 31$)		
	M (SD)	M (SD)	M (SD)	F	p
RURAL SCENARIOS					
Driving error	2.29 (1.47)	4.48 (2.79)	1.22 (1.41)	13.20	<0.01
Top speed	62.02 (5.82)	60.55 (5.30)	64.91 (4.87)	7.08	0.012
Mean speed	40.30 (3.88)	39.89 (3.44)	42.95 (3.34)	19.96	<0.01
Lane accuracy (%)	77.05 (4.50)	77.38 (4.71)	75.96 (4.70)	2.71	<i>n. s.</i>
Lane variability (%)	13.89 (1.47)	13.52 (1.16)	13.59 (1.14)	1.44	<i>n. s.</i>
Reaction time (s)	1.35 (0.13)	1.20 (0.11)	1.15 (0.14)	60.54	<0.01
URBAN SCENARIO					
Driving error	4.48 (2.79)	3.32 (2.31)	2.51 (1.99)	14.01	<0.01
Top speed	35.95 (4.69)	35.88 (5.03)	36.21 (4.62)	0.07	<i>n. s.</i>
Mean speed	16.33 (2.07)	17.07 (1.83)	17.80 (2.15)	16.99	<0.01
Lane accuracy (%)	65.42 (4.01)	65.95 (3.24)	65.65 (3.91)	0.69	<i>n. s.</i>
Lane variability (%)	25.58 (1.40)	24.88 (1.53)	24.80 (1.65)	7.73	<0.01
Reaction time (s)	1.11 (0.19)	1.15 (0.24)	1.05 (0.17)	1.77	<i>n. s.</i>

n. s., not significant; *s*, seconds.

Table 5 | Training progress: cognitive training.

Variable	Friedman χ^2 test (Comparison of levels from all training session)		
	N	χ^2	Asymp. Sig.
Intrinsic alertness	23	145.99	<0.01
Phasic alertness	23	157.05	<0.01
vigilance	23	185.95	<0.01

PROCEDURE

The general study design is a pre-post design. During the pre- and post-test sessions all participants conducted cognitive and on-road tests. Between the pre-test and post-test measurements the participants performed either the training regimes (driving simulator or attention) or simply waited to participate for the post-test (control group). Data acquisition took 25 months (May 2010–June 2012). The 91 participants were assigned to one of 13 training blocks. During every block seven participants took part (three for the simulator training, two for the attention training, and two for the control group). During the entire study, every participant took part in a single setting. Block duration was 7 weeks with two dates per week (two dates pre-tests; 10 training sessions; two dates post-tests; in total: 14 dates). In the first week (pre-test) and last week (post-test) on-road driving performance and cognitive performance was measured. Furthermore, all participants underwent electroencephalography (EEG) recordings during a set of three inhibition tasks (stroop, negative priming, and flanker). These data will be presented elsewhere (in preparation).

Before each computer test and on-road drive participants received an introduction about the test process and conditions (for tests: written instruction by software; for on-road drive: verbal instruction by the DI), but no feedback about their performance. Before all computerized tests (pre-test phase) the corresponding software automatically measured reaction time, correct and wrong answers to evaluate the participants' understanding of the particular test. In this pre-test phase participants were allowed to ask questions about the cognitive tests or in case of any other problems. During the first appointment participants conducted initially the cognitive test battery and thereafter the on-road drive (each lasting 1 h). On the second appointment (not included in this article), inhibition tasks (Stroop, negative Priming, Flanker) and EEG recording was conducted. From the second to the sixth week, both training groups participated in 10 training sessions with two training sessions a week. During this time period, the control group received no intervention. During post-test week on the second but last date, inhibition tasks and EEG recordings were conducted (not included in this article). On the last appointment, participants conducted again the cognitive test battery and thereafter the on-road drive (each lasting 1 h).

To control for mood and motivational changes during the training, participants completed in the first, fifth, and tenth training session an adapted version of the SAM (Self Assessment Manikin) for mood changes (Beeli et al., 2008) and an

adapted version of the FAM (Fragebogen zur Erfassung aktueller Motivation) for motivational changes (Rheinberg et al., 2001). In the driving simulator training group, SS was measured by calculating the mean of the three main subjective symptoms: nausea (N), oculomotor (O), and disorientation (D) (Kennedy et al., 1993) at begin, middle and end of training. Each symptom was scored from 1 to 5 (low SS = 1, severe SS = 3, strong SS = 5).

STATISTICAL ANALYSIS

Statistics were calculated using SPSS 18 for Windows 7 (SPSS Inc., Chicago, IL), with a significance level of $\alpha = 0.05$. Differences in baseline performance and in demographic data between the groups were compared using Kruskal-Wallis and ANOVA tests.

With hierarchical multiple regression analysis with planned group comparisons the training benefits (dependent variables pre- and post-test) were analyzed for the cognitive and on-road performances. For the planned group comparisons orthogonal contrast coding was used. Contrast coding was used in accordance with the hypotheses formulated in the introduction.

We defined a-priori (planned) contrasts allowing us to test interaction effects (Pedhazur, 1982), which are of utmost importance to test our hypothesis formulated in the introduction. First we designed interaction contrasts allowing us to test pre-post differences between both training groups (attention and driving simulator training) vs. the control group. The second contrast was designed as orthogonal to the first contrast allowing us to test for pre-post differences between both training groups. Since we adopted orthogonal contrasts we only can use two contrasts (pre- and post-measures: $df = 1$; number of groups: $df = 2$).

The advantage of this contrast design is that we gain more statistical power to detect even moderately strong effects without increasing sample size too much. In addition, this kind of a priori defined testing is strongly hypothesis-driven. Since we anticipate that training results in improvement we decided to test uni-directionally. According to our hypothesis we are not interested to compare the two training groups separately with the control group since we are not interested in potential differences to the control group. We are mainly interested in differences between the training groups. We also focus statistical testing on the composite measure for on-road driving and cognitive performance. For the sub-measures of which the composite scores are calculated we only report the results on a descriptive basis.

Because of the relative small number of subjects and large number of dependent variables, which we can possibly be used for statistical testing, it is nearly impossible to perform classical statistical inference tests. The reason is the small power even when moderate or even strong effects are present. Thus, when applying corrections for multiple testing, no or only a few of strong effects would have been identified. Because of this, we decided to use a more descriptive statistical approach for most of the variables. For a subset of variables, we performed a strongly hypothesis-driven statistical analysis (the composite scores for on-road performance and cognitive test performance). For these tests we draw stronger conclusions from the analyses. For hypothesis-free analyses (the sub-measures constituting the composite scores), the statistical test results are not interpreted in terms of statistical significance, they are rather used as descriptive measures of between-group

differences. For these analyses, we will be more reluctant in interpreting the findings. The p -values for these comparisons can be taken as measures of effect (Krauth, 1988). Since we have to consider the fact that p -values depend on sample size, we also calculate effect sizes according to Cohen (1988). A $d > 0.3$ and <0.5 is considered as small, a $d > 0.5$ and <0.8 as moderate, while a $d > 0.8$ is considered as large.

RESULTS

DEMOGRAPHIC AND BASELINE GROUP COMPARISON

There were no differences between the groups with regard to daily driving activities, relevant demographic variables, or gender (all $p > 0.05$). Baseline comparisons showed a significant group difference in crystalline intelligence, decision time in the simple reaction task (RT) and district dependent behavior in the on-road assessment (Table 6). These variables did not correlate (as computed with Pearson correlations) with the composite scores for on-road and cognitive performance (all p -values at least < 0.10). In all other measures there were no significant baseline differences (all $p > 0.05$). Furthermore, no baseline differences exist for the overall on-road and cognitive performance.

SIMULATOR SICKNESS, EMOTIONAL, AND MOTIVATIONAL STATUS

Participants in the simulator training reported SS, which significantly changed during the training ($X^2 = 30.98$, $p < 0.001$). Wilcoxon tests revealed a drop in subjectively experienced SS

from the start of the driving simulator session ($median = 2.17$) to half time ($median = 1.43$, $z = -4.21$, $p < 0.001$) and at the end ($median = 1.38$, $z = -3.83$, $p < 0.001$) of the training. Therefore, no SS differences occurred between half time and end of training. Training groups differed in their emotional valence during training participation [$F_{(1, 52)} = 4.56$, $p = 0.038$]. *Post-hoc t-test* showed on average a lower positive valence in the driving simulator training group ($M = 3.35$, $SE = 0.16$) than in the attention training group at beginning ($M = 4.43$, $SE = 0.16$), [$t_{(52)} = -4.67$, $p < 0.01$]. No significant differences existed to half time and at the end of training. Furthermore, there was no group difference in emotional arousal and motivation (all $p > 0.05$).

ON-ROAD TRAINING EFFECT

Descriptive statistics from on-road performance are displayed in Table 7 including Cohen's d for the pre-post differences broken down for the three groups.

Training gains and the regression model on overall on-road performance are displayed in Figure 3 and Table 8. Significantly different training gains for the different groups are displayed in Table 8. Compared to the control group, there was no significant change in the overall on-road performance [$F_{(1, 74)} = 1.59$, $p = 0.11$, $d = 0.35$] as a result of the training, but a significant improvement in the simulator training group compared to the attention training group [$F_{(1, 74)} = 2.86$, $p < 0.05$, $d = 0.48$].

Table 6 | Baseline differences between groups.

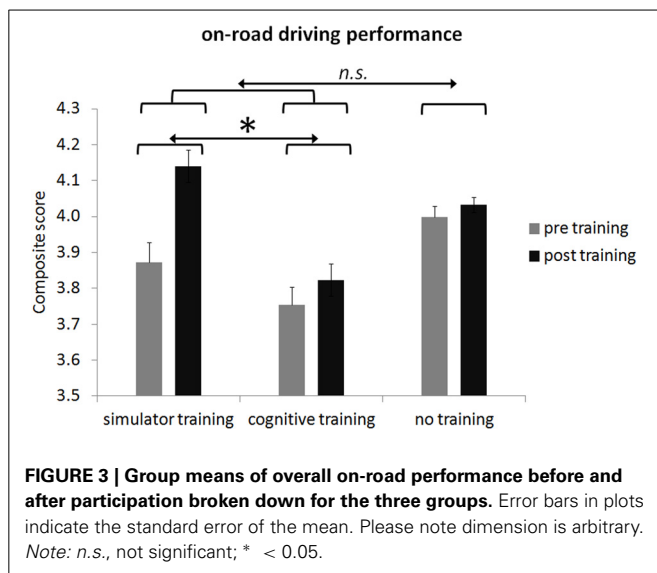
Variable	Simulator training group ($N = 31$)	Cognitive training group ($N = 23$)	Control group ($N = 23$)	p
	M (SD)	M (SD)	M (SD)	
Crystallized IQ	125.6 (11.3)	128.9 (10.3)	118.6 (11.9)	0.009
Simple choice reaction (DS), ms	467.3 (81.5)	475.3 (90.7)	545.3 (91.2)	0.004
District dependent behavior	3.96 (0.78)	3.79 (0.77)	4.29 (0.47)	0.047

ms, milliseconds.

Table 7 | Descriptive statistics for on-road measures.

Variable		Simulator training group (<i>N</i> = 31)		<i>d</i>	Cognitive training group (<i>N</i> = 23)		<i>d</i>	Control group (<i>N</i> = 23)		<i>d</i>
		Pre-test	Post-test		Pre-test	Post-test		Pre-test	Post-test	
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	
On-road driving assessment	Overall driving performance	3.87 (0.63)	4.14 (0.51)	0.56	3.75 (0.57)	3.82 (0.52)	0.15	3.99 (0.36)	4.01 (0.24)	0.08
	Traffic observation	3.85 (0.75)	4.04 (0.77)	0.27	3.64 (0.71)	3.72 (0.73)	0.12	3.94 (0.53)	3.99 (0.61)	0.09
	Speed at intersection	3.95 (0.69)	4.38 (0.68)	0.66	4.08 (0.63)	3.96 (0.79)	0.17	4.29 (0.71)	4.03 (0.69)	0.40
	Gaze behavior	3.71 (0.79)	3.91 (0.61)	0.27	3.60 (0.65)	3.57(0.63)	0.04	3.75 (0.61)	3.82 (0.51)	0.12
	Change of direction	3.95 (0.69)	4.38 (0.68)	0.55	4.08 (0.64)	3.95 (0.79)	0.16	4.29 (0.71)	4.03 (0.69)	0.33
	District dependent behavior	3.96 (0.78)	4.39 (0.67)	0.63	3.79 (0.77)	4.06 (0.64)	0.41	4.29 (0.47)	4.25 (0.43)	0.10
	Use of different speed limits	4.06 (0.79)	4.47 (0.54)	0.65	3.98 (0.60)	4.11 (0.49)	0.25	4.22 (0.47)	4.29 (0.56)	0.14
	Lane behavior	3.48 (0.83)	3.40 (0.78)	0.10	3.17 (0.80)	3.27 (0.50)	0.16	3.15 (0.57)	3.50 (0.48)	0.68

d , Cohen's d effect size (Cohen, 1988) using the pooled SD for both conditions and correcting for dependence between means according to Morris and DeShon (2002).



Please note that the following comparisons are only done on a descriptive basis to prevent an inflation of statistical tests. Increased performance for both training groups compared to the control group were found for the following sub-measures: *change of direction* [$t_{(74)} = 2.24$, $p < 0.05$, 1-tailed, $d = 0.56$], *district dependent behavior* compared to the control group [$t_{(74)} = 2.62$, $p < 0.05$, 1-tailed, $d = 0.68$]. Significantly better performance yielded for the simulator training group compared to the attention training group for the variable *change of direction* [$t_{(74)} = 2.68$, $p < 0.01$, 1-tailed, $d = 0.79$]. For *lane behavior* there was an increase in this measure for the control group compared to both training groups [$t_{(74)} = -1.96$, $p < 0.05$, 1-tailed, $d = 0.54$].

COGNITIVE TRAINING EFFECT

Descriptive statistics from cognitive performance are displayed in **Table 9** including Cohen's d for the pre-post differences broken down for the three groups.

Training gains and the regression model on overall cognitive performance are displayed in **Figure 4** and **Table 10**. Significantly different training gains for the different groups are displayed in **Table 10**. Compared to the control group, there was a significant improvement in the *overall cognitive performance* for both training groups [$F_{(1, 74)} = 8.99$, $p < 0.01$, $d = 0.48$] but no significant improvement in the simulator training group compared to the attention training group [$F_{(1, 74)} = 0.36$, $p = 0.55$, $d = 0.22$].

For the sub-measures of cognitive performance we revealed improved performance in several measures. As explained in the method section the following comparisons are only used on a descriptive basis to prevent an inflation of statistical tests. Increased performance for both training groups compared to the control group were found for the following sub-measures: *motor speed* compared to the control group [$t_{(74)} = -1.98$, $p < 0.05$, 1-tailed, $d = 0.48$]. Significantly better performance yielded for the attention training group compared to the simulator training group for the variable *decision speed* [$t_{(74)} = -1.81$, $p < 0.05$, 1-tailed, $d = 0.56$].

Table 8 | Multiple regression for the interaction between orthogonal contrasts and training gain for the composite score of the on-road performance.

Variable	B	SE	β
Overall driving performance			
Linear interaction AB \times C	0.022	0.018	0.061
Linear interaction A \times B	0.050	0.029	0.082*

A, driving simulator group; B, cognitive training group; C, control group. AB \times C, comparison of the average of the training effect for group A and B vs. the training effect for group C.

A \times B, comparison of the training effect for group A vs. the training effect for group B.

* < 0.05.

DISCUSSION

The main goal of this study was to investigate whether on-road driving in older healthy active drivers can be improved by two different training approaches: training with a driving simulator or training of cognitive functions known to be involved in controlling driving. Based on the current literature we hypothesized that both training approaches would increase on-road driving performance as well as cognitive performance. Secondly, we hypothesized that driving simulator training, in which scenarios are comparable to on-road driving, would induce stronger on-road driving performance gains than attention training because this training requires multitasking. We also anticipated that cognitive performance growth would strongly benefit from the simulator training since this training also induces lots of cognitive functions (e.g., attention, spatial perception, sensory-motor coordination and tracking, executive functions, vision, working memory etc.; Lees et al., 2010; Romoser, 2013; Casutt et al., 2014) and is a kind of multitasking training for which a recent paper has shown strong beneficial effects on cognitive performance and its underlying brain functions especially for older adults (Anguera et al., 2013). It is worth noting that we have used driving simulator training employing different naturalistic virtual realities with increasing complexity and difficulty. In addition, a well-established cognitive training software was used consisting of three consecutively conducted cognitive training approaches (Hauke et al., 2011).

Participants in the driving simulator training group improved their on-road driving performance compared to the attention training group. Cognitive performance, however, improved in both training groups (driving simulator and attention training groups) in comparison to the control group. Thus, the driving simulator group showed improvements in on-road driving performance as well as cognition (near and far transfer) while the attention training group only showed improvement in cognition (near transfer). Thus, the driving simulator training (as an example for a complex training) obviously induces near and far transfer and exerts stronger training gains what was suggested in previous influential reviews (Lustig et al., 2009; Zelinski, 2009).

But what are the reasons for the different learning effects on on-road performance and cognitive functions? Highly interactive and complex cognitive training approaches (not only using

driving simulators but interactive video games) have been shown to exert positive influences on cognitive functions, behavior, emotion, and lots of other actions and functions (Green and Bavelier, 2006a,b; Achtman et al., 2008; Basak et al., 2008; Karbach and Kray, 2009; Marmeleira et al., 2009). In this context it has been argued that virtual environments and scenarios are inherently attractive and motivating (Lees et al., 2010). Many subjects feel a kind of presence when interacting with highly immersive video scenarios especially when they interact with or within the virtual environment (Havranek et al., 2012). These are circumstances enhancing attractiveness of these scenarios, which most likely also enhance motivation and attention, both factors, which are pivotal for learning and memory consolidation (Green and Bavelier, 2006a,b; Green et al., 2010; Bavelier et al., 2012). Therefore, it is most likely that attention and motivation to learn is stronger for those subjects who practice with a driving simulator than for those subjects who only practice more or less abstract cognitive functions. However, the rated motivation in our study did not differ between both training groups. Thus, it might be that the questionnaires measuring subjective motivation and arousal are not sensitive enough to capture fine graded motivation and arousal differences. It is well-known that subjective and physiological measures of motivation and arousal only weakly covariate (Erdmann and Janke, 1978). Thus, it is possible that our subjects participating in the driving simulator group were indeed stronger motivated or aroused (with the accompanying physiological changes) but without noticing it. In addition, it is also possible that all subjects were motivated or aroused to a quite high degree, which cannot be captured due to ceiling effects. Secondly, the driving simulator provides traffic scenarios, which are quite close and partly similar to real traffic situations. Thus, the subjects practicing with the driving simulator, train something, which they directly can use in real situations. Thus, the conceptual and practical “distance” of the learned aspects from a driving simulator context to an on-road driving situation is closer (near transfer) than the “distance” from attention training to on-road driving (far transfer). Similar beneficial effects from simulator training

(even when the used simulators are simple) to real life actions have been demonstrated quite frequently for controlling specific problematic driving skills in older drivers (Romoser and Fisher, 2009; Lavallière et al., 2011, 2012; Romoser, 2013; Romoser et al., 2013). Even when the attention training is designed to be a bit more realistic and dynamic (e.g., visual search strategies at intersection) not only cognitive functions like DA, monitoring, and decision making improve but also on-road driving performance (Romoser and Fisher, 2009). Thus, the realistic and dynamic aspects of driving simulator training are most likely important factors enhancing learning and more importantly enhancing the improvement of on-road driving.

A further aspect of the driving simulator training might enhance improving on-road performance and cognition. Driving simulator training as we have used it in our study is very similar to multitasking training. During simulator driving the

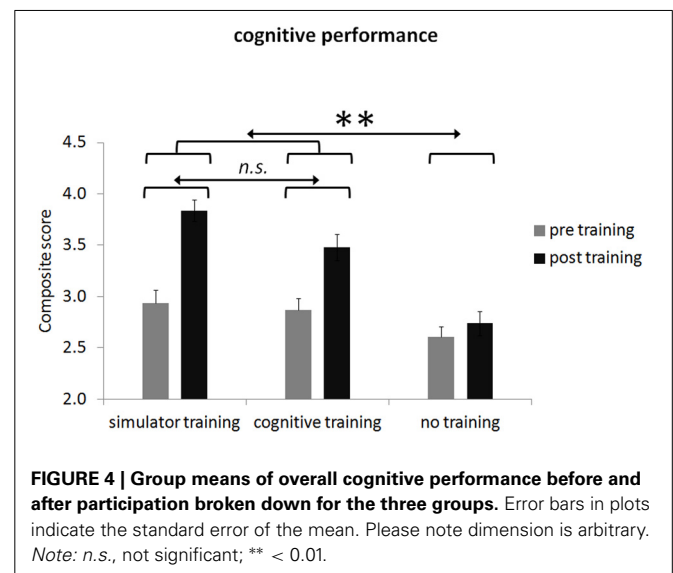


Table 9 | Descriptive statistics for cognitive measures.

Variable	Simulator training group (N = 31)			Cognitive training group (N = 23)			Control group (N = 23)		
	Pre-test		d	Pre-test		d	Pre-test		d
	M (SD)	M (SD)		M (SD)	M (SD)		M (SD)	M (SD)	
Overall cognitive performance	2.94 (1.46)	3.84 (1.15)	0.66	2.87 (1.32)	3.48 (1.44)	0.42	2.61 (1.18)	2.74 (1.32)	0.10
Simple choice reaction (DS), ms ^a	467.3 (81.4)	455.5 (90.9)	0.18	475.2 (90.7)	423.3 (68.0)	0.84	545.3 (91.2)	529.4 (100.6)	0.21
Simple choice reaction (MS), ms ^a	212.7 (65.8)	186.2 (72.1)	0.42	227.0 (82.2)	201.9 (66.9)	0.37	215.3 (78.5)	222.8 (85.8)	0.10
Selective attention (CIAn), s ^a	3.04 (0.64)	3.11 (0.66)	0.13	3.22 (0.76)	3.10 (0.48)	0.24	3.04 (0.51)	2.95 (0.46)	0.23
Complex choice reaction (CR)	187.3 (26.1)	203.1 (29.7)	1.01	183.1 (24.7)	198.3 (27.9)	1.02	181.9 (29.4)	195.0 (33.9)	0.73
Field of vision (FV)	154.7 (18.4)	157.3 (18.0)	0.21	152.8 (21.0)	153.2 (22.1)	0.03	152.3 (14.7)	155.3 (14.6)	0.30
Divided attention (DA) ^a	13.2 (4.1)	12.1 (2.7)	0.31	13.0 (2.8)	13.6 (4.6)	0.16	13.1 (2.8)	13.8 (3.9)	0.20
Perceptual speed (PS)	-0.31 (0.84)	-0.03 (1.06)	0.30	-0.61 (0.88)	-0.26 (0.88)	0.40	-0.48 (0.96)	-0.15 (0.78)	0.38

^aSmaller scores reflect better performance. RT, reaction time; s, seconds; ms, milliseconds; d, Cohen's d effect size (Cohen, 1988) using the pooled SD for both conditions and correcting for dependence between means according to Morris and DeShon (2002).

Table 10 | Multiple regression for the interaction between orthogonal contrasts and training gain for the composite score of the cognitive performance.

Variable	<i>B</i>	<i>SE</i>	β
Overall cognitive performance (NN)			
Linear interaction AB \times C	0.104	0.058	0.106**
Linear interaction A \times B	0.074	0.096	0.045

A, driving simulator group; *B*, cognitive training group; *C*, control group. *AB \times C*, comparison of the average of the training effect for group *A* and *B* vs. the training effect for group *C*.

A \times B, comparison of the training effect for group *A* vs. the training effect for group *B*. ** < 0.01.

trainees have to orchestrate different psychological functions either simultaneously or in an elegant and efficient way sequentially. This kind of orchestration of several and different psychological functions is pivotal for efficiently driving a car. While driving in a driving simulator (and in a real car), the subjects have to control their car (sensorimotor control), scan the scenarios (perception), remember similar situations (memory), and anticipate as well as plan the maneuvers (cognition). Thus, this training has much in common with interactive cognitive multitasking (Basak et al., 2008; Marmeleira et al., 2009; Anguera et al., 2013). Moreover a recent publication showed that multitasking training increases not only performance in different cognitive domains (working memory, attention) but also induces changes in brain activity (Anguera et al., 2013). The authors interpreted their results of brain plasticity as an increased suppression of the default network during task engagement. In line with this evidence our results support the multitasking approach and its brain plasticity in the older adult brain and its positive transfer in cognition and on-road driving.

Additionally, in driving simulator studies it was shown that the level of multitasking costs is associated with driving uncertainty and driving errors (Bélanger et al., 2010) and that the multitasking costs in older drivers are greater than in younger drivers (Cantin et al., 2009). The multitasking nature of the driving simulator training is supported by the improved DA performance for the simulator group. DA is known to be a cognitive function relying on the complex interplay between different brain structures and is also a kind of multitasking. Cognitive training regimes during which one psychological function is practiced more or less isolated without switching lacks this dynamic interaction between different psychological functions (Zelinski, 2009).

Having a closer look at the improved aspects of on-road driving it becomes evident that they correspond to those traffic situations (behavior at crossroads, junctions, and lane change), which are discussed in the literature as typical problematic driving situations causing reduced driving safety and increased driving errors (Braitman et al., 2007; Romoser and Fisher, 2009; Lavallière et al., 2011). Romoser and Fisher (2009) showed that active simulator training improves older driver visual scanning strategies at intersections, which also were observed after a follow-up of 2 years (Romoser, 2013). Furthermore, these problematic driving behaviors are related to declines in executive functions, for

example in decision making (Daigneault et al., 2002; Horswill et al., 2008; Romoser and Fisher, 2009). According to these results, the present study complements the existing research. Interactive and multitask simulator training increases higher order cognitive functions and everyday life abilities in older adults.

LIMITATIONS

First of all it should be kept in mind that SS is still a problem at least for some subjects practicing with the driving simulator. However, reported average sickness diminished during the simulator training and even disappeared entirely for most of the subjects. Thus, between-groups differences with respect to these variables could not account for the improvement in on-road driving and cognitive performance. However, some subjects were excluded from the study when the sickness symptoms did not disappear or attenuate to a strong degree. Although only seven subjects were excluded because of SS, SS might have influenced the present results in several ways. For example, we measured those subjects who could cope with the sickness symptoms. Thus, their training performance might be linked somehow with this coping and struggling. Maybe they employ more self-control and/or self-discipline during training than those who didn't experience these obstacles.

There are also some baseline differences between the groups with respect to the driving performance and the performance in the cognitive tests, which are difficult to explain (e.g., district dependent behavior, or reaction times in some cognitive tests). However, since these baseline differences have only been identified in two measures and did not influence the overall on-road performance and overall cognitive performance we are sure that these differences do not influence training performance.

It should be noted that quality of lane behavior (a sub-variable contributing to on-road driving performance) did not improve as a consequence of the driving simulator training while the control group improved their performance with respect to this measure. This partly paradoxical finding is difficult to explain and we would like to refrain from making too strong and speculative arguments in this case. One tentative explanation could be that lane accuracy or its deviation is not a sensitive measure. In another simulator study comparing young and old drivers, there was no significant between-group difference with respect to this measure (Cantin et al., 2009). Further research is thus needed to study the moderating influences on this variable.

One important limitation in the present study is the absence of a further active control group to control for simple activity (even being unrelated to driving). Since this experiment was extremely demanding for the participating subjects (e.g., they had to travel to the psychological institute several times to practice the cognitive tasks or the driving simulator) it would have caused additional organizational effort to hire additional subjects for our active control group. In addition, it is borderline unethical to let a group of older adults practice something, which is unrelated to on-road driving and from which we anticipate no direct or indirect influence on on-road driving. We are thus sure that the local ethics committee never would have approved a control group like this. However, we used both experimental groups as control groups for the other group. Thus, the attention

training group acted as control group for the simulator group and vice versa.

Training intensity and duration are also issues, which will have substantial impacts on training results, either for the attention or the driving simulator training. The training intensity and frequency used in our study might be too low to induce strong training gains. Thus, it would be interesting to study whether increased training durations and frequencies will result in stronger improvements in cognition and driving behavior.

Additionally, a critical point of our study is the specific sample of older adults. All subjects (irrespective to which group they have been assigned) were highly interested to participate and most of them were active drivers using their car frequently. For example, the average mileage in Switzerland for this age group is 3200 km (Bundesamt für Statistik und Bundesamt für Raumentwicklung, 2012). The mean mileage in the participating subjects varied between 8973 and 11,909 km. Whether subjects who are closer to the average mileage would benefit differently from the driving simulator or cognitive training has to be shown in different experiments. That the particular sample of older adults has an influence on driving improvement has been shown in a study of Roenker et al. (2003). They uncovered a positive influence of a cognitive training on on-road driving in high-risk older adults, which are deemed to perform suboptimal in real driving situations.

A final critical point is that during cognitive training not only cognition is improved but also other functions (e.g., perception). However, the exact nature of the relation between perception and cognition is currently unknown and has to be elucidated in future studies. (for a similar conjecture see Anstey et al., 2003). Thus, we are not in the position to delineate whether the subjects of our experimental group demonstrate improved sensory and perceptual functions as a consequence of our training approaches. However, we can state that cognitive functions are altered due to our training.

CONCLUSION

In this study we directly compared the influence of attention training and simulator training on on-road performance and cognition. Here we showed that only participants practicing to drive in different traffic scenarios using a driving simulator significantly improved their on-road driving performance compared to a group involved in attention training. In addition, both training groups (the driving simulator and the attention training group) showed improved cognitive performance compared to a control group. Thus, the present study shows that driving simulators are useful training tools to improve on-road performance as well as cognition in older adults. Although this study supports the beneficial role of driving simulators to improve on-road driving (and cognition) further studies have to be conducted disentangling the different cognitive processes benefiting most from driving simulator training. In addition, it has to be shown how the measured on-road performance relates to those traffic measures, which are most important for real traffic like traffic safety or crash numbers. It will also be interesting how different samples of older drivers (e.g., at-risk drivers with mild or advanced cognitive problems) will benefit from driving simulator and/or attention training.

AUTHOR CONTRIBUTIONS

Gianclaudio Casutt: Study conception and preparation, acquisition of data, statistical analysis, interpretation of data, drafting manuscript. Nathan Theill: Statistical analysis, interpretation of data, revising manuscript. Mike Martin: Revising manuscript. Martin Keller: Study conception. Lutz Jäncke: Supervision of study conception, statistical analysis, interpretation of data, revising manuscript.

ACKNOWLEDGMENTS

This research project was supported by funds of the “Forschungskredit” of the University of Zurich. We thank Dr. Jacqueline Zoellig for comments and suggestions provided through the planning phase of this project.

REFERENCES

- Achtman, R. L., Green, C. S., and Bavelier, D. (2008). Video games as a tool to train visual skills. *Restor. Neurol. Neurosci.* 26, 435–446.
- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature* 501, 97–101. doi: 10.1038/nature12486
- Anstey, K. J., Hofer, S. M., and Luszcz, M. A. (2003). A latent growth curve analysis of late-life sensory and cognitive function over 8 years: evidence for specific and common factors underlying change. *Psychol. Aging* 18, 714–726. doi: 10.1037/0882-7974.18.4.714
- Anstey, K. J., and Wood, J. (2011). Chronological age and age-related cognitive deficits are associated with an increase in multiple types of driving errors in late life. *Neuropsychology* 25, 613–621. doi: 10.1037/a0023835
- Anstey, K. J., Wood, J., Lord, S., and Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clin. Psychol. Rev.* 25, 45–65. doi: 10.1016/j.cpr.2004.07.008
- Baldock, M. R., Thompson, J. P., and Mathias, J. L. (2008). “Self-regulation of driving behaviour among older drivers: findings from a five year follow-up,” in *Proceedings of the 2008 Australasian Road Safety Research, Policing and Education Conference* (Adelaide).
- Ball, K., Edwards, J. D., Ross, L. A., and McGwin, G., Jr. (2011). Cognitive training decreases motor vehicle collision involvement of older drivers. *J. Am. Geriatr. Soc.* 58, 2107–2113. doi: 10.1111/j.1532-5415.2010.03138.x
- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Bavelier, D., Green, C. S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annu. Rev. Neurosci.* 35, 391–416. doi: 10.1146/annurev-neuro-060909-152832
- Beeli, G., Casutt, G., Baumgartner, T., and Jäncke, L. (2008). Modulating presence and impulsiveness by external stimulation of the brain. *Behav. Brain Func.* 4, 33. doi: 10.1186/1744-9081-4-33
- Bélanger, A., Gagnon, S., and Yamin, S. (2010). Capturing the serial nature of older drivers’ responses towards challenging events: a simulator study. *Accid. Anal. Prev.* 42, 809–817. doi: 10.1016/j.aap.2009.07.010
- Braitman, K. A., Kirley, B. B., Ferguson, S., and Chaudhary, N. K. (2007). Factors leading to older drivers’ intersection crashes. *Traffic Inj. Prev.* 8, 267–274. doi: 10.1080/15389580701272346
- Bundesamt für Statistik und Bundesamt für Raumentwicklung (2012). *Mobilität in der Schweiz, Ergebnisse des Mikrozensus Mobilität und Verkehr 2010*. Neuchâtel: Office fédéral de la statistique.
- Cantin, V., Lavalliere, M., Simoneau, M., and Teasdale, N. (2009). Mental workload when driving in a simulator: effects of age and driving complexity. *Accid. Anal. Prev.* 41, 763–771. doi: 10.1016/j.aap.2009.03.019
- Cassavaugh, N. D., and Kramer, A. F. (2009). Transfer of computer-based training to simulated driving in older adults. *Appl. Ergon.* 40, 943–952. doi: 10.1016/j.apergo.2009.02.001
- Casutt, G., Martin, M., and Jäncke, L. (2013). Alterseffekte auf die Fahrsicherheit bei Schweizer Kraftfahrern im Jahr 2010. *Zeitschrift für Verkehrssicherheit* 59, 162–169.

- Casutt, G., Martin, M., Keller, M., and Jäncke, L. (2014). The relation between performance in on-road driving, cognitive screening and driving simulator in older healthy drivers. *Transp. Res. Part F Traffic Psychol. Behav.* 22, 232–244. doi: 10.1016/j.trf.2013.12.007
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Daigneault, G., Joly, P., and Frigon, J. Y. (2002). Executive functions in the evaluation of accident risk of older drivers. *J. Clin. Exp. Neuropsychol.* 24, 221–238. doi: 10.1076/j.jcen.24.2.221.993
- Edwards, J. D., Lunsman, M., Perkins, M., Rebok, G. W., and Roth, D. L. (2009a). Driving cessation and health trajectories in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* 64A, 1290–1295. doi: 10.1093/gerona/glp114
- Edwards, J. D., Myers, C., Ross, L. A., Roenker, D. L., Cissell, G. M., McLaughlin, A. M., et al. (2009b). The longitudinal impact of cognitive speed of processing training on driving mobility. *Gerontologist* 49, 485–494. doi: 10.1093/geront/gnp042
- Erdmann, G., and Janke, W. (1978). Interaction between physiological and cognitive determinants of emotions: experimental studies on Schachter's theory of emotions. *Biol. Psychol.* 6, 61–74. doi: 10.1016/0301-0511(78)90007-8
- Green, C. S., and Bavelier, D. (2006a). Effect of action video games on the spatial distribution of visuospatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1465–1478. doi: 10.1037/0096-1523.32.6.1465
- Green, C. S., and Bavelier, D. (2006b). Enumeration versus multiple object tracking: the case of action video game players. *Cognition* 101, 217–245. doi: 10.1016/j.cognition.2005.10.004
- Green, C. S., Pouget, A., and Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Curr. Biol.* 20, 1573–1579. doi: 10.1016/j.cub.2010.07.040
- Hakamies-Blomqvist, L. (1994). Compensation in older drivers as reflected in their fatal accidents. *Accid. Anal. Prev.* 26, 107–112. doi: 10.1016/0001-4575(94)90073-6
- Hauke, J., Fimm, B., and Sturm, W. (2011). Efficacy of alertness training in a case of brainstem encephalitis: clinical and theoretical implications. *Neuropsychol. Rehabil.* 21, 164–182. doi: 10.1080/09602011.2010.541792
- Havranek, M., Langer, N., Cheetham, M., and Jancke, L. (2012). Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behav. Brain Func.* 8, 34. doi: 10.1186/1744-9081-8-34
- Holm, A., Lukander, K., Korpela, J., Sallinen, M., and Muller, K. M. (2009). Estimating brain load from the EEG. *Sci. World J.* 9, 639–651. doi: 10.1100/tsw.2009.83
- Hornke, L. F., Etzel, S., and Rettig, K. (2003). *Manual Adaptive Matrices Test (AMT)*. Mödling: SCHUHFRIED GmbH.
- Horswill, M. S., Kemala, C. N., Wetton, M., Scialfa, C. T., and Pachana, N. A. (2010). Improving older drivers' hazard perception ability. *Psychol. Aging* 25, 464–469. doi: 10.1037/a0017306
- Horswill, M. S., Marrington, S. A., McCullough, C. M., Wood, J., Pachana, N. A., McWilliam, J. et al. (2008). The hazard perception ability of older drivers. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 63, P212–P218. doi: 10.1093/geronb/63.4.P212
- Jäncke, L., and Klimmt, C. (2011). Expertise in video gaming and driving skills. *Zeitschrift für Neuropsychologie* 22, 279–284. doi: 10.1024/1016-264X/a000052
- Karbach, J., and Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Dev. Sci.* 12, 978–990. doi: 10.1111/j.1467-7687.2009.00846.x
- Kennedy, R., Lane, N., Berbaum, K., and Lilienthal, M. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi: 10.1207/s15327108ijap0303_3
- Krauth, J. (1988). *Distribution-Free Statistics. An Application-Oriented Approach*. Amsterdam; New York; Oxford: Elsevier.
- Langford, J., Fitzharris, M., Newstead, S., and Koppel, S. (2004). Some consequences of different older driver licensing procedures in Australia. *Accid. Anal. Prev.* 36, 993–1001. doi: 10.1016/j.aap.2003.11.003
- Lavallière, M., Laurendeau, D., Simoneau, M., and Teasdale, N. (2011). Changing lanes in a simulator: effects of aging on the control of the vehicle and visual inspection of mirrors and blind spot. *Traffic Inj. Prev.* 12, 191–200. doi: 10.1080/15389588.2010.548426
- Lavallière, M., Simoneau, M., Tremblay, M., Laurendeau, D., and Teasdale, N. (2012). Active training and driving-specific feedback improve older drivers' visual search prior to lane changes. *BMC Geriatr.* 12:5. doi: 10.1186/1471-2318-12-5
- Lees, M. N., Cosman, J. D., Lee, J. D., Fricke, N., and Rizzo, M. (2010). Translating cognitive neuroscience to the driver's operational environment: a neuroergonomic approach. *Am. J. Psychol.* 123, 391–411. doi: 10.5406/amer-jpsyc.123.4.0391
- Lustig, C., Shah, P., Seidler, R., and Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: a review and future directions. *Neuropsychol. Rev.* 19, 504–522. doi: 10.1007/s11065-009-9119-9
- Lyman, S., Ferguson, S. A., Braver, E. R., and Williams, A. F. (2002). Older driver involvements in police reported crashes and fatal crashes: trends and projections. *Traffic Inj. Prev.* 8, 116–120. doi: 10.1136/ip.8.2.116
- Marmeleira, J. F., Godinho, M. B., and Fernandes, O. M. (2009). The effects of an exercise program on several abilities associated with driving performance in older adults. *Accid. Anal. Prev.* 41, 90–97. doi: 10.1016/j.aap.2008.09.008
- Morris, S. B., and DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol. Methods* 7, 105–125. doi: 10.1037/1082-989X.7.1.105
- OECD. (2001). *Ageing and Transport: Mobility Needs and Safety Issues*. Paris: OECD publications.
- Owsley, C., McGwin, G. Jr., Phillips, J. M., McNeal, S. F., and Stalvey, B. T. (2004). Impact of an educational program on the safety of high-risk, visually impaired, older drivers. *Am. J. Prev. Med.* 26, 222–229. doi: 10.1016/j.amepre.2003.12.005
- Pedhazur, E. J. (1982). *Multiple Regression and Behavioral Science. Explanation and Prediction*. New York, NY: Holt, Rinehart and Winston.
- Rheinberg, F., Vollmeyer, R., and Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica* 47, 57–66. doi: 10.1026/0012-1924.47.2.57
- Risser, R., Chaloupka, C., Grundler, W., Sommer, M., Hausler, J., and Kaufmann, C. (2008). Using non-linear methods to investigate the criterion validity of traffic-psychological test batteries. *Accid. Anal. Prev.* 40, 149–157. doi: 10.1016/j.aap.2006.12.018
- Roenker, D. L., Cissell, G. M., Ball, K. K., Wadley, V. G., and Edwards, J. D. (2003). Speed-of-processing and driving simulator training result in improved driving performance. *Hum. Factors* 45, 218–233. doi: 10.1518/hfes.45.2.218.27241
- Romoser, M. R. (2013). The long-term effects of active training strategies on improving older drivers' scanning in intersections: a two-year follow-up to Romoser and Fisher (2009). *Hum. Factors* 55, 278–284. doi: 10.1177/0018720812457566
- Romoser, M. R., and Fisher, D. L. (2009). The effect of active versus passive training strategies on improving older drivers' scanning in intersections. *Hum. Factors* 51, 652–668. doi: 10.1177/0018720809352654
- Romoser, M. R., Pollatsek, A., Fisher, D. L., and Williams, C. C. (2013). Comparing the glance patterns of older versus younger experienced drivers: scanning for hazards while approaching and entering the intersection. *Transp. Res. Part F Traffic Psychol. Behav.* 16, 104–116. doi: 10.1016/j.trf.2012.08.004
- Schuhfried, (n.d.). Available online at: <http://www.schuhfried.com/cogniplus-cps/the-training-programs> (Accessed August 16, 2013)
- Schuhfried, G. (1998). *Manual Determination Test (DT)*. Mödling: SCHUHFRIED GmbH.
- Schuhfried, G. (2005). *Manual Expert System Traffic (XPSV)*. Mödling: SCHUHFRIED GmbH.
- Schuhfried, G., and Prieler, J. (1997). *Manual Reaction Test (RT)*. Mödling: SCHUHFRIED GmbH.
- Schuhfried, G., Prieler, J., and Bauer, W. (2002). *Manual Peripheral Perception (PP)*. Mödling: SCHUHFRIED GmbH.
- Selander, H., Lee, H. C., Johansson, K., and Falkmer, T. (2011). Older drivers: on-road and off-road test results. *Accid. Anal. Prev.* 43, 1348–1354. doi: 10.1016/j.aap.2011.02.007
- Sheridan, T. B. (2004). Driver distraction from a control theory perspective. *Hum. Factors* 46, 587–599. doi: 10.1518/hfes.46.4.587.56807
- Sommer, M., Heidinger, C., Arendasy, M., Schauer, S., Schmitz-Gielsdorf, J., and Hausler, J. (2010). Cognitive and personality determinants of post-injury driving fitness. *Arch. Clin. Neuropsychol.* 25, 99–117. doi: 10.1093/arclin/acp109
- Sommer, M., Herle, M., Häusler, J., and Arendasy, M. (2009). "Von TAVTMB zu ATAVT: eine Anwendung der automatisierten Itemgenerierung unter einschränkenden Rahmenbedingungen," in *Zweites Österreichisches Symposium für Psychologie im Militär*, eds G. Ebner and G. Fleck (Wien: Schrifreihreihe der Landesverteidigungsakademie), 27–52.

- Sommer, M., Herle, M., Häusler, J., Risser, R., Schützhofer, B., and Chaloupka, C. (2008). Cognitive and personality determinants of fitness to drive. *Transp. Res. Part F Traffic Psychol. Behav.* 11, 362–375. doi: 10.1016/j.trf.2008.03.001
- Stalvey, B. T., and Owsley, C. (2003). The development and efficacy of a theory-based educational curriculum to promote self-regulation among high-risk older drivers. *Health Promot. Pract.* 4, 109–119. doi: 10.1177/1524839902250757
- von Bastian, C. C., Langer, N., Jäncke, L., and Oberauer, K. (2013). Effects of working memory training in young and old adults. *Mem. Cognit.* 41, 611–624. doi: 10.3758/s13421-012-0280-7
- Wagner, M., and Karner, T. (2001). *Manual Cognitrone (COG)*. Mödling: SCHUHFRIED GmbH.
- Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restor. Neurol. Neurosci.* 27, 455–471. doi: 10.3233/RNN-2009-0495

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 November 2013; accepted: 23 April 2014; published online: 13 May 2014.

Citation: Casutt G, Theill N, Martin M, Keller M and Jäncke L (2014) The drive-wise project: driving simulator training increases real driving performance in healthy older drivers. Front. Aging Neurosci. 6:85. doi: 10.3389/fnagi.2014.00085

This article was submitted to the journal Frontiers in Aging Neuroscience.

Copyright © 2014 Casutt, Theill, Martin, Keller and Jäncke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Age-related increases in false recognition: the role of perceptual and conceptual similarity

Laura M. Pidgeon* and Alexa M. Morcom

Department of Psychology, Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Michael A. Yassa, Johns Hopkins University, USA

Wilma Koutstaal, University of Minnesota, USA

*Correspondence:

Laura M. Pidgeon, Department of Psychology, Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK
e-mail: L.M.Pidgeon@sms.ed.ac.uk

Older adults (OAs) are more likely to falsely recognize novel events than young adults, and recent behavioral and neuroimaging evidence points to a reduced ability to distinguish overlapping information due to decline in hippocampal pattern separation. However, other data suggest a critical role for semantic similarity. Koutstaal et al. [(2003) false recognition of abstract vs. common objects in older and younger adults: testing the semantic categorization account, *J. Exp. Psychol. Learn.* 29, 499–510] reported that OAs were only vulnerable to false recognition of items with pre-existing semantic representations. We replicated Koutstaal et al.'s (2003) second experiment and examined the influence of independently rated perceptual and conceptual similarity between stimuli and lures. At study, young and OAs judged the pleasantness of pictures of abstract (unfamiliar) and concrete (familiar) items, followed by a surprise recognition test including studied items, similar lures, and novel unrelated items. Experiment 1 used dichotomous "old/new" responses at test, while in Experiment 2 participants were also asked to judge lures as "similar," to increase explicit demands on pattern separation. In both experiments, OAs showed a greater increase in false recognition for concrete than abstract items relative to the young, replicating Koutstaal et al.'s (2003) findings. However, unlike in the earlier study, there was also an age-related increase in false recognition of abstract lures when multiple similar images had been studied. In line with pattern separation accounts of false recognition, OAs were more likely to misclassify concrete lures with high and moderate, but not low degrees of rated similarity to studied items. Results are consistent with the view that OAs are particularly susceptible to semantic interference in recognition memory, and with the possibility that this reflects age-related decline in pattern separation.

Keywords: cognitive aging, episodic memory, false recognition, pattern separation, gist, perceptual similarity, conceptual similarity

INTRODUCTION

Older adults (OAs) are more prone to false recognition than young adults (YAs), particularly when novel events are similar to those previously encountered (Koutstaal and Schacter, 1997; Yassa et al., 2011a). This suggests that a fundamental feature of recognition impairments in older individuals is reduced ability to discriminate in memory (mnemonic discrimination) between similar representations (Stark et al., 2013; Yeung et al., 2013). Electrophysiological and neuroimaging data support proposals that age-related impairments of mnemonic discrimination reflect reduced capacity to pattern separate (orthogonalise) incoming representations, leading to greater overlap between memory representations (Wilson et al., 2006). Pattern separation accounts predict that many kinds of representational overlap are less efficiently discriminated in older age, without specifying a unique role for semantic overlap. However, it has also been suggested that increased false recognition with age is due to greater emphasis on processing of *gist*, particularly semantic gist (Koutstaal and Schacter, 1997; Tun et al., 1998). Koutstaal et al.'s (2003) *semantic categorization account* makes the specific proposal that increases in false recognition are due to greater emphasis by OAs on semantic processing at encoding. To evaluate these

accounts it is critical to establish whether the age-related increase in false recognition is in fact driven specifically by semantic similarity.

False recognition of words and images increases markedly with age. In the Deese–Roediger–McDermott paradigm (Deese, 1959; Roediger and McDermott, 1995), OAs falsely recognize up to 80% of critical lure words which are strongly associated with lists of studied words, compared to up to 65% among YAs (Norman and Schacter, 1997). In paradigms employing visual images, OAs falsely recognize perceptually similar images from the same basic-level conceptual category as studied images (e.g., *cats*) up to 35% more often than YAs (Koutstaal and Schacter, 1997; Koutstaal et al., 1999a). Such findings have been described in terms of greater reliance among OAs on semantic gist (general representations of meaning; Brainerd and Reyna, 2002), leading to heightened false recognition where lure items overlap with studied items in meaning or conceptual category (Norman and Schacter, 1997; Reyna and Brainerd, 1998; Tun et al., 1998). In support of this, Dennis et al. (2007, 2014) report functional magnetic resonance imaging (fMRI) evidence of increased semantic processing among OAs at encoding and during recognition. Moreover, studying multiple category

exemplars is thought to strengthen semantic gist representations: larger effects of category size on false recognition have been reported in OAs, supporting the notion of increased reliance on semantic gist representations with age (Koutstaal and Schacter, 1997).

Although most gist-based accounts emphasize the role of overlapping semantic information in driving false recognition, they do not discount the notion that other kinds of similarity can contribute to recognition outcomes. Perceptual gist representations (general perceptual representations based on overall shape, color etc.) are also thought to be strengthened with exposure to multiple visually similar items (Koutstaal et al., 1999b), but effects of perceptual similarity on false recognition have been considered to be equivalent in young and OAs (Schacter et al., 1997; Koutstaal et al., 1999b; see also Budson et al., 2003). To our knowledge only one study has compared perceptual and semantic similarity effects on false recognition in healthy aging (Koutstaal et al., 2003). Noting that in recognition studies employing visual images, lures typically show both conceptual (category membership) and perceptual (shape, color) overlap with studied items, Koutstaal et al. (2003) employed unfamiliar abstract shapes grouped into categories based on visual similarity to examine whether age-related increases in false recognition result only from semantic gist, or can be driven by perceptual similarity. When verbal conceptual labels were assigned to abstract categories at both study and test, OAs were more likely to show false recognition than when no labels were provided. In a second experiment, OAs were more likely than YAs to falsely recognize concrete, meaningful images sharing basic-level category membership with studied images, but were not more likely to falsely recognize abstract lures (sharing only perceptual features with studied items). This effect of conceptual information was present even when many perceptually similar abstract category members were presented, suggesting age-related increases in false recognition are influenced by semantic but not perceptual gist, consistent with the semantic categorization account (Koutstaal et al., 2003).

Pattern separation accounts do not place special emphasis on semantic information, instead pointing to age-related impairments in mnemonic discrimination along multiple dimensions of similarity. This is proposed to be due to decline in hippocampal pattern separation; the formation of unique neural representations from incoming sensory input, minimizing overlap with existing representations (Wilson et al., 2003, 2006). Rodent electrophysiological and human fMRI data suggest advancing age is associated with rigidity of hippocampal neuronal responses, with a shift from a tendency to pattern separate incoming representations toward a tendency for pattern completion, i.e., reinstatement of existing representations based on incomplete cues. It has been suggested that this decline in pattern separation contributes to age-related reductions in the capacity for mnemonic discrimination (Wilson et al., 2003; Yassa et al., 2011a,b). Although by definition an encoding process, pattern separation can be elicited at retrieval if there is interference between incoming information and to-be-retrieved representations, and therefore can contribute to mnemonic discrimination outcomes at encoding and/or retrieval (Yassa and Stark, 2011). Behavioral investigations suggest that

OAs are less likely to correctly reject as “similar” lures which are perceptually and conceptually similar to studied items, and more likely to falsely recognize lures as “old” (Toner et al., 2009; Stark et al., 2013). The degree of this impairment in mnemonic discrimination has been found to correlate with age-related structural and functional changes in the hippocampus, supporting assumptions that less efficient pattern separation contributes to decline in mnemonic discrimination (Yassa et al., 2011a,b). Age-related impairments have also been demonstrated in mnemonic discrimination of lure and studied items presented in close temporal or spatial proximity over short retention intervals of less than 1 min (Stark et al., 2010; Holden et al., 2012; Tolentino et al., 2012), and Reagh et al. (2014) showed spatial discrimination deficits over delays up to 12 min, similar to intervals typically employed in recognition studies. Spatial discrimination is also impaired in aged rodents (e.g., Wilson et al., 2003). A study employing visually presented verbal stimuli reported age-related impairments in mnemonic discrimination of perceptually similar, but not conceptually similar words (Ly et al., 2013).

Age-related decline in pattern separation with increased bias toward pattern completion has been proposed as a potential mechanism for the age-related increase in gist-based false recognition as well as for reductions in mnemonic discrimination along multiple dimensions of overlap (Schacter et al., 1998; Yassa et al., 2011a). However, any integration of pattern separation and gist accounts requires specification of the role of semantic information, which is central to gist accounts (Norman and Schacter, 1997; Reyna and Brainerd, 1998; Tun et al., 1998). Moreover, the semantic categorization account (Koutstaal et al., 2003) proposes a specific mechanism for greater gist reliance in aging. If OAs explicitly process conceptual information at encoding at the expense of perceptual detail, false recognition increases will be driven specifically by semantic relatedness, and perceptual similarity will have, if anything, a reduced effect. Conversely, a decline in neural pattern separation predicts a more general impact on mnemonic discrimination and gist reliance in OAs which extends beyond semantic gist. Human behavioral investigations of mnemonic discrimination have suggested such a general decline, unlike Koutstaal et al.'s (2003) findings, but it is possible that use of images of meaningful objects (Stark et al., 2010; Yassa et al., 2011a), or nameable shapes (Tolentino et al., 2012) has meant that estimates of age-related impairments in spatial, temporal, or perceptual discrimination are influenced by semantic content.

The present study investigated whether age-related increases in false recognition are dependent on overlapping conceptual representations, as well as evaluating the relative contributions of perceptual and conceptual similarity. In two experiments, we sought to replicate and extend findings of Koutstaal et al. (2003), using the original study's abstract and concrete images. Experiment 1 employed an “old/new” recognition paradigm in a direct replication of Koutstaal et al.'s (2003) Experiment 2, with one modification. In the earlier study, multiple exemplars from each studied category were presented at test and at study, making it difficult to separate study phase and test phase interference, and possibly leading to elevated false recognition in both age groups by increasing bias toward responding “old,” lessening age differences.

Thus, during the test phase we included a single studied item and single lure for each category which was encountered in the study phase. Age-related increases in false recognition of concrete but not abstract lures would provide support for a uniquely semantic gist-based account of false recognition in aging, such as the semantic categorization account (Koutstaal et al., 2003). Increased false recognition of abstract as well as concrete lures would however indicate impaired mnemonic discrimination along multiple domains of similarity, as proposed by pattern separation models and also consistent with a more generalized gist-based account. If perceptual gist as well as semantic gist influences false recognition in old age, effects of study phase category size would be expected to be larger among OAs for both stimulus types.

Experiment 2 was identical to Experiment 1, but participants were asked to respond “old,” “similar” or “new” to studied, lure and novel items respectively, instead of simply “old/new.” The additional “similar” response option has been found to reduce gist-based false recognition in young and OAs (Koutstaal et al., 1999a), and is thought to place greater demands on pattern separation (Stark et al., 2013). Based on previous findings, it was predicted that OAs would show greater false recognition and reduced correct rejection of lures (Stark et al., 2013). As in Experiment 1, if older age is associated with general mnemonic discrimination decline, this pattern was expected for abstract and concrete items, whereas the semantic categorization account predicted age differences only for concrete items.

We also sought to investigate the prediction from pattern separation models that OAs require greater change in input in order to successfully discriminate lures from studied items (Wilson et al., 2006), using measures of subjectively rated within-category perceptual and conceptual similarity (Konkle et al., 2010). These measures provided a further test of the role of conceptual similarity in age-related increases in false recognition. If this is critical in driving age-related increases in false recognition, as assumed by the semantic categorization account, OAs would be expected to show greater effects of conceptual similarity on false recognition, while perceptual similarity effects would be equivalent in the two age groups.

MATERIALS AND METHODS

PARTICIPANTS

Demographic and cognitive test data for participants in Experiments 1 and 2 are shown in Table 1. We estimated that 24 participants per group were required to replicate Koutstaal et al.’s (2003) critical Stimulus Type × Age interaction with 95% power (G*Power; Erdfelder et al., 1996), following correction for bias true effect size estimation (Tversky and Kahneman, 1971; <http://www.john-uebersax.com/stat/bpower.htm>). Twenty-four YAs (aged 18–26 years) and 24 OAs (aged 60–79) took part in Experiment 1. A further YA was excluded from analyses due to incomplete data. Twenty-six YAs (aged 18–28) and 26 OAs (aged 62–79) participated in Experiment 2. Data from one additional OA were excluded due to failure to use the “similar” response during the test phase. All participants completed the following baseline cognitive tests: the Wechsler Test of Adult Reading (WTAR;

Wechsler, 2001), and the Digit Symbol Coding and Digit Span Forward and Backward subscales of the Wechsler Adult Intelligence Scale IV (WAIS-IV^{UK}; Wechsler, 2008). Raw WTAR scores were converted to Standard Scores based on the UK Standardization Sample.

A separate sample of 24 OAs (aged 60–75) and 24 YAs (aged 18–25) provided subjective ratings of stimuli employed in Experiments 1 and 2. Half gave perceptual similarity ratings and half gave conceptual similarity ratings for all categories. Two YAs and one OA were excluded from analyses of conceptual and perceptual ratings respectively, as the average correlation of their ratings with the remainder of the sample was >2 SDs from the mean sample correlation (Konkle et al., 2010). Conceptual ratings were therefore based on 10 YAs and 12 OAs, and perceptual ratings on 12 YAs and 11 OAs.

All experimental procedures were approved by the Psychology Research Ethics Committee of the University of Edinburgh. Informed consent was obtained, and all participants were fully debriefed following completion of the experiment.

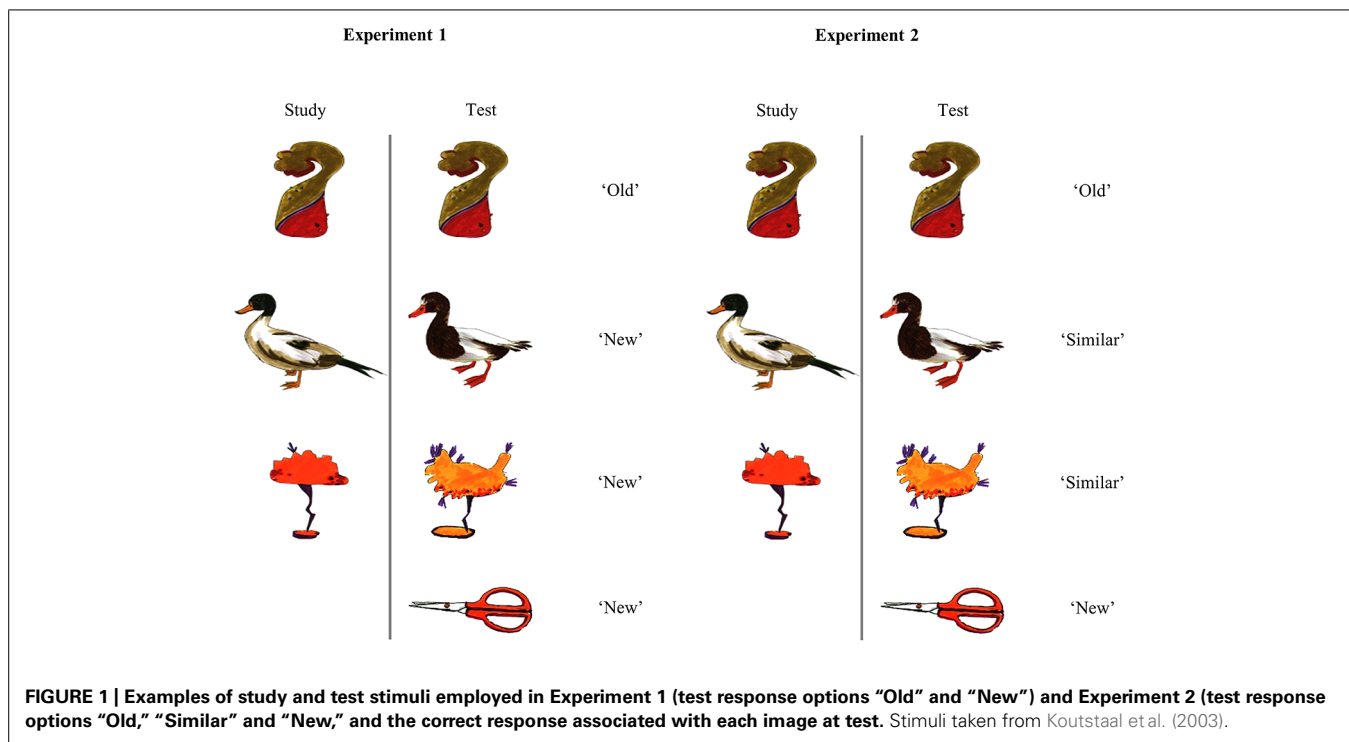
STIMULI

Stimuli were colored line drawings of categorized abstract and concrete items, employed by Koutstaal et al. (2003; see Figure 1 for examples). Abstract items were unfamiliar shapes grouped into categories based on perceptual features, e.g., shape, color. Concrete items were drawings of familiar objects and animals grouped into categories according to basic-level conceptual category, e.g., hats, ducks. Categories consisted of 2 or 13 exemplars. Thirteen-exemplar categories were employed during the study phase as

Table 1 | Demographic and cognitive test data for participants from Experiments 1 and 2.

	Experiment 1		Experiment 2	
	YA (n = 24)	OA (n = 24)	YA (n = 26)	OA (n = 26)
Age	20.6 (2.0) ^a	70.5 (5.6) ^a	21.6 (2.8) ^a	69.2 (4.4) ^a
Sex (N female)	13	12	21 ^a	14 ^a
Years of education	15.6 (1.6)	16.6 (3.5)	16.0 (2.0)	15.9 (4.2)
WTAR	122 (5.7)	120 (7.1)	117 (6.8)	118 (6.3)
(Standard Score)	(n = 19)		(n = 16)	
Digit Symbol	63.5 (10.4) ^a	46.5 (10.8) ^a	67.5 (8.7) ^a	50.7 (13.4) ^a
Digit Span Forward	7.6 (1.0)	7.3 (1.0)	7.2 (1.1)	7.5 (1.1)
Digit Span Backward	5.8 (1.1) ^b	5.6 (1.1)	4.9 (1.2) ^b	5.3 (1.2)

Mean (standard deviation). ^a Denotes significant within-experiment age difference ($p < 0.05$). ^b Denotes significant difference between Experiments 1 and 2. WTAR, Wechsler Test of Adult Reading. WTAR scores for non-native English speakers were excluded. See section “Cognitive Tests” for details of statistical analyses.



either single or large categories, and exemplars were presented as studied items and lures during the test phase. Two-exemplar categories were employed as novel categories during the test phase.

In both experiments, concrete items presented at study included 12 large categories (nine items presented per category) and 12 single-item categories (one exemplar). Thus, a total of 108 large category concrete items (nine exemplars from 12 categories) and 12 single category concrete items (one exemplar from 12 categories), were presented at study. The same distribution applied to abstract items (108 large category items; 12 single category items). Test phase lists comprised 48 studied items (of which 24 had been presented as part of large categories at study; 24 as single exemplars); 48 similar lures, i.e., novel images from studied categories (24 from large categories; 24 from single categories), and 48 novel items from 48 novel categories. Half of the items in each of these stimulus conditions were abstract, half were concrete. The stimulus condition of exemplars (studied or lure) was counterbalanced across participants. During study and test phases, abstract and concrete stimuli were intermixed and a unique pseudorandom order of presentation was generated for each participant.

PROCEDURE

Experiment 1

Experiment 1 consisted of a study phase, followed by a 10 min filled interval, before a surprise recognition test. Study and test procedure followed Koutstaal et al.'s (2003) Experiment 2. During the study phase, participants viewed 240 images, and were asked to rate how pleasant they found each image from 1 to 5 (1 = very unpleasant; 5 = very pleasant). Images were $\sim 350 \times 350$ pixels and were viewed at a distance of around 50 cm on a PC screen,

against a white background. Images were presented for 3 s, with a black fixation cross presented for 1.5 s between trials.

The test phase followed a 10 min interval during which cognitive tests were completed (see Participants). During the test phase, participants viewed 144 images, and judged each as “old” or “new” using key presses (**Figure 1**). Images were presented for 3.5 s, followed by a fixation cross for 1 s. Following each trial, participants were prompted to rate their confidence in their response, from 1 (just guessing) to 5 (very confident) using number keys. The prompt was presented for 4 s followed by a fixation cross for 0.8 s before the start of the next trial.

Experiment 2

Experiment 2 differed from Experiment 1 only in the test phase task instructions and response options; stimuli and study phase procedure were as in Experiment 1. Participants were informed that during the recognition test, items would either be identical to studied items, novel but similar to studied items, or novel and unrelated to studied items. Participants were asked to respond “old,” “similar,” or “new” accordingly (**Figure 1**). As in Experiment 1, participants then rated their confidence in their responses. Study phase timings were as in Experiment 2. During the test phase, stimulus presentation timings were as in Experiment 1 for all YAs. However, as the initial three OAs struggled to give recognition responses within 3.5 s, on average responding to only 65% of trials, presentation time was increased to 4.5 s for the remaining OAs. For the first 3 OAs, responses given within 4.5 s were retrieved from log files and recorded: their pattern of performance did not differ from the other OAs. For both age groups, the confidence rating prompt was presented for up to 4 s, or until 0.8 s after a response

was made, followed by a fixation cross for 0.8 s before the next trial.

Ratings task

Participants in the ratings task gave either perceptual or conceptual similarity ratings for all categories (abstract and concrete). Images were reduced to $\sim 150 \times 150$ pixels, and all exemplars of a given category were presented simultaneously, against a white background.

During the perceptual task, participants rated the overall visual similarity of items in each category from 1 (very similar) to 5 (very distinctive) using the number keys. Participants were asked to base judgments on visual features only, e.g., shape, color. Abstract and concrete categories were presented together in a single block, with the order of category presentation randomized across participants. Images remained on screen until a response was made.

In the conceptual task, concrete and abstract images were presented in separate blocks, the order of which was counterbalanced across participants. For concrete images, participants were asked to rate categories according to how many *kinds* of object were present (1 = few kinds; 5 = many kinds), following Konkle et al. (2010). For example, a category of ducks comprising several distinct breeds of duck would be considered to contain more kinds than a category of apples containing only red apples. This provided a measure of within-category conceptual similarity. As abstract items by definition were not necessarily conceptually meaningful, for these items a modified conceptual ratings task was used. Participants were presented with abstract categories in sequence, and were asked firstly to provide a verbal label of a concrete object which they perceived some or all category members to resemble, and secondly to rate from 1 to 5 the ease of assigning a label that fit all category exemplars. This measure was assumed to reflect the fit of conceptual labels within each abstract category, equivalent to conceptual similarity. Ratings were inverted so that low scores reflected greater similarity, in line with the perceptual scale.

The 13-exemplar categories used as studied or lure items at test were divided into tertiles representing high, medium, and low conceptual and perceptual similarity on the basis of average ratings, separately for abstract and concrete items. For both experiments, mean proportion of lures falsely recognized by each participant at each level of similarity was calculated, separately for abstract and concrete stimuli and for perceptual and conceptual ratings. In each experiment, proportions for each participant were calculated across nine high, eight medium, and seven low perceptual similarity abstract lures, and across seven high, nine medium, and eight low perceptual similarity concrete lures. For conceptual similarity, mean proportions were calculated from 8 abstract and eight concrete lures at each level of similarity.

RESULTS

In all analyses of variance (ANOVAs), Greenhouse–Geisser corrected degrees of freedom and p values are reported in cases where Mauchly's test for violation of the assumption of sphericity was significant. In analyses of recognition performance, "highly confident" refers to responses receiving a confidence rating of 4 or 5.

To allow comparison of the critical effects of age on concrete and abstract false recognition, Cohen's d (Cohen, 1988) measures of effect size are given for differences in mean false recognition between YAs and OAs for each stimulus type. Effect sizes are also reported for F -tests of similarity effects, permitting comparison of the magnitude of perceptual and conceptual similarity effects, using η_p^2 (Cohen, 1973). Large effects are defined as d of >0.8 and η_p^2 of >0.14 .

SIMILARITY RATINGS

The raw ratings data are not included in this report but are available from the first author on request. ANOVAs examined effects of Rater Age (young, older) and Category Type (abstract, concrete) on perceptual and conceptual ratings for each category. Within categories, exemplars were rated as more perceptually similar by YAs than OAs, and concrete items were rated as more perceptually similar than abstract items (Age: $F_{(1,46)} = 63.34$; Category Type: $F_{(1,46)} = 71.35$, $ps < 0.001$). Effects of Age and Category Type interacted ($F_{(1,46)} = 34.40$, $p < 0.001$), such that only abstract items were rated as more perceptually similar by YAs (abstract: $t_{(23)} = 9.57$, $p < 0.001$; concrete: $t_{(23)} = 0.72$, $p = 0.48$). Conceptual ratings did not differ by Rater Age or Category Type (Age: $F_{(1,46)} = 0.02$, $p = 0.89$; Category Type: $F_{(1,46)} = 0.05$, $p = 0.83$; interaction: $F_{(1,46)} = 0.88$, $p = 0.35$), although it should be noted that as the conceptual task differed for concrete and abstract items, these ratings are not directly comparable. Perceptual and conceptual ratings of abstract categories did not correlate reliably in young or older raters (young: $r = 0.36$, $p = 0.08$; older: $r = 0.23$, $p = 0.29$), indicating the scales were indeed measuring distinct stimulus qualities. For concrete categories, perceptual and conceptual ratings were positively correlated in both age groups (young: $r = 0.66$, $p < 0.001$; older: $r = 0.45$, $p = 0.03$).

Across both item types, YAs and OAs showed high, positive correlations for both perceptual and conceptual ratings (perceptual: $r = 0.83$, $p < 0.001$; conceptual: $r = 0.6$, $p < 0.001$). Therefore, to obtain perceptual and conceptual similarity scales which were equally applicable to both age groups, ratings from YAs and OAs were averaged. In the averaged scale, perceptual and conceptual ratings again did not correlate for abstract items ($r = 0.32$, $p = 0.12$), but were positively correlated for concrete items ($r = 0.69$, $p < 0.001$). Concrete items were rated more perceptually similar than abstract items ($M_{\text{concrete}} = 2.89$, $SD = 0.50$; $M_{\text{abstract}} = 3.65$, $SD = 0.37$; $t_{(46)} = 8.45$, $p < 0.001$; lower figures represent greater similarity), while no difference in mean conceptual rating was found between item types ($M_{\text{concrete}} = 2.92$, $SD = 0.40$; $M_{\text{abstract}} = 2.77$, $SD = 0.25$; $t_{(46)} = 0.22$, $p = 0.83$). Proportions of participants assigning the same conceptual label to abstract categories ranged from 0 (each gave a different label) to 0.45 (10/22 gave the same label), with a mean proportion of 0.23 (5/22).

COGNITIVE TESTS

Cognitive test results for participants in Experiments 1 and 2 are summarized in **Table 1**. In Experiment 1, one OA did not complete the Digit Span Backward task, and WTAR Standard Scores were excluded for five YAs who were non-native speakers of English.

YAs and OAs did not differ in years of education, WTAR (Standard Scores), or Digit Span Forward or Backward (education: $t_{(41)} = 1.10$, $p = 0.28$; WTAR: $t_{(41)} = 1.21$, $p = 0.23$; Digit Span Forward: $t_{(46)} = 1.02$, $p = 0.31$; Digit Span Backward: $t_{(46)} = 0.82$, $p = 0.42$). YAs outperformed OAs on the Digit Symbol task ($t_{(46)} = 5.56$, $p < 0.001$) as expected. Chi-squared test of independence confirmed that the sex distribution did not differ between age groups [$\chi^2(1, N = 48) = 0.08$, $p = 0.77$].

In Experiment 2, one OA was unable to complete the Digit Symbol test due to an injury, and one YA was excluded from this test due to failure to follow procedure. WTAR scores were disregarded for 10 YAs who were not native speakers of English. The proportion of females was higher in the young group [Chi-squared test of independence: $\chi^2(1, N = 52) = 4.28$, $p < 0.05$]. However, rates of veridical recognition, lure false recognition, or lure correct rejections did not differ by sex for either abstract or concrete stimuli, suggesting bias of results was unlikely. YAs again scored more highly on the Digit Symbol task ($t_{(48)} = 5.28$, $p < 0.001$). No age differences were observed in years of education, WTAR (Standard Score), or Digit Span Forward or Backward (education: $t_{(50)} = 0.06$, $p = 0.95$; WTAR: $t_{(40)} = 0.71$, $p = 0.48$; Digit Span Forward: $t_{(50)} = 0.74$, $p = 0.46$; Digit Span Backward: $t_{(50)} = 1.1$, $p = 0.29$).

Comparing samples for the two experiments, 2 (Age) \times 2 (Experiment) ANOVAs showed an effect of Experiment for Digit Symbol Backward, with participants in Experiment 1 outperforming those in Experiment 2 ($F_{(1,95)} = 6.77$, $p = 0.011$). No further differences between samples were observed in age, years of education, sex or cognitive test performance, and there were no interactions of Experiment \times Age (max $F = 1.82$; Table 1).

In both experiments, memory performance results (including effects of similarity) were equivalent when non-native speakers of English were excluded from analyses, and so only results from the full samples are reported.

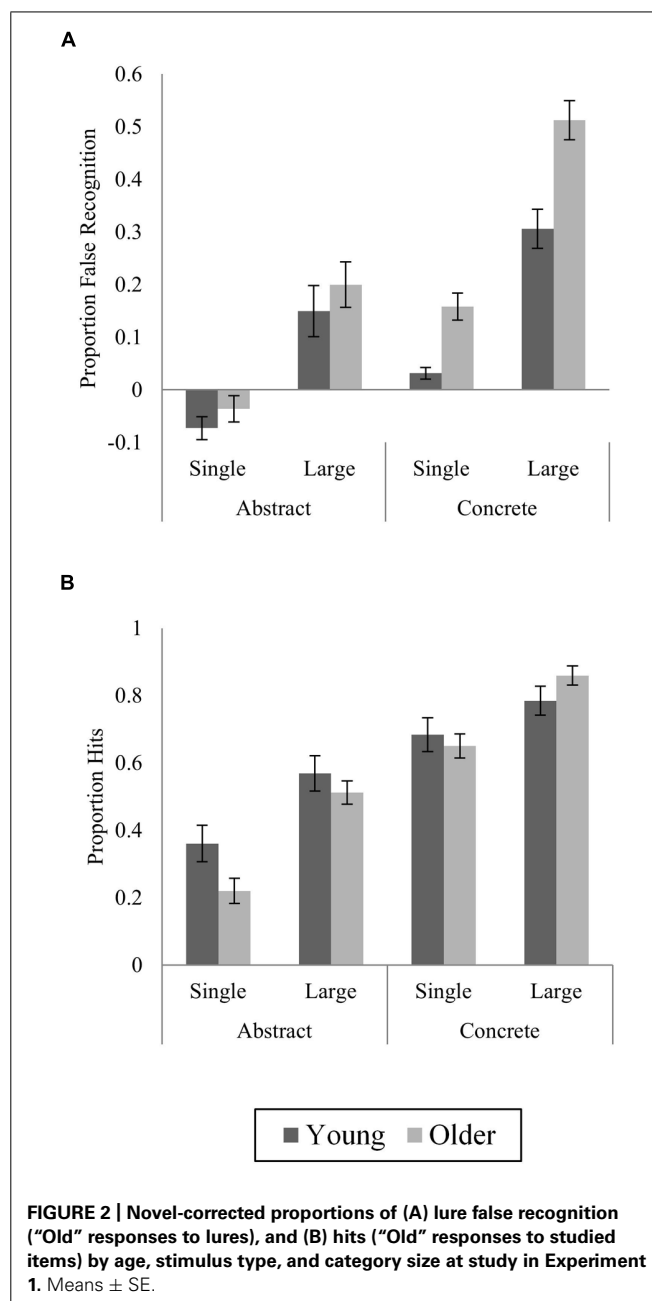
EXPERIMENT 1 – MEMORY PERFORMANCE

Baseline false recognition of novel items

We examined effects of Stimulus Type (abstract, concrete) and Age (young, old) on baseline novel false recognition. More abstract items were falsely recognized than concrete ($F_{(1,46)} = 52.41$, $p < 0.001$; $M_{\text{abstract}} = 0.12$, $SD = 0.17$; $M_{\text{concrete}} = 0.03$, $SD = 0.04$), but this effect did not differ by Age ($F_{(1,46)} = 0.06$, $p = 0.82$), and no overall effect of Age was observed ($F < 1$). Findings did not differ when analyses were restricted to novel items falsely recognized with high confidence, (Stimulus Type: $F_{(1,46)} = 12.53$, $p = 0.001$; Age: $F_{(1,46)} = 0.32$, $p = 0.57$; interaction: $F_{(1,46)} = 0.56$, $p = 0.46$; $M_{\text{abstract}} = 0.04$, $SD = 0.07$; $M_{\text{concrete}} = 0.01$, $SD = 0.03$).

Corrected false recognition of lures

Following Koutstaal et al. (2003), lure false recognition was corrected for baseline false recognition of novel items. Proportions of false recognition of abstract and concrete novel items were subtracted from proportions of false recognition of abstract and concrete lures, respectively. Corrected false recognition is shown in Figure 2A. ANOVA with factors of Age (young, older), Category Size at study (single, large) and Stimulus Type (abstract,



concrete) revealed significant main effects of all three variables (Age: $F_{(1,46)} = 13.88$, $p = 0.001$; Category Size: $F_{(1,46)} = 123.41$, $p < 0.001$; Stimulus Type: $F_{(1,46)} = 81.48$, $p < 0.001$), reflecting greater false recognition among OAs, for large category items, and for concrete items. Crucially, the predicted Age \times Stimulus Type interaction was significant ($F_{(1,46)} = 8.41$, $p = 0.006$), with greater false recognition among OAs compared to YAs for concrete items, but no difference for abstract items (concrete: $t_{(46)} = 5.29$, $p < 0.001$, $d = 1.53$; abstract: $t_{(46)} = 1.12$, $p = 0.27$, $d = 0.32$). The effect of Category Size interacted with Stimulus Type ($F_{(1,46)} = 4.84$, $p = 0.03$), with a larger effect of Category Size (greater false recognition of large category lures) for concrete items, but neither this interaction or the main effect

of Category Size differed by Age (Category Size \times Stimulus Type \times Age: $F_{(1,46)} = 0.73$, $p = 0.4$; Category Size \times Age: $F_{(1,46)} = 0.92$, $p = 0.34$). Following Koutstaal et al. (2003), we also assessed novel-corrected false recognition of single and large category abstract items alone, to determine whether age-related differences were truly unique to concrete items. As predicted by the semantic categorization account, no age differences were observed (single: $t_{(46)} = 1.1$, $p = 0.28$; large: $t_{(46)} = 0.78$, $p = 0.44$).

Restricting analyses to highly confident false recognition, corrected for highly confident novel false recognition (Table 2), similar results were observed. OAs showed greater confident false recognition than YAs ($F_{(1,46)} = 25.94$, $p < 0.001$), concrete lures were more often confidently falsely recognized than abstract lures ($F_{(1,46)} = 66.09$, $p < 0.001$), and large category lures attracted more highly confident false recognition responses than single-item categories ($F_{(1,46)} = 92.14$, $p < 0.001$). A Stimulus Type \times Category Size interaction ($F_{(1,46)} = 13.71$, $p = 0.001$) reflected a greater effect of Category Size for concrete items. A greater effect of Category Size was also observed in OAs compared to YAs ($F_{(1,46)} = 11.34$, $p = 0.002$). The predicted Stimulus Type \times Age interaction was again observed ($F_{(1,46)} = 7.34$, $p = 0.009$; Category Size \times Stimulus Type \times Age: $F_{(1,46)} = 0.55$, $p = 0.46$) however, unlike in the overall analysis OAs showed greater false recognition for abstract ($t_{(46)} = 2.93$, $p = 0.005$, $d = 0.85$) as well as concrete items ($t_{(46)} = 5.28$, $p < 0.001$, $d = 1.52$) with age effects larger for concrete items. Planned tests of age differences among abstract items alone revealed greater highly confident false recognition in OAs of large but not single category abstract lures (single: $t_{(46)} = 1.14$, $p = 0.26$; large: $t_{(46)} = 2.66$, $p = 0.01$).

Corrected veridical recognition

As for false recognition, proportions of hits (correctly identified old items) were corrected for baseline novel false recognition (Figure 2B) equivalent to Snodgrass and Corwin's (1988) P_t measure. Proportion of novel-corrected hits did not vary by age ($F_{(1,46)} = 0.75$, $p = 0.39$), but effects of Stimulus Type and Category Size were greater in OAs (Stimulus Type \times Age: $F_{(1,46)} = 4.39$, $p = 0.04$; Category Size \times Age: $F_{(1,46)} = 4.41$, $p = 0.04$). Main effects of Category Size ($F_{(1,46)} = 79.21$, $p < 0.001$) and Stimulus Type ($F_{(1,46)} = 132.46$, $p < 0.001$) were modified by a Stimulus

Type \times Category Size interaction ($F_{(1,46)} = 7.27$, $p = 0.01$), reflecting a greater increment in hit rate with larger Category Size for abstract (25%) compared to concrete (15.5%) items. This interaction did not however, differ by Age (3-way interaction: $F_{(1,46)} = 0.12$, $p = 0.73$).

For highly confident hits corrected for highly confident novel false recognition (Table 2), the Stimulus Type \times Age interaction was not reliable ($F_{(1,46)} = 2.41$, $p = 0.13$), nor was the Category Size \times Stimulus Type interaction ($F_{(1,46)} = 2.47$, $p = 0.12$). However, the remaining effects were equivalent to those for total hits (Age: $F_{(1,46)} = 0.72$, $p = 0.4$; Stimulus Type: $F_{(1,46)} = 119.64$, $p < 0.001$; Category Size: $F_{(1,46)} = 38.85$, $p < 0.001$; Category Size \times Age: $F_{(1,46)} = 13.18$, $p = 0.001$).

Effects of stimulus perceptual and conceptual similarity

Novel-corrected false recognition of abstract and concrete images according to within-category conceptual and perceptual similarity is illustrated in Figure 3. For each rating (perceptual and conceptual), ANOVAs examined effects of Age (young, older) and Similarity (high, medium, low) on novel-corrected false recognition of abstract and concrete lures, combining across both category sizes. As the perceptual similarity ratings task was equal for abstract and concrete stimuli, Stimulus Type was included as a factor in perceptual similarity analyses. However, as conceptual ratings tasks differed for abstract and concrete images, separate ANOVAs were conducted for each stimulus type. Planned comparisons of age effects at each level of similarity were conducted to test pattern separation predictions in all cases where there were significant effects of similarity.

Perceptual similarity. False recognition rates differed according to within-category Perceptual Similarity ($F_{(2,92)} = 32.67$, $p < 0.001$, $\eta_p^2 = 0.42$), with highly perceptually similar lures attracting more false recognition responses than the most perceptually distinctive lures. This effect was modified by Stimulus Type ($F_{(2,92)} = 14.92$, $p < 0.001$) and was reliable for concrete lures only (concrete: $F_{(2,92)} = 39.08$, $p < 0.001$, $\eta_p^2 = 0.46$; abstract: $F_{(2,92)} = 1.39$, $p = 0.26$, $\eta_p^2 = 0.03$). Across both age groups, false recognition was greater for high vs. both medium and low similarity concrete lures, but did not differ between medium and low similarity concrete lures (high vs. medium: $t_{(47)} = 7.28$, $p < 0.001$; high vs. low: $t_{(47)} = 8.16$, $p < 0.001$; medium vs. low: $t_{(47)} = 0.56$, $p = 0.58$, adjusted $\alpha = 0.017$; Perceptual Similarity \times Age: $F_{(2,92)} = 1.28$, $p = 0.28$; 3-way interaction: $F_{(2,92)} = 0.43$, $p = 0.65$). In the main ANOVA, the interaction of Stimulus Type \times Similarity did not differ by Age (3-way interaction: $F_{(2,92)} = 0.43$, $p = 0.65$).

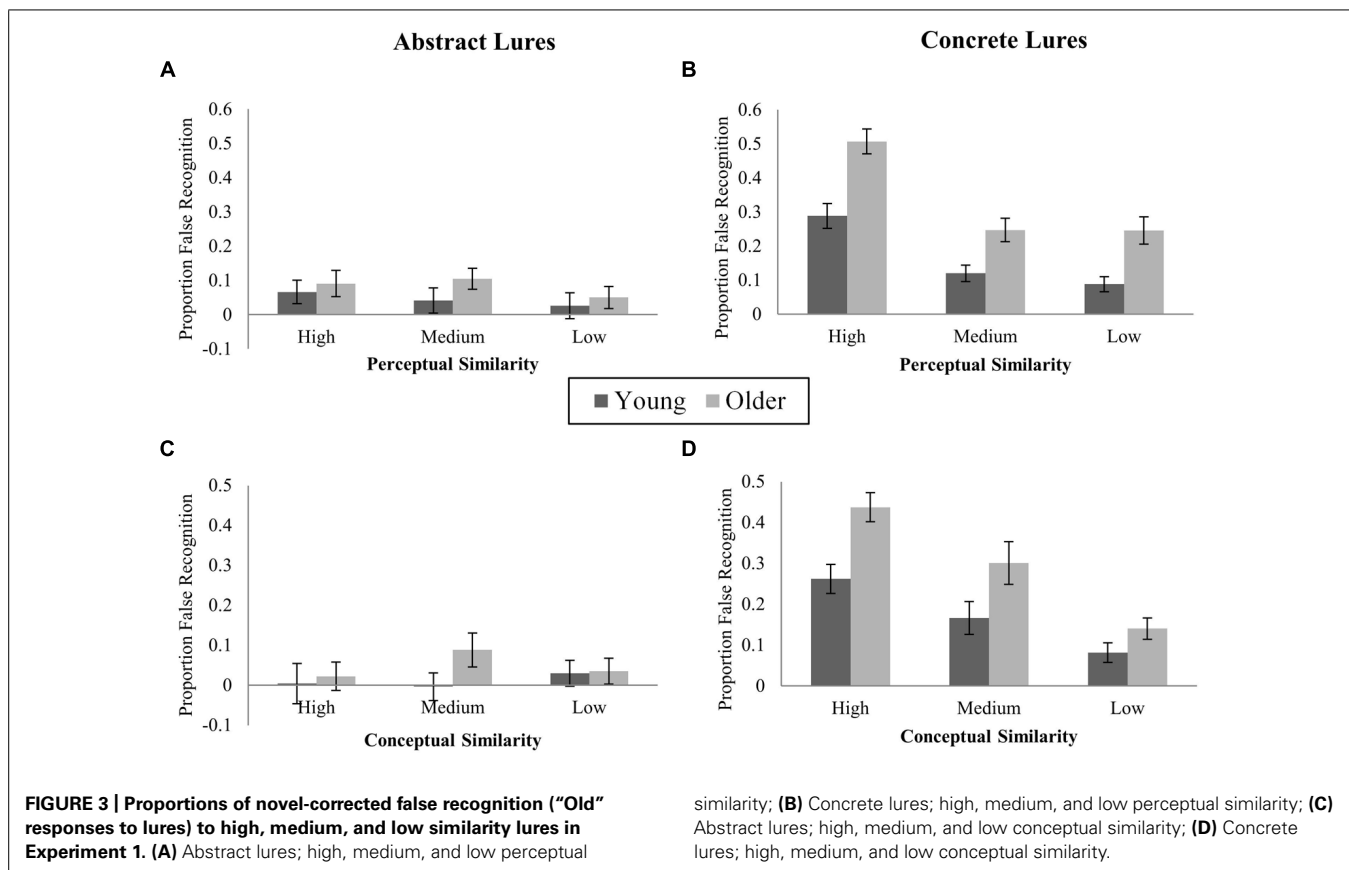
Predictions from pattern separation models that OAs requires greater change in input (i.e., less similarity) to support successful discrimination were tested among concrete items (for which a significant effect of Perceptual Similarity was observed) via planned contrasts of group differences at each level of similarity. OAs were more likely than YAs to falsely recognize concrete lures at all levels of Perceptual Similarity (high: $t_{(46)} = 4.24$, $p < 0.001$; medium: $t_{(46)} = 3.01$, $p = 0.004$; low: $t_{(46)} = 3.43$, $p = 0.001$, adjusted $\alpha = 0.017$).

Conceptual similarity. For abstract items, no reliable effect of Conceptual Similarity was observed ($F_{(2,92)} = 0.48$, $p = 0.62$,

Table 2 | Mean proportions (SD) of novel-corrected highly confident lure false recognition and novel-corrected highly confident hits to studied items in Experiment 1.

	Abstract				Concrete			
	Single		Large		Single		Large	
	YA	OA	YA	OA	YA	OA	YA	OA
Lure FR	−0.02 (0.05)	−0.003 (0.05)	0.05 (0.10)	0.16 (0.19)	0.02 (0.04)	0.09 (0.12)	0.16 (0.11)	0.37 (0.16)
Hits	0.35 (0.23)	0.20 (0.18)	0.46 (0.22)	0.44 (0.24)	0.65 (0.26)	0.54 (0.23)	0.65 (0.26)	0.75 (0.26)

FR = false recognition.



$\eta_p^2 = 0.01$; Age \times Similarity: $F_{(2,92)} = 1.22$, $p = 0.30$). For concrete items, false recognition varied according to Conceptual Similarity ($F_{(2,92)} = 28.44$, $p < 0.001$, $\eta_p^2 = 0.38$). Lures from highly conceptually similar categories were falsely recognized more often than medium or low similarity lures, and more medium than low similarity lures were falsely recognized (high vs. medium: $t_{(47)} = 3.45$, $p = 0.001$; high vs. low: $t_{(47)} = 8.63$, $p < 0.001$; medium vs. low: $t_{(47)} = 3.61$; $p = 0.001$; adjusted $\alpha = 0.017$). Although OAs showed a numerically greater increment in false recognition from low to high similarity (OA: 29.6%, YA: 18%), the Similarity \times Age interaction was not reliable ($F_{(2,92)} = 1.74$, $p = 0.18$). For concrete items, planned contrasts revealed greater false recognition in OAs than YAs for highly conceptually similar lures ($t_{(46)} = 3.5$, $p = 0.001$), but not medium or low similarity lures (medium: $t_{(46)} = 2.05$, $p = 0.046$; low: $t_{(46)} = 1.67$, $p = 0.1$; adjusted $\alpha = 0.017$).

EXPERIMENT 2

Memory performance

With the added "similar" response option in Experiment 2, there were nine possible response outcomes. Studied items could be correctly recognized (hits), judged "similar," or missed (judged "new"). Lures could be falsely recognized as "old," correctly rejected as "similar," or incorrectly judged "new." Novel items could be falsely recognized as "old" (novel false recognition); incorrectly judged "similar," or correctly rejected as "new."

Raw proportions of responses in each of these categories are displayed in **Table 3**. Analyses focused on novel-corrected hits, and lure false recognition and correct rejection, illustrated in **Figure 4**.

Baseline false recognition of novel items. Analysis of variance examining effects of Stimulus Type and Age on proportions of baseline false recognition (**Table 3**) showed that this was more likely for abstract than concrete novel items ($F_{(1,50)} = 8.78$, $p = 0.005$). There was no main effect of Age ($F_{(1,50)} = 0.67$, $p = 0.42$), and no Stimulus Type \times Age interaction ($F_{(1,50)} < 0.001$, $p = 1$). Among highly confident responses, rates of novel false recognition no longer differed according to Stimulus Type ($F_{(1,50)} = 1.12$, $p = 0.30$), but there remained no effects involving Age (Age: $F_{(1,50)} = 0.48$, $p = 0.49$; Stimulus Type \times Age: $F_{(1,50)} = 0.50$, $p = 0.48$).

Corrected false recognition of lures. Proportions of abstract and concrete lure false recognition were corrected for baseline false recognition of novel abstract and concrete items as in Experiment 1, and are displayed in **Figure 4A**. For lure false recognition, ANOVA with factors of Stimulus Type, Category Size, and Age revealed significant main effects of each (Stimulus Type: $F_{(1,50)} = 55.07$; Category Size: $F_{(1,50)} = 94.04$; Age: $F_{(1,50)} = 17.93$, $ps < 0.001$), with greater false recognition of concrete lures relative to abstract, of large category items vs. single, and in OAs. As expected, the effect of Stimulus Type was modulated by Age

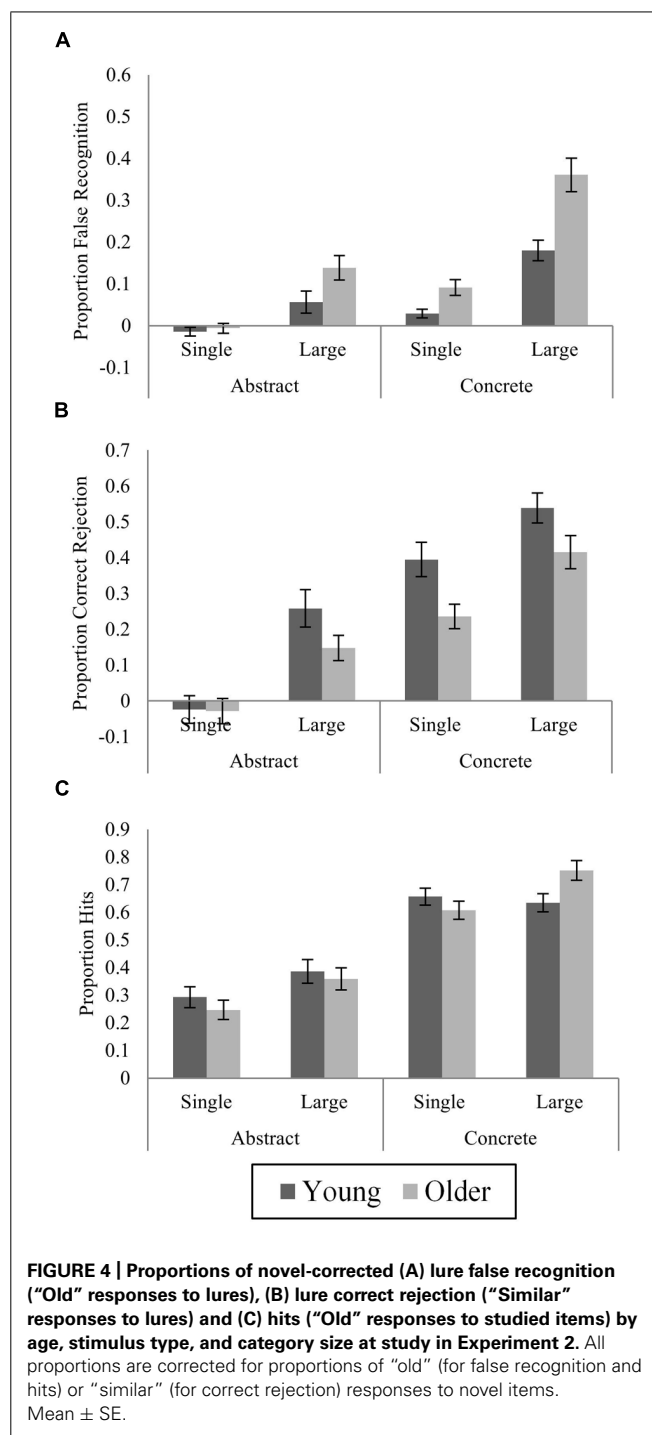
Table 3 | Proportions of raw “Old,” “Similar” and “New” responses to Studied, Lure and Novel items, collapsed across single and large category conditions.

		Abstract		Concrete	
		YA	OA	YA	OA
Studied	“Old”	0.38 (0.17)	0.35 (0.20)	0.66 (0.11)	0.71 (0.15)
	“Similar”	0.35 (0.13)	0.41 (0.19)	0.19 (0.08)	0.17 (0.11)
	“New”	0.20 (0.14)	0.19 (0.08)	0.08 (0.05)	0.08 (0.08)
Lure	“Old”	0.06 (0.09)	0.12 (0.11)	0.12 (0.08)	0.25 (0.15)
	“Similar”	0.43 (0.18)	0.42 (0.14)	0.53 (0.21)	0.37 (0.18)
	“New”	0.44 (0.19)	0.41 (0.17)	0.30 (0.19)	0.33 (0.13)
Novel	“Old”	0.04 (0.06)	0.05 (0.06)	0.02 (0.04)	0.03 (0.07)
	“Similar”	0.32 (0.22)	0.36 (0.18)	0.06 (0.10)	0.04 (0.06)
	“New”	0.59 (0.24)	0.54 (0.21)	0.89 (0.16)	0.91 (0.14)

Mean (standard deviation).

($F_{(1,50)} = 5.49$, $p = 0.023$). There was a clear age difference for concrete items ($t_{(50)} = 4.25$, $p < 0.001$, $d = 1.18$), with OAs 12.2% more likely to falsely recognize concrete lures, and a smaller but significant age difference for abstract items (4.5%; $t_{(50)} = 2.02$, $p = 0.049$, $d = 0.56$). The effect of Category Size also differed by Age ($F_{(1,50)} = 8.64$, $p = 0.005$). In both age groups, large category lures attracted more false recognition responses than single category lures (YAs: $t_{(25)} = 5.65$; OAs: $t_{(25)} = 8.95$, $ps < 0.001$), but the effect of Category Size was larger in OAs. The effect of Stimulus Type was modulated by Category Size ($F_{(1,50)} = 13.24$, $p = 0.001$), with a larger effect of Category Size for concrete items. No three-way interaction was observed ($F_{(1,50)} = 0.63$, $p = 0.43$). Age differences among abstract lures alone were again tested. OAs were more likely than YAs to falsely recognize large but not single category abstract lures (single: $t_{(50)} = 0.51$, $p = 0.61$; large: $t_{(50)} = 2.07$, $p = 0.043$).

Similar effects were observed for novel-corrected highly confident false recognition (Table 4), i.e., main effects of Stimulus Type, Category Size, and Age (Stimulus Type: $F_{(1,50)} = 47.16$; Category Size: $F_{(1,50)} = 73.89$; Age: $F_{(1,50)} = 18.01$, $ps < 0.001$), and interactions of Category Size with Stimulus Type and Age (Category Size \times Stimulus Type: $F_{(1,50)} = 12.51$, $p = 0.001$; Category Size \times Age: $F_{(1,50)} = 11.04$, $p = 0.002$; Category Size \times Stimulus Type \times Age: $F_{(1,50)} = 0.63$, $p = 0.43$). As in Experiment 1, the effect of Stimulus Type was modulated by Age ($F_{(1,50)} = 8.07$, $p = 0.007$), but the age difference in false recognition was now



reliable for concrete items only (concrete: $t_{(50)} = 4.30$, $p < 0.001$, $\eta_p^2 = 1.19$; abstract: $t_{(50)} = 1.93$, $p = 0.059$, $\eta_p^2 = 0.54$). However, in planned analysis of abstract lures, OAs again showed higher confident false recognition of large but not single abstract lures (single: $t_{(50)} = 0.50$, $p = 0.62$; large: $t_{(50)} = 2.28$, $p = 0.027$).

Correct rejection of lures. Proportions of lure correct rejections were corrected by subtracting proportions of “similar” responses to novel items of the same stimulus type (equal to Stark et al.’s

Table 4 | Mean proportions (SD) of novel-corrected highly confident lure false recognition, novel-corrected highly confident correct rejections and novel-corrected highly confident hits in Experiment 2.

	Abstract				Concrete			
	Single		Large		Single		Large	
	YA	OA	YA	OA	YA	OA	YA	OA
Lure FR	−0.002 (0.04)	−0.02 (0.03)	0.04 (0.09)	0.12 (0.15)	0.03 (0.04)	0.09 (0.10)	0.14 (0.11)	0.33 (0.21)
Lure CR	0.01 (0.12)	0.01 (0.11)	0.25 (0.23)	0.12 (0.14)	0.37 (0.24)	0.14 (0.12)	0.47 (0.22)	0.34 (0.24)
Hits	0.27 (0.19)	0.24 (0.17)	0.33 (0.20)	0.31 (0.19)	0.64 (0.15)	0.53 (0.19)	0.60 (0.18)	0.74 (0.18)

FR = false recognition (lures judged “old”); CR = correct rejection (lures judged “similar”).

(2013) behavioral pattern separation score), and are displayed in **Figure 4B**. Main effects of Stimulus Type, Category Size and Age were observed (Stimulus Type: $F_{(1,50)} = 108.21$, $p < 0.001$; Category Size: $F_{(1,50)} = 79.31$, $p < 0.001$; Age: $F_{(1,50)} = 5.56$, $p = 0.022$), with YAs 14.1% more likely than OAs to correctly reject concrete lures, and only 5.7% more likely to correctly reject abstract lures, but the predicted Age \times Stimulus Type interaction was non-significant ($F_{(1,50)} = 1.98$, $p = 0.17$), and did not vary by Category Size (3-way interaction: $F_{(1,50)} = 3.14$, $p = 0.08$), suggesting OAs were impaired in correct rejection of both abstract and concrete lures. Planned contrasts for abstract items alone did not show reliable age-related differences for correct rejection of either single or large category abstract lures (single: $t_{(50)} = 0.09$, $p = 0.93$; large: $t_{(50)} = 1.75$, $p = 0.09$).

Rates of novel-corrected highly confident correct rejection are shown in **Table 4**. A three-way Stimulus Type \times Category Size \times Age interaction was observed ($F_{(1,50)} = 11.21$, $p = 0.002$), as well as the predicted Stimulus Type \times Age interaction ($F_{(1,50)} = 4.88$, $p = 0.03$). Lure correct rejection was more likely in YAs than OAs for both large and single category concrete lures (large: $t_{(50)} = 2.07$, $p = 0.04$; single: $t_{(50)} = 4.31$, $p < 0.001$), and for large category abstract lures ($t_{(50)} = 2.67$, $p = 0.01$), but not for single category abstract lures ($t_{(50)} = 0.05$, $p = 0.96$).

Corrected veridical recognition. Novel-corrected hits to studied items are shown in **Figure 4C**. Concrete images were correctly recognized more often than abstract ($F_{(1,50)} = 186.26$, $p < 0.001$), and more large than single category items were recognized ($F_{(1,50)} = 14.19$, $p < 0.001$), though there was no main effect of Age ($F_{(1,50)} = 0.002$, $p = 0.96$). A marginally significant 3-way interaction ($F_{(1,50)} = 4.0$, $p = 0.05$) reflected presence of a Category Size \times Age interaction for concrete ($F_{(1,50)} = 8.66$, $p = 0.005$) but not abstract items ($F_{(1,50)} = 0.11$, $p = 0.74$). OAs were more likely than YAs to recognize large category concrete items ($t_{(50)} = 2.43$, $p = 0.019$), but no age difference was present for single-items ($t_{(50)} = 1.10$, $p = 0.28$).

Similar effects were found for novel-corrected highly confident responses, shown in **Table 4** (Age: $F_{(1,50)} = 0.04$, $p = 0.84$; Stimulus

Type: $F_{(1,50)} = 198.45$, $p < 0.001$; Category Size: $F_{(1,50)} = 14.58$, $p < 0.001$; Category Size \times Age: $F_{(1,50)} = 11.06$, $p = 0.002$). A three-way interaction ($F_{(1,50)} = 8.64$, $p = 0.005$) again reflected a Category Size \times Age interaction among concrete items only (concrete: $F_{(1,50)} = 18.60$, $p < 0.001$; abstract: $F_{(1,50)} = 0.13$, $p = 0.72$), with OAs more likely than YAs to recognize large category concrete items ($t_{(50)} = 2.72$, $p = 0.009$).

Effects of stimulus perceptual and conceptual similarity

Proportions of novel-corrected lure false recognition and correct rejection according to input similarity are shown in **Figures 5** and **6**, respectively. Analyses followed the same strategy as in Experiment 1, but examined lure correct rejections as well as false recognition.

Perceptual similarity. Analysis of variance revealed a significant effect of rated Perceptual Similarity on false recognition ($F_{(2,100)} = 39.81$, $p < 0.001$, $\eta_p^2 = 0.44$) modified by an interaction with Stimulus Type ($F_{(2,100)} = 19.39$, $p < 0.001$). The effect of Perceptual Similarity was significant for concrete items only (concrete: $F_{(2,100)} = 42.16$, $p < 0.001$, $\eta_p^2 = 0.46$; abstract: $F_{(2,100)} = 2.15$, $p = 0.12$, $\eta_p^2 = 0.04$). Highly perceptually similar concrete lures were more often falsely recognized than medium or low similarity lures, but false recognition of medium and low similarity lures did not differ (high vs. medium: $t_{(51)} = 8.46$; high vs. low: $t_{(51)} = 7.22$; $ps < 0.001$; medium vs. low: $t_{(51)} = 0.77$, $p = 0.45$, adjusted $\alpha = 0.017$). The effect of Perceptual Similarity did not vary with Age ($F_{(2,100)} = 1.98$, $p = 0.14$; 3-way interaction, $F_{(2,100)} = 2.27$, $p = 0.11$). As in Experiment 1, planned contrasts examined age-related differences at each level of Perceptual Similarity for concrete items, for which the overall effect of similarity was reliable. OAs were more likely to falsely recognize high and medium perceptual similarity concrete lures, but not the most distinctive lures (high: $t_{(50)} = 2.74$, $p = 0.008$; medium: $t_{(50)} = 3.49$, $p = 0.001$; low: $t_{(50)} = 1.18$, $p = 0.24$, adjusted $\alpha = 0.017$).

For lure correct rejections, we observed a main effect of Perceptual Similarity and a 3-way interaction (Perceptual Similarity: $F_{(1,50)} = 3.17$, $p = 0.046$; Perceptual Similarity \times Stimulus

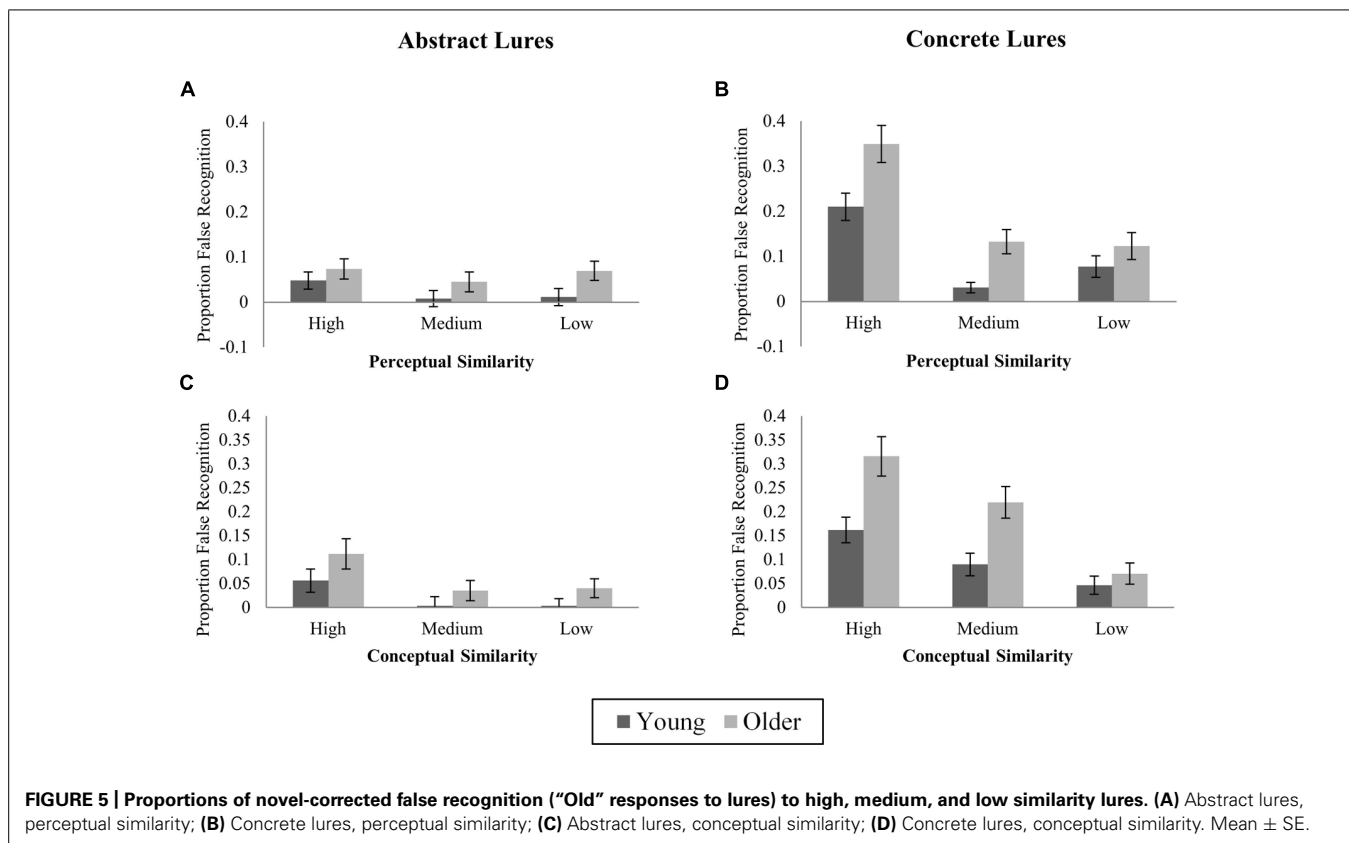


FIGURE 5 | Proportions of novel-corrected false recognition ("Old" responses to lures) to high, medium, and low similarity lures. (A) Abstract lures, perceptual similarity; **(B)** Concrete lures, perceptual similarity; **(C)** Abstract lures, conceptual similarity; **(D)** Concrete lures, conceptual similarity. Mean \pm SE.

Type \times Age: $F_{(2,100)} = 3.09$, $p = 0.05$; Similarity \times Stimulus Type, $F_{(2,100)} = 0.60$, $p = 0.55$). *Post hoc* tests showed an Age \times Stimulus Type interaction for highly perceptually similar lures only (high: $F_{(1,50)} = 10.53$, $p = 0.002$; medium: $F_{(1,50)} = 0.18$, $p = 0.67$; low: $F_{(1,50)} = 2.10$, $p = 0.15$). This reflected higher correct rejection among YAs of highly similar concrete lures, but no age difference for highly similar abstract lures (concrete: $t_{(50)} = 3.09$, $p = 0.003$; abstract: $t_{(50)} = 0.90$, $p = 0.37$). Planned tests of age differences at each level of similarity were conducted for both abstract and concrete lures. Abstract lures did not show reliable age differences at any level of similarity (high: $t_{(50)} = 0.90$, $p = 0.37$; medium: $t_{(50)} = 0.96$, $p = 0.34$; low: $t_{(50)} = 1.30$, $p = 0.20$; $\alpha = 0.017$). For concrete lures, YAs showed greater correct rejection than OAs of the most perceptually similar and most perceptually distinctive lures, but no difference for medium similarity lures (high: $t_{(50)} = 3.09$, $p = 0.003$; medium: $t_{(50)} = 1.40$, $p = 0.17$; low: $t_{(50)} = 2.77$, $p = 0.008$; adjusted $\alpha = 0.017$).

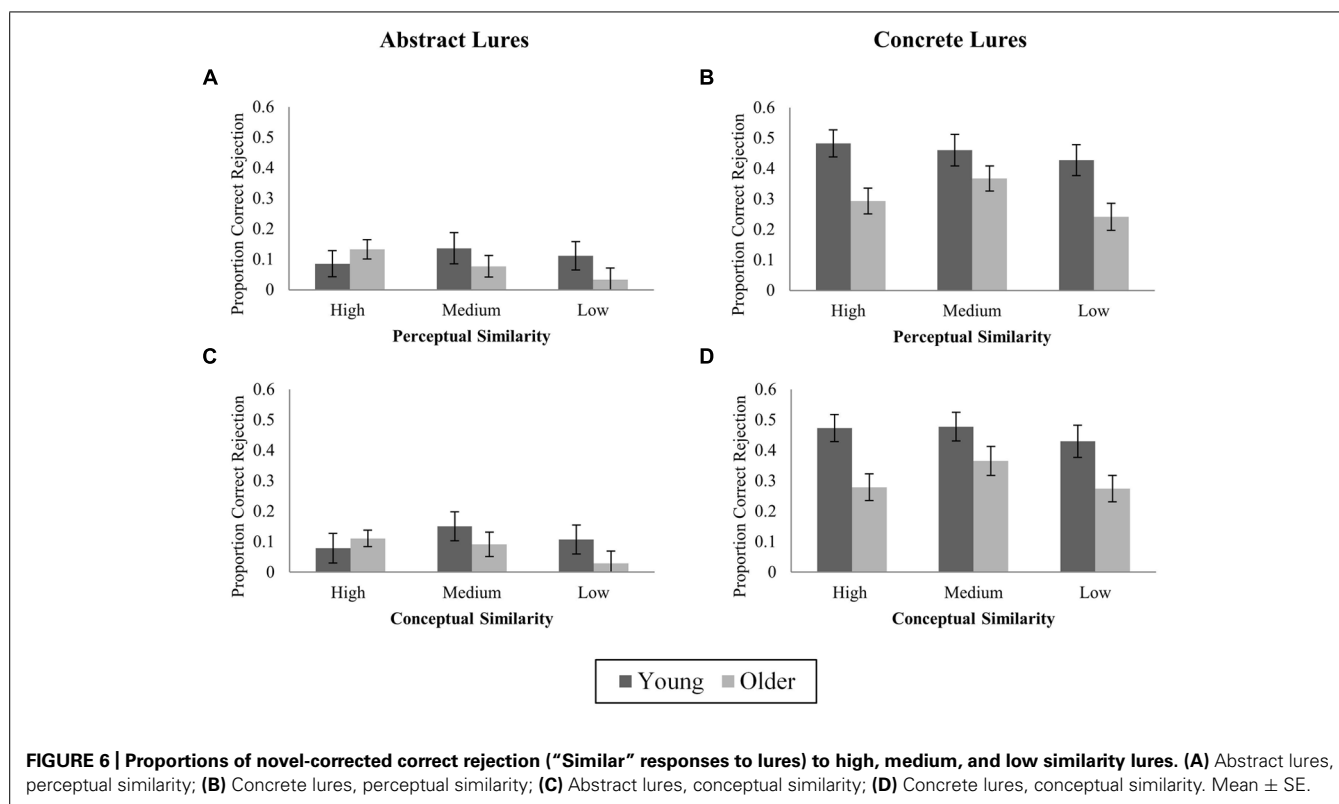
Conceptual similarity. Conceptual Similarity had a significant effect on novel-corrected false recognition of abstract lures ($F_{(2,100)} = 6.75$, $p = 0.002$, $\eta_p^2 = 0.12$). Highly conceptually similar lures were falsely recognized more often than medium or low similarity lures (high vs. medium: $t_{(51)} = 2.66$, $p = 0.01$; high vs. low: $t_{(51)} = 3.01$, $p = 0.004$). The similarity effect did not differ by age ($F_{(2,100)} = 0.20$, $p = 0.82$), and there were no age differences at any level of similarity (high: $t_{(50)} = 1.40$, $p = 0.17$; medium: $t_{(50)} = 1.12$, $p = 0.27$; low: $t_{(51)} = 1.50$, $p = 0.14$, adjusted $\alpha = 0.017$).

False recognition of concrete lures also differed according to Conceptual Similarity ($F_{(2,100)} = 26.41$, $p < 0.001$, $\eta_p^2 = 0.35$), with the most conceptually similar lures attracting the highest levels of false recognition. Crucially, Age interacted with Conceptual Similarity ($F_{(2,100)} = 3.87$, $p = 0.024$), as there was a significantly greater drop in false recognition from high to low similarity lures among OAs (24.5%) compared to YAs (11.5%; $t_{(50)} = 2.98$, $p = 0.004$; **Figure 5**). OAs showed higher false recognition than YAs for high and medium, but not low similarity lures (high: $t_{(50)} = 3.13$, $p = 0.003$; medium: $t_{(50)} = 3.17$, $p = 0.003$; low: $t_{(50)} = 0.82$, $p = 0.42$; adjusted $\alpha = 0.017$).

Conceptual Similarity did not reliably affect correct rejection of either abstract or concrete lures (abstract: $F_{(2,100)} = 1.39$, $p = 0.25$; concrete: $F_{(2,100)} = 2.12$, $p = 0.13$), and did not interact with Age for either Stimulus Type (abstract: $F_{(2,100)} = 1.73$, $p = 0.18$; concrete: $F_{(2,100)} = 0.71$, $p = 0.50$).

COMPARISON OF FALSE RECOGNITION IN EXPERIMENTS 1 AND 2

We conducted combined analyses of false recognition in Experiments 1 and 2 in order to assess whether addition of a "similar" response option altered false recognition rates, and whether age effects on abstract false recognition were robust across all participants. ANOVAs examining effects of Stimulus Type, Category Size, and Age in addition to a between-subjects factor of Experiments (1, 2) were conducted for total and for highly confident false recognition. Only significant effects involving Experiment,



or Stimulus Type and Age are reported, and age effects for abstract items alone.

Corrected false recognition was lower in Experiment 2 for concrete items only (Experiment: $F_{(1,96)} = 9.26$, $p = 0.003$; Experiment \times Stimulus Type: $F_{(1,96)} = 6.93$, $p = 0.01$; abstract: $t_{(98)} = 0.75$, $p = 0.45$; concrete: $t_{(98)} = 3.38$, $p = 0.001$). Effects of Experiment also differed by Category Size ($F_{(1,96)} = 15.18$, $p < 0.001$): false recognition was lower for large category lures in Experiment 2 than in Experiment 1, but Experiment had no effect on false recognition of single lures (large: $t_{(98)} = 3.67$, $p < 0.001$; single: $t_{(98)} = 0.12$, $p = 0.90$). None of these effects differed by Age (Experiment \times Age: $F_{(1,96)} = 0.41$, $p = 0.52$; Experiment \times Stimulus \times Age: $F_{(1,96)} = 0.76$, $p = 0.39$; Experiment \times Category Size \times Age: $F_{(1,96)} = 0.72$, $p = 0.40$). For highly confident false recognition, no effects involving Experiment were significant (max $F = 2.36$).

The three-way interactions of Stimulus Type \times Category Size \times Age were not significant across experiments for either total ($F_{(1,96)} = 1.37$, $p = 0.25$) or highly confident corrected false recognition ($F_{(1,96)} = 1.05$, $p = 0.31$). Across experiments, the predicted Stimulus Type \times Age interaction was reliable for both total and highly confident corrected false recognition (total: $F_{(1,96)} = 14.14$; confident: $F_{(1,96)} = 15.23$, $ps < 0.001$). Although larger age effects were observed for concrete lures (total: $t_{(98)} = 6.27$, $p < 0.001$, $d = 1.25$; confident: $t_{(98)} = 6.74$, $p < 0.001$, $d = 1.35$), age effects were also significant for abstract lures (total: $t_{(98)} = 2.03$, $p = 0.045$, $d = 0.41$; confident: $t_{(98)} = 3.48$, $p = 0.001$, $d = 0.70$). Stimulus Type \times Age interactions did not differ by Experiment (total: $F_{(1,96)} = 0.76$, $p = 0.39$; confident: $F_{(1,96)} = 0.08$, $p = 0.78$).

Consistent with findings from each experiment, age differences in abstract false recognition were not reliable for either total or highly confident false recognition of lures from single-item categories (total: $t_{(98)} = 1.18$, $p = 0.24$; confident: $t_{(98)} = 0.66$, $p = 0.51$), but either approached significance or were highly significant for large category abstract lures (total: $t_{(98)} = 1.77$, $p = 0.08$; confident: $t_{(98)} = 3.52$, $p = 0.001$).

DISCUSSION

The current study investigated whether age-related increases in false recognition are driven by overlapping semantic information between studied items and lures (Koutstaal et al., 2003), or whether as suggested by models of pattern separation decline, OAs are impaired in discrimination along multiple dimensions of similarity (Wilson et al., 2006; Stark et al., 2010). In two experiments, age differences in false recognition of exemplars of previously studied categories of concrete and abstract images were examined, replicating, and extending Koutstaal et al.'s (2003; Experiment 2) earlier study. Despite equivalent or greater veridical recognition of studied items, OAs were less able than YAs to discriminate between studied items and similar lures, showing heightened lure false recognition. The age-related increase in false recognition was particularly evident for concrete images, in line with results of Koutstaal et al. (2003), and as predicted by their semantic categorization account which proposes that OAs emphasize semantic information at encoding, to the detriment of item-specific information. However, in contrast with Koutstaal et al.'s (2003) findings, abstract lure false recognition was also significantly increased in older relative to YAs in both

experiments, particularly when multiple abstract category exemplars had been encountered at study. This is consistent with an age-related reduction in mnemonic discrimination across multiple dimensions of similarity, as predicted by models of pattern separation decline (e.g., Yassa et al., 2011a). We consider below how semantic gist and pattern separation views can account for these results.

Findings of larger effects of age on false recognition of concrete images are consistent with proposals that semantic overlap leads to particularly heightened false recognition in OAs (Koutstaal et al., 2003). OAs' increased false recognition of concrete relative to abstract lures was a robust finding in both experiments and was reflected in highly confident as well as overall recognition responses, again replicating Koutstaal et al.'s (2003) earlier study. In Experiment 2, the requirement to explicitly classify lures as "similar" led to reduced overall false recognition, but this effect was equivalent in both age groups (Koutstaal et al., 1999a), and importantly did not affect the magnitude or direction of the critical Age \times Stimulus Type interaction.

We found evidence of increased veridical as well as false recognition in older relative to YAs for large category concrete items, implying comparable effects of semantic similarity on both true and false recognition in older age. In this regard our findings differ from those of Koutstaal et al. (2003), who report non-significant age effects on true recognition, with trends to a slight reduction in veridical recognition in OAs. However, parallel effects of gist on veridical and false recognition are predicted by Fuzzy Trace Theory and other gist accounts, which assume that gist traces can support both veridical and false memory, but that verbatim traces are often required for veridical memory (Koutstaal and Schacter, 1997; Brainerd and Reyna, 2002; Verfaellie et al., 2002). According to gist accounts, including the semantic categorization account, encoding of multiple meaningful exemplars leads to stronger semantic gist representations (Koutstaal et al., 1999a). If OAs rely to a greater extent on gist, this should result in the tendency we observed to endorse more studied items as well as semantically similar lures as "old" on the basis of accurately recognized semantic gist. This is also consistent with evidence that intact semantic knowledge can facilitate episodic memory in older age (Reder et al., 2007; see Umanath and Marsh, 2014 for review). From a pattern separation perspective, presence of multiple overlapping representations in memory (as when multiple similar category exemplars have been encoded), results in increased likelihood of pattern completion of further similar representations, particularly in OAs (Wilson et al., 2006). This too may result in age-related increases in both true and false recognition. Unlike gist accounts however, pattern separation models as currently specified do not predict a specific impact of semantic overlap on true or false recognition.

Despite replication of Koutstaal et al.'s (2003) findings regarding differential effects of age on concrete and abstract false recognition, our findings diverge from theirs in another important respect. In the present study OAs showed reliably increased confident false recognition of abstract lures in both experiments, as well as increased total false recognition of abstract lures in Experiment 2. These effects were smaller than those among concrete items, and were restricted to abstract lures for which multiple category

exemplars were presented at study. In analyses collapsing across both experiments, the age effects were highly robust for confident responses. This difference between the two studies may be due to increased power in the current experiments (Experiment 1: $n = 24$ per age group; Experiment 2: $n = 26$, combined analysis: $n = 50$; Koutstaal et al., 2003; $n = 18$). Koutstaal et al. (2003) found numerically (though not significantly) greater large category abstract false recognition in OAs, and referred to unpublished data showing a similar pattern. It is therefore likely that our findings reflect a genuine trend toward heightened false recognition of large category abstract lures, and this suggests age-related increases in false recognition may not be entirely driven by greater reliance on semantic information. This finding presents a challenge to the semantic categorization account, and suggests that OAs may show pervasive effects of similarity on memory, consistent with a decline in pattern separation along multiple dimensions. However, the pattern separation account as currently specified does not explain why age-related differences for abstract items were restricted to large category lures.

One possibility is that OAs indeed rely more on gist representations than YAs but that these gist representations extend beyond the semantic domain. Representations of image perceptual features (e.g., overall shape, color) can form a perceptual gist representation corresponding to the average of an image's global perceptual features (Oliva, 2005; Oliva and Torralba, 2006). Perceptual similarity between exemplars of abstract categories can drive false recognition in a similar manner to semantic similarity (Koutstaal et al., 1999b; Budson et al., 2001), as can phonological similarity (Budson et al., 2003), suggesting OAs may be more vulnerable to interference from non-semantic as well as semantic gist (Koutstaal and Schacter, 1997). Our false recognition findings are consistent with this: OAs showed larger effects of category size on false recognition than YAs, across both stimulus types. However, alternative possibilities must be considered in interpreting findings for abstract lure false recognition. One is that repeated exposure to abstract category exemplars may prompt formation of a concept for these categories, i.e., they become meaningful. This is consistent with the observation that age differences in abstract false recognition emerged only for large categories. The null result for single abstract items might however, have reflected a floor effect, or a failure on the part of OAs to perceive the similarity between studied items and lures without presentation of multiple exemplars, consistent with their lower ratings of perceptual similarity among abstract items.

A related alternative is that abstract images resembled real-world objects sufficiently to drive increased false recognition directly through increased semantic categorization (Koutstaal et al., 2003). Although abstract categories were designed and normed to be novel and not conceptually meaningful (Koutstaal et al., 2003), our ratings data demonstrated that raters were able to assign conceptual labels to abstract categories to some degree, with modest inter-rater agreement in conceptual labels (0.23). These explanations are compatible with the semantic categorization account, in that semantic gist representations are either formed for, or extracted from, abstract category exemplars, and both may have contributed to current findings.

Regardless of whether non-semantic overlap played a role, enhanced veridical and false recognition for concrete items is assumed to reflect their stronger and pre-existing conceptual representations, resulting in stronger or more readily extracted gist representations (see Umanath and Marsh, 2014). Our replication of Koutstaal et al.'s (2003) finding of larger age differences among concrete items suggests that age-related increases in false recognition are driven principally by strong, pre-existing, rather than recently formed or weaker, semantic representations (Buchler and Reder, 2007). Future studies examining mnemonic discrimination in young and OAs before and after learning of membership criteria for novel categories of objects, e.g., "species" of *greebles* and *fribbles* (Gauthier and Tarr, 1997; Barry et al., 2014), compared with familiar categories, may aid in elucidating whether pre-existing conceptual representations are necessary to elicit age-related deficits in discrimination, and whether formation of a concept leads to similar patterns. It may also be informative to examine age-related differences in the relative influence of semantic and perceptual similarity on false recognition for further classes of stimuli, e.g., scenes, words, and faces (see Smith et al., 1990; Ly et al., 2013; Lee et al., 2014), particularly in light of recent evidence that OAs show a benefit of prior experience in discrimination of perceptually similar faces (Lee et al., 2014).

A caveat to the present findings is that the ratings sample judged concrete items as more perceptually similar than abstract items, unlike raters in Koutstaal et al.'s (2003) iterative procedure used to match complexity and perceptual comparability between abstract and concrete items. Although this discrepancy may be due to differences between ratings samples, it is perhaps more likely that here, despite ratings task instructions, concrete items were rated as more perceptually similar due to their being both perceptually and conceptually similar. In future studies it will be important to reduce the correlation between ratings of the two dimensions, for example using separate ratings of color and shape similarity (Konkle et al., 2010). It is important to note that older raters were not more prone to perceive category members as more similar than YAs; the only age difference in ratings was that YAs rated abstract stimuli as more perceptually similar than OAs. This implies that current findings are not attributable to OAs being less able to discriminate similar items perceptually.

As outlined in the Introduction, pattern separation models predict age-related decline in discrimination ability across multiple dimensions of similarity, but previous investigations of mnemonic discrimination using "old/similar/new" recognition tasks have typically employed meaningful stimuli or familiar shapes. Use of this task with concrete and abstract stimuli in Experiment 2 permitted examination of whether age-related reductions in lure correct rejection and increases in false recognition in this task depend on conceptual overlap. Consistent with age-related decline in discrimination across multiple dimensions, findings of greater overall correct rejection performance among YAs did not reliably differ according to stimulus type, suggesting OAs were impaired in discrimination of both concrete and abstract images. Across both levels of confidence, there were numerically larger effects of age (14 vs. 5.7%) on correct rejection of concrete lures, which may suggest a tendency to a greater age-related reduction in discrimination of meaningful items, although the Stimulus Type \times Age interaction

was not significant. However, this tendency was driven by effects for the single categories, as reflected in the 3-way interaction for highly confident correct rejection. Unlike for false recognition, OAs were equally impaired in correctly rejecting abstract and concrete lures if multiple category exemplars had been encountered. When a single category exemplar had been studied they were significantly more impaired in concrete than abstract correct rejection. However, this latter finding should be interpreted with caution as lure rejection as "similar" was at floor in both groups in the single category abstract condition. Therefore the data are not conclusive with respect to whether semantic overlap had parallel effects on lure rejection and false recognition, but as in the pattern separation studies, overall age-related differences were present for both.

Parametric measures of perceptual and conceptual similarity permitted testing of a specific prediction of the pattern separation account; that OAs require greater reduction in similarity before lures can be successfully discriminated (Wilson et al., 2006). Findings for concrete items were in line with this: OAs showed greater false recognition than the young for lures with high and medium conceptual (Experiment 1) and conceptual and perceptual (Experiment 2) similarity to studied items, while for the most distinctive lures, false recognition did not differ according to age. Although the predicted pattern dominated for false recognition, group differences were present at all levels of perceptual similarity in Experiment 1. In Experiment 2, group differences for lure correct rejection also did not follow the predicted pattern. Across both age groups, overall effects of perceptual and conceptual similarity on concrete false recognition were of comparable magnitude, but as conceptual and perceptual ratings were correlated among concrete items, it is difficult to determine whether the reduction in the effectiveness of pattern separation was driven by one or both dimensions of similarity. We also note that although planned tests followed practice in the earlier pattern separation studies (e.g., Lacy et al., 2011), effects were relatively modest, an interaction of Similarity with Age being observed only in Experiment 2. Future studies using a similar manipulation can maximize ability to detect age differences in effects across the range of possible item similarity by using a larger number of levels of input similarity (Lacy et al., 2011; Reagh et al., 2014). It would be of interest also to examine mnemonic discrimination of abstract and concrete lures parametrically varied in perceptual features such as angle of rotation or spatial location (Stark et al., 2010; Motley and Kirwan, 2012).

Lack of clear similarity effects for abstract items in both experiments may be due to a combination of substantial variance in abstract false recognition at each level of similarity (compared to concrete lures), very low rates of false recognition of single category abstract lures, and the need to combine single and large category items for similarity analyses to obtain sufficient trials in each bin. However, similarity effects for concrete items are generally consistent with the pattern separation prediction that OAs requires greater reduction in similarity in order to successfully discriminate lures from studied items. It should be noted that similarity ratings were based on raters' perception of the perceptual/conceptual similarity of all thirteen exemplars presented concurrently, whereas for participants in the recognition

experiments, representations of within-category similarity were formed gradually over the course of the study phase. It is possible that greater correspondence between subjectively rated similarity and false recognition rates, and thus clearer age differences, would be obtained if ratings were based on pairs of images and their corresponding lures.

Our findings contrast with those of Ly et al. (2013), who report age-related deficits in mnemonic discrimination of perceptually (phonologically) but not conceptually similar words. However, this apparent discrepancy may be due to use of a different measure of lure discrimination. Ly et al. (2013) examined age differences in the proportion of “new” responses to lures minus the proportion of “new” responses to studied items, a measure which, as dichotomous old/new responses were employed, could not differentiate “new” responses to lures resulting from forgetting of studied items from those resulting from their successful discrimination as similar and therefore “new.” We instead opted to examine novel-corrected false recognition and correct rejection measures which are more typically employed in studies of false recognition (e.g., Schacter et al., 1997; Abe et al., 2011) and pattern separation (e.g., Yassa et al., 2011a; Stark et al., 2013), respectively, and which are arguably more able to isolate the cognitive process under examination (i.e., unsuccessful or successful mnemonic discrimination, respectively). It in fact appears from raw results reported by Ly et al. (2013) that examination of novel-corrected false recognition would reveal a trend in the opposite direction, with OAs showing numerically increased conceptual false recognition, and perceptual false recognition reduced relative to YAs.

The current results are largely consistent with predictions derived from models of declining hippocampal pattern separation ability in older age (Wilson et al., 2006). However, we did not measure neural pattern separation directly: future neuroimaging investigations are needed to assess modulation of hippocampal and cortical functional activity by semantic and non-semantic overlap independently, and assess whether this changes with age. Converging neuroimaging studies are also essential to test predictions about the specific roles of pattern separation and completion during episodic encoding and retrieval (Yassa and Stark, 2011). This would also aid in clarifying whether our findings reflect age-related differences during initial encoding of category exemplars, during the explicit retrieval phase, or both, and in testing specific predictions that semantic categorization at encoding is more pronounced in OAs (Koutstaal et al., 2003). Although behavioral studies are relatively poor at distinguishing between encoding and retrieval effects (Fletcher et al., 1997), task manipulations unique to each phase such as those used by Koutstaal et al. (2003; Experiment 1) may also yield useful information about the locus of age-related differences.

According to dual process models of recognition, OAs are impaired in recollection, and rely to a greater extent on a general feeling of familiarity (Yonelinas, 1999). It has been proposed that strengthened gist representations lead to increased familiarity (Yonelinas, 2002; Duarte et al., 2010), and that increased gist-based false recognition with age reflects greater influence of familiarity (Koutstaal et al., 1999a; Pierce et al., 2005; Dennis et al., 2014). However, medial temporal lobe amnesics showing intact

familiarity alongside severely impaired recollection (Turriziani et al., 2008; Addante et al., 2012) have demonstrated reduced conceptual and perceptual gist-based false recognition relative to age-matched controls (Schacter et al., 1997; Verfaellie et al., 2005) suggesting gist-based false recognition is associated with recollection rather than familiarity (Dodson and Krueger, 2006). Consistent with this, conceptually driven false recognition in older age has been more strongly linked to recollection than familiarity (Schacter et al., 1997; Dennis et al., 2014). In YAs, there is increasing evidence of the importance of misrecollection in false memory paradigms (see Arndt, 2012 for review), and a recent study showed lure false recognition in a mnemonic discrimination task to be largely mediated by recollection (Kim and Yassa, 2013). Although an imperfect measure of recollection (Wixted and Mickes, 2010), confident responses are more strongly linked to recollection than familiarity (Mickes et al., 2009). In the current study, as in Koutstaal et al.'s (2003) experiment, age differences in false recognition were substantial and robust for high confidence responses as well as overall, suggesting the impact of semantic interference on age-related increases in false recognition cannot be explained by increased reliance on familiarity, and instead may be mediated by recollection.

Our results support the view that OAs are more susceptible than YAs to memory errors based on conceptual, and to a lesser extent, perceptual gist. They suggest that explicit semantic categorization cannot fully explain this effect, although it may contribute to gist formation. As indicated in the Introduction, although gist accounts assume a central role of semantic similarity in driving increased false recognition with age, while pattern separation models currently do not, the two types of account are likely complementary in other respects (Schacter et al., 1998; Yassa et al., 2011a). A synthesis between these views may explain our findings of substantial age differences in both false recognition and correct rejection of concrete lures, as well as OAs' greater sensitivity to input interference, and greater false recognition of abstract lures for which multiple perceptually similar images have been viewed. For meaningful items, pre-existing conceptual representations are assumed to be supported by traces stored in semantic memory. Wilson et al. (2006) proposed that prior memories contribute to OAs' bias toward pattern completion. The present results suggest the presence of semantic overlap between incoming and existing representations may specifically enhance this bias. This would however, not apply in the same way to abstract items, which possess no or few links to existing traces and as such may be less likely to drive pattern completion. However, if recent encoding of multiple similar abstract items permits formation of a perceptual or weak conceptual gist representation, the resulting overlapping traces may drive pattern completion, particularly among OAs. This view is consistent with the complementary learning systems account of memory, which describes the hippocampus as engaged in pattern separation, while the neocortex extracts commonalities between episodes by integrating overlap over experiences (McClelland et al., 1995; Norman and O'Reilly, 2003). As hippocampal function declines with age, the older brain may be more likely to rely on neocortical contributions to memory, emphasizing overlap with previous episodes via pattern completion (Wilson et al., 2006; see also Buchler and Reder, 2007).

CONCLUSION

In the present study, OAs showed impaired mnemonic discrimination, evidenced by reduced lure correct rejection and heightened lure false recognition. These impairments were particularly heightened when lures were conceptually as well as perceptually similar to studied items. However, increased false recognition was also observed for abstract lures for which multiple perceptually similar images had been viewed. Convergent patterns of results were observed in a typical “old/new” recognition paradigm and an “old/similar/new” recognition task. Our data support the view that OAs are particularly vulnerable to conceptual similarity, as proposed by the semantic categorization account, but are not fully consistent with this view. They also suggest that their false recognition may be increased by perceptual or conceptual gist representations formed for previously unfamiliar “abstract” items. In line with predictions that OAs require greater change in input to successfully pattern separate similar representations, age-related increases in concrete item false recognition were most likely to be observed for highly perceptually or conceptually similar lures, but OAs often performed at the same level as YAs for the most distinctive lures. Together, findings are consistent with a view that the shift in the older brain from a tendency for pattern separation toward pattern completion of input is particularly evident where strong, easily extracted similarities exist between incoming and existing traces, as in the case of frequently encountered common concepts.

ACKNOWLEDGMENTS

Laura M. Pidgeon was supported by a PhD studentship from the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), part of the cross-council Lifelong Health and Wellbeing Initiative, Grant number G0700704/84698. Alexa M. Morcom is a member of CCACE and was supported by an RCUK Academic Fellowship at the University of Edinburgh. The authors are grateful to Wilma Koutstaal for use of the picture stimuli. The authors thank Andrew McIntyre, Jamie McGhee, Katrina Rowe, and Hui-Qing. Chim for assistance with data collection.

REFERENCES

- Abe, N., Fujii, T., Nishio, Y., Iizuka, O., Kanno, S., Kikuchi, H., et al. (2011). False item recognition in patients with Alzheimer's disease. *Neuropsychologia* 49, 1897–1902. doi: 10.1016/j.neuropsychologia.2011.03.015
- Addante, R. J., Ranganath, C., and Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: evidence of accurate source recognition without recollection. *Neuroimage* 62, 439–450. doi: 10.1016/j.neuroimage.2012.04.031
- Arndt, J. (2012). False recollection: empirical findings and their theoretical implications. *Psychol. Learn. Motiv.* 56, 81–124. doi: 10.1016/b978-0-12-394393-4.00001-7
- Barry, T. J., Griffith, J. W., De Rossi, S., and Hermans, D. (2014). Meet the Fribbles: novel stimuli for use within behavioural research. *Front. Psychol.* 5:103. doi: 10.3389/fpsyg.2014.00103
- Brainerd, C. J., and Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Curr. Dir. Psychol. Sci.* 11, 164–169. doi: 10.1111/1467-8721.00192
- Buchler, N. E. G., and Reder, L. M. (2007). Modeling age-related memory deficits: a two-parameter solution. *Psychol. Aging* 22, 104–121. doi: 10.1037/0882-7974.22.1.104
- Budson, A. E., Desikan, R., Daffner, K. R., and Schacter, D. L. (2001). Perceptual false recognition in Alzheimer's disease. *Neuropsychology* 15, 230–243. doi: 10.1037/0894-4105.15.2.230
- Budson, A. E., Sullivan, A. L., Daffner, K. R., and Schacter, D. L. (2003). Semantic versus phonological false recognition in aging and Alzheimer's disease. *Brain Cogn.* 51, 251–261. doi: 10.1016/S0278-2626(03)00030-7
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educ. Psychol. Meas.* 33, 107–112. doi: 10.1177/001316447303300111
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Deese, J. (1959). Influence of inter-item associative strength upon immediate free-recall. *Psychol. Rep.* 5, 305–312. doi: 10.2466/pr0.1959.5.3.305
- Dennis, N. A., Bowman, C. R., and Peterson, K. M. (2014). Age-related differences in the neural correlates mediating false recollection. *Neurobiol. Aging* 35, 395–407. doi: 10.1016/j.neurobiolaging.2013.08.019
- Dennis, N. A., Kim, H., and Cabeza, R. (2007). Effects of aging on true and false memory formation: an fMRI study. *Neuropsychologia* 45, 3157–3166. doi: 10.1016/j.neuropsychologia.2007.07.003
- Dodson, C. S., and Krueger, L. E. (2006). I misremember it well: why older adults are unreliable eyewitnesses. *Psychon. Bull. Rev.* 13, 770–775. doi: 10.3758/bf03193995
- Duarte, A., Graham, K. S., and Henson, R. N. (2010). Age-related changes in neural activity associated with familiarity, recollection and false recognition. *Neurobiol. Aging* 31, 1814–1830. doi: 10.1016/j.neurobiolaging.2008.09.014
- Erdfeiler, E., Faul, F., and Buchner, A. (1996). GPower: a general power analysis program. *Behav. Res. Methods Instrum. Comput.* 28, 1–11. doi: 10.3758/BF03203630
- Fletcher, P. C., Frith, C. D., and Rugg, M. D. (1997). The functional neuroanatomy of episodic memory. *Trends Neurosci.* 20, 213–218. doi: 10.1016/s0166-2236(96)01013-2
- Gauthier, I., and Tarr, M. J. (1997). Becoming a “greeble” expert: exploring mechanisms for face recognition. *Vision Res.* 37, 1673–1682. doi: 10.1016/s0042-6989(96)00286-6
- Holden, H. M., Hoebel, C., Loftis, K., and Gilbert, P. E. (2012). Spatial pattern separation in cognitively normal young and older adults. *Hippocampus* 22, 1826–1832. doi: 10.1002/hipo.22017
- Kim, J., and Yassa, M. A. (2013). Assessing recollection and familiarity of similar lures in a behavioral pattern separation task. *Hippocampus* 23, 287–294. doi: 10.1002/hipo.22087
- Konkle, T., Brady, T. F., Alvarez, G. A., and Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *J. Exp. Psychol. Gen.* 139, 558–578. doi: 10.1037/a0019165
- Koutstaal, W., Reddy, C., Jackson, E. M., Prince, S., Cendan, D. L., and Schacter, D. L. (2003). False recognition of abstract versus common objects in older and younger adults: testing the semantic categorization account. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 499–510. doi: 10.1037/0278-7393.29.4.499
- Koutstaal, W., and Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *J. Mem. Lang.* 37, 555–583. doi: 10.1006/jmla.1997.2529
- Koutstaal, W., Schacter, D. L., Galluccio, L., and Stofer, K. A. (1999a). Reducing gist-based false recognition in older adults: encoding and retrieval manipulations. *Psychol. Aging* 14, 220–237. doi: 10.1037/0882-7974.14.2.220
- Koutstaal, W., Schacter, D. L., Verfaellie, M., Brenner, C., and Jackson, E. M. (1999b). Perceptually based false recognition of novel objects in amnesia: effects of category size and similarity to category prototypes. *Cogn. Neuropsychol.* 16, 317–341. doi: 10.1080/026432999380816
- Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., and Stark, C. E. L. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn. Mem.* 18, 15–18. doi: 10.1101/lm.197111
- Lee, Y., Smith, C. R., Grady, C. L., Hoang, N., and Moscovitch, M. (2014). Broadly tuned face representation in older adults assessed by categorical perception. *J. Exp. Psychol. Hum.* 40, 1060–1071. doi: 10.1037/a0035710
- Ly, M., Murray, E., and Yassa, M. A. (2013). Perceptual versus conceptual interference and pattern separation of verbal stimuli in young and older adults. *Hippocampus* 23, 425–430. doi: 10.1002/hipo.22110
- McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex – insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457. doi: 10.1037/0033-295x.102.3.419

- Mickes, L., Wais, P. E., and Wixted, J. T. (2009). Recollection is a continuous process: implications for dual-process theories of recognition memory. *Psychol. Sci.* 20, 509–515. doi: 10.1111/j.1467-9280.2009.02324.x
- Motley, S. E., and Kirwan, C. B. (2012). A parametric investigation of pattern separation processes in the medial temporal lobe. *J. Neurosci.* 32, 13076–13378. doi: 10.1523/jneurosci.5920-11.2012
- Norman, K. A., and O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* 110, 611–646. doi: 10.1037/0033-295x.110.4.611
- Norman, K. A., and Schacter, D. L. (1997). False recognition in younger and older adults: exploring the characteristics of illusory memories. *Mem. Cognit.* 25, 838–848. doi: 10.3758/bf03211328
- Oliva, A. (2005). "Gist of the Scene," in *Neurobiology of Attention*, eds L. Itti, G. Rees, and J. K. Tsotsos (San Diego, CA: Elsevier Academic Press Inc.), 251–256. doi: 10.1016/B978-012375731-9/50045-8
- Oliva, A., and Torralba, A. (2006). Building the gist of a scene: the role of global image features in recognition. *Prog. Brain Res.* 155, 23–36. doi: 10.1016/s0079-6123(06)55002-2
- Pierce, B. H., Sullivan, A. L., Schacter, D. L., and Budson, A. E. (2005). Comparing source-based and gist-based false recognition in aging and Alzheimer's disease. *Neuropsychology* 19, 411–419. doi: 10.1037/0894-4105.19.4.411
- Reagh, Z. M., Roberts, J. M., Ly, M., DiProspero, N., Murray, E., and Yassa, M. A. (2014). Spatial discrimination deficits as a function of mnemonic interference in aged adults with and without memory impairment. *Hippocampus* 24, 303–314. doi: 10.1002/hipo.22224
- Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., and Dickison, D. (2007). "Experience is a double-edged sword: a computational model of the encoding/retrieval trade-off with familiarity," in *The Psychology of Learning and Motivation*, eds B. Ross and A. S. Benjamin (Waltham, MA: Academic Press), 271–312.
- Reyna, V. E., and Brainerd, C. J. (1998). Fuzzy-trace theory and false memory: new frontiers. *J. Exp. Child Psychol.* 71, 194–209. doi: 10.1006/jecp.1998.2472
- Roediger, H. L., and McDermott, K. B. (1995). Creating false memories - remembering words not presented in lists. *J. Exp. Psychol. Learn.* 21, 803–814. doi: 10.1037/0278-7393.21.4.803
- Schacter, D. L., Norman, K. A., and Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annu. Rev. Psychol.* 49, 289–318. doi: 10.1146/annurev.psych.49.1.289
- Schacter, D. L., Verfaellie, M., and Anes, M. D. (1997). Illusory memories in amnesic patients: conceptual and perceptual false recognition. *Neuropsychology* 11, 331–342. doi: 10.1037//0894-4105.11.3.331
- Smith, A. D., Park, D. C., Cherry, K., and Berkovsky, K. (1990). Age differences in memory for concrete and abstract pictures. *J. Gerontol.* 45, 205–209. doi: 10.1093/geronj/45.5.P205
- Snodgrass, J. G., and Corwin, J. (1988). Pragmatics of measuring recognition memory - applications to dementia and amnesia. *J. Exp. Psychol. Gen.* 117, 34–50. doi: 10.1037//0096-3445.117.1.34
- Stark, S. M., Yassa, M. A., Lacy, J. W., and Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: data from healthy aging and mild cognitive impairment. *Neuropsychologia* 51, 2442–2449. doi: 10.1016/j.neuropsychologia.2012.12.014
- Stark, S. M., Yassa, M. A., and Stark, C. E. L. (2010). Individual differences in spatial pattern separation performance associated with healthy aging in humans. *Learn. Mem.* 17, 284–288. doi: 10.1101/lm.1768110
- Tolentino, J. C., Pirogovsky, E., Luu, T., Toner, C. K., and Gilbert, P. E. (2012). The effect of interference on temporal order memory for random and fixed sequences in nondemented older adults. *Learn. Mem.* 19, 251–255. doi: 10.1101/lm.026062.112
- Toner, C. K., Pirogovsky, E., Kirwan, C. B., and Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learn. Mem.* 16, 338–342. doi: 10.1101/lm.1315109
- Tun, P. A., Wingfield, A., Rosen, M. J., and Blanchard, L. (1998). Response latencies for false memories: gist-based processes in normal aging. *Psychol. Aging* 13, 230–241. doi: 10.1037//0882-7974.13.2.230
- Turriziani, P., Serra, L., Fadda, L., Caltagirone, C., and Carlesimo, G. A. (2008). Recollection and familiarity in hippocampal amnesia. *Hippocampus* 18, 469–480. doi: 10.1002/hipo.20412
- Tversky, A., and Kahneman, D. (1971). Belief in law of small numbers. *Psychol. Bull.* 76, 105–110. doi: 10.1037/h0031322
- Umanath, S., and Marsh, E. J. (2014). Understanding how prior knowledge influences memory in older adults. *Perspect. Psychol. Sci.* 9, 408–426. doi: 10.1177/1745691614535933
- Verfaellie, M., Page, K., Orlando, F., and Schacter, D. L. (2005). Impaired implicit memory for gist information in amnesia. *Neuropsychology* 19, 760–769. doi: 10.1037/0894-4105.19.6.760
- Verfaellie, M., Schacter, D. L., and Cook, S. P. (2002). The effect of retrieval instructions on false recognition: exploring the nature of the gist memory impairment in amnesia. *Neuropsychologia* 40, 2360–2368. doi: 10.1016/s0028-3932(02)00074-x
- Wechsler, D. (2001). *Wechsler Test of Adult Reading (WTAR)*. San Antonio, TX: Harcourt Assessment.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV)*. San Antonio, TX: Pearson.
- Wilson, I. A., Gallagher, M., Eichenbaum, H., and Tanila, H. (2006). Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends Neurosci.* 29, 662–670. doi: 10.1016/j.tins.2006.10.002
- Wilson, I. A., Ikonen, S., McMahan, R. W., Gallagher, M., Eichenbaum, H., and Tanila, H. (2003). Place cell rigidity correlates with impaired spatial learning in aged rats. *Neurobiol. Aging* 24, 297–305. doi: 10.1016/s0197-4580(02)00080-5
- Wixted, J. T., and Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychol. Rev.* 117, 1025–1054. doi: 10.1037/a0020874
- Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., and Stark, C. E. L. (2011a). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus* 21, 968–979. doi: 10.1002/hipo.20808
- Yassa, M. A., Mattfeld, A. T., Stark, S. M., and Stark, C. E. L. (2011b). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8873–8878. doi: 10.1073/pnas.1101567108
- Yassa, M. A., and Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends Neurosci.* 34, 515–525. doi: 10.1016/j.tins.2011.06.006
- Yeung, L.-K., Ryan, J. D., Cowell, R. A., and Barense, M. D. (2013). Recognition memory impairments caused by false recognition of novel objects. *J. Exp. Psychol. Gen.* 142, 1384–1397. doi: 10.1037/a0034021
- Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1415–1434. doi: 10.1037/0278-7393.25.6.1415
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 July 2014; accepted: 26 September 2014; published online: 17 October 2014.

Citation: Pidgeon LM and Morcom AM (2014) Age-related increases in false recognition: the role of perceptual and conceptual similarity. *Front. Aging Neurosci.* 6:283. doi: 10.3389/fnagi.2014.00283

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Pidgeon and Morcom. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Holistic face perception in young and older adults: effects of feedback and attentional demand

Bozana Meinhardt-Injac*, Malte Persike and Günter Meinhardt

Department of Psychology, Johannes Gutenberg University Mainz, Mainz, Germany

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Sven Obermeyer, Goethe University Frankfurt Am Main, Germany
David Andrew Ross, Vanderbilt, USA

*Correspondence:

Bozana Meinhardt-Injac, Section for Developmental and Educational Psychology, Department of Psychology, Johannes Gutenberg University Mainz, Wallstr. 3, 55099 Mainz, Germany
e-mail: meinharb@uni-mainz.de

Evidence exists for age-related decline in face cognition ability. However, the extents to which attentional demand and flexibility to adapt viewing strategies contribute to age-related decline in face cognition tests is poorly understood. Here, we studied holistic face perception in older (age range 65–78 years, mean age 69.9) and young adults (age range 20–32 years, mean age 23.1) using the complete design for a sequential study-test composite face task (Richler et al., 2008b). Attentional demand was varied using trials that required participants to attend to both face halves and to redirect attention to one face half during the test (high attentional demand), and trials that allowed participants to keep a pre-adjusted focus (low attentional demand). We also varied viewing time and provided trial-by-trial feedback or no feedback. We observed strong composite effects, which were larger for the elderly in all conditions, independent of viewing time. Composite effects were smaller for low attentional demand, and larger for high attentional demand. No age-related differences were found in this respect. Feedback also reduced the composite effects in both age groups. Young adults could benefit from feedback in conditions with low and high attentional demands. Older adults performed better with feedback only in trials with low attentional demand. When attentional demand was high, older adults could no longer use the feedback signal, and performed worse with feedback than without. These findings suggest that older adults tend to use a global focus for faces, albeit piecemeal analysis is required for the task, and have difficulties adapting their viewing strategies when task demands are high. These results are consistent with the idea that elderly rely more on holistic strategies as a means to reduce perceptual and cognitive load when processing resources are limited (Konar et al., 2013).

Keywords: age-related decline, holistic face perception, composite effect, attentional focus, attentional demand

1. INTRODUCTION

As it is found for other cognitive abilities, face cognition performance also undergoes age-related decline (Bartlett et al., 1989; Crook and Larrabee, 1992; Searcy et al., 1999; Pfutze et al., 2002; Chaby et al., 2003; Hildebrandt et al., 2010; Germine et al., 2011). Clearly, the well-documented decline in memory function with age may, at least partly, underlie the decline observed in face recognition tests (Fulton and Bartlett, 1991), and loss in general perceptual functions (Sekular and Sekular, 2000; Lott et al., 2005), speed limitations (Salthouse, 1996, 2000), and top-down suppression and attentional control (Gazzaley et al., 2005a,b, 2008) may contribute to these effects. Because there are multiple sources of age-related decline it is hard to judge whether impaired performance of the elderly is due to a decline in face-specific mechanisms or to impairment in general cognitive functioning, which is necessarily involved in face cognition tests.

Recent cross-sectional studies (Wilhelm et al., 2010; Hildebrandt et al., 2011) revealed that face cognition ability is predicted by non-facial general ability in memory function, speed, and object cognition with about 50% explained variance. The degree of predictability proved to be relatively stable across young, middle, and late adulthood, indicating no age-related

dedifferentiation of face and non-face cognition (Hildebrandt et al., 2011). These findings suggest that the special status of face cognition, as a set of distinct abilities, is preserved in late adulthood.

A key feature that characterizes face cognition as a distinct and highly developed ability is its holistic nature. Processing of face parts is highly sensitive for the facial context such that a change of parts usually changes the overall appearance of a face. Striking demonstrations are the “part-to-whole effect” (Tanaka and Farah, 1993; Tanaka and Sengco, 1997) and the “composite effect” (Young et al., 1987). The part-to-whole effect shows that facial features are more easily identified when they appear in their natural face contexts. The composite face effect shows that upper and lower face halves interact perceptually, and cannot be judged independently. When two composite faces are shown that combine upper and lower face halves from different persons, observers have difficulty matching the identity of only the upper or lower halves (see example in Figure 3). Meanwhile, the composite effect is frequently used to assess holistic face processing (for an overview, see Rossion, 2013).

Aging studies have addressed whether age-related decline exists in the ability to apply holistic viewing strategies for faces.

Corresponding to Wilhelm et al. (2010) and Hildebrandt et al. (2011), recent studies have corroborated that the integrative nature of face processing is not affected by aging. A study with young and older adults (mean age 68.6) found age-related decline in face identification was reported, but, no decline in the composite effect (Konar et al., 2013). Further, the composite effect predicted face identification performance to the same degree in both age groups, which indicates that the association of holistic face perception and face cognition is maintained at mature ages.

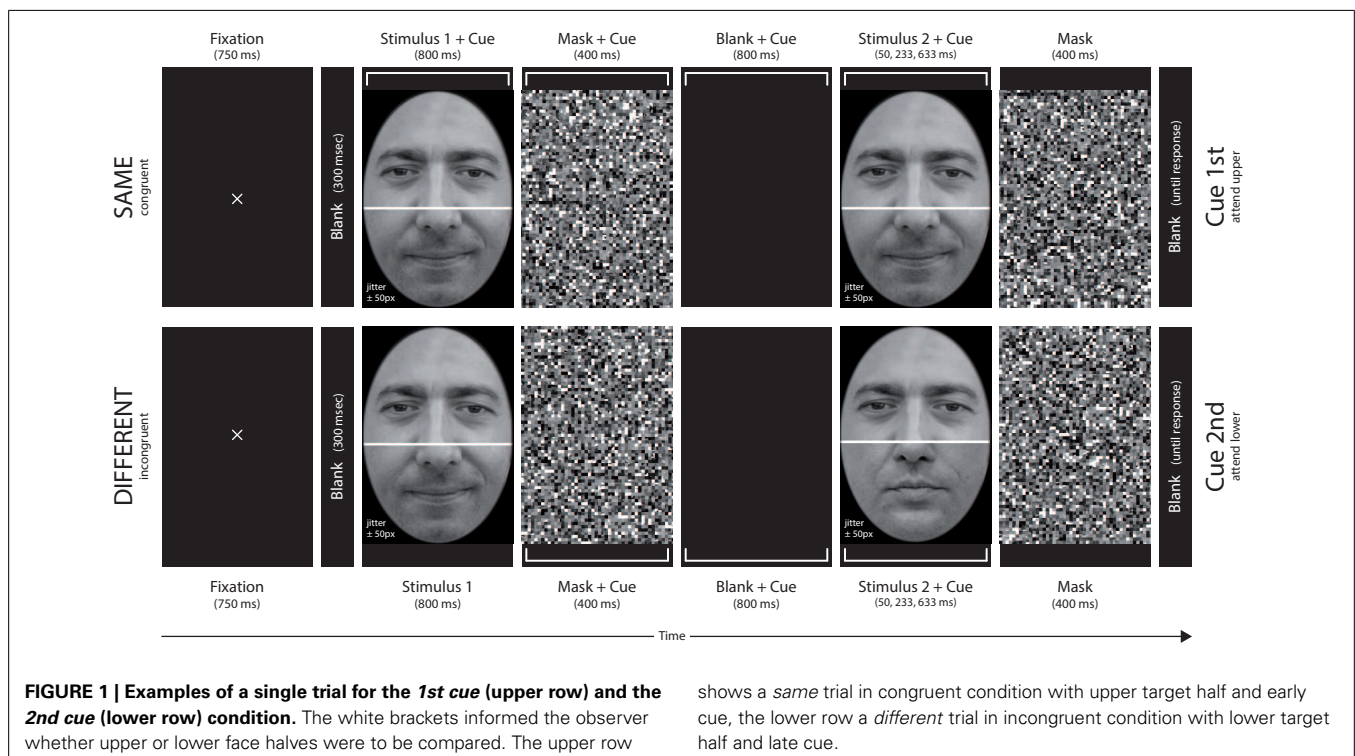
Meinhardt-Injac et al. (2014b) studied how external features modulate the perception of internal features in young and older adults (mean age 70.4), and tested the accuracy of sequentially matching faces by attending either feature class. They found about equally strong holistic effects in both age groups. However, older adults performed better with external features, while the accuracy of assessing the inner face details declined. Daniel and Bentin (2012) recorded the face specific N170 potential and the P300 component to assess global, configural and featural face-processing strategies in younger and older adults (mean age 77.1). They found that older adults processed faces by relying on global features, which shows deficits in tasks that require local configural cues. Taken together, present evidence suggests that the elderly do not suffer from deficits in the ability to process faces holistically, but may have difficulties attending diagnostic facial cues. These findings point to the role of attentional capabilities.

Ample evidence exists that attentional selection of older adults is impaired in tasks with simultaneous presentation of target and non-target stimuli (Quigley et al., 2010; Schmitz et al., 2010). Further, serious age-related deficits have been reported for tasks that require one to change attentional focus during a trial (Georgiou-Karistianis et al., 2006). However, in tests of holistic

face processing, researchers have assessed how unattended facial features affect the judgments of attended facial features. Holistic face processing is concluded from the failure to selectively attend to face parts (Richler et al., 2008b). Therefore, the sensitivity of holistic face perception should be controlled in regard to the variation of attentional task demands. If higher attentional demands yield stronger holistic effects for older adults, then their preference for global and holistic viewing strategies would, at least partly, be due to age-related decline in attentional mechanisms (see Discussion).

The effect of attentional demand can be examined in a sequential study-test composite face task by varying the temporal position of the cue that indicates which of both face halves, the upper or lower, have to be attended (see **Figure 1**). If the cue comes with the study image (**Figure 1**, upper panel) the observer can try to attend just the cued half and maintain the attentional focus throughout the trial. If the cue comes after the study image (**Figure 1**, lower panel) she/he must attend to both halves, and switch attention toward the target test half within the trial.

The effects of the unattended face halves are expected to be stronger in the late cue condition because the whole study face is attended. Both conditions differ in two respects, which are relevant for comparison among age groups. First, the late cue condition requires one to change the attentional focus from the whole study face toward only one half during the test. If age-related differences in attentional control and reallocation of resources modulate performance, the effect of cue position should be different in both age groups. Second, varying the temporal cue position alters not only attentional requirements, but also memory demands. In the late cue condition, features from the upper and lower halves must be encoded and held in memory until the



test. If working memory load is crucial for performance in the composite face task, a differential effect of cue position in both age groups should also exist because working memory capacity differs strongly among young and old adults (Brockmole and Logie, 2013). Hence, varying the temporal cue position can reveal whether age-related differences in coping with increased task demands in the composite face task.

A further aspect is the role of cognitive control that may be used to regulate the influence of the unattended on the attended facial features. Meinhardt-Injac et al. (2011) observed that young adults could enhance accuracy in judging the identities of internal features in the presence of incongruent external features by about 10% when trial-by-trial feedback about correctness is provided. This result was stable for exposure durations between 200 and 650 ms, indicating that young adults are able to replace holistic features by piecemeal face processing if they have sufficient temporal resources and the opportunity to adjust their viewing strategy with the help of feedback. For older adults, the role of feedback for optimizing the face viewing strategy, to date, has not been addressed.

The focus of the present study was threefold. First, we remeasured the composite effect for young and older adults because the current state of evidence for maintenance of holistic face perception at mature ages is not yet settled. In recent studies (Boutet and Faubert, 2006; Konar et al., 2013), the composite face effect was examined by comparing aligned and misaligned composite face arrangements, as in the seminal study on the composite effect (Young et al., 1987). However, in the last years, there was progress in the methodological development of the composite face paradigm, leading to a fully balanced and complete design (Gauthier and Bukach, 2007; Cheung et al., 2008). We decided to use this new design because of its methodological advantages (see Methods) to first add results on aging effects with the complete design to the literature. Second, we varied task demands, allowing the observer to select the attentional focus in advance and maintain it throughout the trial, or to force her/him to reallocate attentional resources during a brief time interval. Comparing across age should dismantle age-related capabilities and limitations in coping with higher task demands. Third, we provided trial-by-trial feedback, or not, to reveal whether older adults are able to use higher-level cognitive control to learn and refine their viewing strategies in the same way as young adults.

2. MATERIALS AND METHODS

2.1. EXPERIMENTAL OUTLINE

We used a variety of the sequential composite face tasks (Richler et al., 2008b). In the experimental trials, subjects first fixated on the screen center and then saw a composite study face. The image remained on the screen for 800 ms. After masking with a carefully designed mask pattern (see below), another blank screen interval followed, and then the composite test face was presented for one of three possible presentation times chosen at random. Subjects then decided by button press whether the study and test agreed or disagreed in the face halves that were being attended (upper or lower). In the first cue condition, a large white bracket marking the face half to be attended was shown with the study image. In

the second cue condition, the bracket appeared after the study image, together with its subsequent mask (see **Figure 1**).

Cue position conditions were run in separate experimental blocks because pilot measurements showed that the task was too hard for the elderly if the target cue position was varied randomly interleaved. Each experimental block was run with acoustical trial-by-trial feedback about correctness and without. Three exposure durations were chosen for the test image, one brief timing precluding saccades and serial scans (50 ms), an intermediate timing (233 ms), and a relaxed timing (633 ms) to allow for detailed image scrutiny.

2.2. EXPERIMENTAL DESIGN

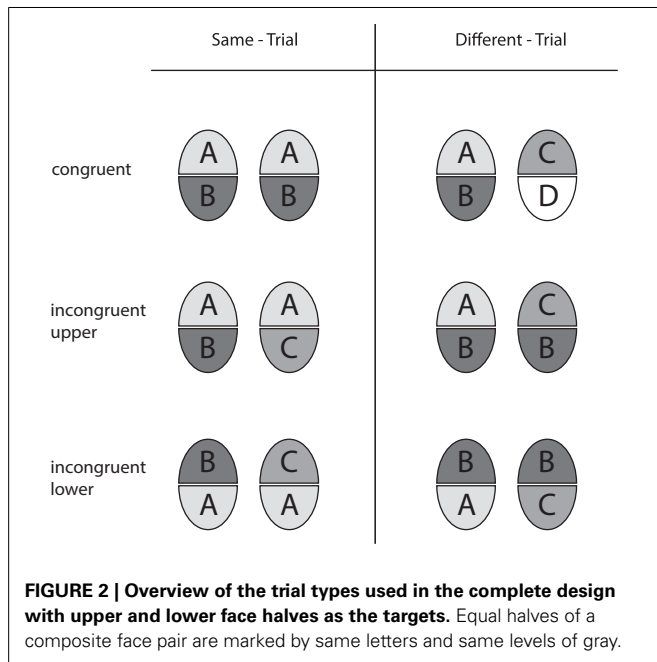
We employed the “complete design” (CD) of the composite face paradigm (Gauthier and Bukach, 2007; Cheung et al., 2008). In contrast to a former variety (called the “partial design,” PD, by Cheung and colleagues) congruent and incongruent face half pairings are fully balanced in the CD, and performance in terms of accuracy as well as holistic effects are calculated from both response categories in order to avoid confounds with a possible preference (bias) toward either response category. The design is illustrated in **Figure 2**. Same-trials and different-trials are realized in the *congruent* and the *incongruent* variety. In congruent trials, the non-attended halves agree when the attended halves agree (same-trial), and disagree when the attended ones disagree (different-trial). This means that attended and non-attended halves are *congruent* with respect to the correct decision. In incongruent trials, however, the unattended halves disagree when the attended halves agree (same-trial), and agree when the attended ones disagree (different-trial). Hence, attended and unattended halves are *incongruent* with regard to the correct decision. Holistic effects are operationally defined as *congruency effects*, reflecting the performance difference achieved in congruent and incongruent trials (see Performance Measures)¹.

2.3. STIMULI

Photographs of 20 male models were used for stimulus construction (see **Figure 3** for examples). These were frontal view shots of the whole face, captured in a professional photo studio under controlled lighting conditions. The original images were edited with Adobe Photoshop CS4 to generate the set of stimuli used in the experiment. Photographs were initially converted to 8 bit grayscale pictures and superimposed with an elliptical frame mask to obliterate all external facial features such as hair, ears, or chin line. The elliptical cutouts were then split horizontally

¹We chose the complete design because of its methodological advantages (see Cheung et al., 2008; Richler et al., 2011). In the PD same-trials are used only with incongruent pairings of upper and lower halves, while different-trials are used only with congruent relations. This means that different-trials are, in principle, easier than same-trials, and might lead to artificial composite effects because they are concluded from the relative frequency observed for erroneously classifying same-trials as different. Further, different face halves occur more frequent than same face halves, which might induce a response bias toward “different” responses. In criticism of the PD, Cheung and colleagues proposed to measure the composite effect in terms of the congruency effect, which considers the judgments for both congruency relations and both response categories. For a discussion of the design issue see Richler and Gauthier (2013) and Rossion (2013), for an alternative position.

at the bridge of the nose, thus yielding 20 upper and 20 lower face halves. Each upper half was recombined with three lower halves to constitute the final set of 60 compound faces. The cutline between the face halves was concealed with a white bar of 5 pixels thickness. It was warranted that any upper face part was never recombined with the lower half of the same original face. In addition, each of the twenty lower and upper halves appeared exactly three times in the final set of stimuli. Stimulus size was 250×350 pixels (width \times height), which corresponded to 10×12.5 cm of the screen. For each face stimulus a corresponding mask was constructed by sampling randomly ordered 5×5 pixel blocks from the face image. Masks subtended 350×450 pixels (width \times height), and covered the whole region where two subsequent face stimuli were displayed.



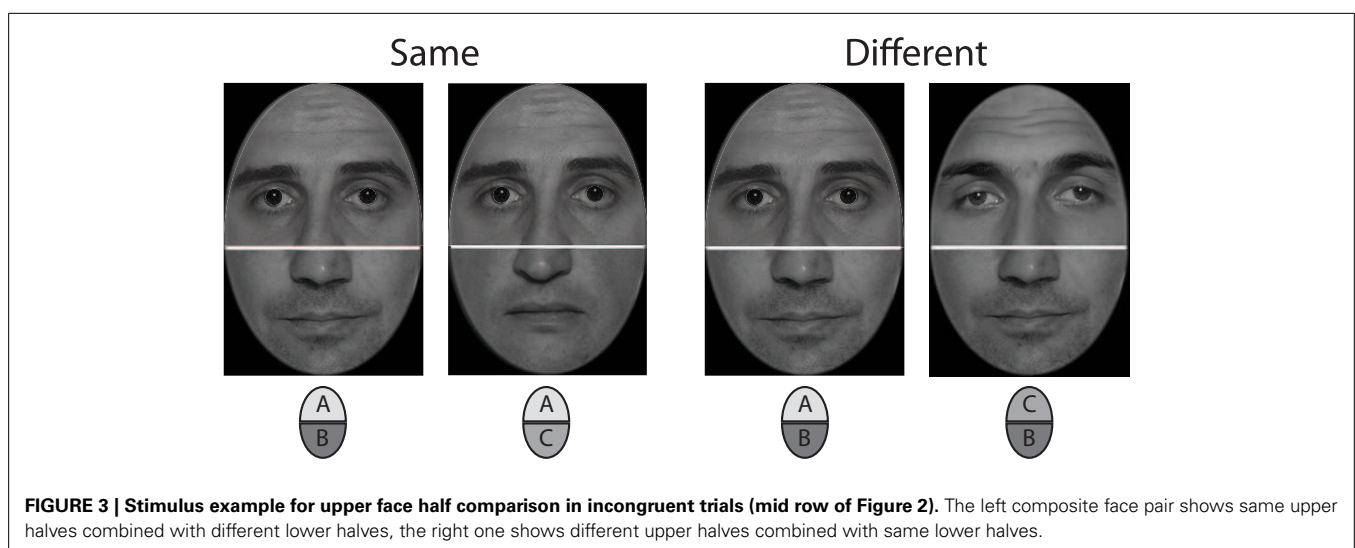
2.4. SUBJECTS

Overall, 46 young adults and 40 senior subjects participated in the present study. The two samples were halved, one group participated in the experiment with feedback, the other without feedback. All participants had normal or corrected to normal vision and reported normal neurological and psychiatric status. Senior subjects lived independent lives and were paid for participation.

The mini-mental state examination (MMSE; Folstein et al., 1975) was used to evaluate mental status. Young adult subjects were undergraduate students, 20% were male and 80% female. The mean age of the student group was 23.1 (range 20–32). These participants were given course credit points for participation, or received payment. Senior subjects were assigned to the feedback and no-feedback groups in a pseudo-random procedure with the constraint to keep the age structure of the groups equivalent. Feedback group: 20 subjects (11 female; mean age = 69.7; range 65–78 years), and No feedback group: 20 subjects (14 female; mean age = 70.1; range 65–77 years). All subjects were naive with respect to the purpose of the experiment. The study was conducted in accordance with the Declaration of Helsinki. In detail, subjects participated voluntarily and gave written informed consent to their participation. In addition, participants were informed that they were free to stop the experiment at any time without negative consequences. The data were analyzed anonymously.

2.5. APPARATUS

The experiment was executed with Inquisit runtime units. Stimuli were displayed on NEC Spectra View 2040 TFT displays in 1280×1024 resolution at a refresh rate of 60 Hz. Screen mean luminance L_0 was 100 cd/m^2 at a michelson contrast of $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min}) = 0.98$, therefore the background was practically dark (about 1.4 cd/m^2 , measured with a Cambridge Research Systems ColorCAL colorimeter). No gamma correction was used. The room was darkened so that the ambient illumination approximately matched the illumination on the screen. Stimuli were viewed binocularly at a distance



of 70 cm. Subjects used a distance marker but no chin rest throughout the experiment. Stimuli were viewed at 70 cm viewing distance. Subjects responded via an external key-pad, and wore light headphones for acoustical feedback in the feedback condition.

2.6. PREPARATION AND PRELIMINARY MEASUREMENTS

Preliminary measurements were taken with four senior subjects to assure that the task could, in principle, be executed by the elderly, and to determine the proper exposure durations for the test stimuli. Several exposure durations were probed to find a relaxed timing that allowed for maximum performance of senior subjects under the experimental conditions with the lowest attentional and perceptual demands (i.e., for the target cue with the study image, providing feedback, and for congruent trials). It turned out that senior subjects could respond to these trials with about 90% correctness at test stimulus exposure durations of half a second and longer.

Enlarging exposure duration to about a second did not increase accuracy any further. Note that 90% correct corresponded to only three errors out of 32 replications. We then presented incongruent and congruent trials mixed in random order, which did not lead to a stronger decline in accuracy for the congruent trials when exposure durations of well beyond 500 ms were used. We decided to use 633 ms (36 frames of the monitor at 60 Hz refresh rate) as the largest exposure duration.

2.7. PROCEDURE

Subjects were informed that face pairs could differ in the cued halves, but also in non-cued halves, and face half comparison was to be done for just the cued halves. They were also instructed to compare the face halves as accurately as possible, without speed pressure for the response. The temporal order of events in a trial sequence was: fixation mark (750 ms), blank (300 ms), study face stimulus (800 ms), mask (400 ms), blank (800 ms), test face stimulus (50, 233, or 633 ms), mask (400 ms), and blank frame until response (see **Figure 1**).

In the *1st cue* condition a rectangular bracket marking the target face half was shown simultaneously with the study face, and remained until the test face was masked. In the *2nd cue* condition the cue presentation began with the mask of the study face. Stimulus position jittered randomly within a region of ± 50 pixels around the center of the screen to preclude image region matching strategies between two subsequent stimulus presentations.

Young adults were made familiar with the task by going through randomly selected probe trials to ensure that the instruction was understood and could be put into practice. Senior subjects were carefully prepared for the experiment. First, the researcher explained the sequential composite face task using paper print examples of the stimulus pairings. To ensure that subjects understood the composite face task with incongruent face half pairings, the experimenter displayed paper prints of 10 stimulus pairs, and asked participants to name the five pairs showing objects with the same upper (lower) halves and five showing different upper (lower) halves. Subjects were given as

much time as needed to label the 10 pairs. If errors occurred, the experimenter adverted to the wrongly labeled pairs and drew attention to just the halves to be compared. The first minutes at the computer were spent on just congruent trials presented with the longest viewing time (633 ms), which all subjects could do with good accuracy. They then saw probe trials of the experiment with congruent and incongruent trials for about 8 min. After the preparation phase, the experimental blocks started.

Each subject went through 2 (cue position) $\times 2$ (congruency) $\times 3$ (duration) = 12 conditions. Each condition was measured with 16 same- and 16 different- trials. Eight of these 16 replications were done with upper half, and 8 with lower half as the target, resulting in 384 trials. These were subdivided into a block of 192 trials where the target cue came at the first position and a block of 192 trials where the cue came at the second position. Going through a block took about 20 min. Interleaved by a brief pause, the two blocks were administered on a single day, one with 1st cue, and one with 2nd cue, in random order across subjects.

2.8. PERFORMANCE MEASURES

Accuracy was measured in terms of the proportion of correct judgments, P_c . The rates were calculated from the frequencies of correct “same” [h_S] and correct “different” [h_D] judgments, i.e., $P_c = (h_S + h_D)/(n_S + n_D)$. With $n_S = n_D = 16$ replications per trial, each proportion correct datum rested on $n = 32$ trials. Congruency effects were calculated as the difference

$$CE = P_c(\text{congruent}) - P_c(\text{incongruent}). \quad (1)$$

Originally, Cheung et al. (2008) referred to the d' measure as a bias-free measure. We used the proportion correct measure, because, as d' , proportion correct also derives from the performance achieved for both response alternatives. However, it avoids hypothetical assumptions about sensory mapping of face stimuli, and the distribution of the corresponding sensory states. Further, it reflects task difficulty on a direct and intuitive scale.

Moreover, a direct and intuitive measure of response bias can be defined by referring to the relative frequencies for the errors of both kinds (Meinhardt-Injac et al., 2014a). For the same/different experiment the “same” response category is commonly defined as the target category (e.g., Richler et al., 2011). Accordingly, hit-rate (Hit) was defined as the rate of correctly identifying same target halves and correct rejection rate (CR) was defined as the rate of correctly identifying different target halves. False alarm rate (FA) and the rate of misses (Miss) were defined as being the complementary rates to CR and Hit, respectively. We measured response bias in terms of the error proportion, Q , which indicates which of both errors is more likely:

$$Q = \frac{\text{Miss}}{\text{Miss} + \text{FA}}. \quad (2)$$

If $Q = 0.5$, then both kinds of errors are made with the same frequency. A ratio of $Q > 0.5$ indicates a tendency to say “different” while $Q < 0.5$ indicates a preference toward “same” responses. The Q -measure has the advantage that it is easy to interpret.

For example, a value of $Q = 0.7$ means that 70% of all errors are wrong “different” responses and 30% are wrong “same” responses².

2.9. DATA ANALYSIS

The proportion correct data and the Q -measure were analyzed with ANOVA, having feedback (FB) and age group (Age) as grouping factors and cue position (Cuepos), congruency (Congru) and exposure duration (Time) as repeated measurement factors. We do not report ANOVA results for the CE measure, since the results for the difference measure are already included in the results for all interactions involving congruency at the original P_c data.

3. RESULTS

Figure 4 shows the mean proportion of correct responses as a function of exposure duration for all experimental conditions. Generally, both younger and older adults reached good accuracy levels above 90% correct at intermediate (233 ms) and large (633 ms) viewing times for congruent trials. For incongruent trials, performance did not come close to these levels, and even declined. Hence, a large congruency effect was found in all experimental conditions, which became obvious by the space between the black and gray curves.

Data analysis using ANOVA revealed main effects of exposure duration [$F_{(3, 164)} = 69.08, p < 0.001$], congruency [$F_{(1, 82)} = 191.10, p < 0.001$], cue position [$F_{(1, 82)} = 89.65, p < 0.001$], and age group [$F_{(1, 82)} = 64.04, p < 0.001$], but no main effect of feedback [$F_{(1, 82)} = 3.02 \times 10^{-4}, p = 0.986$]. We explored these effects further explored by analyzing first and higher order interactions.

3.1. EFFECTS OF FEEDBACK AND CUE POSITION

There was no main effect of feedback, and no interaction of feedback with age [$F_{(1, 82)} = 0.62, p = 0.434$]. Hence, feedback did not change the general level of performance in both age groups. However, feedback substantially modified the effect of congruency [congruency \times feedback, $F_{(1, 82)} = 10.18, p < 0.002$, see below], and the effect of cue position [cue position \times feedback, $F_{(1, 82)} = 4.66, p < 0.04$]. However, the latter was further moderated by age group [cue position \times feedback \times age group, $F_{(1, 82)} = 6.28, p < 0.02$]. **Figure 5** illustrates this interaction.

For young adults the effects of feedback were the same in both cue positions. For older adults, performance in the 2nd cue condition was disproportionately worse *with* feedback. This finding was confirmed by pairwise Fisher LSD *post-hoc* tests. Older adults did not perform significantly different in both feedback conditions at the 1st cue position ($\Delta P_c = 0.023, p = 0.373$), but significantly worse with feedback at the 2nd cue position ($\Delta P_c = 0.052, p < 0.04$). Exploring the role of trial type showed the same performance for both feedback conditions in incongruent trials ($\Delta P_c = 0.019, p = 0.582$), but worse performance with feedback in congruent trials ($\Delta P_c = 0.085, p < 0.02$; see also right panels of **Figure 4**). At the 1st cue position older adults performed

better with feedback than without in incongruent trials ($\Delta P_c = 0.064, p < 0.05$), and the same in both feedback conditions in congruent trials ($\Delta P_c = 0.018, p = 0.328$). This finding indicates a paradox effect of feedback in the old age group for the condition with high attentional demand. For young adults, the same results scheme for the effects of feedback was found for the 1st cue and the 2nd cue position. These participants performed better with feedback in incongruent trials (1st cue: $\Delta P_c = 0.041, p < 0.05$; 2nd cue: $\Delta P_c = 0.054, p < 0.04$) and the same with and without feedback in congruent trials (1st cue: $\Delta P_c = 0.015, p < 0.40$; 2nd cue: $\Delta P_c = 0.02, p = 0.516$).

Figure 5 illustrates that cue position modified performance strongly, which led to significantly lower performance in the 2nd cue condition. The effect of cue position was not modulated by age [cue position \times age group, $F_{(1, 82)} = 2.71, p = 0.104$], however, it was by age and feedback (see above). **Table 1** shows the effects of cue position for both age groups and feedback conditions and their effect sizes (Cohen's d effect size measure). The data show that cue position had large effects of comparable sizes in both age groups in the no feedback condition. Adding feedback did not affect much for young adults, but more than doubled the effect for older adults, both in the accuracy measure, and in effect size.

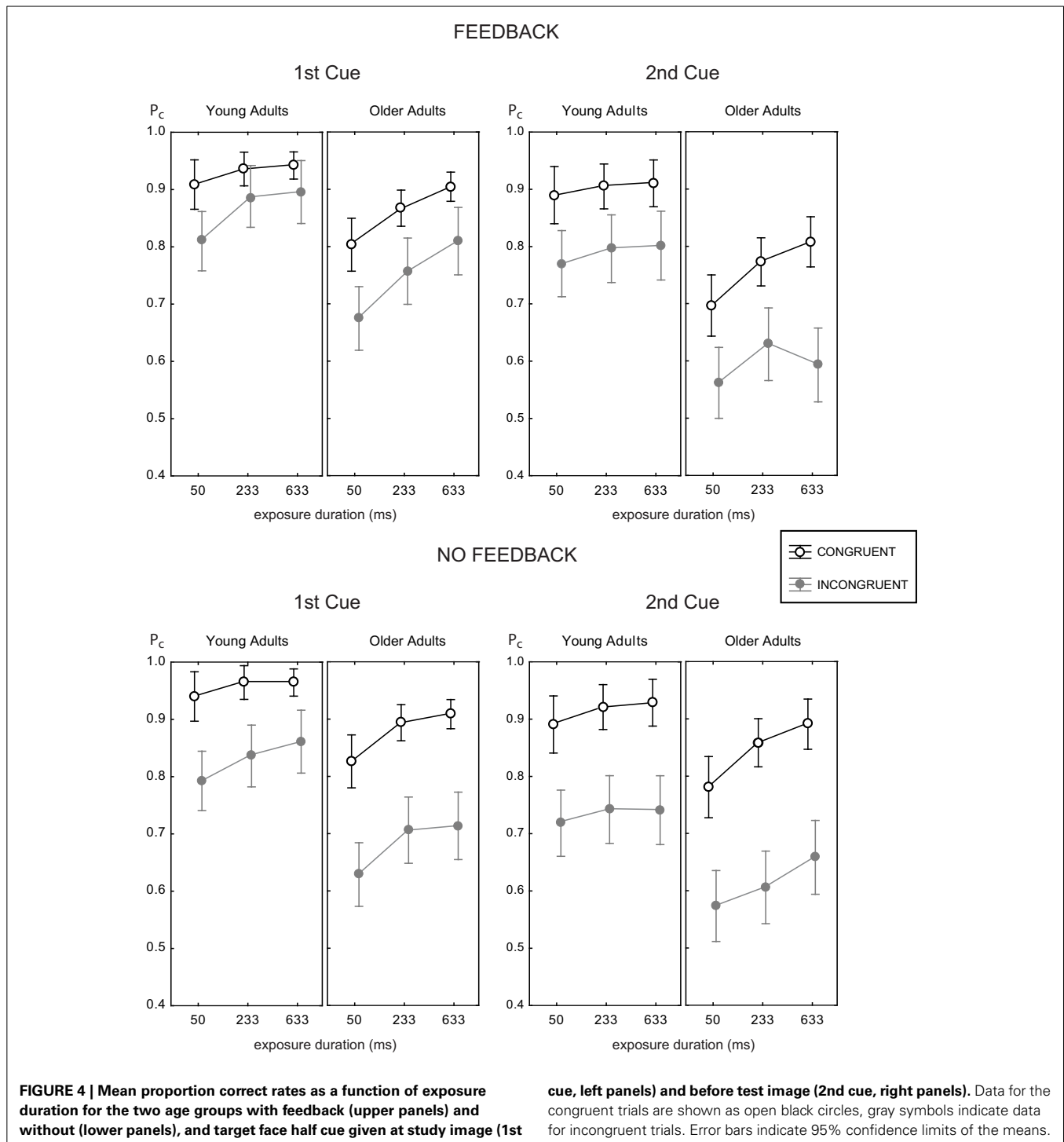
3.2. CONGRUENCY EFFECTS

Variation of the congruency relation (congruent/incongruent) among face halves strongly modulated performance. With respect to age group we found larger congruency effects for older adults [congruency \times age group, $F_{(1, 82)} = 5.34, p < 0.02$]. Comparing across age for congruent and incongruent trials separately with Fisher LSD *post-hoc* tests showed that young adults were better than older adults particularly in incongruent trials (congruent: $\Delta P_c = 0.096, p < 0.001$; incongruent: $\Delta P_c = 0.146, p < 0.001$). This finding indicates age-related differences in the ability to suppress incongruent facial context.

Feedback strongly modified the effect of congruency [congruency \times feedback, $F_{(1, 82)} = 10.18, p < 0.002$]; the congruency effect was strongly attenuated when feedback was provided, which is readily seen when the space between the black and the gray curves shown in **Figure 4** is compared among the upper and the lower data panels. Further, cue position strongly modulated the congruency effect [cue position \times congruency, $F_{(1, 82)} = 13.48, p < 0.001$], which reflects larger congruency effects for the 2nd than for the 1st cue position.

Interestingly, no higher interactions were found with age group, indicating that the congruency effect was modulated by feedback and cue position in the same way for younger and older adults [congruency \times feedback \times age group, $F_{(1, 82)} = 0.05, p = 0.819$; cue position \times congruency \times age group, $F_{(1, 82)} = 0.03, p = 0.853$]. **Table 2** lists the congruency effects of both age groups, for both cue positions and feedback conditions. The data reflect that older adults had consistently larger congruency effects than did the young adults, in the order of magnitude of 5% (see last column). The table also shows that the modulating effects of feedback and cue position on the congruency effect were the same in both age groups, and in the range of 4–6% (cue position), and 5–8% (feedback), respectively.

²Results in terms of d' and the response criterion, c , both calculated using standard formulae, can be found in the Supplementary Material.



3.3. EFFECTS OF EXPOSURE DURATION

The effect of exposure duration was different in the two age groups [exposure duration \times age group, $F_{(2, 164)} = 9.14$, $p < 0.001$], with smoothly rising performance across viewing times for young adults, while performance showed stronger improvement with viewing time for older adults. There were no time-related effects of feedback, cue position, or congruency, which indicates that all these effects were relatively constant across

exposure duration. There was time related effect that concerned the congruency effect at the two cue positions [cue position \times congruency \times exposure duration, $F_{(2, 164)} = 3.45, p < 0.04$]. This effect reflected that congruency effects tended to decline with increasing viewing time when the cue came at the 1st position, while congruency effects tended to increase with exposure duration when the cue came at the 2nd position. No age-related differences were indicated for this effect by statistical testing

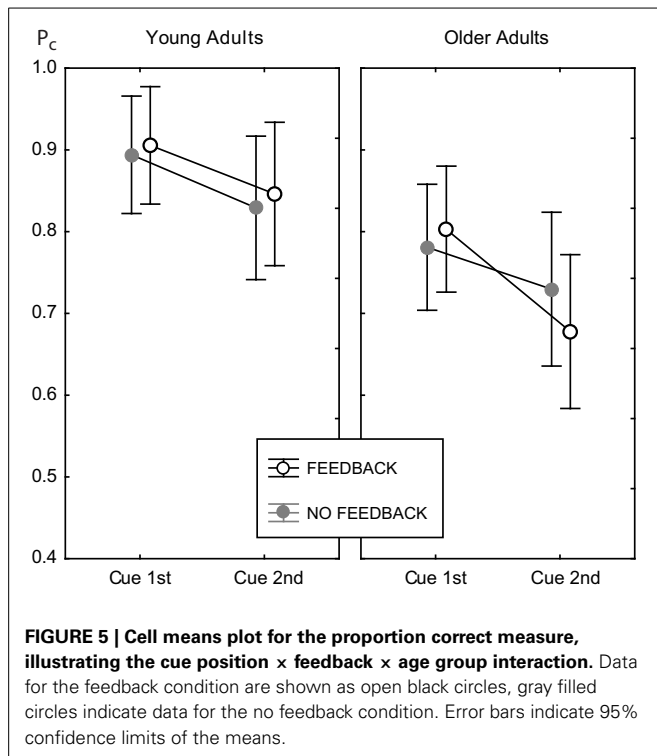


FIGURE 5 | Cell means plot for the proportion correct measure, illustrating the cue position \times feedback \times age group interaction. Data for the feedback condition are shown as open black circles, gray filled circles indicate data for the no feedback condition. Error bars indicate 95% confidence limits of the means.

Table 1 | Effects of cue position for the two age groups and feedback conditions.

Age group	Feedback	ΔP_c	$F_{(1,82)}$	p	d
Older adults	No FB	0.051	9.72	0.003	0.70
Older adults	FB	0.125	58.18	0.001	1.71
Young adults	No FB	0.065	17.94	0.001	0.88
Young adults	FB	0.059	15.02	0.001	0.81

The table shows the accuracy difference for 1st and 2nd cue position, F - value, significance level, and Cohen's d measure.

[cue position \times congruency \times exposure duration \times age group, $F_{(2,164)} = 0.29$, $p = 0.746$].

3.4. RESPONSE BIAS

Figure 6 shows the data for the Q -measure. ANOVA revealed main effects of age group [$\Delta Q = 0.08$, $F_{(1,82)} = 12.05$, $p < 0.001$], congruency [$\Delta Q = 0.09$, $F_{(1,82)} = 101.6$, $p < 0.001$], and feedback [$\Delta Q = 0.07$, $F_{(1,82)} = 9.55$, $p < 0.001$], but no effects of cue position [$F_{(1,82)} = 1.38$, $p = 0.24$] and exposure duration [$F_{(1,82)} = 0.36$, $p = 0.70$]. Young adults tended to prefer the “different” response category [$\bar{Q} = 0.53$, $CI = [0.50, 0.56]$], while older adults preferred “same” responses [$\bar{Q} = 0.45$, $CI = [0.42, 0.48]$]. The Q -measure was consistently larger in incongruent trials, compared to congruent trials [$\bar{Q}(IC) = 0.54$, $CI = [0.51, 0.56]$, $\bar{Q}(CC) = 0.45$, $CI = [0.42, 0.47]$], and also consistently larger in the no feedback condition, compared to the feedback condition [$\bar{Q}(NoFB) = 0.53$, $CI = [0.49, 0.56]$],

Table 2 | Congruency effects for the two target cue positions and the two feedback conditions, for both age groups.

Age group	Feed back	Cuepos	CE	Δ (Cuepos)	Δ (Feedback)	Δ (Age)
Older adults	No FB	Cue 2nd	0.231	0.038	0.066	0.043
Older adults	No FB	Cue 1st	0.193		0.082	0.067
Older adults	FB	Cue 2nd	0.165	0.053		0.052
Older adults	FB	Cue 1st	0.111			0.038
Young adults	No FB	Cue 2nd	0.188	0.061	0.075	
Young adults	No FB	Cue 1st	0.126		0.053	
Young adults	FB	Cue 2nd	0.112	0.039		
Young adults	FB	Cue 1st	0.073			

The table shows mean congruency effects, the difference of the mean congruency effects for both cue positions, both feedback conditions, and the difference of congruency effects across age group for the same conditions.

$\bar{Q}(FB) = 0.45$, $CI = [0.42, 0.48]$). There was a significant interaction of congruency and feedback [$F_{(1,82)} = 4.87$, $p < 0.03$], which indicated that the difference in the Q -measure for congruent and incongruent trials was stronger in the no feedback condition, compared to the feedback condition (see Figure 6). Young adults showed a strong response bias toward “different” responses in incongruent trials when there was no feedback (see Figure 6). The bias vanished when feedback was provided. Older adults did not prefer “different” responses in any experimental condition.

4. DISCUSSION

We studied holistic face perception with the complete design of the composite face paradigm, to explore the particular role of attentional demand, feedback, and viewing time, and to compare these factors across younger and older adults. Younger adults could do the study-test composite face task at brief timings (50 ms exposure duration) and at good performance levels. Older adults started at lower levels for the shortest timing, but well above chance, and reached good performance of about 90% accuracy at relaxed viewing times (633 ms).

We obtained strong congruency effects in all experimental conditions, which were consistently larger for older adults. Age-related differences were particularly pronounced for incongruent trials. However, the modulation of congruency effects by feedback and attentional demand was highly similar in both age groups. Generally, congruency effects were strongest when subjects were forced to change their attentional focus within a trial, and when no feedback was provided. A strong interaction of attentional demand, feedback, and age group was observed. Young adults could exploit trial-by-trial feedback to improve performance in incongruent trials with high and low attentional demand. Older adults could do so only in trials with low attentional demand. When participants were forced to reallocate attentional resources within a trial, performance was worse with feedback than without.

Analysis of response bias revealed a tendency of older adults toward preferring “same” responses, while young adults were

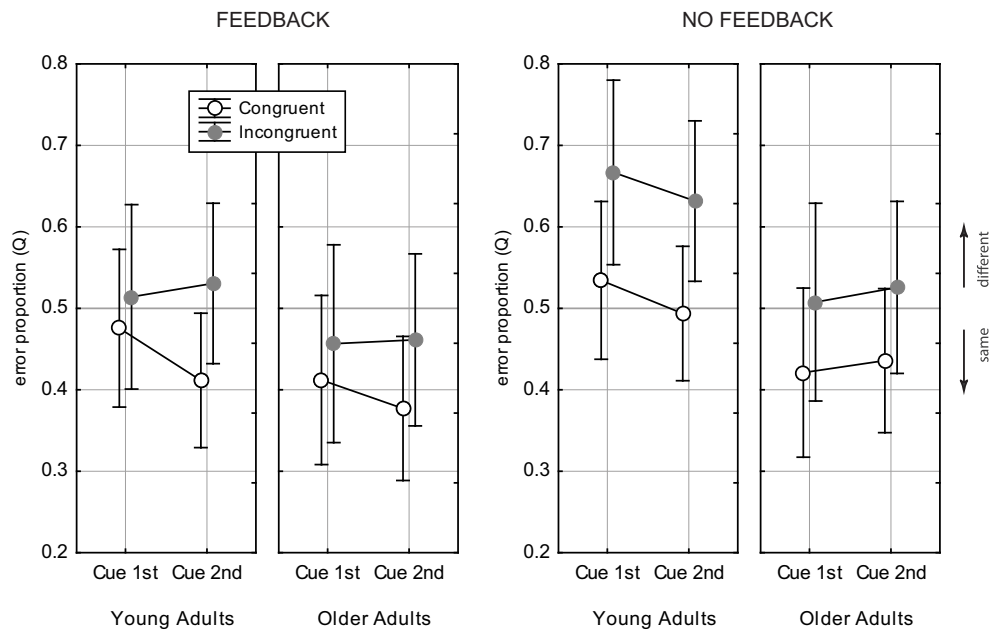


FIGURE 6 | Error proportion, Q , for assessing response bias. Data for congruent trials are shown as open black circles, gray filled circles indicate data for incongruent trials. Values of $Q > 0.5$

indicate a bias toward “different” responses, and values of $Q < 0.5$ toward “same” responses. Error bars indicate 95% confidence limits of the means.

slightly biased toward “different” responses. Feedback led toward more frequent “same” responses in both age groups.

4.1. NO AGE-RELATED DECLINE IN CONGRUENCY EFFECTS

One aim of this study was to re-examine age-related changes in the congruency effect as an important hallmark of perceptual integration in face perception. We obtained consistently larger congruency effects for the elderly in all experimental conditions. Further, the strong performance difference of congruent and incongruent trials was observed in both age groups at brief timings of 50 ms, and remained for more relaxed timings. Hence, no indication was found of age-related decline in the general capabilities to view faces holistically. In line with recent results (Daniel and Bentin, 2012; Konar et al., 2013; Meinhardt-Injac et al., 2014b), our results suggest the elderly prefer global and holistic viewing strategies, albeit part-based viewing strategies are more effective for task success.

4.2. EFFECTS RELATED TO TASK DEMANDS

Face half comparisons are more difficult in the 2nd cue condition, since a late cue enforces fast reallocation of resources (Greenwood and Parasuraman, 1999; Georgiou-Karistianis et al., 2006). A second reason for higher task difficulty in the late cue condition is enhanced demand for encoding and fast retrieval from working memory. When the cue comes with the study image observers can encode only the face half of interest and compare it to the target test half, while trying to ignore the non-target half. When the cue comes late it is not possible to proceed this way, and the observers must encode information of both halves at study.

Both attentional control and working memory are known to be affected by aging. Several studies have shown that the elderly operate much worse than young adults in tasks that require attentional switch (Lincourt et al., 1997; Greenwood and Parasuraman, 1999; Vanneste and Pouthas, 1999; Georgiou-Karistianis et al., 2006). Using Navon-like stimuli and task (Navon, 1977), Georgiou-Karistianis and colleagues showed that older adults exhibited a similar global precedence effect as young adults, but they performed worse when a switch from global to local or from local to global was required. In contrast, young adults exhibited only moderate or no switching costs. Age-related decline in working memory is a well-established finding that is substantiated by many studies (for a review, see Rajah and D’Esposito, 2005). Both the decline in working memory function and loss of attentional control can be understood within the framework of the frontal lobe hypothesis of aging (West, 1996), because divided attention, attentional and executive control, and working and episodic memory were found to be mediated by frontal brain areas (Goldman-Rakic, 1995; Cabeza et al., 1997; Fink et al., 1997; Rajah and D’Esposito, 2005; Prakash et al., 2009). From these results it can be expected that the combined effects of higher attentional demands and stronger working memory requirements in the late cue condition should disproportionately affect the performance of older adults. Interestingly, our results do *not* support a disproportionate age-related decline of performance in the 2nd cue condition.

As outlined in the Results section (see **Figure 5** and **Table 1**) the effect of cue position was the same in both age groups, as long as there was no feedback. The effect of cue position was larger for older adults *only* in the feedback condition, for specific

reasons (see below). Hence, the increase of task demands in the 2nd cue condition compared to the 1st cue condition affected performance of young and older adults to the same degrees. This finding indicates that younger and older adults handled increased task demands equally well. In view of the fact that cue position modulated task difficulty strongly, this finding is at odds with expectation from the known aging effects on working memory function and attentional control.

We also found that the effect of cue position on the congruency effect was not different for young and older adults (see Congruency Effects). Increased task demands strengthened the influence of the unattended face halves, in the same way for both age groups. The surprising fact that both performance and congruency effects of older adults were not disproportionately affected by the much higher task demands in the late cue condition points to a potential benefit of holistic encoding, which might have been used as a strategy. Holistic encoding spares the costs of divided attention to lower and upper halves at study, which precludes the effects of restricted capabilities in divided attention to become effective (Greenwood and Parasuraman, 1999). However, the encoding advantage is at the costs of having to recall the diagnostic features of just one half from a holistic representation, which results in stronger interference among target the half and incongruent non-target half. Accordingly, an increase of contextual interference for 2nd cue trials should result, which was indeed observed.

The composite face task was generally much more difficult for the elderly, as indicated by the strong main effect of age. One likely reason why face comparisons were more difficult for older adults is the use of elliptical frames that leave only the inner face parts and mask global face shape and further external features. Meinhardt-Injac et al. (2014b) used full and intact face stimuli, and had subjects attend to either the internal or external features. They found that older adults were nearly as good as young adults in comparing external features, but were much worse when internal features were the focus. This finding indicates that global face shape is a relevant face identity cue for the elderly (see below).

4.3. RESPONSE BIAS

A considerable advantage of the CD compared to the PD is that the CD is fully balanced with respect to congruency relation and the number of same and different face halves (Richler et al., 2011). Thus, the CD avoids that response bias is induced due to methodological artifacts. Analysis of response preferences can therefore reveal true age-related differences, as well as influence of experimental conditions on decision behavior. In this study we found evidence for different response behavior in both age groups, and modulatory influence of feedback and congruency relation, but no influence of task demands and exposure duration. Young adults strongly preferred the “different” response in incongruent trials when there was no feedback. The bias toward “different” responses was found in several studies using the CD (Cheung et al., 2008; Richler et al., 2008a; Gao et al., 2011), and might indicate that the difference of the wholes and the unattended parts bias the observer toward responding “different,” albeit the attended parts are same (Gao et al., 2011). Interestingly, trial-by-trial feedback canceled this effect. With the help of feedback

young observers noticed that they relied on the wrong features, and they could revise their decisional strategy. This is in line with the observation that feedback helped to improve young adults’ performance in incongruent trials. Older adults, in contrast, did not show a “different” bias in any experimental condition. While they responded “different” more often in incongruent trials, compared to congruent trials, they stayed generally biased toward “same” responses. With feedback the overall preference toward “same” responses even increased. The general “same” bias might indicate that elderly tend to overlook local diagnostic features that are crucial for facial comparisons. This is supported by earlier and recent findings which show that older adults tend to more likely identify new faces as previously seen ones (Bartlett et al., 1989; Fulton and Bartlett, 1991; Lee et al., 2014). In a recent aging study of Konar et al. (2013) no response bias was found for young and older adults. However, the authors used the PD and concluded holistic processing from the difference achieved with aligned and misaligned presentation. This might account for differences of their results and the findings of this study.

4.4. THE PARADOX EFFECT OF FEEDBACK IN THE OLDER ADULTS GROUP

Perceptual learning studies have found that feedback enables observers to revise and to optimize their viewing strategies (Herzog and Fahle, 1997, 1999). Face perception studies have also found that young adults identify diagnostic facial features and regulate the influence of irrelevant context with the help of feedback (Meinhardt-Injac et al., 2011). The results obtained here show that feedback had exactly this effect for young and older adults, as long as task demands were moderate. In the late cue condition young adults were still able to benefit from feedback, particularly in the incongruent trials. In contrast, the performance of older adults was not better with feedback in incongruent trials, while performance in the easier congruent trials declined (see Results). Seemingly, older adults were confused by the feedback signal in the late cue condition, and failed to establish a correlation of strategy revision and success. At the same time, the lower performance levels of older adults indicate that they experienced high task difficulty (see Figure 5). This finding corresponds to an interaction of task difficulty and learning observed in perceptual learning (Ahissar and Hochstein, 1997, 2004). When task difficulty is high, learning usually does not occur, even when external markers are provided. Subjects need some easy trial instances to initiate learning (“eureka effect,” see Ahissar and Hochstein, 2004). Hence, the inability to benefit from feedback in the condition with the highest task demands may indicate an interaction of learning and task difficulty for the elderly. This effect should not be over-estimated, as it is observed for the first time in the context of the composite face task. However, it would be interesting to see whether the effect is also obtained with non-face objects because older adults do not seem to apply global viewing strategies (Meinhardt-Injac et al., 2014b). As the stronger congruency effects for older adults indicate, it is adherence to global viewing strategies that is in conflict with feedback. The difficulties of elderly to replace a global viewing strategy with a more effective piecemeal strategy when task demands are high is in line with recent claims that older adults use holistic processing as a

strategy to reduce perceptual and cognitive load (Dror et al., 2005; Konar et al., 2013).

4.5. HOW DO ELDERLY LOOK AT FACES?

Looking at the composite effects for the elderly (see **Table 2**) shows that the influence of unattended face halves in the feedback condition is still as great as for young adults in the no feedback condition. Therefore, the general level of contextual influence remains high for older adults, even in conditions that are optimal for setting up a piecemeal viewing strategy.

The large global-contextual influence for older adults indicates that age-related decline in face perception does not concern mechanisms of perceptual integration. Rather, the elderly suffer from deficits when analytical processing of faces and control of facial context is required. Further evidence that face-specific processing is intact in older adults comes from the face inversion effect (FIE, Yin, 1969). Comparing across the life span, Germine et al. (2011) reported that the FIE in a face recognition task gradually increases up to ages 62 years, indicating that the experience dependent advantage of upright face processing is not lost in mature ages. Murray et al. (2010) found that elderly were much more vulnerable to face rotation than were young adults, which indicates that they strongly rely on configural information of facial features. Similar findings were reported by Creighton et al. (submitted). For older adults accuracy, response latency, and intensity rating for facial expressions of anger, happiness, fear and sadness were notably impaired when faces were turned upside down. Inversion effects for young adults were much smaller (fear, sadness) or even absent (anger, happiness).

Comparing the FIE for horizontal (eye distance) and vertical (eye-mouth distance) relational face manipulations across age, Chaby et al. (2011) observed that the FIE for vertical-relational manipulations was preserved in the elderly, while the FIE for horizontal-relational manipulations was lost. However, the overall accuracy level was lower than for young adults in detecting vertical relational changes. Obermeyer and colleagues obtained similar findings concerning age-related decline in face recognition with images that contained only horizontal spatial frequency information (Obermeyer et al., 2012). They also found a strong FIE of more than one d' unit in both age groups for this type of image manipulation. The strong FIE for vertical-relational manipulations, together with the loss of the FIE for horizontal-relational manipulations is diagnostic of the facial cues preferred by older adults. Eye distance (horizontal) is a local-relational feature judged relatively independent of facial context (Leder et al., 2001). In contrast, eye height (vertical) is defined in terms of its distance to the mouth, forehead and face outline, and is a global, long-range relational feature (Sekunova and Barton, 2008; Meinhardt-Injac et al., 2011). Chaby et al. (2011) reported a strong age-related decline in assessing local-configural facial features, while global-configural features could still be assessed. This finding is in-line with Daniel and Bentin (2012), who recorded the face specific N170 potential and the P300 component to reveal global, configural and featural face-processing strategies. Daniel and Bentin (2012) found that older adults relied on distal global information, and tended to process faces merely at the basic level of categorization until identification was required. Moreover, the

elderly did not apply configural information by default, and showed deficits in subordinate categorization (gender classification based on internal features), which strongly relies on local-configural cues. Recent results from sequential same/different tasks with whole or just part-based agreement in external and internal features showed that the elderly rely more on global shape information than do young adults, and they experience deficits in judging inner face details (Meinhardt-Injac et al., 2014b). Also the finding of a global bias toward “same” responses indicates that elderly have difficulties to focus the diagnostic features when they compare faces. These results, together with the findings of a less flexible handling of viewing strategies show that the elderly generally process faces holistically, but suffer from losses in assessing local-configural features, particularly when maintenance of attentional focus is impeded by the complexity of the visual task.

AUTHOR CONTRIBUTIONS

All authors contributed equally to the conceptualization of the study. Bozana Meinhardt-Injac set up the basic design. Malte Persike conducted the experiments and data preparation. Günter Meinhardt contributed data analysis and interpretation. All authors were involved in writing, preparation of the manuscript and final approval. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are investigated and resolved appropriately.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnagi.2014.00291/abstract>

REFERENCES

- Ahissar, M., and Hochstein, S. H. (1997). Task difficulty and the specificity of perceptual learning. *Nature* 378, 401–406. doi: 10.1038/387401a0
- Ahissar, M., and Hochstein, S. H. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464. doi: 10.1016/j.tics.2004.08.011
- Bartlett, J. C., Leslie, J. E., Tubbs, A., and Fulton, A. (1989). Aging and memory for pictures of faces. *Psychol. Aging* 4, 276–283. doi: 10.1037/0882-7974.4.3.276
- Boutet, I., and Faubert, I. (2006). Recognition of faces and complex objects in younger and older adults. *Mem. Cogn.* 34, 854–864. doi: 10.3758/BF03193432
- Brockmole, J. R., and Logie, R. H. (2013). Age-related change in visual working memory: a study of 55,753 participants aged 8–75. *Front. Psychol.* 4:12. doi: 10.3389/fpsyg.2013.00012
- Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., et al. (1997). Age-related differences in neural activity during memory encoding and retrieval: a positron emission tomography study. *J. Neurosci.* 17, 391–400.
- Chaby, L., George, N., Renault, B., and Fiori, N. (2003). Age-related changes in brain responses to personally known faces: an event-related potential (erp) study in humans. *Neurosci. Lett.* 349, 125–129. doi: 10.1016/S0304-3940(03)00800-0
- Chaby, L., Narme, P., and George, N. (2011). Older adults' configural processing of faces: role of second-order information. *Psychol. Aging* 26, 71–79. doi: 10.1037/a0020873
- Cheung, O. S., Richler, J. J., Palmeri, T. J., and Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1327–1336. doi: 10.1037/a0011752
- Crook, T. H., and Larrabee, G. J. (1992). Changes in face recognition memory across the adult life span. *J. Gerontol.* 47, 138–141. doi: 10.1093/geronj/47.3.P138
- Daniel, S., and Bentin, S. (2012). Age-related changes in processing faces from detection to identification: Erp evidence. *Neurobiol. Aging* 33, 206.e1–28. doi: 10.1016/j.neurobiolaging.2010.09.001

- Dror, I. E., Schmitz-Williams, I. C., and Smith, W. (2005). Older adults use mental representations that reduce cognitive load: mental rotation utilizes holistic representations and processing. *Exp. Aging Res.* 31, 409–420. doi: 10.1080/03610730500206725
- Fink, G., Halligan, P., Marshall, J., Frith, C., Frackowiak, R., and Dolan, R. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain* 120, 1779–1791. doi: 10.1093/brain/120.10.1779
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). “Mini-mental state.” A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.
- Fulton, A., and Bartlett, J. C. (1991). Young and old faces in young and old heads: the factor of age in face recognition. *Psychol. Aging* 6, 623–630. doi: 10.1037/0882-7974.6.4.623
- Gao, Z., Flevaris, A. V., Robertson, L. C., and Bentin, S. (2011). Priming global and local processing of composite faces: revisiting the processing-bias effect on face perception. *Attent. Percept. Psychophys.* 73, 1477–1486. doi: 10.3758/s13414-011-0109-7
- Gauthier, I., and Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition* 103, 322–330. doi: 10.1016/j.cognition.2006.05.003
- Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., and D’Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13122–13126. doi: 10.1073/pnas.0806074105
- Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., and D’Esposito, M. (2005a). Top-down enhancement and suppression of the magnitude and speed of neural activity. *J. Cogn. Neurosci.* 17, 505–517. doi: 10.1162/0898929053279522
- Gazzaley, A., Cooney, J. W., Rissman, J., and D’Esposito, M. (2005b). Top-down suppression deficit underlies working memory impairment in normal aging. *Nat. Neurosci.* 8, 1298–1300. doi: 10.1038/nn1543
- Georgiou-Karistianis, N., Tang, J., Mehmedbegovic, F., Farrow, M., Bradshaw, J., and Sheppard, D. (2006). Age-related differences in cognitive function using a global local hierarchical paradigm. *Brain Res.* 1124, 86–95. doi: 10.1016/j.brainres.2006.09.070
- Germin, L. T., Duchaine, B., and Nakayama, K. (2011). Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition* 118, 201–210. doi: 10.1016/j.cognition.2010.11.002
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6
- Greenwood, P. M., and Parasuraman, R. (1999). Scale of attentional focus in visual search. *Percept. Psychophys.* 61, 837–859. doi: 10.3758/BF03206901
- Herzog, M., and Fahle, M. (1997). The role of feedback in learning a vernier discrimination task. *Vis. Res.* 37, 2133–2141. doi: 10.1016/S0042-6989(97)00043-6
- Herzog, M., and Fahle, M. (1999). Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vision Res.* 39, 4232–4243. doi: 10.1016/S0042-6989(99)00138-8
- Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G., and Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychol. Aging* 26, 701–715. doi: 10.1037/a0023056
- Hildebrandt, A., Sommer, W., Herzmann, G., and Wilhelm, O. (2010). Structural invariance and age-related performance differences in face cognition. *Psychol. Aging* 25, 794–810. doi: 10.1037/a0019774
- Konar, Y., Bennett, P. J., and Sekuler, A. B. (2013). Effects of aging on face identification and holistic face processing. *Vision Res.* 88, 38–46. doi: 10.1016/j.visres.2013.06.003
- Leder, H., Candrian, G., Huber, O., and Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception* 30, 73–83. doi: 10.1068/p2911
- Lee, Y., Smith, C. R., Grady, C. L., Hoang, N., and Moscovitch, M. (2014). Broadly tuned face representation in older adults assessed by categorical perception. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1060–1071. doi: 10.1037/a0035710
- Lincourt, A. E., Folk, C. L., and Hoyer, W. (1997). Effects of aging on voluntary and involuntary shifts of attention. *Aging Neuropsychol. Cogn.* C4, 290–303. doi: 10.1080/13825589708256654
- Lott, L. A., Haegerstrom-Portnoy, G., Schneek, M. E., and Brabyn, J. A. (2005). Face recognition in the elderly. *Optom. Vis. Sci.* 82, 874–881. doi: 10.1097/01.opx.0000180764.68737.91
- Meinhardt-Injac, B., Persike, M., and Meinhardt, G. (2011). The context effect in face matching: effects of feedback. *Vision Res.* 51, 2121–2131. doi: 10.1016/j.visres.2011.08.004
- Meinhardt-Injac, B., Persike, M., and Meinhardt, G. (2014a). Development of visual systems for faces and objects: further evidence for prolonged development of the face system. *PLoS ONE* 9:e99942. doi: 10.1371/journal.pone.0099942
- Meinhardt-Injac, B., Persike, M., and Meinhardt, G. (2014b). Holistic processing and reliance on global viewing strategies in older adults’ face perception. *Acta Psychol.* 151, 155–163. doi: 10.1016/j.actpsy.2014.06.001
- Murray, J. E., Halberstadt, J., and Ruffman, T. (2010). The face of aging: sensitivity to facial feature relations changes with age. *Psychol. Aging* 25, 846–850. doi: 10.1037/a0019864
- Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383. doi: 10.1016/0010-0285(77)90012-3
- Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between old and young adults’ ability to recognize human faces underlie processing of horizontal information. *Front. Aging Neurosci.* 4:3. doi: 10.3389/fnagi.2012.00003
- Pfütze, E. M., Sommer, W., and Schweinberger, S. R. (2002). Age-related slowing in face and name recognition: evidence from event-related brain potentials. *Psychol. Aging* 17, 140–160. doi: 10.1037/0882-7974.17.1.140
- Prakash, R. S., Erickson, K. I., Colcombe, S. J., Kim, J. S., Voss, M. W., and Kramer, A. F. (2009). Age-related differences in the involvement of the prefrontal cortex in attentional control. *Brain Cogn.* 71, 328–335. doi: 10.1016/j.bandc.2009.07.005
- Quigley, C., Andersena, S., Schulze, L., Grunwald, M., and Müller, M. (2010). Feature-selective attention: evidence for a decline in old age. *Neurosci. Lett.* 474, 5–8. doi: 10.1016/j.neulet.2010.02.053
- Rajah, M. N., and D’Esposito, M. (2005). Region-specific changes in prefrontal function with age: a review on pet and fmri studies on working and episodic memory. *Brain* 128, 1964–1983. doi: 10.1093/brain/awh608
- Richler, J. J., Cheung, O. S., and Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychol. Sci.* 22, 464–471. doi: 10.1177/0956797611401753
- Richler, J. J., and Gauthier, I. (2013). When intuition fails to align with data: a reply to rossion. *Vis. Cogn.* 21, 1–13. doi: 10.1080/13506285.2013.796035
- Richler, J. J., Gauthier, I., Wenger, M. J., and Palmeri, T. L. (2008a). Holistic processing of faces: perceptual and decisional components. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 328–342. doi: 10.1037/0278-7393.34.2.328
- Richler, J. J., Tanaka, J. W., Brown, D. D., and Gauthier, I. (2008b). Why does selective attention to parts fail in face processing? *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1356–1368. doi: 10.1037/a0013080
- Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Vis. Cogn.* 21, 139–253. doi: 10.1080/13506285.2013.772929
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403
- Salthouse, T. A. (2000). Aging and measures of processing speed. *Biol. Psychol.* 54, 35–54. doi: 10.1016/S0301-0511(00)00052-1
- Schmitz, T., Cheng, F., and De Rosa, E. (2010). Failing to ignore: paradoxical neural effects of perceptual load on early attentional selection in normal aging. *J. Neurosci.* 30, 14750–14758. doi: 10.1523/JNEUROSCI.2687-10.2010
- Searcy, J. H., Bartlett, J. C., and Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Mem. Cogn.* 27, 538–552. doi: 10.3758/BF03211547
- Sekular, R., and Sekular, A. B. (2000). “Visual perception and cognition,” in *Oxford Textbook of Geriatric Medicine*, eds J. G. Evans, T. F. Williams, J.-P. Michel, and L. Beattie (Oxford: Oxford University Press), 874–880.
- Sekunova, A., and Barton, J. J. S. (2008). The effects of face inversion on the perception of longrange and local spatial relations in eye and mouth configuration. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 91129–91135. doi: 10.1037/0096-1523.34.5.1129
- Tanaka, J. W., and Farah, M. J. (1993). Parts and wholes in face recognition. *Q. J. Exp. Psychol.* 45, 34–79.
- Tanaka, J. W., and Sengco, J. A. (1997). Features and their configuration in face recognition. *Mem. Cogn.* 25, 583–592. doi: 10.3758/BF03211301
- Vanneste, S., and Pouthas, V. (1999). Timing in aging: the role of attention. *Exp. Aging Res.* 25, 49–67. doi: 10.1080/036107399244138
- West, R. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychol. Bull.* 120, 272–292. doi: 10.1037/0033-2909.120.2.272
- Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., and Sommer, W. (2010). Individual differences in perceiving and recognizing faces - one element of social cognition. *J. Pers. Soc. Psychol.* 90, 530–548. doi: 10.1037/a0019972

Yin, R. K. (1969). Looking at upside-down faces. *J. Exp. Psychol.* 81, 141–145. doi: 10.1037/h0027474

Young, A. M., Hellawell, D., and Hay, D. C. (1987). Configural information in face perception. *Perception* 16, 747–759. doi: 10.1068/p160747

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 May 2014; accepted: 05 October 2014; published online: 22 October 2014.

Citation: Meinhardt-Injac B, Persike M and Meinhardt G (2014) Holistic face perception in young and older adults: effects of feedback and attentional demand. *Front. Aging Neurosci.* 6:291. doi: 10.3389/fnagi.2014.00291

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Meinhardt-Injac, Persike and Meinhardt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spread of activation and deactivation in the brain: does age matter?

Brian A. Gordon¹*, Chun-Yu Tse², Gabriele Gratton³ and Monica Fabiani³

¹ Department of Radiology, Washington University in St. Louis, St. Louis, MO, USA

² Department of Psychology, Chinese University of Hong Kong, Shatin, Hong Kong

³ Department of Psychology and Beckman Institute, University of Illinois, Urbana, IL, USA

Edited by:

Harriet Ann Allen, University of Nottingham, UK

Reviewed by:

Alexa Morcom, University of Edinburgh, UK

Manuel De Vega, Universidad de La Laguna, Spain

*Correspondence:

Brian A. Gordon, Department of Radiology, Washington University in St. Louis, 660 South Euclid Avenue, Campus Box 8225, St. Louis, MO 63110, USA
e-mail: bagordon@wustl.edu

Cross-sectional aging functional MRI results are sometimes difficult to interpret, as standard measures of activation and deactivation may confound variations in signal amplitude and spread, which however, may be differentially affected by age-related changes in various anatomical and physiological factors. To disentangle these two types of measures, here we propose a novel method to obtain independent estimates of the peak amplitude and spread of the BOLD signal in areas activated (task-positive) and deactivated (task-negative) by a Sternberg task, in 14 younger and 28 older adults. The *peak* measures indicated that, compared to younger adults, older adults had increased activation of the task-positive network, but similar levels of deactivation in the task-negative network. Measures of *signal spread* revealed that older adults had an increased spread of activation in task-positive areas, but a starkly reduced spread of deactivation in task-negative areas. These effects were consistent across regions within each network. Further, there was greater variability in the anatomical localization of peak points in older adults, leading to reduced cross-subject overlap. These results reveal factors that may confound the interpretation of studies of aging. Additionally, spread measures may be linked to local connectivity phenomena and could be particularly useful to analyze age-related deactivation patterns, complementing the results obtained with standard peak and region of interest analyses.

Keywords: aging, functional magnetic resonance imaging (fMRI), task-negative network, default mode network (DMN), task-positive network, spread, activation, deactivation

INTRODUCTION

Functional MRI (fMRI) provides a powerful tool for investigating brain activity. However, inherent to many of the measures typically used for fMRI analyses, estimates of the magnitude of activation of particular voxels and of the area over which the signal spreads are conflated with one another. Changes in signal amplitude and spread may have different theoretical interpretations. While signal amplitude is supposed to reflect the degree of involvement of very precise cortical areas, signal spread may reflect the extent to which more diffuse local inhibitory or excitatory networks are involved (Tolias et al., 2005; Tehovnik et al., 2006). Importantly, separating these two properties may provide additional information to understand cognitive theories of aging. For instance, theories investigating changes in brain activity during working memory performance in aging may invoke constructs such as compensatory mechanisms that respond to increasing task difficulty (Reuter-Lorenz and Cappell, 2008; Schneider-Garcés et al., 2010), which could predict a focal increase in activity, or in turn may focus on a broad loss of specialization, or dedifferentiation, of tissue (e.g., Park et al., 2004), which may lead to an increased spread of activation. Current analysis methods do not provide an accurate way to dissociate these phenomena.

Advancing age leads to cognitive decline, even in populations of healthy older adults, and it is also characterized by altered patterns of neural activity (see Kramer et al., 2006; Park and Reuter-Lorenz,

2009; Fabiani, 2012). Such changes are typically explained as up-regulation of resources, or alternatively as the reduced suppression of distracting mental processes. Importantly, these altered functional patterns greatly depend on the networks of brain areas being considered. fMRI studies indicate that during task performance, not only are some brain areas “activated” (i.e., their blood oxygen-level dependent, or BOLD, signal is higher than that observed during a baseline period), but also that others are “deactivated” (i.e., their BOLD signal is below that observed during a baseline period). For instance, attention-demanding tasks are typically associated with activation of a set of areas encompassing dorsal and lateral frontal and parietal regions forming a dorsal attentional network (DAN), with concurrent deactivation of more medial and ventral regions (the default-mode network, DMN; e.g., Shulman et al., 1997b; Mazoyer et al., 2001; Raichle et al., 2001; Raichle and Snyder, 2007). It is not clear whether aging affects focal blood flow modulations and the spread of such activations and deactivations differently. In this paper we present a novel approach whose purpose is to disambiguate peak amplitude and spread, and show how this may help understand some of the brain activation and deactivation patterns that occur with aging.

Regions within the DAN are considered to be centrally involved in controlling attention and supporting working memory and executive functions (Corbetta and Shulman, 2002). Activation of regions within the DMN has been linked to monitoring

the environment (Raichle et al., 2001), stimulus-independent thoughts (Mason et al., 2007), self-referential thinking (Gusnard et al., 2001), social cognition (Harrison et al., 2008), and mental projection (Buckner and Carroll, 2007). Conversely, deactivations within the DMN during externally driven tasks suggest the suppression of these distracting mental processes, and a shift of resources to task-relevant processes (McKiernan et al., 2003, 2006; Binder et al., 2005; Sonuga-Barke and Castellanos, 2007; Castellanos et al., 2008). Failure to suppress the DMN is associated with attentional lapses (Weissman et al., 2006) and forgetting (Otten and Rugg, 2001; Daselaar et al., 2004). The DMN is strongly anti-correlated with attentional areas (Fox et al., 2005; Fransson, 2006; Toro et al., 2008) and the strength of this anti-correlation is predictive of behavioral performance on tasks requiring attentional control (Kelly et al., 2008). Thus, the literature suggests a diametric opposition and an active competition for attentional resources between these two networks (Fox et al., 2005, 2009; Fransson, 2006). For the purposes of this paper, and to avoid still-debated interpretation issues, we will label the DAN the “task-positive network,” and the DMN the “task-negative network.”

Interestingly, aging appears to impact these two networks differentially. Substantial evidence demonstrates that areas comprising the task-positive network are often up-regulated in older adults, especially during working memory and executive control studies (e.g., Jonides et al., 2000; Reuter-Lorenz et al., 2000; Cabeza et al., 2002; Grady et al., 2010; Schneider-Garces et al., 2010). In contrast, several studies find that during task performance older adults deactivate regions in the task-negative network to a lesser extent than younger adults (e.g., Lustig et al., 2003; Grady et al., 2006, 2010; Persson et al., 2007; Sambataro et al., 2010).

When examining such age-related differences, the vast majority of functional fMRI research focuses on the *amplitude* of the BOLD response. This is done using whole-brain group-level maps, or by measuring values extracted from a peak point in a region of interest (ROI). Group-level maps can be problematic as they are dependent upon spatial overlap across subjects. In cases where inter-subject topographic variability is high, different results can be obtained when examining subject-specific rather than group-level maps (see Feredoes and Postle, 2007). Further, there is an underlying assumption that the observed differences reflect variations in the magnitude of activation as a function of age rather than as a function of confounding variables that could be altering spatial properties of the BOLD signal (e.g., increased anatomical variability or smaller spread of activation).

Region of interest analyses provide the flexibility to extract signal change using a location specific to each subject or group. Still, it is not uncommon to implement a ROI peak-based approach that uses a fixed point for all subjects. A second concern relates to how the peak value is quantified. A common approach is to define ROIs using spherical kernels whose diameters can vary from quite small (3–4 mm) to quite large (8–10 mm). Such discrepancies between studies are worth further consideration, as these BOLD effects are not pure measures of amplitude. In standard ROI analyses the results are a product of both the amplitude of the signal and of how consistently that signal spreads through the volume that

is being sampled. If peak amplitudes are similar, but the spread varies across two populations, drastically different results could be obtained depending on a researcher's choice of ROI size for the measurement.

If the BOLD response around a peak is conceptualized as a Gaussian kernel, it contains two important characteristics—its height (amplitude) and its width (spread). While a great deal of the literature focuses on perceived age differences in amplitude measures, age effects on the spread of activation or deactivation have only been cursorily explored. This limited body of work usually finds that older adults have reduced spatial extents (D'Esposito et al., 1999; Buckner et al., 2000; Hesselmann et al., 2001; Huettel et al., 2001; Stebbins et al., 2002; Aizenstein et al., 2004) or that the extent varies across brain regions (Grady et al., 2010). Although intriguing, previous examinations have been limited by a combination of small sample sizes (Huettel et al., 2001; Aizenstein et al., 2004), using group rather than individual extents (Buckner et al., 2000; Cabeza et al., 2002; Mattay et al., 2002; Stebbins et al., 2002; Grady et al., 2010), approximate voxel-counting metrics (D'Esposito et al., 1999; Hesselmann et al., 2001; Huettel et al., 2001; Stebbins et al., 2002; Aizenstein et al., 2004; Sambataro et al., 2010), examining only one attentional network (all but Grady et al., 2010), or using qualitative rather than quantitative estimates of spatial extent (Cabeza et al., 2002; Grady et al., 2010).

In the papers that do take a quantitative approach, typically these “spatial extent” analyses are conducted by considering the number of voxels within a region that pass some statistical threshold-level of activation. With this approach the amplitude of the peak and its spread are confounded and no independent estimation is possible. In other words, the current literature suggests an age-related modification in the spatial extent of the BOLD signal, but such an effect has not been thoroughly explored. This is crucially important as such changes on an individual level would impact both group-level whole-brain maps, as well as the results obtained from peak ROI measures where the size of the kernel varies across studies.

Here we introduce a new quantitative approach to estimate the spatial extent, or spread, of the BOLD response (measured in mm), around peak activations and deactivations in task-positive and task-negative networks. This approach is based on estimating a parameter (*signal spread*) that reflects the rate of decay of the BOLD signal as a function of distance from its peak. As the decay is expressed *relative* to the peak, this measure does not confound signal amplitude with its spread. This spread measure could be conceptualized as reflecting how coherently local connections are engaged around peak areas. In other words, to the extent that the signal spreads further within modulated areas, it could be thought that local connections are more consistently engaged. Conversely, a reduction in spread may be associated with a loss or reduced consistency in local connections. In this way, measures of spread of the BOLD signal may provide information about the local connectivity within a particular region, separately from measures of peak amplitude, which instead are typically interpreted as estimates of the up- or down-regulation of a particular cortical region.

MATERIALS AND METHODS

PARTICIPANTS

The participants were 14 younger (range = 18–27; mean = 23.3; females = 6) and 28 older adults (range = 65–80; mean = 70.6; females = 12)¹. Subjects were screened for psychological and neurological problems, medications, and vision. To participate in the experiment individuals had to be cognitively unimpaired, as indicated by score at least 51 on the modified Mini-Mental Status exam (mMMSE; Mayeux et al., 1981)², and show no signs of depression on Beck’s Depression Scale (BDI; Beck et al., 1996). Participants were also administered the Vocabulary subtest of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) and the operation word span task (O-Span, Engle et al., 1999). The university’s institutional review board approved all procedures, and participants provided written informed consent. Participants were part of a larger project and a subset of these data, involving completely independent analyses from those reported here, have been presented elsewhere (Schneider-Garces et al., 2010). Demographics are presented in **Table 1**.

PROCEDURES

Subjects performed a modified version of Sternberg’s memory search task (Sternberg, 1966) with memory load varying from two to six items. Subjects saw an initial display of letters, and

then had to indicate whether a subsequently presented probe was included in the array. The letters were uppercase (B, D, F, G, H, J, M, R, and T). Corresponding lower-case probes were used to avoid a direct visual match. Each letter subtended approximately 1.4° of visual angle in the diagonal. This task and design were selected because they produce robust activation of the attentional network and deactivation of the task-negative network.

The stimuli were presented across five runs. Each run consisted of five rest intervals (20 s each) and four task blocks (48 s each) alternating with each other. Within each block subjects were presented with eight trials, each beginning with the presentation of a memory set for 3 s above a fixation-cross. A 1-s maintenance interval followed where only the fixation-cross remained on screen. The probe letter was then presented for 500 ms, followed by another 1.5-s fixation-period. During this 2-s interval, subjects indicated via button press whether the probe was new or part of the preceding memory set.

Each memory set was composed of randomly chosen letters, with the constraint that no identical letters were allowed within the same set. The probe letter was present (yes response) on 50% of the trials. Load was parametrically manipulated (2, 3, 4, 5, or 6 letters) across the five runs in either ascending (2–6) or descending (6–2) order, with each run containing only one set size, yielding a total of 32 trials per load. The random assignment to an ascending or descending order was made for counterbalancing purposes and did not significantly affect either the behavioral or BOLD results.

DATA ACQUISITION AND PREPROCESSING

Participants’ fMRI data were obtained on a Siemens Allegra 3T scanner. Data were recorded with a fast echo-planar imaging sequence with BOLD contrast (TR = 2000 ms, TE = 25 ms, flip angle = 80°, FOV = 220 mm, 64 × 64 acquisition matrix). The scans consisted of 38 slices interleaved, 3-mm-thick axial slices (3-mm in-plane resolution, 0.3-mm gap). T1-weighted anatomical scans (MPRAGE, 192 slices, 1 mm × 1 mm × 1 mm voxel size) were obtained to enable accurate anatomical coregistration.

The data were analyzed using FMRIB’s Software Library 4.1.4 (FSL; Smith, 2004; Woolrich et al., 2009). Preprocessing included motion correction, brain extraction, spatial smoothing with a Gaussian kernel of FWHM 6.0-mm, and the application of a 70-s high-pass temporal filter. Brain-extracted functional images were transformed into Montreal Neurological Institute (MNI) space through a two-stage process between the subject’s functional and T1 scans, and the subject’s T1 to the MNI template with affine transformations of 6 and 12° of freedom, respectively.

Each run was modeled as a boxcar design convolved with a gamma hemodynamic response function. Runs within a subject were combined using a fixed-effects model. Both activity positively and negatively correlated with the predicted model were considered for the current analyses. Group-level statistics were calculated using FSL with a mixed-effects design and adjusted for multiple comparisons using a cluster correction determined by $z > 2.3$ and a (corrected) cluster significance threshold of $p = 0.05$.

¹This study was part of a multi-session project on the effects of age and fitness on neurovascular coupling. Fitness was expected to be particularly important in the older adults. For this reason, more older than younger adults were recruited. See Fabiani et al. (2014).
²This version of the MMSE includes picture naming and forward and backward span in addition to standard questions.

Table 1 | Demographic, behavioral, and peak variability data.

Variable	Younger adults (N = 14)	Older adults (N = 28)	t(40) or F(1,38)
Age	23.3 (2.3)	70.6 (4.3)	46.41**
Education (years)	16.0 (1.7)	16.2 (3.4)	0.40
Modified MMSE	56.7 (1.3)	55.5 (1.3)	7.84*
Vocabulary (WAIS-R)	13.0 (2.4)	13.3 (2.4)	0.15
O-Span	23.8 (9.7)	13.4 (10.1)	2.80**
Sternberg accuracy (avg) ⁺	0.94 (0.05)	0.87 (0.08)	4.71*
Sternberg RT (avg)	926(176)	1068(168)	5.35*
Peak location variability (log)			
Task-positive	1.93 (0.09)	1.97 (0.02)	1.74
Task-negative	1.77 (0.08)	1.85 (0.08)	8.03*

Mean (SD); MMSE, modified mini-mental status exam; O-Span, Operation-span task. For the MMSE, Vocabulary, O-Span, and peak location variability age differences were tested using an ANCOVA in which gender and education were entered as covariates. Group significantly differ at * $p < 0.05$; ** $p < 0.005$.
⁺ F-statistic calculated on average Fisher transformed accuracy data.

ROI PEAK ANALYSES

A series of anatomical regions were selected to further examine the two networks (**Figure 1**). Peak coordinates from published studies (Shulman et al., 1997a; Raichle et al., 2001; Greicius et al., 2003; Fransson, 2005; Uddin et al., 2009) were placed into the Harvard-Oxford atlases included with FSL to select a series of ROIs. Subdivisions within a region (e.g., anterior and posterior superior temporal gyrus) were combined to yield a solitary ROI. The spatial pattern of resulting ROIs was highly congruent with visual depictions of default and attentional control networks published in the literature (Beckmann et al., 2005; Damoiseaux et al., 2006, 2008; Buckner and Carroll, 2007; Smith et al., 2009).

The ROIs were used to mask a subject's first level analysis. In this way it was possible to extract subject- and run-specific peaks of activation or deactivation within each ROI. Allowing the peak to vary across subjects accounts for individual anatomical variability and avoids biases that might be introduced if a singular peak location was used for all subjects (Swallow et al., 2003; Devlin and Poldrack, 2007; Feredoes and Postle, 2007).

As ROIs sizes vary within the literature, a cross-section of sphere sizes was selected to determine whether different results would have been obtained with different choices of ROI sizes. Spheres with a radius of 3, 5, and 10 mm were placed around each ROI

peak to assess percent signal change. Contrasts were such so that this assessed activation for the task-positive regions, and deactivation for the task-negative ones. These values were averaged across their respective networks. As the current focus is not on the load manipulation, values were also averaged across the five varying levels of difficulty. For each network and sphere size data were entered into an analysis of covariance (ANCOVA) controlling for years of education and gender.

SPATIAL VARIABILITY

As older adults are conceptualized as being more anatomically variable due to atrophy, intra-group consistency in spatial localization is an important issue to consider (Swallow et al., 2003; Devlin and Poldrack, 2007). Within each ROI, the average Euclidian distance between each individual's peak location in MNI space and every other subject in their group (young or old) was calculated. This captures how tightly loci of blood flow are clustered and is a measure of within group consistency in peak locations. These measures were subsequently averaged across set sizes and across all regions in a network. This resulted in every individual having one summary measure per network representing mean spatial deviation in peak locations from their respective cohort. These scores were compared between the age groups using an ANCOVA controlling for years of education and gender. Finally, a

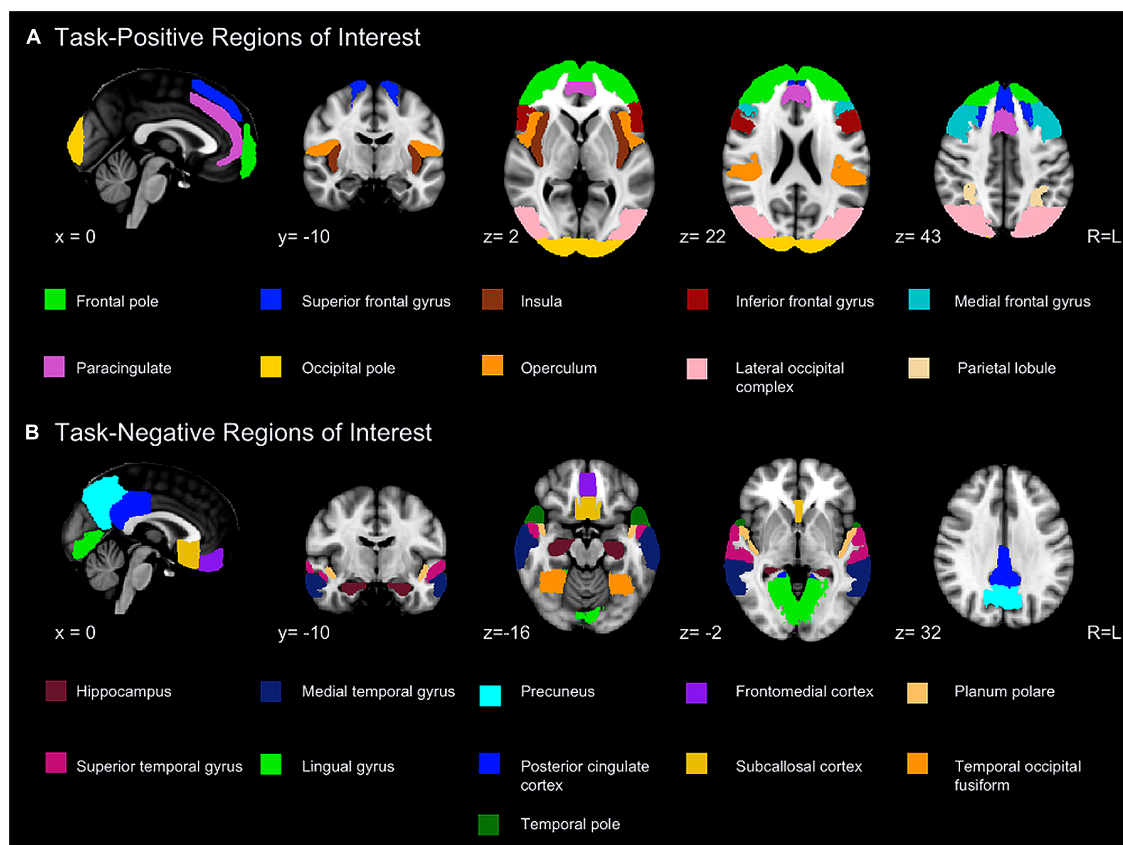


FIGURE 1 | Regions included in region of interest (ROI) analyses for the task-positive (A) and task-negative network (B).

repeated-measures ANCOVA was performed examining potential network by group interactions, also controlling for education and gender.

SPREAD OF ACTIVATION/DEACTIVATION

This measurement involved obtaining estimates of the relative amplitude of the signal at various distances from the peak point of activation or deactivation (averaged across all directions). The process of spatial smoothing acts as a spatial filter and modifies the distribution of the BOLD signal across the functional volume. As a result the BOLD data were reprocessed *without spatial smoothing*, to avoid any interaction between sphere size and smoothing kernel. Peak locations for each area and subject were then identified, and the percent signal-change values were extracted using a series of spherical ROIs with radii ranging from 3 mm (voxel size) to 10 mm, with 1-mm steps, placed around the extracted peak location. Values for the task-positive regions were obtained from the contrast positively correlated with the task, and representative of activation, while those of the task-negative regions were drawn from the negatively correlated contrast, and thus representative of deactivation. Voxels outside the brain were excluded from analysis.

Observations were then individually *normalized* for each subject, run, and ROI by dividing them by the value obtained using the 3-mm sphere. In this way subsequent measures were transformed into a proportion of the initial 3-mm sphere. This was done to account for baseline variations in the magnitude of percent signal-change data across subjects (D'Esposito et al., 2003; Ances et al., 2009). This transformation makes changes assessed with the increasing sphere sizes purely a function of the *relative* spread of the BOLD signal and therefore independent from amplitude. Values were averaged across all runs to yield one pattern per ROI per subject.

We estimated the BOLD signal changes occurring in the unique voxels added with each subsequently larger sphere (e.g., when going from 3 mm to 4 mm) using the following procedure: (a) We multiplied the volume of a sphere (e.g., $268 \text{ mm} \times 268 \text{ mm} \times 268 \text{ mm}$) by the normalized percent signal change value obtained with that sphere (e.g., 0.83) to compute the overall signal in each sphere adjusted for its size (e.g., $0.83 \times 268 = 222.44$); (b) we subtracted the overall amount of signal change obtained for a particular sphere (e.g., 4 mm) from that obtained for the next larger sphere (e.g., 5 mm); and (c) we divided the results by the difference in volume between the two. The resulting value corresponds to the average amount of activation (or deactivation) in those *unique voxels* added when going up a step from a smaller to a larger sphere, relative to the intensity of the 3 mm sphere.

When repeated across all pairs of consecutive spheres, this analysis characterizes the decay of the BOLD response as a function of distance from the peak. This curve was then fitted using a function that assumes that signal should decay proportionally to the square of the distance from the peak ($1/\text{radius}^2$) value. The slope of this relationship was considered a measure of the speed of signal “decay” around a peak point. Larger slopes correspond to a more focal pattern of activation or deactivation and faster dissipation of a signal in surrounding tissue. Conversely, smaller

values as representing a slower decay or broader “spread” of the BOLD response. As the slopes were derived from “normalized” values (i.e., values relative to the peak amplitude), they should be considered as measures of spread irrespective of peak amplitude. However, we also directly examined the degree of independence of the spread and peak measures by analyzing the amount of shared variance between the two measures. Slopes were first averaged across all regions in a network and entered into an ANOVA. If these omnibus tests were significant, individual ROIs, collapsed across hemispheres, were examined.

RESULTS

BEHAVIORAL RESULTS

Due to a response-box malfunction, full behavioral data were unavailable for four younger adults. The mean reaction time (RT), Fisher-corrected accuracy, and Cowan's K data³ were entered in repeated-measure ANCOVAs controlling for education and gender. The RT data indicated main effects of set size ($F_{4,136} = 77.80$, $p < 0.001$, $\epsilon = 0.7$) and age group ($F_{1,34} = 4.20$, $p < 0.05$), but not a set-size by age interaction ($F_{4,136} = 0.52$, *n.s.*, $\epsilon = 0.7$). Similarly, the Fisher-corrected accuracy data indicated main effects of set size ($F_{4,136} = 8.60$, $p < 0.001$, $\epsilon = 0.7$) and age ($F_{1,36} = 8.69$, $p < 0.01$), but no significant interaction ($F_{4,136} = 1.90$, $p = 0.141$, $\epsilon = 0.7$). Crucially, Cowan's K data showed a main effect of age ($F_{1,136} = 22.20$, $p < 0.0001$), indicating that span decreased with age from 5.09 in the younger adults to 3.89 in the older adults. The average RT and accuracy data are presented in **Table 1**.

MEAN ACTIVATION AND DEACTIVATION ANALYSES

The mean contrasts for the task-positive and task-negative networks are presented in **Figure 2**. Both groups produced robust activation, with foci of recruitment in prefrontal, parietal, and occipital areas. Foci of deactivation were located in medial prefrontal cortex, precuneus, posterior cingulate cortex, medial temporal lobe, and bilateral parietal cortex. These latter brain regions are representative of areas belonging to the DMN (e.g., Raichle et al., 2001; Smith et al., 2009). Older adults showed an expanded pattern of activation, but more limited patterns of deactivation localized to medial prefrontal cortex, posterior cingulate cortex, and precuneus. These results are consistent with previous work (Lustig et al., 2003; Grady et al., 2006; Persson et al., 2007).

Whole-brain analyses were expanded by using ROI peak analyses with spheres with radii of 3, 5, and 10 mm. These values were chosen to encompass a variety of sizes that may be used in typical ROI analyses in the literature. The data were collapsed across different areas within each network, keeping task-positive and task-negative regions separate, and submitted to an ANCOVA controlling for gender and years of education. The inclusion of education did not materially alter the results of these or subsequent analyses. Education was maintained within all models to be consistent with prior work in the literature.

Grand mean values for each network and sphere size, controlling for years of education and gender, are presented in

³ K was calculated according to the method reported in Schneider-Garces et al. (2010), where these analyses are reported in a more extensive fashion.

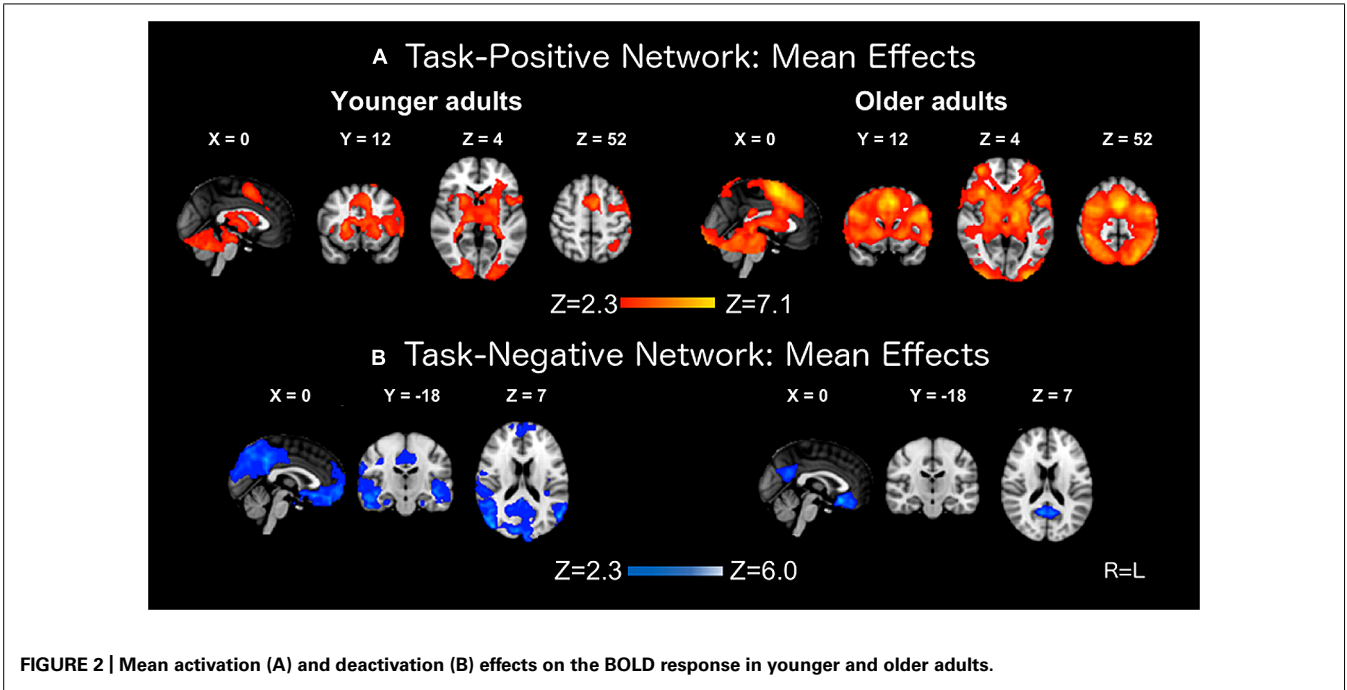


FIGURE 2 | Mean activation (A) and deactivation (B) effects on the BOLD response in younger and older adults.

Table 2. Older adults had significantly greater activation of the task-positive network areas than younger adults, as measured with the 3 mm ($F_{1,38} = 8.32, p < 0.01$), 5 mm ($F_{1,38} = 9.40, p < 0.01$), and 10 mm ($F_{1,38} = 10.30, p < 0.01$) spheres. In contrast, older adults demonstrated significantly reduced deactivation of task-negative areas compared to younger adults only when assessed with the 10 mm sphere ($F_{1,38} = 6.98, p < 0.05$) but similar levels of deactivation using the 3 mm ($F_{1,38} = 0.33, p = 0.57$) and 5 mm ($F_{1,38} = 0.08, p = 0.78$) spheres. These data provide initial indication that increases in activation in the older adults are present for both the peak locations and the immediately surrounding tissue, while decreases are located only close to peak points of deactivation and rapidly weaken in surrounding tissue. This phenomenon will be examined in greater detail with the following analyses.

SPATIAL VARIABILITY

This analysis assessed the variability of the peak point across individuals and age groups. Data for this measure, controlling for years of education and gender, are presented in **Table 1**. The distribution in space of task-evoked activation peak points was not significantly

different between the two age groups ($F_{1,38} = 1.74, p = 0.195$). Points of deactivation were significantly more variable in the older adults ($F_{1,38} = 8.03, p < 0.01$). The network by group interaction was not significant ($F_{1,38} = 1.70, p = 0.20$).

ANALYSES OF SIGNAL SPREAD

To measure signal spread, we looked at the amount of task-related activation (or de-activation) present in voxels at increasing distances from peak foci. The relative signal changes for these voxels were normalized with respect to the peak value measured with a 3 mm sphere to account for variability in amplitude across subjects. Data for task-positive and task-negative networks was averaged across ROIs with a network. Group results are presented in **Figure 3**.

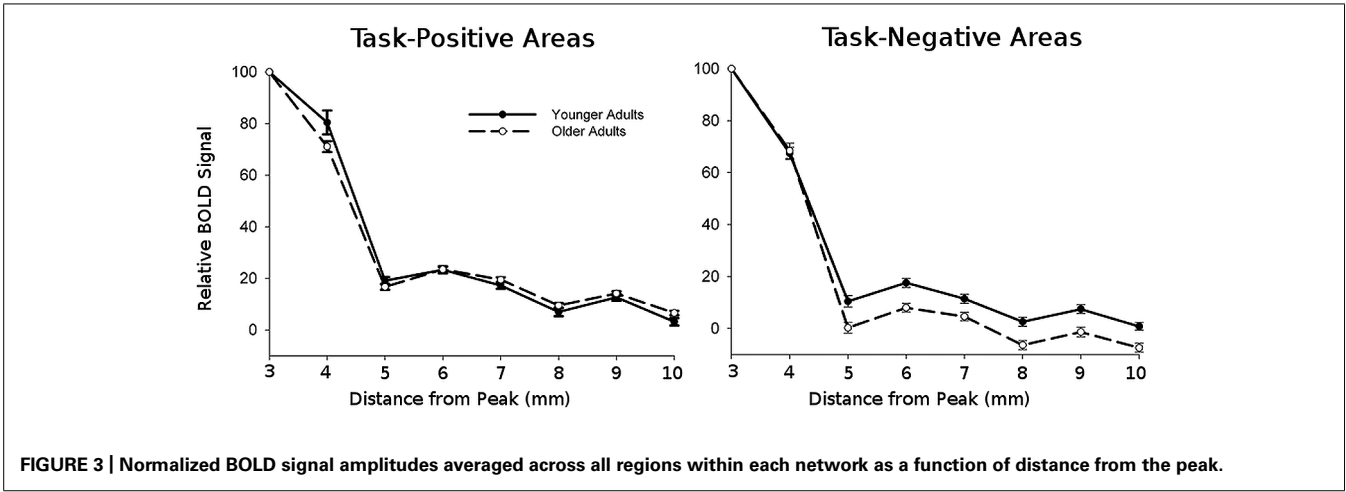
This figure shows that, as expected, the amplitude of the signal decays with distance. To quantify this decay, we fitted a quadratic decay function to the activation (or deactivation) values separately for each location and subject. The fits of this function with the data were typically good, with r 's > 0.5 in all cases⁴. The slope of this

⁴We also fitted other functions – cubic, exponential — with similar, but slightly lower, fits.

Table 2 | Mean percent change values observed using 3 mm, 5 mm, and 10 mm spheres for both the task-positive and task-negative networks.

	3 mm sphere		5 mm sphere		10 mm sphere	
	Positive*	Negative	Positive**	Negative	Positive*	Negative*
Old	2.20	−1.35	1.49	−0.780	0.641	−0.14
Young	1.57	−1.25	1.02	−0.812	0.378	−0.31

Groups significantly differ at * $p < 0.05$ ** $p < 0.005$.



function indicates the decay/spread of the signal around the area of peak; larger values represent a more focal spread and thus a more rapid decay. All the statistical analyses were then conducted on these slope estimates. For ease of presentation, we labeled the slope of the quadratic function as “spread.”

The decay parameters for each task-positive and task-negative area, averaged across subjects separately for younger and older adults, are presented in **Figure 4** (bottom row) and **Table 3**. The values presented are estimated grand means for task-positive and task-negative networks. The omnibus ANCOVA (controlling

for education and gender) performed on data averaged across all areas and networks revealed a significant group by network interaction ($F_{1,38} = 17.82, p < 0.001$). Separate planned analyses for task-positive ROIs indicated a main effect of group ($F_{1,38} = 6.99, p < 0.05$), with the signal decaying faster in younger than older adults (left bottom graph in **Figure 4**). The opposite was true for task-negative areas (right bottom graph in **Figure 4**); the average signal decayed faster (i.e., spread less) in older than in younger adults ($F_{1,38} = 13.82, p < 0.001$). For both task-positive and task-negative areas,

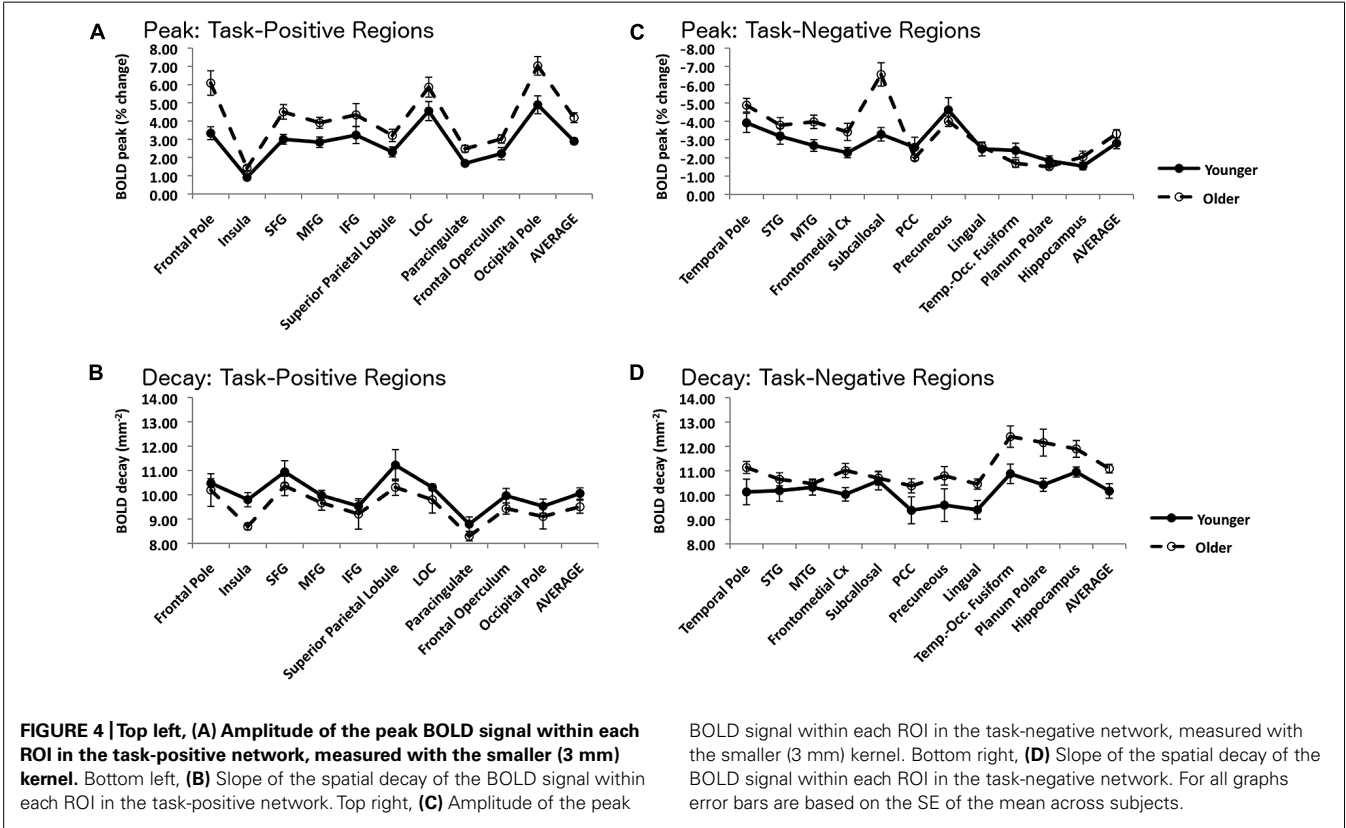


Table 3 | Decay slopes for each ROI, and probability of *t*-test of the decay functions for younger and older adults.

	Decay Young	Decay Old	<i>p</i> -value
Task positive regions			
Frontal pole	10.477	10.193	0.319
Insula	9.797	8.697	0.006
Superior frontal gyrus	10.941	10.363	0.255
Middle frontal gyrus	9.971	9.665	0.365
Inferior frontal gyrus	9.533	9.209	0.245
Superior parietal lobule	11.218	10.305	0.126
Lateral occipital complex	10.303	9.795	0.116
Paracingulate cortex	8.791	8.290	0.097
Frontal operculus	9.960	9.432	0.142
Occipital pole	9.530	9.103	0.136
Task negative regions			
Temporal pole	10.131	11.135	0.014
Superior temporal gyrus	10.185	10.650	0.267
Middle temporal gyrus	10.321	10.473	0.626
Fronto-medial cortex	10.033	11.012	0.074
Subcallosal cortex	10.587	10.700	0.795
Posterior cingulate cortex	9.378	10.382	0.030
Precuneous	9.591	10.793	0.038
Lingual gyrus	9.398	10.456	0.005
Temporooccipital gyrus	10.875	12.403	0.023
Planum polare	10.425	12.157	0.042
Hippocampus	10.949	11.896	0.113

there was also a significant effect of area within a network (respectively, $F_{9,360} = 4.00$, $p < 0.0001$, and $F_{10,400} = 4.125$, $p < 0.0001$).

As the omnibus test was significant, a series of analyses examining individual ROIs within the task-positive network was performed. All regions demonstrated similar directional trends, with the younger adults having a more pronounced decay than the older adults, although this effect reached significance only in the insula ($F_{1,38} = 9.00$, $p < 0.05$). The opposite was true for task-negative areas; the average signal decayed faster (i.e., spread less) in older than in younger adults. This effect was significant in several regions, including the temporal pole ($F_{1,38} = 6.49$, $p < 0.05$), the posterior cingulate cortex (PCC; $F_{1,38} = 5.69$, $p < 0.05$), the precuneus ($F_{1,38} = 4.29$, $p < 0.05$), the lingual gyrus ($F_{1,38} = 9.28$, $p < 0.005$) and the planum polare ($F_{1,38} = 4.40$, $p < 0.05$), whereas it was marginal in fronto-medial ($F_{1,38} = 4.02$, $p = 0.052$) and temporo-occipital cortex ($F_{1,38} = 2.92$, $p = 0.10$). Thus, in general, the activation signal decayed more slowly and the deactivation signal decayed faster in older compared to younger adults.

To provide a more intuitive idea of the significance of these phenomena, we also computed the signal spread in volumetric terms. To this end, we estimated the distance at which the

signal decays by 50%, and then computed the associated volume of signal spread, separately for each subject and brain region. This transformation indicates that the signal spreads to a volume that is 8.8% bigger in task-positive regions [12.45 vs. 11.44 cubic mm, $t(40) = 2.667$, $p < 0.02$] and 10.7% smaller in task-negative regions [9.98 vs. 11.17 cubic mm, $t(40) = -3.48$, $p < 0.002$] in the older compared to the younger adults.

Comparison of peak and decay/spread measures

The graphs in the top row of **Figure 4** show the peak measures obtained with the 3-mm sphere for each ROI and network. Note that this sphere size was chosen to separate the effects of peak and spread, which are confounded when using larger spheres. Note also that for the task-negative regions the most negative peak point was chosen. These graphs indicate that for the task-positive network (top left) there was a similar pattern across ROIs, with the older adults showing significantly larger peaks in the frontal pole ($F_{1,40} = 7.78$, $p < 0.01$), insula ($F_{1,40} = 7.45$, $p < 0.01$), superior frontal gyrus ($F_{1,40} = 6.33$, $p < 0.05$), middle frontal gyrus ($F_{1,40} = 5.08$, $p < 0.05$), paracingulate cortex ($F_{1,40} = 7.89$, $p < 0.01$) and occipital pole ($F_{1,40} = 7.06$, $p < 0.05$). There was also a trend in the same direction in superior parietal cortex ($F_{1,40} = 3.10$, $p < 0.10$) and frontal operculum ($F_{1,40} = 4.07$, $p < 0.10$). For the task-negative ROIs, however, the results were less consistent, with two regions showing a larger (negative) peak for older adults (subcallosal cortex: $F_{1,40} = 12.06$, $p < 0.005$; MTG: $F_{1,40} = 5.21$, $p < 0.05$), while several others showing trends in the opposite direction.

In order to test the utility of using spread measures in addition to standard measures of peak we entered peak and spread measures as simultaneous predictors in a multiple regression analysis, using age as the criterion variable. For both the task-positive and task-negative networks, the overall multiple regression results were significant [respectively, $R(2,39) = 0.541$, $p < 0.005$ for the task-positive network and $R(2,39) = 0.512$, $p < 0.005$ for the task-negative network]. For the task-positive network, the beta value was only significant for the peak measure ($\beta = 0.431$, $p < 0.05$), but not for the spread measure ($\beta = -0.21$ n.s.). For the task-negative network, the beta value was only significant for the spread measure ($\beta = 0.452$, $p < 0.05$), but not for the peak measure ($\beta = 0.132$ n.s.). This suggests that in the task-positive network the peak amplitudes are being modulated by age above and beyond changes in the spread of blood flow. Conversely in the task-negative networks there are residual age effects on the spread of deactivations after controlling for changes in amplitude.

Independence of peak and spread estimates

An important issue for the purposes of this paper is how independent the spread estimates are from the magnitude of the peak value. Both measures may be considered indices of the degree of cortical activation (or deactivation) during the task. It is important to know, therefore, whether they provide similar or different information. We used an intra/inter-class correlation analysis approach to assess the degree of independence of spread

and peak measures. For each region we compared the average amount of variance (across subjects) that was shared between different measures. Specifically, for each region we computed four types of shared variances: (a) the average shared variance between measures of spread in one region and measures of spread in different regions of the same network (SS); (b) the average shared variance of measures of peak in one region and measures of peak in different regions of the same network (PP); (c) the average shared variance between measures of spread in one region and measures of peak in different regions of the same network (SP-all); and (d) the average shared variance between measures of spread and measures of peak taken from the same region (SP-same).

The expectation is that all correlations share the network as a common source of variance. In addition both SS and PP will have one other source of variance in common (i.e., the same type of measure), whereas SP-same will have the same region in common. SP-all correlations will have no other common sources of variance (different measures and regions) and therefore will provide an estimate of the baseline level for shared variance.

These data were submitted to a mixed-design ANOVA, with one fixed between-cases factor (network), one random factor (region, nested within network), and a four-level repeated-measure factor (correlation type). The results, averaged across task-positive, task-negative and all regions (see **Table 4**) indicated a significant effect of correlation type ($F_{3,57} = 29.16, p < 0.0001$). Importantly, all the intra-class correlations (i.e., PP and SS) were significant (all $F_{s_{19,40}} > 2.20, p < 0.05$) even when Bonferroni-corrected. However, none of the inter-class correlations (SP, including measures of spread and peak from the same regions) were significant (all $F_{s_{19,40}} < 1.66$), with the exception of peak-spread measures for task-positive networks ($F_{19,40} = 1.93, p < 0.05$), which however, would not reach significance when Bonferroni-corrected. Planned comparisons showed that the same modality (PP and SS) within a network were more highly intercorrelated than measures across modalities from the same regions (SP same)

($t_{20} = 2.39, p < 0.05$, and $t_{20} = 5.38, p < 0.001$, respectively). This indicates a high degree of independence across measures. We also computed the intraclass/interclass correlations separately in younger and older adults. The results were essentially identical for the two groups. Within-measures (PP and SS) shared-variance between areas were significant ($F_{19,13} = 4.85, p < 0.005$ for the young group, and $F_{19,27} = 3.44, p < 0.01$ for the old group), whereas across-measures (SP) shared-variance between areas were not significant ($F_{19,13} = 2.09$ n.s. for the young group, and $F_{19,27} = 0.88$, n.s. for the old group). There is therefore evidence that the two measures are independent in both groups.

DISCUSSION

The quantification of BOLD fMRI data is typically carried out on a three-dimensional volume extending over a number of voxels. As such the observed effects are not a pure measure of signal amplitude, but are a combination of both the peak strength of local blood flow changes as well as the spread of such changes throughout that volume. Anything that modulates the spread of this signal (e.g., changing the size of the smoothing kernel) can drastically impact the observed strength and localization of functional effects (see White et al., 2001; Jo et al., 2008; Mikl et al., 2008). The goal of the analyses reported in this paper was to investigate whether there were systematic age-related differences in the spreading of activation and deactivation during an attention-demanding task, and whether the spread measures provided additional information compared to the measures of peak activity.

The whole-brain analyses indicated that older adults over-activate areas positively associated with the task, while simultaneously failing to fully deactivate areas of the task-negative (DMN) network, replicating previous findings (e.g., Lustig et al., 2003; Persson et al., 2007; Park and Reuter-Lorenz, 2009). These analyses were supplemented by ROI analyses using three different sized spheres to approximate ROI analyses that are often performed in the literature. Compared to the younger adults, the older adults

Table 4 | Analysis of shared variance (intra-class correlation, r^2) between spread and peak measures.

Brain networks	Spread–spread (SS)	Peak–peak (PP)	Spread–peak (SP-all)	Spread–peak (SP-same)
Shared sources of variance	Measure type network	Measure type network	Network	Region network
Task-positive	0.141	0.220	0.051	0.094
Task-negative	0.104	0.182	0.050	0.068
All regions	0.122	0.201	0.050	0.080
p-value vs. SP-same	0.026	0.000	0.023	

SS = the average shared variance between measures of spread of one region and measures of spread in different regions of the same network. PP = the average shared variance between measures of peak of one region and measures of peak in different regions of the same network. SP-all = the average shared variance between measures of spread of one region and measures of peak in different regions of the same network. SP-same = the average shared variance between measures of spread and measures of peak taken from the same region.

had greater levels of activation at all three sizes (3 mm, 5 mm, and 10 mm), but only demonstrated reduced deactivation when using the 10 mm sphere. These results are generally consistent with previous work, but also suggest that the size of the kernel used for quantification impacts the results. This is likely because, in this analysis, peak amplitude, and spread of activity are confounded.

To address this concern, we introduced a novel technique to assess the spread of activation and deactivation of the BOLD signal around its peak. This new approach shows that older adults have alterations in the *spread* of the BOLD response compared to younger adults. By measuring the *relative* (normalized) size of the BOLD response at various distances from the peak point, we could evaluate signal spread for both activation and deactivation *separately from peak amplitude*. When examining activations, the older adults had a shallower average slope of decay from the peak point. This supports previous notions that older adults possess broader (or less focused) areas of activation. Within the task-positive network this was particularly true for the insula while other regions only showed a trend for age-related differences. This suggests a systematic, but relatively subtle, increase in the spread of activations in older adults. This also suggests that the expanded activations seen in group level maps of older adults are a product of both greater peak amplitudes as well as a broader spread of such activity to surrounding tissue.

This pattern was highly significant but *inverted* for the deactivation of tissue. Older adults showed a rapid decay of deactivation with increasing distance from the peak point. This indicates that the deactivation patterns are relatively focal, and then quickly dissipate. This finding was significant across a wide range of areas including core areas of the task-negative network such as the posterior cingulate cortex, precuneus, and temporal pole. The results from this analysis indicate that older adults have a significant reduction in the spread of deactivation. They also suggest that the reduced group level deactivation maps of older adults are not due to changes in focal deactivations, but rather in how these deactivations propagate to surrounding tissue.

For voxel-wise analyses, statistics at the group level are dependent upon spatial overlap across subjects. The peak points for the older adults were significantly more spatially variable for task-negative areas. Although our analyses of the task-positive network did not reveal significant differences in spatial clustering, the numeric directions, as well as the lack of a network by group interaction, are consistent with increased spatial variability occurring throughout the brain but being more pronounced in the task-negative regions. This pattern alone would affect group-level maps, and such issues have been considered in the literature (e.g., Swallow et al., 2003; Devlin and Poldrack, 2007). This effect would compound systematic age differences in the amplitudes and spread of activations and deactivations. In areas where older adults have stronger and broader activations, such as the task-positive network, increased anatomical variability could lead to a more diffuse group-level pattern of activity. In areas where activity is narrower, such as the task-negative network, an increase in spatial variability would lead to a spatially underestimated group-level map. The type of spatial normalization could also interact

with such phenomenon. By their very nature non-linear registrations warp tissue differently across the brain. Selective atrophy in aging or disease populations may exacerbate such phenomenon relative to younger adults. This could induce an artificial broadening or narrowing of spread of blood flow within a cortical region.

The general problems of spatial overlap are readily known and are a good argument in favor of ROI analyses, which provide more flexibility. Still, as seen in our typical ROI peak analyses using three different sized spheres, the choice of kernel size can interact with differences in the spread around peaks. As clearly seen in **Table 2**, the selection of a 3 or 5 mm compared to a 10 mm sphere would alter our interpretation of the data. Using the 3 or 5 mm sphere we would have concluded that the older adults had stronger activations than the younger adults in task-positive areas, while the two groups did not differ in the strength of their deactivations in task-negative areas. Using the 10 mm spheres the results would now reveal a significant group effect for both the task-positive and task-negative modulations. This does not mean that ROI analyses are inappropriate, just that interpretations must be considered in terms of both the area of tissue being modulated as well as the strength of this modulation. Such considerations are particularly important when comparing two groups that may systematically differ from each other, rather than when examining a manipulation within the same subject.

Some possible limitations to the approach proposed here should also be considered. Potential confounds when comparing younger and older adults could arise from either cortical atrophy or head motion. Atrophy would reduce the total volume that a given cortical region encompasses. Due to such shrinkage, one would expect a narrower focus where blood flow is modulated. In contrast, head motion could lead to a smearing of activity to produce a more diffuse locus of activity or deactivation. Our current data demonstrated dissociations, with older adults showing a broader extent of activity in the task-positive areas and a reduced spread in task-negative ones. As illustrated in **Figure 4**, the vast majority of regions within each network displayed consistent age-related patterns of spread despite having a range of spatial locations throughout the brain. It is highly unlikely that atrophy or head motion alone could produce such a dissociation and consistency within networks rather than manifesting as a global effect on the brain.

Another well-known concern in aging studies is the potential occurrence of age-related differences in neurovascular coupling. Neural activity is inferred from the BOLD signal based upon the relationship between neuronal firing, metabolic consumption of oxygen, and the subsequent increase in blood perfusion. The coupling between the hemodynamic response and neural activity is thought to be impaired in older age (D'Esposito et al., 1999; Buckner et al., 2000; Hesselmann et al., 2001; Huettel et al., 2001; Aizenstein et al., 2004; Fabiani et al., 2014). Most studies examine coupling in terms of activation profiles, yet the same sluggish vascular response should also impair the down-regulation of blood flow. Hence the spread of activation and deactivation should be equally (or at least similarly) impaired by a reduction in local vasculature. Therefore, the observation of a selective deficit in the deactivation spread appears inconsistent with this account.

To the extent that the patterns of activity reported here for the younger adults represent the gold standard for optimal brain function, we could speculate about possible interpretations of the alterations of signal spread in older adults. In fact, the increased spread of activation in the task-positive network in older adults is inherent to the idea of de-differentiation (Park et al., 2004) and is also consistent with notions of compensation (e.g., Cabeza et al., 2002; Persson et al., 2004; Reuter-Lorenz and Cappell, 2008; Schneider-Garces et al., 2010). It should be noted, however, that we do not mean to imply that subjects deliberately compensate for poor performance by varying the amount of spread of the activated brain areas. We are only stating that age is associated with an increase in the spread of activation in areas up-regulated during the task. The measurement approach presented in this paper may allow researchers to further explore the dissociations and overlaps existing between different models of cognitive aging.

The observation of reduced spread of deactivation in the task-negative network had not been previously characterized, and can be interpreted in several ways. Both age groups are deactivating tissue focally to the same degree, but this signal does not spread as far to neighboring tissue in the older participants. This is particularly true for core regions of the DMN such as the precuneus, posterior cingulate cortex, and fronto-medial cortex. One possible interpretation is that *local* connections are less efficient in controlling the deactivation, either mediating or possibly compounding the widely reported reduction in top-down attention control over sensory areas in older adults (e.g., Fabiani et al., 1998, 2006; Gazzaley et al., 2005, 2008).

Another interpretation is that the task-negative (DMN) network has properties that make it uniquely vulnerable to age-related declines. For example, this network has an elevated susceptibility to disrupted metabolic processes and preferentially accumulates amyloid beta (Klunk et al., 2004; Buckner et al., 2005). Reduced levels of deactivation are associated with Alzheimer's disease (Petrella et al., 2007; Persson et al., 2008; Sperling et al., 2009) further supporting the idea that the task-negative network may be selectively linked to cognitive health. In addition, amyloid plaques in DMN regions may cause functional disruption even in older adults classified as normal (Hedden et al., 2009). Age-related structural damage in these regions may therefore be the substrate for drops in local connectivity as a function of age, which may in turn result in drops in the spread of the BOLD signal.

Although this evidence suggests that the task-negative network may be particularly sensitive to age-related decline, it should also be considered that this network is not ubiquitously less responsive in older adults (but see Grady, 2012). In a tests of emotional memory by Kensinger and Schacter (2008), older and younger adults possessed comparable levels of functional activation in regions within this network. In fact, older adults slightly over-activated these regions during encoding. Similar work using emotional stimuli has found preserved or enhanced activation in older adults (Gutchess et al., 2007). This suggests that the task-negative network is not always impaired, but rather that age differences may be more evident when it must be suppressed. It may be that older adults have difficulty inhibiting the activation of any networks, but the design of most

functional studies requires a disengagement of processes that recruit the DMN during rest. Ultimately, an examination of BOLD signal spread across the task-negative network in a task that specifically activates this network, such as that reported by Grady et al. (2010), is needed. This will help determine whether the age-related effects observed in the current study are due to failing deactivation/top-down control that can affect multiple networks or if problems are specific to the task-negative network.

A third possibility is that a common mechanism may account for both the increased spread in the task-positive network and the decreased spread in the task-negative network occurring in aging. It is thought that deactivation may involve a relative inhibition or suppression of a particular cortical region. Inhibition in the cortex is carried out through GABAergic interneurons (Chagnac-Amitai and Connors, 1989) whose genetic modulators are down-regulated with age (Loerch et al., 2008; Bishop et al., 2010). Thus, a reduction in the expression of GABA receptors in the cortex may lead to a reduction of the deactivation process. The same mechanism may also account for the increased spread of the activation signal observed in older adults, as the spread of activation may be limited in younger adults by the action of inhibitory interneurons, which may be reduced in aging. This age-related change could potentially alter the balance between activation and deactivation signals in the brain, as well as the spread of these signals. Thus the same mechanism – reduction of GABAergic inhibition in the cortex – could potentially account for both reduced spread of deactivation and greater spread of activation.

Currently each proposed interpretation is plausible but speculative. It may be that no single interpretation can entirely account for these findings, but rather that a combination of multiple mechanisms drives the observed modulations in spread. For example there could be a down-regulation of GABA interneurons, but this deficit may be non-uniform across the brain. A multimodal approach combining fMRI, electrophysiological measures, and positron emission tomography tailored to the investigation of GABA receptors (e.g., Heiss and Herholz, 2006) may begin to address these questions. Improved understanding of the mechanisms that drive the changes in spread may inform studies of aging, and provide a new avenue of research to explore the brain.

In summary, many aging studies that utilize fMRI data focus on perceived differences in the amplitude of activation. Many of these analyses, particularly those drawing upon group-level maps, are actually conflating differences in amplitude with changes in the spread of blood flow. Uniquely within the field of cognitive aging, the current work independently examines both the amplitude and the spread of the BOLD response in younger and older adults. Understanding both of these properties is important when interpreting differences between these age groups. The current experiment supports previous work demonstrating over-activation and under-deactivation in older adults, while using an innovative approach to assess the spread of functional blood flow changes. This metric revealed that older adults have a broader extent of activation while simultaneously having a narrower focus of deactivation, independent of amplitude differences. These results provide a novel measure that illustrates the two-fold pattern

of differences in both amplitude and spread of functional blood flow changes with increasing age.

AUTHOR CONTRIBUTIONS

Brian A. Gordon collected the functional data, performed statistical analyses, and drafted and revised the manuscript. Chun-Yu Tse assisted in data analysis and revising the manuscript. Monica Fabiani and Gabrielle Gratton assisted in data analysis, drafting, and revising the manuscript.

ACKNOWLEDGMENTS

We wish to acknowledge the support of NIMH grant 5R56MH097973 and NIA grant 1RC1AG035927 to Drs. Gratton, Fabiani, and Gordon's support on training grant 5T32AG00035 to Dave Balota. This work was completed in partial fulfillment of the Ph.D. requirements of the University of Illinois of the first author. We thank Carrie Brumback-Peltz and Yukie Lee for assisting collect subject data and providing general support for the project. We also wish to acknowledge the comments of Drs. Lynn Hasher, Art Kramer, and Greg Miller on an earlier draft of this manuscript.

REFERENCES

- Aizenstein, H. J., Clark, K. A., Butters, M. A., Cochran, J., Stenger, V. A., Meltzer, C. C., et al. (2004). The BOLD hemodynamic response in healthy aging. *J. Cogn. Neurosci.* 16, 786–793. doi: 10.1162/089892904970681
- Ances, B. M., Liang, C. L., Leontiev, O., Perthen, J. E., Fleisher, A. S., Lansing, A. E., et al. (2009). Effects of aging on cerebral blood flow, oxygen metabolism, and blood oxygenation level dependent responses to visual stimulation. *Hum. Brain Mapp.* 30, 1120–1132. doi: 10.1002/hbm.20574
- Beck, A., Steer, R., and Brown, G. (1996). *Manual for the Beck Depression Inventory*. San Antonio: Psychological Corporation.
- Beckmann, C. F., DeLuca, M., Devlin, J. T., and Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1001–1013. doi: 10.1098/rstb.2005.1634
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., and Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *J. Cogn. Neurosci.* 17, 905–917. doi: 10.1162/0898929054021102
- Bishop, N. A., Lu, T., and Yankner, B. A. (2010). Neural mechanisms of ageing and cognitive decline. *Nature* 464, 529–535. doi: 10.1038/nature08983
- Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004
- Buckner, R. L., Snyder, A. Z., Sanders, A. L., Raichle, M. E., and Morris, J. C. (2000). Functional brain imaging of young, nondemented, and demented older adults. *J. Cogn. Neurosci.* 12 (Suppl. 2), 24–34. doi: 10.1162/089892900564046
- Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., et al. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J. Neurosci.* 25, 7709–7717. doi: 10.1523/JNEUROSCI.2177-05.2005
- Cabeza, R., Anderson, N. D., Locantore, J. K., and McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Castellanos, F. X., Margulies, D. S., Kelly, C., Uddin, L. Q., Ghaffari, M., Kirsch, A., et al. (2008). Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 63, 332–337. doi: 10.1016/j.biopsych.2007.06.025
- Chagnac-Amitai, Y., and Connors, B. W. (1989). Horizontal spread of synchronized activity in neocortex and its control by GABA-mediated inhibition. *J. Neurophysiol.* 61, 747–758.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- D'Esposito, M., Deouell, L. Y., and Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat. Rev. Neurosci.* 4, 863–872. doi: 10.1038/nrn1246
- D'Esposito, M., Zarahn, E., Aguirre, G. K., and Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. *Neuroimage* 10, 6–14. doi: 10.1006/nimg.1999.0444
- Damoiseaux, J. S., Beckmann, C. F., Arigita, E. J. S., Barkhof, F., Scheltens, P., Stam, C. J., et al. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cereb. Cortex* 18, 1856–1864. doi: 10.1093/cercor/bhm207
- Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., et al. (2006). Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13848–13853. doi: 10.1073/pnas.0601417103
- Daselaar, S. M., Prince, S. E., and Cabeza, R. (2004). When less means more: deactivations during encoding that predict subsequent memory. *Neuroimage* 23, 921–927. doi: 10.1016/j.neuroimage.2004.07.031
- Devlin, J. T., and Poldrack, R. A. (2007). In praise of tedious anatomy. *Neuroimage* 37, 1033–1041; discussion 1050–1058. doi: 10.1016/j.neuroimage.2006.09.055
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309
- Fabiani, M. (2012). It was the best of times, it was the worst of times: a psychophysiology's view of cognitive aging. *Psychophysiology* 49, 283–304. doi: 10.1111/j.1469-8986.2011.01331.x
- Fabiani, M., Friedman, D., and Cheng, J. C. (1998). Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology* 35, 698–708. doi: 10.1111/1469-8986.3560698
- Fabiani, M., Gordon, B. A., Maclin, E. L., Pearson, M. A., Brumback-Peltz, C. R., Low, K. A., et al. (2014). Neurovascular coupling in normal aging: a combined optical, ERP and fMRI study. *Neuroimage* 85, 592–607. doi: 10.1016/j.neuroimage.2013.04.113
- Fabiani, M., Low, K. A., Wee, E., Sable, J. J., and Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *J. Cogn. Neurosci.* 18, 637–650. doi: 10.1162/jocn.2006.18.4.637
- Feredoes, E., and Postle, B. R. (2007). Localization of load sensitivity of working memory storage: quantitatively and qualitatively discrepant results yielded by single-subject and group-averaged approaches to fMRI group analysis. *Neuroimage* 35, 881–903. doi: 10.1016/j.neuroimage.2006.12.029
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102
- Fox, M. D., Zhang, D., Snyder, A. Z., and Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* 101, 3270–3283. doi: 10.1152/jn.90777.2008
- Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* 26, 15–29. doi: 10.1002/hbm.20113
- Fransson, P. (2006). How default is the default mode of brain function? Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia* 44, 2836–2845. doi: 10.1016/j.neuropsychologia.2006.06.017
- Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., and D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13122–13126. doi: 10.1073/pnas.0806074105
- Gazzaley, A., Cooney, J. W., Rissman, J., and D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nat. Neurosci.* 8, 1298–1300. doi: 10.1038/nn1543
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505. doi: 10.1038/nrn3256
- Grady, C. L., Protzner, A. B., Kovacevic, N., Strother, S. C., Afshin-Pour, B., Wojtowicz, M., et al. (2010). A multivariate analysis of age-related differences in default mode and task-positive networks across multiple cognitive domains. *Cereb. Cortex* 20, 1432–1447. doi: 10.1093/cercor/bhp207
- Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., and Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *J. Cogn. Neurosci.* 18, 227–241. doi: 10.1162/jocn.2006.18.2.227
- Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258. doi: 10.1073/pnas.0135058100

- Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264. doi: 10.1073/pnas.071043098
- Gutchess, A. H., Kensinger, E. A., and Schacter, D. L. (2007). Aging, self-referencing, and medial prefrontal cortex. *Soc. Neurosci.* 2, 117–133. doi: 10.1080/17470910701399029
- Harrison, B. J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., et al. (2008). Consistency and functional specialization in the default mode brain network. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9781–9786. doi: 10.1073/pnas.0711791105
- Hedden, T., Van Dijk, K. R. A., Becker, J. A., Mehta, A., Sperling, R. A., Johnson, K. A., et al. (2009). Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. *J. Neurosci.* 29, 12686–12694. doi: 10.1523/JNEUROSCI.3189-09.2009
- Heiss, W.-D., and Herholz, K. (2006). Brain receptor imaging. *J. Nucl. Med.* 47, 302–312.
- Hesselmann, V., Zaro Weber, O., Wedekind, C., Krings, T., Schulte, O., Kugel, H., et al. (2001). Age related signal decrease in functional magnetic resonance imaging during motor stimulation in humans. *Neurosci. Lett.* 308, 141–144. doi: 10.1016/S0304-3940(01)01920-6
- Huetzel, S., Singerman, J., and McCarthy, G. (2001). The effects of aging upon the hemodynamic response measured by functional MRI. *Neuroimage* 13, 161–175. doi: 10.1006/nimg.2000.0675
- Jo, H. J., Lee, J.-M., Kim, J.-H., Choi, C.-H., Gu, B.-M., Kang, D.-H., et al. (2008). Artificial shifting of fMRI activation localized by volume- and surface-based analyses. *Neuroimage* 40, 1077–1089. doi: 10.1016/j.neuroimage.2007.12.036
- Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., and Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *J. Cogn. Neurosci.* 12, 188–196. doi: 10.1162/089892900561823
- Kelly, A., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. *Neuroimage* 39, 527–537. doi: 10.1016/j.neuroimage.2007.08.008
- Kensinger, E. A., and Schacter, D. L. (2008). Neural processes supporting young and older adults' emotional memories. *J. Cogn. Neurosci.* 20, 1161–1173. doi: 10.1162/jocn.2008.20080
- Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., et al. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann. Neurol.* 55, 306–319. doi: 10.1002/ana.20009
- Kramer, A. F., Fabiani, M., and Colcombe, S. J. (2006). "Contributions of cognitive neuroscience to the understanding of behavior and aging," in *Handbook of the Psychology of Aging*, eds J. E. Birren and K. Schaie (New York, NY: Academic Press), 57–83.
- Loerch, P. M., Lu, T., Dakin, K. A., Vann, J. M., Isaacs, A., Geula, C., et al. (2008). Evolution of the aging brain transcriptome and synaptic regulation. *PLoS ONE* 3:e3329. doi: 10.1371/journal.pone.0003329
- Lustig, C., Snyder, A. Z., Bhakta, M., O'Brien, K. C., McAvoy, M., Raichle, M. E., et al. (2003). Functional deactivations: change with age and dementia of the Alzheimer type. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14504–14509. doi: 10.1073/pnas.2235925100
- Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395. doi: 10.1126/science.1131295
- Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., Das, S., Callicott, J. H., et al. (2002). Neurophysiological correlates of age-related changes in human motor function. *Neurology* 58, 630–635. doi: 10.1212/WNL.58.4.630
- Mayeux, R., Stern, Y., Rosen, J., and Leventhal, J. (1981). Depression, intellectual impairment, and Parkinson disease. *Neurology* 32, 211–227. doi: 10.1212/WNL.31.6.645
- Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., et al. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* 54, 287–298. doi: 10.1016/S0361-9230(00)00437-8
- McKiernan, K. A., D'Angelo, B. R., Kaufman, J. N., and Binder, J. R. (2006). Interrupting the "stream of consciousness": an fMRI investigation. *Neuroimage* 29, 1185–1191. doi: 10.1016/j.neuroimage.2005.09.030
- McKiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., and Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J. Cogn. Neurosci.* 15, 394–408. doi: 10.1162/089892903321593117
- Mikl, M., Marecek, R., Hlustik, P., Pavlicová, M., Drastich, A., Chlebus, P., et al. (2008). Effects of spatial smoothing on fMRI group inferences. *Magn. Reson. Imaging* 26, 490–503. doi: 10.1016/j.mri.2007.08.006
- Otten, L. J., and Rugg, M. D. (2001). When more means less?: neural activity related to unsuccessful memory encoding. *Curr. Biol.* 11, 1528–1530. doi: 10.1016/S0960-9822(01)00454-7
- Park, D. C., Polk, T. A., Park, R., Minear, M., Savage, A., and Smith, M. R. (2004). Aging reduces neural specialization in ventral visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13091–13095. doi: 10.1073/pnas.0405148101
- Park, D. C., and Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* 60, 173–196. doi: 10.1146/annurev.psych.59.103006.093656
- Persson, J., Lind, J., Larsson, A., Ingvar, M., Slegers, K., Van Broeckhoven, C., et al. (2008). Altered deactivation in individuals with genetic risk for Alzheimer's disease. *Neuropsychologia* 46, 1679–1687. doi: 10.1016/j.neuropsychologia.2008.01.026
- Persson, J., Lustig, C., Nelson, J. K., and Reuter-Lorenz, P. A. (2007). Age differences in deactivation: a link to cognitive control? *J. Cogn. Neurosci.* 19, 1021–1032. doi: 10.1162/jocn.2007.19.6.1021
- Persson, J., Sylvestre, C.-Y. C., Nelson, J. K., Welsh, K. M., Jonides, J., and Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: differential recruitment in older and younger adults. *Neuroimage* 23, 1382–1390. doi: 10.1016/j.neuroimage.2004.08.004
- Petrella, J. R., Prince, S. E., Wang, L., Hellegers, C., and Doraiswamy, P. M. (2007). Prognostic value of posteromedial cortex deactivation in mild cognitive impairment. *PLoS ONE* 2:e1104. doi: 10.1371/journal.pone.0001104
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Raichle, M. E., and Snyder, A. Z. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage* 37, 1083–1090; discussion 1097–1099. doi: 10.1016/j.neuroimage.2007.02.041
- Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x
- Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12, 174–187. doi: 10.1162/089892900561814
- Sambataro, F., Murty, V. P., Callicott, J. H., Tan, H.-Y., Das, S., Weinberger, D. R., et al. (2010). Age-related alterations in default mode network: impact on working memory performance. *Neurobiol. Aging* 31, 839–852. doi: 10.1016/j.neurobiolaging.2008.05.022
- Schneider-Garces, N. J., Gordon, B. A., Brumback-Peltz, C. R., Shin, E., Lee, Y., Sutton, B. P., et al. (2010). Span, CRUNCH, and beyond: working memory capacity and the aging brain. *J. Cogn. Neurosci.* 22, 655–669. doi: 10.1162/jocn.2009.21230
- Shulman, G. L., Corbetta, M., Buckner, R. L., Fiez, J. A., Miezin, F. M., Raichle, M. E., et al. (1997a). Common blood flow changes across visual tasks: I. increases in subcortical structures and cerebellum but not in nonvisual cortex. *J. Cogn. Neurosci.* 9, 624–647. doi: 10.1162/jocn.1997.9.5.624
- Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997b). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cogn. Neurosci.* 9, 648–663. doi: 10.1162/jocn.1997.9.5.648
- Smith, S. M. (2004). Overview of fMRI analysis. *Br. J. Radiol.* 77, S167–S175. doi: 10.1259/bjr/33553595
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Sonuga-Barke, E. J. S., and Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. *Neurosci. Biobehav. Rev.* 31, 977–986. doi: 10.1016/j.neubiorev.2007.02.005

- Sperling, R. A., Laviolette, P. S., O'Keefe, K., O'Brien, J., Rentz, D. M., Pihlajamäki, M., et al. (2009). Amyloid deposition is associated with impaired default network function in older persons without dementia. *Neuron* 63, 178–188. doi: 10.1016/j.neuron.2009.07.003
- Stebbins, G. T., Carrillo, M. C., Dorfman, J., Dirksen, C., Desmond, J. E., Turner, D. A., et al. (2002). Aging effects on memory encoding in the frontal lobes. *Psychol. Aging* 17, 44–55. doi: 10.1037/0882-7974.17.1.44
- Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652
- Swallow, K. M., Braver, T. S., Snyder, A. Z., Speer, N. K., and Zacks, J. M. (2003). Reliability of functional localization using fMRI. *Neuroimage* 20, 1561–1577. doi: 10.1016/S1053-8119(03)00436-1
- Tehovnik, A. S., Tolias, A. S., Sultan, F., Slocum, W. M., and Logothetis, N. K., (2006). Direct and indirect activation of cortical neurons by electrical microstimulation. *J. Neurophysiol.* 96, 512–521. doi: 10.1152/jn.00126.2006
- Tolias, A. S., Sultan, F., Augath, M., Oeltermann, A., Tehovnik, E. J., Schiller, P. H., et al. (2005). Mapping cortical activity elicited with electrical microstimulation using fMRI in the macaque. *Neuron* 48, 901–911. doi: 10.1016/j.neuron.2005.11.034
- Toro, R., Fox, P. T., and Paus, T. (2008). Functional coactivation map of the human brain. *Cereb. Cortex* 18, 2553–2559. doi: 10.1093/cercor/bhn014
- Uddin, L. Q., Kelly, A. M., Biswal, B. B., Xavier Castellanos, F., and Milham, M. P. (2009). Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Hum. Brain Mapp.* 30, 625–637. doi: 10.1002/hbm.20531
- Wechsler, D. (1981). *WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.
- Weissman, D. H., Roberts, K. C., Visscher, K. M., and Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nat. Neurosci.* 9, 971–978. doi: 10.1038/nn1727
- White, T., O'Leary, D., Magnotta, V., Arndt, S., Flaum, M., and Andreasen, N. C. (2001). Anatomic and functional variability: the effects of filter size in group fMRI data analysis. *Neuroimage* 13, 577–588. doi: 10.1006/nimg.2000.0716
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., et al. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173–S186. doi: 10.1016/j.neuroimage.2008.10.055

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 June 2014; accepted: 29 September 2014; published online: 16 October 2014.

Citation: Gordon BA, Tse C-Y, Gratton G and Fabiani M (2014) Spread of activation and deactivation in the brain: does age matter? *Front. Aging Neurosci.* 6:288. doi: 10.3389/fnagi.2014.00288

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Gordon, Tse, Gratton and Fabiani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read,
for greatest visibility



COLLABORATIVE PEER-REVIEW

Designed to be rigorous
– yet also collaborative,
fair and constructive



FAST PUBLICATION

Average 85 days from
submission to publication
(across all journals)



COPYRIGHT TO AUTHORS

No limit to article
distribution and re-use



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



SUPPORT

By our Swiss-based
editorial team



IMPACT METRICS

Advanced metrics
track your article's impact



GLOBAL SPREAD

5'100'000+ monthly
article views
and downloads



LOOP RESEARCH NETWORK

Our network
increases readership
for your article

Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on

