


# Functional Near-Infrared Spectroscopy as a Measure of Listening Effort in Older Adults Who Use Hearing Aids

Trends in Hearing  
Volume 23: 1–22  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/2331216519886722  
journals.sagepub.com/home/tia  


Joseph Rovetti<sup>1</sup>, Huiwen Goy<sup>1</sup> , M. Kathleen Pichora-Fuller<sup>2</sup>, and Frank A. Russo<sup>1,3</sup>

## Abstract

Listening effort may be reduced when hearing aids improve access to the acoustic signal. However, this possibility is difficult to evaluate because many neuroimaging methods used to measure listening effort are incompatible with hearing aid use. Functional near-infrared spectroscopy (fNIRS), which can be used to measure the concentration of oxygen in the prefrontal cortex (PFC), appears to be well-suited to this application. The first aim of this study was to establish whether fNIRS could measure cognitive effort during listening in older adults who use hearing aids. The second aim was to use fNIRS to determine if listening effort, a form of cognitive effort, differed depending on whether or not hearing aids were used when listening to sound presented at 35 dB SL (flat gain). Sixteen older adults who were experienced hearing aid users completed an auditory n-back task and a visual n-back task; both tasks were completed with and without hearing aids. We found that PFC oxygenation increased with n-back working memory demand in both modalities, supporting the use of fNIRS to measure cognitive effort during listening in this population. PFC oxygenation was weakly and nonsignificantly correlated with self-reported listening effort and reaction time, respectively, suggesting that PFC oxygenation assesses a dimension of listening effort that differs from these other measures. Furthermore, the extent to which hearing aids reduced PFC oxygenation in the left lateral PFC was positively correlated with age and pure-tone average thresholds. The implications of these findings as well as future directions are discussed.

## Keywords

hearing loss, aging, neuroimaging, working memory, n-back task

Received 4 June 2019; revised 25 September 2019; accepted 1 October 2019

## Introduction

The assessment of hearing acuity typically involves measures such as pure-tone or speech audiometry. While these methods are crucial to clinical diagnosis and intervention, they do not measure an important dimension of the experience of listening: the level of effort expended by a listener to achieve a listening goal. Indeed, in both clinical and real-world settings, individuals with hearing loss sometimes report that listening is effortful even when sounds are loud enough (Pichora-Fuller et al., 2016). The term listening effort has thus been defined as “the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a [listening] task” (Pichora-Fuller et al., 2016, p. 10S).

## Hearing Loss and Listening Effort

When auditory input is unclear, listeners are forced to recruit additional cognitive resources in order to extract the intended message (Pelle, 2018). For instance, a listener may need to focus attention on the auditory stimulus in order to understand speech (Pelle, 2018;

<sup>1</sup>Department of Psychology, Ryerson University, Toronto, ON, Canada

<sup>2</sup>Department of Psychology, University of Toronto, ON, Canada

<sup>3</sup>Toronto Rehabilitation Institute, ON, Canada

### Corresponding Author:

Frank A. Russo, Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3.

Email: russo@psych.ryerson.ca



Pichora-Fuller et al., 2016). Verbal working memory is also required, as unclear auditory stimuli must be held in memory to allow for ongoing processing until recognition is achieved (Peelle, 2018; Rönnberg, Holmer, & Rudner, 2019; Signoret, Johnsrude, Classon, & Rudner, 2018). Indeed, verbal working memory capacity is correlated with the ability of both normal-hearing and hearing-impaired listeners to process degraded speech (Ward, Rogers, Van Engen, & Peelle, 2015). Thus, sustained listening effort requires the listener to utilize both attention and working memory.

Several of the negative outcomes resulting from hearing loss may be attributable to the cognitive demands imposed by listening in adverse conditions. As more cognitive resources are consumed during listening, fewer resources remain available for other purposes (Kahneman, 1973). For instance, for people with hearing loss, the effort invested in listening may limit their ability to multitask (Picou, Ricketts, & Hornsby, 2013) or to rehearse the auditory input, thereby worsening recall of what has been heard (McCoy et al., 2005). Moreover, effortful listening is linked to negative psychosocial consequences, such as psychological distress (Hétu, Riverin, Lalonde, Getty, & St-Cyr, 1988), and people with hearing loss may withdraw from social interaction to avoid this distress (Pichora-Fuller, Mick, & Reed, 2015). Such social withdrawal may ultimately hasten age-related cognitive decline (Lin et al., 2013).

### *Measuring Listening Effort*

Listening effort has been measured in various research contexts, but standardized measures for clinical use have not yet been developed (Pichora-Fuller et al., 2016). Nonetheless, there are several advantages to measuring listening effort in audiological practice (Kiessling et al., 2003; McGarrigle et al., 2014; Pichora-Fuller, 2003). These advantages include identifying cases where an individual patient may have hearing difficulties that are not obvious from typical hearing assessments (e.g., pure-tone and speech audiometry) as well as assessing the benefit that an individual patient may derive from using hearing aids, assuming that hearing aids can improve ease of listening for people with hearing loss (Pichora-Fuller et al., 2016). Thus, a clinical measure of listening effort could be used to assess the extent to which a patient's hearing aids reduce listening effort (see Wild et al., 2012) as well as changes in how a patient allocates cognitive capacity in demanding listening situations. In addition to comparing listening effort with and without hearing aids, a clinical measure of listening effort could be used to compare listening effort when using different signal-processing algorithms, which could in turn guide the selective use of the advanced features available in modern hearing aids

(Kiessling et al., 2003; Pichora-Fuller, 2003). Such comparisons would be of use not only to the treatment of individual patients but also to the design of new hearing aid technologies.

In research contexts, current measures of listening effort fall into three broad classes: subjective, behavioral, and physiological. Subjective measures, used during or after a listening task, require participants to self-report the level of effort expended (e.g., Feuerstein, 1992). The assumption underlying this approach is that the experience of listening effort can be accurately perceived, remembered, and reported by listeners (McGarrigle et al., 2014). However, Lemke and Besser (2016) argued that the subjective experience of listening effort is only weakly related to its objective measurement. A possible explanation for this difference comes from Moore and Picou (2018), who noted the tendency for participants to report their perceived level of accuracy in performing a task rather than their perceived level of effort during the task.

Behavioral measures of listening effort can be divided into two subclasses: single-task and dual-task paradigms. Single-task paradigms typically involve measuring participants' reaction times during a listening task (e.g., Gatehouse & Gordon, 1990). The assumption underlying this approach is that participants will respond more slowly when completing tasks that impose a higher cognitive demand, as more processing is needed to select a response (Houben, Doorn-Bierman, & Dreschler, 2013). However, this assumption is far from obvious given the unclear relationship between processing speed and the allocation of mental resources as well as the confounding influence of factors such as attention (McGarrigle et al., 2014). Moreover, reaction time is only suitable to measure listening effort during tasks that require a categorical decision as the response.

In a dual-task paradigm, participants are asked to prioritize performance on a primary task (such as a listening task) while concurrently completing a secondary task (such as a word recall task; e.g., Downs, 1982). The assumption underlying this approach is that cognitive capacity is finite, and that increased cognitive resources spent on the primary task will come at the expense of accurate performance on the secondary task (Kahneman, 1973). Thus, when the primary listening task requires more effort, this will be reflected in less accurate performance on the secondary task. However, secondary task performance is typically regarded as an unreliable measure of listening effort, in part due to the lack of clarity as to what constitutes a suitable secondary task (e.g., whether its modality should match that of the primary task; Gagné, Besser, & Lemke, 2017; Ohlenforst et al., 2017). Furthermore, participants may not always follow instructions to prioritize performance on the

primary task (e.g., Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008; Irwin-Chase & Burns, 2000).

The final class of listening effort measures includes physiological indices. These measures can be divided into two subclasses: measures of the autonomic nervous system and measures of the brain (McGarrigle et al., 2014). The use of autonomic measures assumes that increased effort elevates stress levels, reflected in increased sympathetic activity and decreased parasympathetic activity (Mackersie & Calderon-Moultrie, 2016). These effects drive several physiological changes that have previously been used as indices of the stress resulting from effortful listening, including increased pupil dilation (Kahneman, 1973; Koelewijn, Zekveld, Festen, & Kramer, 2012; Van der Wel & van Steenbergen, 2018), reduced heart rate variability (Mackersie & Cones, 2011; Mackersie, MacPhee, & Heldt, 2015), and increased skin conductance levels (Mackersie & Cones, 2011; Mackersie et al., 2015). An inescapable criticism that can be lodged against autonomic measures is that listening effort involves changes in the brain, whereas these measures do not index brain activity directly. Indeed, few studies have concurrently measured brain and autonomic responses and thus their relationship is not entirely clear (Beissner, Meissner, Bär, & Napadow, 2013). The correspondence of these systems may also be confounded by other variables, both internal (e.g., emotion; Mauss & Robinson, 2009) and external (e.g., ambient light affecting pupil dilation; Ong, Hutch, & Smirnakis, 2018).

Finally, the most direct index of listening effort is the measurement of brain activity. This includes methods such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), which are able to not only measure effort but also provide insight into its underlying mechanisms. fMRI is based on changes in the hemodynamic response, with the assumption being that increased activity in a particular brain region causes an increase in blood flow to that region (Glover, 2011). In the case of effortful listening, attentional and working memory demands cause increased activity in frontoparietal regions involved in these functions (Eckert, Teubner-Rhodes, & Vaden, 2016; Peelle, 2018). For instance, one of the first studies using fMRI to measure effort during listening was that of Wild et al. (2012), who identified the left inferior frontal gyrus (IFG; Broca's area) as a site of activity in participants listening to degraded speech. A greater number of studies have utilized EEG to measure listening effort, with specific indices including power in alpha-band frequency oscillations (e.g., Obleser, Wostmann, Hellbernd, Wilsch, & Maess, 2012) and amplitude of the N1 event-related potential (e.g., Obleser & Kotz, 2011). In Ohlenforst et al.'s (2017) review of listening effort, only studies employing brain-based measures (and not

subjective or behavioral measures) were able to demonstrate a higher level of listening effort attributable to hearing impairment, and therefore these were deemed the most reliable measures.

Despite the directness and reliability of brain-based measures of listening effort, they are severely limited by their inherent lack of compatibility with the use of hearing aids. In the case of fMRI, no ferromagnetic metals are permitted near fMRI scanners, precluding the wearing of most hearing instruments during testing. Similarly, EEG signals are susceptible to electromagnetic field interference from nearby devices, including cochlear implants (Li, Nie, Karp, Tremblay, & Rubinstein, 2010) and hearing aids (Ohlenforst et al., 2017). Although EEG signal-processing measures have been developed to deal with stimulus transduction artifacts (Campbell, Kerlin, Bishop, & Miller, 2012; K. Kim et al., 2015), questions regarding stimulus artifacts can limit interpretation of results and it is not practical to implement these techniques in a clinical setting. Additional obstacles faced by both fMRI and EEG include their high-cost and time-intensive setup. Thus, these measures are sub-optimal for general clinical use or applied research in most development contexts. As a result, it would be difficult to use fMRI or EEG for purposes such as measuring listening effort or assessing the outcomes of hearing aid use.

### *Functional Near-Infrared Spectroscopy*

Currently, there is a need for a reliable measure of listening effort that can be used when listeners are using hearing aids and that is practical for use in both clinical and applied research contexts. One method that could potentially meet this need is functional near-infrared spectroscopy (fNIRS), an optical neuroimaging method that, like fMRI, is sensitive to the overabundance of oxygenated hemoglobin (HbO) and the scarcity of deoxygenated hemoglobin (HbR) that results from brain activity (Buxton, Uludağ, Dubowitz, & Liu, 2004; Izzetoglu, 2012). Unlike fMRI, fNIRS is able to separately measure the concentrations of HbO and HbR (Scarapicchia, Brown, Mayo, & Gawryluk, 2017), although its depth of penetration is limited approximately to 1.5 to 2.5 cm beneath the skull (Quaresima & Ferrari, 2019). fNIRS works by emitting near-infrared light into the scalp and measuring the intensity of the light that returns, from which the relative concentrations of HbO and HbR can be derived using the modified Beer-Lambert law (MBLL; Kocsis, Herman, & Eke, 2006).

Crucially, fNIRS can be used when people are wearing hearing instruments (Scarapicchia et al., 2017). Relative to most other brain-based measures, the equipment needed for fNIRS is also relatively portable,

tolerant of motion artifacts, and inexpensive (Quaresima & Ferrari, 2019). These qualities, along with its noninvasiveness and safety, all render fNIRS suitable to contexts where other brain-based measures are not suitable, such as work with children, older adults, and clinical populations (Ferreri, Bigand, Perrey, & Bugaiska, 2014), including people who use hearing instruments. Unlike fMRI, fNIRS is also quiet, which furthers its suitability for auditory applications. In addition, fNIRS can have a spatial resolution of up to 1 cm (Quaresima & Ferrari, 2019) and the temporal resolution of up to 250 Hz (Scholkmann et al., 2014). Thus, fNIRS represents a compromise between the strengths of fMRI, which possesses superior spatial resolution but inferior temporal resolution to fNIRS, and EEG, which possesses a superior temporal resolution but inferior spatial resolution to fNIRS.

A considerable body of research has already demonstrated fNIRS to be useful for measuring activity of the prefrontal cortex (PFC), a region involved in effortful listening. Indeed, previous studies have successfully used fNIRS to measure PFC activity related to attention (e.g., Harrivel, Weissman, Noll, & Peltier, 2013), working memory (e.g., Fishburn, Norr, Medvedev, & Vaidya, 2014), as well as general cognitive effort (e.g., Ayaz et al., 2012). While most fNIRS studies use visual rather than auditory stimuli, a small number of recent studies have commenced using fNIRS to measure listening effort. For instance, Wijayasiri, Hartley, and Wiggins (2017) replicated the fMRI results of Wild et al. (2012), finding activity of the left IFG in normal-hearing younger adults while they were attending to degraded speech. More recently, Lawrence, Wiggins, Anderson, Davies-Thompson, and Hartley (2018) used fNIRS to measure cortical activity in younger and middle-aged adults in response to varying levels of noise-vocoding in speech. Both sets of authors also highlighted the opportunity that lies in future use of fNIRS to measure listening effort in users of hearing instruments.

### *The Current Study*

This study had two aims: (a) to establish the feasibility of fNIRS as a measure of cognitive effort during listening in older adults who use hearing aids and (b) to use fNIRS to assess differences in listening effort depending on whether or not hearing aids are used when listening to sound presented at 35 dB SL. To achieve both of these aims, older adults who use hearing aids completed a verbal n-back task with visual stimuli as well as a verbal n-back task with auditory stimuli. In both cases, stimuli consisted of a sequence of consonants (presented visually or aurally). Participants completed both of these tasks at four levels of working memory demand: 0-back

(0B; the least difficult), 1-back (1B), 2-back (2B), and 3-back (3B; the most difficult). During the 0B condition, participants indicated when the current consonant was the same as the first trial that was presented in that condition; and in the 1B, 2B, and 3B conditions, participants indicated when the current consonant was the same as the one that was presented *n* trials ago. Participants completed these two versions of the n-back task once with their own hearing aids and once unaided. Concurrently, PFC oxygenation was measured using fNIRS. PFC oxygenation was primarily considered as an average across the entire PFC. However, given that specific regions of the PFC have been implicated in verbal working memory and listening effort (in particular the dorsolateral PFC [DLPFC] and ventrolateral PFC [VLPFC]), it was possible that changes in oxygenation would be greater in some regions of the PFC than in others. Therefore, oxygenation was also measured separately in four subregions of the PFC, defined with regard to their relative position along the coronal plane: the left lateral PFC (L-LPFC), left medial PFC (L-MPFC), right medial PFC (R-MPFC), and right lateral PFC (R-MPFC).

*Aim 1: fNIRS to measure cognitive effort during listening.* The first aim of this study was to establish the feasibility of fNIRS as a measure of cognitive effort during listening in older adults who use hearing aids. In the n-back task, completed by participants in this study, cognitive effort must be increased to meet the increasing demands on verbal working memory and attention. The majority of neuroimaging studies employing the n-back task have tested younger adults with normal hearing using visual stimuli, with several studies of this description measuring brain activity as a function of visual n-back demand (e.g., Braver et al., 1997). Only a handful of studies have compared brain activation during the visual n-back task to that during the auditory n-back task, and these studies used only one level of n-back demand (e.g., Crottaz-Herbette, Anagnoson, & Menon, 2004). Thus, no study has compared brain activity during the visual and auditory n-back tasks as a function of n-back demand.

To accomplish this first aim, we considered the effect of n-back demand on PFC oxygenation during the visual n-back task as well as during the auditory n-back task. Previous fMRI research has reported an increase in PFC activity as the visual n-back task becomes more demanding (e.g., Braver et al., 1997; Callicott et al., 1999). The increase in PFC activity likely reflects the recruitment of additional cognitive resources to maintain task performance as the cognitive demands of the task increase (Nyberg, Dahlin, Neely, & Bäckman, 2009). It is assumed that the increase in PFC activity occurs until cognitive capacity is reached, at which point no more



resources are available to be invested in the task. Brain activity during the auditory n-back task is similar to that during the visual n-back task, with only minor differences being reported in the few studies that have compared the two modalities (Crottaz-Herbette et al., 2004; Rodriguez-Jimenez et al., 2009; Schumacher et al., 1996).

Based on this prior work, we predicted that as the visual verbal n-back task became more difficult, older adults with hearing loss would exhibit an increase in PFC oxygenation until their cognitive capacity was reached, at which point a leveling or decline in PFC oxygenation would be observed. We expected that this leveling or decline would occur after the 2B condition, since the only other fNIRS study of older adults completing the visual n-back task (Vermeij et al., 2012) found that effort increased up to the 2B condition. Moreover, we predicted that the same nonlinear pattern of PFC oxygenation as a function of n-back condition would be measured during the auditory verbal n-back task. Importantly, a conceptual replication of the visual n-back results that have been found using fMRI, and an extension of these findings to the auditory modality, would lend strong support to the use of fNIRS to measure cognitive effort during listening in older adults who use hearing aids. In turn, this finding would support the possible use of fNIRS to measure listening effort in this population.

**Aim 2: Hearing aids to reduce listening effort.** The second aim of this study was to use fNIRS to assess whether using hearing aids reduces listening effort in older adults with hearing loss when listening to sound presented at 35 dB SL. While fNIRS has previously been used to measure listening effort in younger and middle-aged adults with normal hearing (Lawrence et al., 2018; Wijayasiri et al., 2017), it has yet to be used to measure listening effort in older adults with hearing loss. Furthermore, while fNIRS has been applied to listeners who use cochlear implants (e.g., Olds et al., 2016; Van de Rijt et al., 2016), it has yet to be applied to listeners who use hearing aids. Moreover, no fNIRS study has yet considered how hearing instruments affect listening effort.

To accomplish this second aim, we considered the effect of hearing aid use on PFC oxygenation during the auditory n-back task. Several studies have found that the use of hearing aids reduces listening effort (e.g., Ahlstrom, Horwitz, & Dubno, 2014; Gatehouse & Gordon, 1990; Picou et al., 2013). Fewer cognitive resources should be needed to process auditory stimuli (e.g., speech) as listening becomes less demanding with improvement in the quality of sound input, and therefore the ability of hearing aids to increase audibility would be expected to reduce listening effort (Hornsby, 2013; Peelle, 2018; Picou et al., 2013). Nonetheless,

Ohlenforst et al. (2017) argued that the quality of evidence supporting the ability of hearing aids to reduce listening effort is poor.

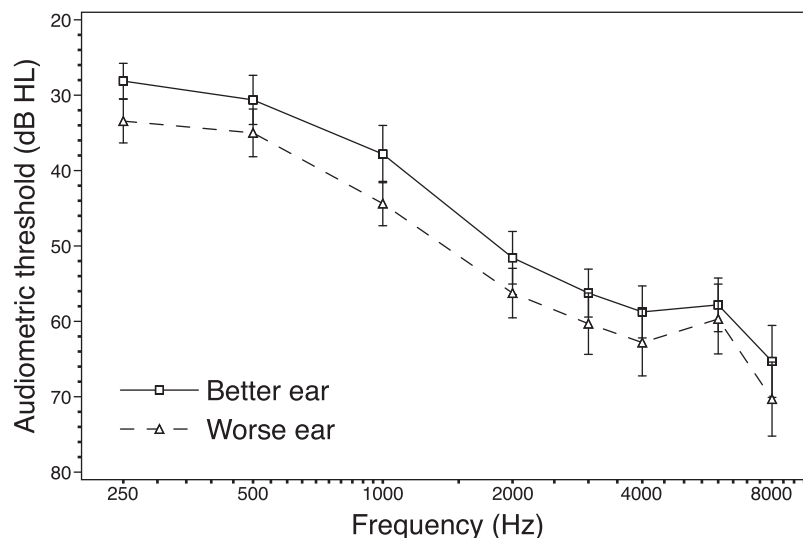
Of all studies conducted on the question of whether hearing aids reduce listening effort, only one study used a brain-based measure to address this question (EEG; Korczak, Kurtzberg, & Stapells, 2005), with the finding being that hearing aids reduced listening effort for people with hearing loss. Therefore, despite the contradictory nature of the literature, we predicted that hearing aids would reduce listening effort during the auditory n-back task. Furthermore, we predicted that there would be no effect of hearing aids on effort during the visual n-back task, which served as a control condition. If hearing aids were found to reduce PFC oxygenation during the auditory n-back task, it would suggest that, in addition to their ability to improve audibility and listening performance, hearing aids also have the potential to reduce listening effort.

## Methods

### Participants

Eighteen community-dwelling older adults were recruited from the Ryerson University Hearing Database. Eligible participants were 60 years of age or older and were experienced with and regularly used hearing aids in both ears. Participants were excluded if they had learned English after the age of seven, experienced tinnitus that interfered with their ability to understand speech, had been diagnosed with a neurological disorder, or had suffered a traumatic brain injury. One participant was excluded for not completing the experiment, and another participant was excluded for using only one hearing aid rather than two.

Data obtained from the 16 remaining participants were analyzed. They included 10 males and six females who ranged in age from 62 to 83 years ( $M = 72.06$ ,  $SD = 6.63$ ). All participants but one reported being right-handed. All were experienced hearing aid users, with two participants (12.5%) reporting 1 to 4 hours of daily use, three (18.75%) reporting 4 to 8 hours of daily use, and 11 (68.75%) reporting 8 to 16 hours of daily use. These participants had participated in previous studies, and thus all were known to have hearing loss based on audiometric testing. Their pure-tone average (PTA) thresholds were calculated by averaging their thresholds at 500, 1000, and 2000 Hz from both ears. PTA ranged from 23.33 to 58.33 dB HL ( $M = 42.60$ ,  $SD = 9.64$ ), with interaural differences in PTA ranging from 0 to 25 dB HL ( $M = 5.21$ ,  $SD = 6.29$ ; see Figure 1). Participants' scores on the Montreal Cognitive Assessment (MoCA), a screening assessment for mild cognitive impairment, ranged from 23 to 30



**Figure 1.** The average audiogram of 16 participants. Error bars show standard errors of the mean.

( $M = 26.44$ ,  $SD = 2.03$ ). These scores are indicative of normal cognitive ability based on relaxed clinical criteria (Carson, Leach, & Murphy, 2017). Participants were compensated at a rate of \$20 per hour. The study protocol was approved by the Research Ethics Board at Ryerson University.

### Design

The study was based on a three-factor within-subject design. The independent variables were n-back condition (0B, 1B, 2B, 3B), n-back modality (auditory, visual), and hearing aid use (aided, unaided). To mitigate order effects, all three variables were counterbalanced. The experiment took place over two testing sessions. Eight participants completed the n-back task aided in the first session and unaided in the second, while the other eight were unaided in the first session and aided in the second. Within each session, eight participants completed the auditory n-back task followed by the visual n-back task, while the other eight completed the visual n-back task followed by the auditory n-back task; for each participant, this order was held constant across both sessions. In addition, a modified Latin square was used to produce four n-back condition orders, with each completed by four participants; for each participant, this order was held constant across both n-back modalities as well as both testing sessions.

### Tasks and Measures

Participants completed a visual verbal n-back task and an auditory verbal n-back task. During the completion of these tasks, four dependent measures were considered: PFC oxygenation (across the entire PFC and in four

subregions of the PFC), n-back percent-correct, n-back reaction time, and n-back self-reported effort.

**N-back task.** Two versions of an n-back working memory task were administered to participants. The first of these was a visual verbal n-back task, adapted from Ragland et al. (2002). The stimuli consisted of a sequence of consonants (all letters besides A, E, I, O, U, and Y) presented visually on a computer screen. The consonants were chosen pseudorandomly from all possible 20 and presented one at a time, with a height on the screen of 5.5 cm. The stimulus duration was 500 ms and the inter-stimulus interval was 2,000 ms.

The task consisted of four conditions presented as blocks: the 0B, 1B, 2B, and 3B. During the 0B condition, participants were asked to indicate when the current consonant was the same as the first consonant that had been presented in that block; during the 1B condition, participants were asked to indicate when the current consonant was the same as the one that had been presented one trial earlier; during the 2B condition, participants were asked to indicate when the current consonant was the same as the one that had been presented two trials earlier; and finally, during the 3B condition, participants were asked to indicate when the current consonant was the same as the one that had been presented three trials earlier. Stimuli that met the criterion of the condition in which they were presented (e.g., during the 2B, a stimulus that was the same as the stimulus that had been presented two trials earlier) were referred to as targets, whereas other stimuli were referred to as non-targets. Participants had 2,500 ms after the stimulus onset to respond to each stimulus before the next stimulus was presented.

Of the stimuli presented, one third were targets and two thirds were nontargets, with the order of targets and nontargets chosen pseudorandomly. In each n-back condition, 41 trials were presented that had the potential to be targets. In addition to these trials, the 0B and 1B conditions required one additional trial at the start of the n-back condition so that a judgment could be made regarding the second trial; the 2B condition required two such additional trials, and the 3B condition required three. Thus, the 0B and 1B consisted of 42 trials, the 2B consisted of 43 trials, and the 3B consisted of 44 trials, with conditions therefore ranging from 105 to 110 s in length. Practice blocks for each n-back condition, completed prior to the task proper, consisted of nine trials with the potential to be targets.

Participants also completed an auditory verbal n-back task in which stimuli were recordings of letter names presented through a loudspeaker. The letter names were restricted to consonants and spoken with a male voice in a North American accent and included the pronunciation *zed* (as opposed to *zee*), as is used by Canadian English speakers. These stimuli were obtained from AudioMicro Stock Audio Library (<https://www.audiomicro.com/>).

For each n-back condition, performance was calculated as the proportion of trials in which the participant correctly chose to respond or not respond. Reaction time for each condition was calculated as the average time to indicate a target, thus ignoring trials in which participants did not respond. Trial reaction times of less than 100 ms were considered too fast to constitute genuine responses and were therefore excluded from this average (Whelan, 2008).

**fNIR oxygenation.** Oxygenation data were collected using the fNIR Imager 1100 (fNIR Devices, LLC, Potomac, MD), a continuous wave optical neuroimaging system (see Figure 2), at a sampling rate of 2 Hz. The software platform Cognitive Optical Brain Imaging (COBI) Studio version 1.4 was used for data collection (Ayaz & Onaral, 2005). The device included a flexible silicon sensor pad containing 4 light sources and 10 light detectors. The sources emitted light into the scalp above the PFC at wavelengths of 730 and 850 nm. Sixteen paths existed between sources and adjacent detectors, defining 16 channels (voxels) from which data could be analyzed independently. The source-detector distance was 2.5 cm, which allowed for a penetration depth of approximately 1.25 cm (León-Carrión & León-Domínguez, 2012).

In each of the 16 fNIRS channels, oxygenation was calculated as the difference between the concentrations of oxygenated and deoxygenated hemoglobin ( $HbO - HbR$ ), as has been done in other fNIRS studies conducted with the n-back task (e.g., Ayaz et al., 2012; Kuruvilla, Green, Ayaz, & Murman, 2013). This

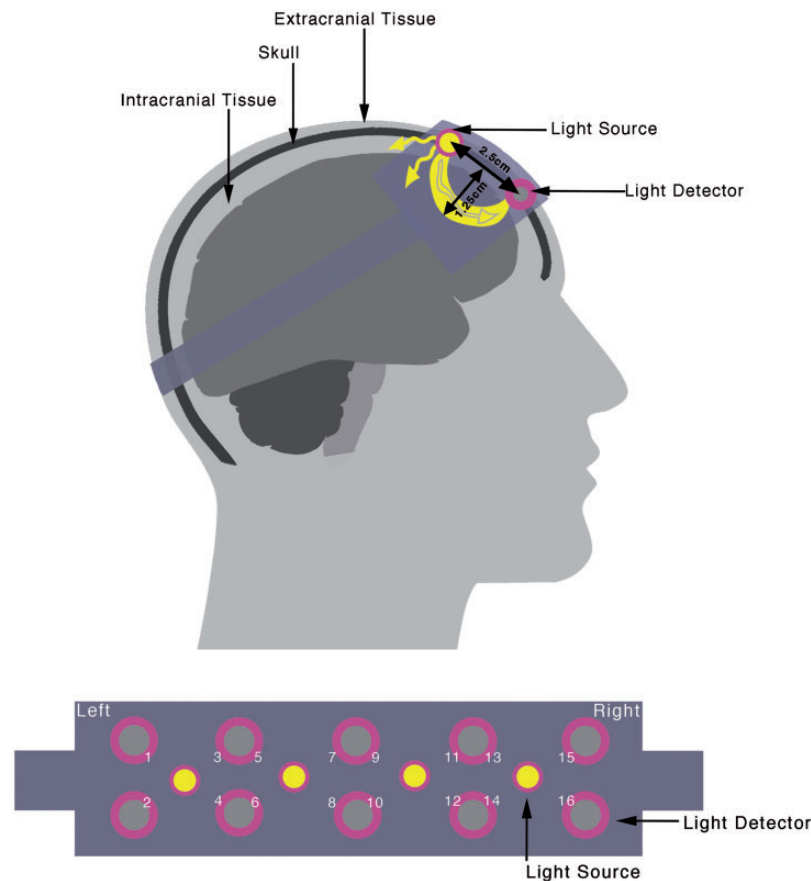
difference calculation as well as  $HbO$  on its own are more sensitive measures of brain activity than  $HbR$  or total hemoglobin ( $HbO + HbR$ ; Liang, Shewokis, & Getchell, 2016). The primary dependent measure used in this study was oxygenation averaged across all 16 channels of the PFC, which we referred to as overall PFC oxygenation. In addition to this, four subregions of the PFC were also considered, which were defined as follows: the R-LPFC included Channels 1 to 4, the R-MPFC included Channels 5 to 8, the L-MPFC included Channels 9 to 12, and the L-LPFC included Channels 13 to 16. Several studies using the same or similar devices have examined these PFC subregions as defined earlier (e.g., Anderson et al., 2018; Aranyi, Charles, & Cavazza, 2015; Cavazza, Aranyi, & Charles, 2015; Liang et al., 2016; Montgomery, Fisk, & Roberts, 2017).

**Self-reported effort.** The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) was used, in which participants were asked to self-report their subjective levels of six states on a 20-point scale: mental demand, physical demand, temporal demand, performance, effort, and frustration. This has been used in other studies involving listening effort (e.g., Ahlstrom et al., 2014) as well as in studies using fNIRS to measure cortical oxygenation during an n-back task (e.g., Montgomery et al., 2017).

### Procedure and Apparatus

Participants came into the lab for two sessions. To start the first session, participants provided written informed consent to take part in the study. Participants then completed an eligibility and demographic questionnaire and were introduced to how the n-back task works using a pencil-and-paper practice task. Once participants expressed an understanding of the task, the experimenter seated them in a double-walled sound-attenuated chamber (Industrial Acoustics Corp., Bronx, NY) in front of an Elo 1515 L 15-in. Touchscreen LCD Monitor, placed 60 cm centrally in front of participants, and a keyboard. The visual angle of the stimuli was  $5.25^\circ$ . A Yamaha MSP5 Studio Monitor loudspeaker was also placed 115 cm centrally in front of participants and at head height.

The fNIRS sensor pad (see Figure 2) was affixed to and centered on the forehead above the eyebrows, aligned with the FpZ location according to the international 10-20 system (Homan, Herman, & Purdy, 1987), as has been done in other studies using the same or similar devices (e.g., Anderson et al., 2018; Liu et al., 2017; McKendrick, Parasuraman, & Ayaz, 2015). Two straps were tightened behind the head to secure the sensor pad in place, and physical adjustments were made to maximize its contact with the scalp. The data collection software displayed the light intensity of all 16 channels,



**Figure 2.** The fNIRS sensor pad (bottom) that we used in this study was secured to the forehead like a headband (top). The system contains four light sources and 10 light detectors. Each light source emits photons of two wavelengths into the scalp. Some of the photons reflect at the scalp-air interface, while others are absorbed by extracranial or intracranial tissues, and the remainder are scattered through the layers of cortex beneath the skull. Some of these scattered photons will follow a banana-shaped path to an adjacent light detector. This source-detector pair constitutes a channel. The proportions of photons of each wavelength that reach this adjacent light detector are used to calculate the concentration of oxygenated and deoxygenated hemoglobin in that channel. The top of the figure shows one source-detector pair (channel) as an example of this concept, although the same concept applies to all 16 channels, shown at the bottom of the figure (numbered in white).

which represented how much of the light that was emitted into the scalp returned to the light detectors. An optimal signal had light intensity between 400 and 4,000 mV for all 16 channels, for both wavelengths (Propper, Patel, Christman, & Carlei, 2017). Light intensity of less than 400 mV suggested that light detectors were obstructed, whereas intensity of greater than 4,000 mV suggested that detectors were saturated (Orihuela-Espina, Leff, James, Darzi, & Yang, 2010). To maximize the number of channels in the optimal intensity range, adjustments were made to the default LED current and detector gain settings (10 mA and 10 mA, respectively). The process of fitting the sensor pad was usually completed in about 5 min.

With an optimal signal achieved, the experimenter began recording a 10-s baseline, instructing participants to relax and minimize movement during this time. With this completed, participants were then presented a series

of 10 pseudorandomly chosen auditory stimuli (spoken consonant names) via loudspeaker while not wearing their hearing aids. Participants were asked to rate, on a 5-point scale, how well they could understand these stimuli (“How well were you able to hear the letters presented to you [i.e., to tell what they are]?”). This was not part of a task and was instead used to ensure that participants could understand the stimuli while not wearing their hearing aids. All of their responses ranged from 1 (*Perfectly*) to 3 (*Somewhat well*;  $M = 2.44$ ,  $SD = 0.61$ ), indicating that stimuli were generally understood. This was done before the completion of the n-back tasks for all participants, regardless of whether they completed the visual or auditory n-back task first. The experimenter next used Inquisit version 4.0 (Millisecond Software, LLC, Seattle, WA) to start an n-back task corresponding to the appropriate counterbalancing condition.



Participants then completed the first n-back modality (e.g., visual). For the visual n-back task, participants had a chance to practice all four n-back conditions once before the start of the task proper. Visual stimuli (written consonants) were presented on the screen. Practice n-back conditions were always presented in the following order: 0B, 1B, 2B, and 3B. Participants were able to repeat the practice. Targets were identified by pressing the “A” key on the keyboard. Following the completion of the practice, participants were instructed to relax and minimize movement for 40 s. This allowed them to take a rest and for their brain activity to return to a baseline state (i.e., a state in which brain activity was no longer influenced by hemodynamic changes induced by the previous n-back condition). With this rest period completed, the experimenter instructed participants to begin the first n-back condition when ready. After the first n-back condition was completed, a message appeared on the screen prompting participants to fill out a printed copy of the NASA-TLX that was provided on a clipboard next to them. Participants were then able to begin the next n-back condition when ready. This continued until all four n-back conditions were completed once for that n-back modality. At the start of all n-back conditions, the experimenter placed a marker manually in the data collection software to indicate the time at which participants started each condition.

Participants then completed the second n-back modality (e.g., auditory). For the auditory n-back task, participants once again had the chance to practice each n-back condition before starting the task proper. Auditory stimuli (spoken consonant names) were presented via loudspeaker. For all participants, the sound presentation level in both the aided and unaided conditions was set to 35 dB SL above their PTA, which had previously been measured using a Grason-Stadler GSI 61 audiometer. Presentation levels were set for each participant at the start of a session using a Brüel & Kjær Hand-Held Analyzer Type 2250-S. This procedure served two purposes. First, it ensured that audibility was approximately equated for all participants when not wearing their hearing aids, avoiding a problem described by Gatehouse and Gordon (1990) who recognized that presenting stimuli at the same level for all participants could result in overestimating the benefit of using hearing aids for some participants and underestimating it for others. Second, following the recommendation of McGarrigle et al. (2014), by presenting the stimuli at 35 dB SL, it was possible for all participants to understand the stimuli, but listening was sufficiently challenging that hearing aids could still have a positive effect by reducing listening effort. Furthermore, if an initial gain of 35 dB SL had not been provided, the relationship between hearing aid use and listening effort could potentially have become more complicated. For

instance, if some participants were unable to hear the stimuli when not wearing their hearing aids, then they may have expended little or no listening effort in the unaided condition; in contrast, their motivation might have increased and they might have exerted more listening effort once it became possible for them to achieve their listening goals.

The second session began with participants completing the 25-item Satisfaction with Amplification in Daily Life Questionnaire, used to assess experience and satisfaction with hearing aids (Cox & Alexander, 1999). As in Session 1, participants were reintroduced to the n-back as a pencil-and-paper task. Participants then practiced and completed the n-back task in both modalities again, with each n-back condition followed by the NASA-TLX. While the same n-back condition and modality orders were used as in Session 1, the opposite hearing aid condition was used (i.e., if participants wore their hearing aids in the first session, they did not wear them in the second session, and vice versa). During the session in which participants wore their hearing aids, they did so at settings of their choosing. At the end of the second session, participants were debriefed and compensated before exiting the lab.

### *Signal Preprocessing*

Optical imaging data were preprocessed using fNIRSoft Professional version 4.6 (Ayaz, 2010). Preprocessing steps were applied independently to each session. A finite impulse response linear phase low-pass filter (order = 20, cutoff frequency = 0.1 Hz) was applied to reduce noise at higher frequencies than the hemodynamic response, such as physiological and equipment noise (Ayaz, 2010; Izzetoglu, Bunce, Izzetoglu, Onaral, & Pourrezaei, 2007). Data compromised by motion artifacts or channel saturation were removed using the Sliding-window Motion Artifact Removal algorithm (window size = 10 s, upper threshold = 0.025 nm, lower threshold = 0.003 nm; Ayaz, Izzetoglu, Shewokis, & Onaral, 2010).

Channels were then rejected if their light intensity fell below 400 mV or exceeded 4,000 mV at any point during the session (Propper et al., 2017). All remaining channels had their light data converted to oxygenation data in fNIRSoft using the MBLL (Ayaz, 2010). The resulting data were baselined channel-wise according to the 10-s baseline measured at the start of data collection: For instance, mean baseline oxygenation data from Channel 1 were subtracted from all subsequent oxygenation data in Channel 1 this process was repeated for all 16 channels. Finally, linear detrending was applied to each channel (Ayaz, 2010). With preprocessing complete, blocks of interest were defined and exported as CSV files.

### Signal Postprocessing

A custom script in MATLAB version 9.3 (The MathWorks Inc., Natick, MA) was used to further process the data. For each session, in addition to rejecting individual channels according to light levels during pre-processing, channels were also rejected if more than 50% of their data were missing across all auditory blocks and/or all visual blocks. Of the 16 channels in total, the number of channels rejected per session ranged from 0 to 6 ( $M=1.81$ ,  $SD=1.57$ ). Individual n-back condition blocks (and all data within them) were rejected if, after their completion, a participant had expressed uncertainty as to whether they had completed the task correctly. According to this criterion, 9 of 256 blocks (3.52%) were rejected. Furthermore, if a participant indicated their responses during the n-back task with the wrong key (in which cases their responses could not be recovered), reaction time and percent-correct data from that block were rejected, although PFC oxygenation and self-reported effort data were retained. A further four blocks of reaction time and percent-correct data were rejected by this criterion, and thus a total of 13 of 256 blocks of reaction time and percent-correct data (5.08%) were rejected.

For each block, participants had mean oxygenation levels calculated for all four subregions of the PFC. This was done by averaging the four channels comprising each region. These block means were calculated using data from 10 s into the block (to minimize hemodynamic carryover from the previous block; see Cavazza et al., 2015) until the end of the block. If all four channels comprising a PFC subregion were rejected, fewer than four subregions were analyzed. This occurred in only one case, in the L-MPFC, and thus while other PFC subregions had 32 sessions of data included in the analysis, the L-MPFC had only 31. The time required to fully process a participant's data was usually about 5 min.

### Statistical Analyses

**Confirmatory analyses.** All statistical analyses were done with R version 3.5.2 (R Core Team, 2018) using  $\alpha=.05$ . Planned analyses concerned the hypotheses of the study and required assessing the effects of n-back condition, modality, hearing aid use, and the interaction of Modality  $\times$  Hearing Aid Use on overall PFC oxygenation. These relationships were assessed using linear mixed-effects modeling with the "nlme" package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2018). Linear mixed-effects modeling was used because some participants had specific blocks rejected from the analysis, and this form of modeling is well suited to cope with such missing data (Baayen, Davidson, & Bates, 2008).

A linear mixed-effects model of overall PFC oxygenation, beginning with a baseline model that included only the intercepts, was built as follows: (a) intercepts were allowed to vary across participants; (b) n-back condition was added as a fixed effect; (c) slopes, representing the effect of n-back condition on overall PFC oxygenation, were allowed to vary across participants; (d) modality was added as a fixed effect, with visual as the reference condition; (e) hearing aid use was added as a fixed effect, with unaided as the reference condition; and (f) Modality  $\times$  Hearing Aid Use was added as a fixed effect. Random slopes were included for n-back condition to account for individual differences in participants' responses to change in n-back condition but were not included for modality or hearing aid use as no main effect was predicted for these variables. Furthermore, while random intercepts were always included in the final model, random slopes only remained in the final model if a likelihood-ratio test revealed that their inclusion explained significantly more variance than a model that included only random intercepts and n-back condition. All fixed effects remained in the final model regardless of significance, due to their relevance to the hypotheses.

Additional linear mixed-effects models were built, using the same procedure, with reaction time, percent-correct, and self-reported effort used as outcome measures. In all cases, multiple comparison testing for n-back condition was done using pairwise  $t$  tests with Holm–Bonferroni correction (Holm, 1979).

**Exploratory analyses.** To determine whether the pattern of PFC oxygenation during the n-back task varied across subregions of the PFC, linear mixed-effects models were created, using the same procedure, with oxygenation in each of the four subregions of the PFC used as outcome measures. A repeated-measures correlational analysis was also done using the "rmcorr" package (Bakdash & Marusich, 2017) to compare the patterns of oxygenation in each PFC subregion to one another.

To determine the association between PFC oxygenation and other measures of effort during the n-back task, a repeated-measures correlational analysis was used. This analysis assessed the relationship between self-reported effort and reaction time, between self-reported effort and oxygenation in each PFC subregion, as well as between reaction time and oxygenation in each PFC subregion.

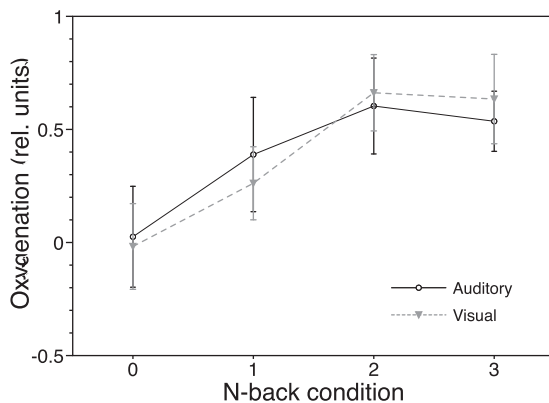
To determine whether personal characteristics predicted the benefit derived from hearing aids, a measure of hearing aid benefit was calculated for each participant. This was done by calculating, for each auditory n-back condition, the extent to which PFC oxygenation was reduced when participants were wearing their hearing aids (unaided PFC oxygenation – aided PFC

oxygenation) and averaging the resulting values. We refer to this value as hearing aid benefit for PFC oxygenation. This calculation was done for each PFC subregion. A correlational analysis was used to determine the relationship between hearing aid benefit for PFC oxygenation in each PFC subregion and age, PTA, and MoCA score. Partial correlational analyses were done when necessary with the “ppcor” package (S. Kim, 2015).

## Results

### Confirmatory Analyses

**Overall PFC oxygenation.** Figure 3 plots the effects of both n-back condition and modality on overall PFC oxygenation (averaged over the aided and unaided conditions). Intercepts varied significantly across participants, meaning that participants differed in their average levels of overall PFC oxygenation. However, slopes did not vary across participants, meaning that participants exhibited similar changes in overall PFC oxygenation as a function of n-back condition. There was a significant main effect of n-back condition on overall PFC oxygenation, with overall PFC oxygenation increasing as n-back condition increased. Overall PFC oxygenation was significantly greater in the 3B compared to the 0B or 1B, in the 2B compared to the 1B or 0B, and in the 1B compared to the 0B ( $p$  values  $< .002$ ). Furthermore, there was no main effect of modality on overall PFC oxygenation. We also used a likelihood-ratio test to assess whether we could improve the fit of the model of overall PFC oxygenation



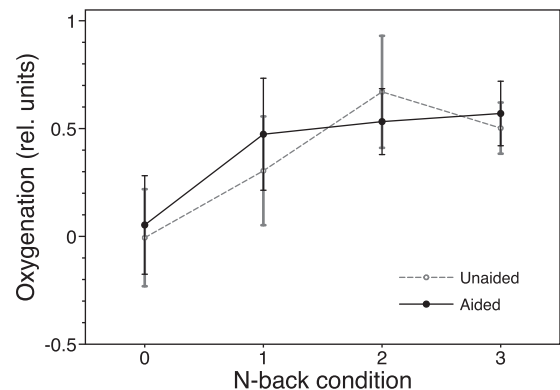
**Figure 3.** Overall PFC oxygenation as a function of n-back condition (working memory demand) and modality, averaged over the aided and unaided conditions. Error bars show standard errors of the mean. Oxygenation is reported in relative units: the device used was unable to determine the absolute concentration of oxygen in the PFC in each condition, but rather only the concentration of oxygen in each condition relative to other conditions (with oxygenation of zero representing the baseline). For both the visual and auditory n-back tasks, overall PFC oxygenation increased with task demand until the 2B, at which point it leveled.

that included random intercepts, n-back condition, and modality. However, the addition of an n-back Condition  $\times$  Modality did not improve the fit of the model,  $\chi^2(8) = 0.39$ ,  $p = .53$ .

Figure 4 plots the effects of both n-back condition and hearing aid use on overall PFC oxygenation during the auditory n-back task. There was no main effect of hearing aid use on overall PFC oxygenation as well as no interaction between modality and hearing aid use on overall PFC oxygenation. See Table 1 for a summary of the final model of overall PFC oxygenation.

**N-back reaction time.** Figure 5 plots the effects of both n-back condition and hearing aid use on reaction time during the auditory n-back task. Intercepts did not vary across participants, and neither did slopes. There was a significant main effect of n-back condition on reaction time, with reaction time becoming slower as n-back condition increased. Reaction times were significantly slower in the 3B compared to the 0B or 1B, in the 2B compared to the 1B or 0B, and in the 1B compared to the 0B ( $p$  values  $< .002$ ). There was also a significant main effect of modality on reaction time, with reaction times faster during the visual n-back task than during the auditory n-back task. There was no main effect of hearing aid use on reaction time as well as no interaction between modality and hearing aid use on reaction time. See Table 1 for a summary of the final model of reaction time.

**N-back self-reported effort.** Figure 6 plots the effects of both n-back condition and hearing aid use on self-reported effort during the auditory n-back task. Intercepts varied significantly across participants, as did slopes. There was a significant main effect of

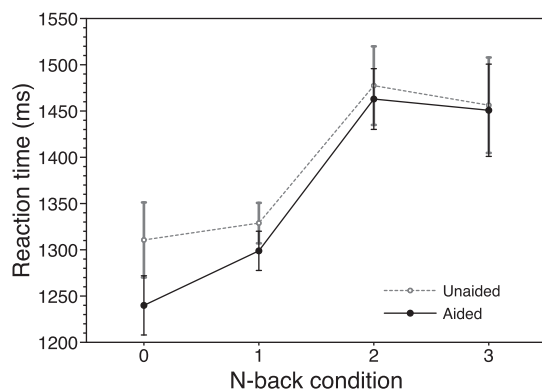


**Figure 4.** Overall PFC oxygenation during the auditory n-back task as a function of n-back condition and hearing aid use. Error bars show standard errors of the mean. The lack of an interaction between modality and hearing aid use on overall PFC oxygenation suggested that hearing aids did not reduce overall PFC oxygenation during the auditory n-back task.

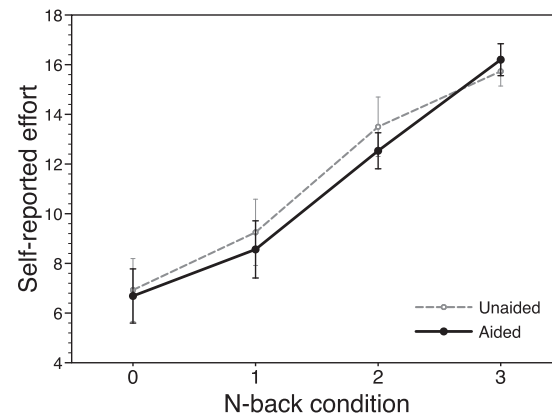
**Table 1.** A Summary of the Linear Mixed-Effects Models of Overall PFC Oxygenation, Reaction Time, Self-Reported Effort, and Percent-Correct.

	Overall PFC oxygenation	Reaction time	Self-reported effort	Percent-correct
Random intercepts	<b><math>SD = 0.25 (0.15, 0.44)</math></b> <b><math>\chi^2(3) = 6.58</math></b> <b><math>p = .010</math></b>	$SD = 65.2 (40.2, 106)$ $\chi^2(3) = 0.0000006$ $p = 1$	<b><math>SD = 3.93 (2.69, 5.72)</math></b> <b><math>\chi^2(3) = 25.8</math></b> <b><math>p &lt; .0001</math></b>	<b><math>SD = 0.014 (0.004, 0.054)</math></b> <b><math>\chi^2(3) = 6.86</math></b> <b><math>p = .009</math></b>
Random slopes	$\chi^2(6) = 3.11$ $p = .21$	$\chi^2(6) = 0.006$ $p = 1$	<b><math>SD = 1.32 (0.87, 2.00)</math></b> <b><math>\chi^2(6) = 40.5</math></b> <b><math>p &lt; .0001</math></b>	<b><math>SD = 0.015 (0.009, 0.028)</math></b> <b><math>\chi^2(6) = 18.3</math></b> <b><math>p &lt; .001</math></b>
N-back condition	<b><math>b = 0.21</math></b> <b><math>t(227) = 4.94</math></b> <b><math>p &lt; .0001</math></b>	<b><math>b = 107</math></b> <b><math>t(223) = 11.8</math></b> <b><math>p &lt; .0001</math></b>	<b><math>b = 3.40</math></b> <b><math>t(227) = 9.33</math></b> <b><math>p &lt; .0001</math></b>	<b><math>b = -0.066</math></b> <b><math>t(223) = -13.1</math></b> <b><math>p &lt; .0001</math></b>
Modality	$b = -0.002$ $t(227) = -0.012$ $p = .99$	<b><math>b = 623</math></b> <b><math>t(223) = 21.6</math></b> <b><math>p &lt; .0001</math></b>	$b = -0.48$ $t(227) = -1.02$ $p = .31$	$b = -0.008$ $t(223) = -0.74$ $p = .46$
Hearing aid use	$b = 0.002$ $t(227) = 0.015$ $p = .99$	$b = 11.0$ $t(223) = 0.38$ $p = .70$	$b = -0.54$ $t(227) = -1.15$ $p = .25$	$b = -0.001$ $t(223) = -0.12$ $p = .90$
Modality $\times$ Hearing Aid Use	$b = 0.039$ $t(227) = 0.21$ $p = .84$	$b = -38.99$ $t(223) = -0.96$ $p = .34$	$b = 0.29$ $t(227) = 0.44$ $p = .66$	$b = 0.015$ $t(223) = 1.08$ $p = .28$

Note. Significant effects are bolded.



**Figure 5.** Reaction time during the auditory n-back task as a function of n-back condition and hearing aid use. Error bars show standard errors of the mean. The lack of an interaction between modality and hearing aid (visual data not shown) on reaction time suggested that hearing aids did not reduce reaction time during the auditory n-back task.



**Figure 6.** Self-reported effort during the auditory n-back task as a function of n-back condition and hearing aid use. Error bars show standard errors of the mean. The lack of an interaction between modality and hearing aid use (visual data not shown) on self-reported effort suggested that hearing aids did not reduce self-reported effort during the auditory n-back task.

n-back condition on self-reported effort, with self-reported effort increasing as n-back condition increased. Self-reported effort was significantly greater in the 3B compared to the 0B, 1B, or 2B; in the 2B compared to the 1B or 0B; and in the 1B compared to the 0B ( $p$  values

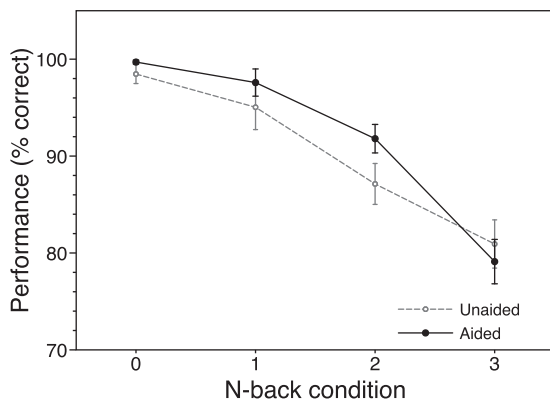
$< .0001$ ). There was no main effect of modality or hearing aid use on self-reported effort as well as no interaction between modality and hearing aid use on self-reported effort. See Table 1 for a summary of the final model of self-reported effort.



*N-back percent-correct.* Figure 7 plots the effects of both n-back condition and hearing aid use on percent-correct during the auditory n-back task. Intercepts varied significantly across participants, as did slopes. There was a significant main effect of n-back condition on percent-correct, with percent-correct decreasing as n-back condition increased. Percent-correct was significantly lower in the 3B compared to the 0B, 1B, or 2B; in the 2B compared to the 1B or 0B; and in the 1B compared to the 0B ( $p$  values < .002). There was no main effect of modality or hearing aid use on percent-correct as well as no interaction between modality and hearing aid use on percent-correct. See Table 1 for a summary of the final model of n-back percent-correct.

**Exploratory Analyses**

*Comparing oxygenation in each PFC subregion.* Figure 8 plots the effects of both n-back condition and hearing aid use on oxygenation in each PFC subregion during the auditory n-back. Oxygenation in all PFC subregions was significantly (strongly) positively correlated with one another ( $r$  values > .85,  $p$  values < .0001). Nonetheless, linear mixed-effects modeling was used to describe oxygenation in each PFC subregion separately. In each case,



**Figure 7.** Percent-correct during the auditory n-back task as a function of n-back condition and hearing aid use. Error bars show standard errors of the mean. The lack of an interaction between modality and hearing aid use (visual data not shown) on percent-correct suggested that hearing aids did not increase percent-correct during the auditory n-back task.

the addition of random intercepts improved a baseline model that included only the intercepts, but the addition of random slopes did not improve a model that included random intercepts and n-back condition. Furthermore, in each case, n-back condition was the only significant effect on oxygenation. See Table 2 for a summary of the final models of oxygenation in each PFC subregion.

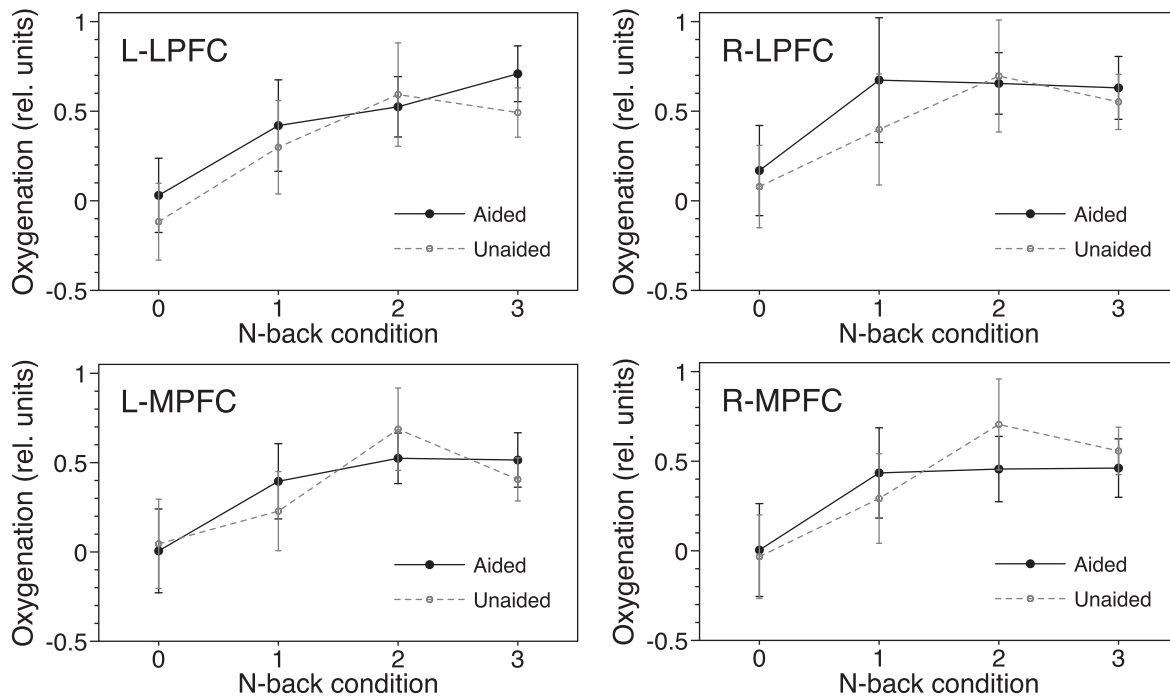
*Comparing measures of cognitive effort during listening.* A repeated-measures correlational analysis found that self-reported effort during the n-back task was significantly (weakly) positively correlated with reaction time,  $r(226) = .25, p < .001$ . Furthermore, self-reported effort was significantly (weakly) positively correlated with L-LPFC oxygenation,  $r(230) = .23, p < .001$ , R-LPFC oxygenation,  $r(230) = .13, p = .041$ , L-MPFC oxygenation,  $r(222) = .23, p < .001$ , and R-MPFC oxygenation,  $r(230) = .21, p = .002$ . There were no correlations between reaction time and oxygenation in any PFC subregion.

*Individual variability in hearing aid benefit for PFC oxygenation.* Figure 9 plots the association between hearing aid benefit for L-LPFC oxygenation and participant PTA as well as participant age. A correlational analysis found that hearing aid benefit for L-LPFC oxygenation was positively (moderately) correlated with PTA,  $r(14) = .54, p = .030$ , as well as with age,  $r(14) = .51, p = .043$ , but not with MoCA score,  $r(14) = .18, p = .49$ . We attempted to use a partial correlational analysis to better understand the influence of age and PTA on hearing aid benefit for L-PFC oxygenation. PTA and age were found to be significantly (moderately) positively correlated,  $r(14) = .51, p = .043$ , and the effects of both PTA and age on hearing aid benefit for L-LPFC oxygenation were nonsignificant when partialling out the other. PTA, age, and MoCA score were not correlated with hearing aid benefit for L-LPFC oxygenation in the other three subregions of the PFC. As a control, we also calculated the extent to which hearing aids reduced oxygenation during the visual n-back task. This was not correlated with participant PTA, age, or MoCA score in any PFC subregion.

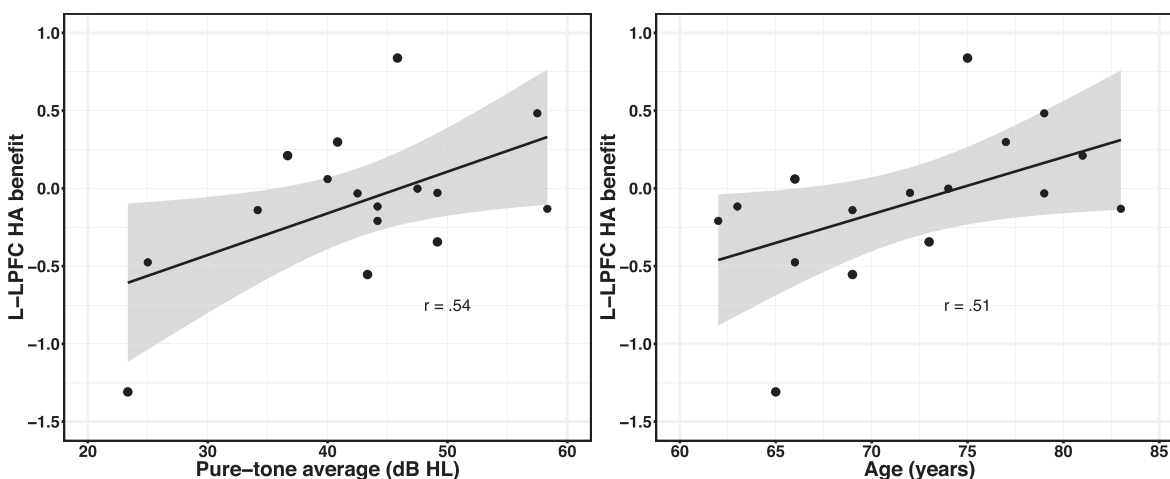
**Table 2.** The Main Effect of n-Back Condition on Oxygenation in Each PFC Subregion.

PFC subregion	$b$	$t$	$df$	$p$	Multiple comparisons
L-LPFC	0.25	5.65	227	<.0001	3B > 1B/0B, 2B > 1B/0B, 1B > 0B
R-LPFC	0.20	3.87	227	<.001	3B > 0B/1B, 2B > 0B, 1B > 1B
L-MPFC	0.19	4.55	277	<.0001	3B > 0B/1B, 2B > 1B/0B, 1B > 0B
R-MPFC	0.22	4.79	277	<.0001	3B > 0B/1B, 2B > 1B/0B, 1B > 0B

Note. R-MPFC = right medial prefrontal cortex; L-MPFC = left medial prefrontal cortex; R-LPFC = right lateral prefrontal cortex; L-LPFC = left lateral prefrontal cortex.



**Figure 8.** Oxygenation during the auditory n-back task as a function of n-back condition and hearing aid use, in the L-LPFC (top-left), R-LPFC (top-right), L-MPFC (bottom-left), and R-MPFC (bottom-right). Error bars show standard errors of the mean. The lack of an interaction between modality and hearing aid use (visual data not shown) on oxygenation in all PFC subregions suggested that hearing aids did not reduce oxygenation in any PFC subregion during the auditory n-back task. R-MPFC = right medial prefrontal cortex; L-MPFC = left medial prefrontal cortex; R-LPFC = right lateral prefrontal cortex; L-LPFC = left lateral prefrontal cortex.



**Figure 9.** During the auditory n-back, hearing aid benefit for oxygenation in the L-LPFC increased with participant PTA (left panel) and participant age (right panel). Shaded regions indicate the 95% confidence interval of the regression line. L-LPFC = left lateral prefrontal cortex; HA benefit = hearing aid benefit for PFC oxygenation (unaided oxygenation – aided oxygenation).

## Discussion

### Aim 1: fNIRS to Measure Cognitive Effort

*Effect of n-back condition on overall PFC oxygenation.* The results of this study showed that as the n-back task

became more difficult, overall PFC oxygenation increased. In particular, we found that overall PFC oxygenation increased up to the 2B condition before leveling. This pattern was found for both the visual and auditory n-back tasks, with the similarity of these modalities supported by the observations that modality

did not affect overall PFC oxygenation, and that the interaction between n-back condition and modality did not improve the fit of our final model of overall PFC oxygenation. These results were consistent with our hypothesis. They also agreed with a previous fNIRS study finding that older adults exert more effort up until the 2B condition (Vermeij et al., 2012) as well as several neuroimaging studies that report comparable PFC recruitment during the visual and auditory n-back tasks (Crottaz-Herbette et al., 2004; Rodriguez-Jimenez et al., 2009; Schumacher et al., 1996). Our findings thus support the use of fNIRS to measure cognitive effort in older adults who use hearing aids, both during visual and auditory tasks.

*Comparing oxygenation in each PFC subregion.* In all four PFC subregions, oxygenation increased from the 0B to the 2B and then leveled. All n-back condition comparisons were significant in each PFC subregion, with the exception of the R-LPFC, in which the greater oxygenation measured during the 2B compared to the 1B did not reach significance. Nonetheless, these results suggested that all PFC subregions responded to changes in n-back condition similarly. This is also true of hearing aid use and modality, both of which did not affect oxygenation in any PFC subregion. The similar activity of PFC subregions is further supported by their strong positive correlations with one another.

When considering the effect of n-back condition in each PFC subregion, the effect size was largest in the L-LPFC, Cohen's  $d_z = .75$ ; the second largest was in the R-MPFC: Cohen's  $d_z = .64$ . Thus, the activity of the L-LPFC was most sensitive to n-back condition. This PFC subregion contains the DLPFC and VLPFC, which a meta-analysis by Owen, Mcmillan, Laird, and Bullmore (2005) found to be among the brain areas most activated by the n-back task. There is also precedent for preferential recruitment of left DLPFC during a verbal n-back task, including from studies using fMRI (e.g., Smith, Jonides, & Koeppe, 1996), fNIRS (e.g., Izzetoglu, 2012), and transcranial magnetic stimulation (e.g., Mull & Seyal, 2001).

*Comparing measures of cognitive effort during listening.* To contextualize overall PFC oxygenation, we included two other measures of cognitive effort: n-back reaction time and n-back self-reported effort. Like overall PFC oxygenation, these measures increased with n-back condition and neither was affected by hearing aid use. Only reaction time was affected by modality, with faster reaction times measured during the visual n-back task than during the auditory n-back task. However, this is likely because reaction time measurement began at the stimulus onset, and during the auditory n-back task, participants had to wait to hear an entire stimulus before

indicating a response (see Crottaz-Herbette et al., 2004). Overall PFC oxygenation and reaction time exhibited similar patterns as a function of n-back condition, with both increasing until the 2B before leveling. The nonlinear pattern exhibited by these two measures suggests that both are sensitive to cognitive effort (rather than performance accuracy or a related concept), as cognitive effort is known to level once it exceeds cognitive capacity (Nyberg et al., 2009).

Self-reported effort was noticeably different from overall PFC oxygenation and reaction time; in particular, it increased linearly from the 0B to the 3B, with no leveling. This is consistent with the findings of Moore and Picou (2018), who found that participants tended to report their perceived level of performance accuracy on a task rather than the effort expended. This tendency could be an example of a heuristic that Kahneman and Frederick (2002) termed attribute substitution, in which a difficult question (e.g., How much effort did you put into this task?) is replaced by one that is easier to answer (e.g., How well do you think you performed on this task?). Thus, subjective measures of cognitive effort appear to be unreliable compared to objective measures of cognitive effort, including reaction time and overall PFC oxygenation.

During the n-back task, self-reported effort was significantly positively correlated with reaction time as well as oxygenation in all subregions of the PFC. However, while significant, these correlations were weak. Furthermore, reaction time was not correlated with oxygenation in any subregion of the PFC. These results are consistent with the results of listening effort studies that have used multiple measures of listening effort concurrently (e.g., Alhanbali, Dawes, Millman, & Munro, 2019). For instance, other studies have found weak or nonsignificant correlations between subjective and behavioral measures (e.g., Hornsby, 2013), between subjective and physiological measures (e.g., Mackersie et al., 2015), and between behavioral and physiological measures (e.g., Strand, Brown, Merchant, Brown, & Smith, 2018). Based on the lack of a consistent relationship between different measures of listening effort, it has been suggested that these measures may be tapping into different dimensions of listening effort rather than the same, unitary construct (Alhanbali et al., 2019; Strand et al., 2018).

## ***Aim 2: Hearing Aids to Reduce Listening Effort***

*Effect of hearing aids on listening effort.* There was no interaction between n-back modality and hearing aid use on any measure of effort (PFC oxygenation [overall and in each PFC subregions], reaction time, and self-reported effort), which suggests that hearing aids did not reduce listening effort during the auditory n-back task. While

contrary to our hypothesis, this finding aligns with the review by Ohlenforst et al. (2017), which found that the quality of evidence supporting the ability of hearing aids to reduce listening effort was poor. Thus, despite hearing aids often being assumed to reduce listening effort due to their ability to increase audibility (Hornsby, 2013; Peelle, 2018; Picou et al., 2013), our results suggest that, at least under some conditions, this is not the case. Furthermore, there was no interaction between modality and hearing aid use on n-back percent-correct, suggesting that hearing aids did not improve performance on the auditory n-back task.

One explanation for these findings is that, in addition to increasing the audibility of the signal of interest, hearing aids can also amplify background noise, distort the signal, and generate internal noise or feedback (Agnew, 1998). To overcome these obstacles, additional cognitive resources are likely required, potentially canceling out any reduction in listening effort that hearing aids would otherwise confer (Picou et al., 2013). Supporting this explanation, McCormack and Fortnum (2013) found that *poor sound quality* was one of the most common reasons that many owners of hearing aids chose not to wear them. Alternatively, the task that was used in this study, the auditory n-back, may have limited the effect of hearing aids. For instance, the working memory demand of the task may have overwhelmed the effect of listening demand and therefore hearing aids. Other qualities of the task may have lessened the need for hearing aids to decrease listening effort or improve performance in the first place, such as the use of simple stimuli from a closed set (spoken consonant names, rather than sentences), the presentation of stimuli in quiet (rather than in noise), and the decision to set presentation levels at 35 dB SL above PTA, which may have artificially decreased listening effort or increased percent-correct in the unaided condition.

#### *Individual variability in hearing aid benefit for listening effort.*

While there was no effect of hearing aids on PFC oxygenation on the whole, we considered whether hearing aid benefit for PFC oxygenation (interpreted as hearing aid benefit for listening effort) exhibited individual differences. Indeed, several studies have found individual differences in hearing aid benefit for listening effort (e.g., Baer, Moore, & Gatehouse 1993; Downs, 1982; Picou et al., 2013; Sarampalis, Kalluri, Edwards, & Hafter, 2009), finding that some participants experience a reduction in listening effort when using hearing aids, some experience no change in listening effort, and others experience an increase in listening effort. Similar variability in hearing aid benefit for listening effort was found in this study. Furthermore, variability in hearing aid benefit for listening effort in the L-LPFC was predicted by participant characteristics, with participants who were

older and participants who had greater hearing loss experiencing a greater reduction in L-LPFC oxygenation during the auditory n-back task from the use of hearing aids.

To our knowledge, the only study to consider the factors that predict hearing aid benefit for listening effort was Picou et al. (2013), which found that benefit was moderately negatively correlated with verbal processing speed. The authors interpreted this finding using the Ease of Language Understanding model (Rönnberg, 2003; Rönnberg et al., 2019), which states that listeners need to match the working memory representation of a signal of interest to an item in long-term memory to determine the identity of the signal, which may not have been immediately clear due to hearing loss. For listeners with a slower processing speed, it is especially effortful to match the working memory representation to an item in long-term memory before the representation decays. Thus, these listeners are helped more by hearing aids, which enhance the quality of the representation and postpone its decay.

The correlation that we observed between participant age and hearing aid benefit for listening effort in the L-LPFC could be partially accounted for by the correlation found by Picou et al. (2013), since verbal processing is known to become slower with age (Eckert, 2010). However, in this study, no measure of verbal processing speed was included, and therefore this explanation could not be evaluated. Furthermore, Hornsby (2013) proposed that one's experience with hearing aids, which could also be correlated with age, may predict the extent to which hearing aids reduce listening effort. However, several studies have found that experience with hearing aids does not predict speech recognition ability (e.g., Dawes, Munro, Kalluri, & Edwards, 2014), and therefore experience is unlikely to contribute to the correlation between age and hearing aid benefit for listening effort in the L-LPFC.

The correlation that we observed between participant PTA and hearing aid benefit for listening effort in the L-LPFC is consistent with another suggestion by Hornsby (2013), in particular that hearing aids are less likely to reduce listening effort in participants with mild to moderate (rather than severe) hearing loss. This suggestion was based on the premise that listening effort would be lower in individuals with mild to moderate hearing loss to begin with, leaving less room to improve. In this study, we attempted to ensure that stimuli were audible for all participants by presenting them at 35 dB SL above participants' PTA. While this manipulation was intended to equate listening effort across participants in the unaided condition, a correlation between PTA and hearing aid benefit for listening effort in the L-LPFC was nonetheless observed. One possible explanation for this could be that participants with a higher



PTA were less able to hear the artifacts that hearing aids produced, meaning that, for these participants, the benefits of hearing aids were more likely to outweigh the costs. Alternatively, hearing aids may have helped participants in ways that were not adequately addressed by 35 dB of flat gain, such as more pronounced amplification in mid- to high-frequency regions.

The above correlations were only significant in the L-LPFC, the same subregion that exhibited the strongest relationship with n-back condition. The L-LPFC encompasses the left DLPFC and VLPFC. Older adults with superior hearing-in-noise abilities have greater left DLPFC and VLPFC volume (Wong, Ettliger, Sheppard, Gunasekera, & Dhar, 2010). Furthermore, the left VLPFC overlaps substantially with left IFG (Broca's area), the area found by Wild et al. (2012) as well as Wijayasiri et al. (2017) to be associated with effortful listening. These findings also conform to Hickok and Poeppel's (2004, 2007) dual-stream model of speech processing, which argued that motor planning areas along the dorsal stream of speech processing are activated under challenging listening conditions.

### Limitations

This study had a number of limitations. First, the manipulation of n-back demand was an effective means to change the cognitive effort exerted to meet increased working memory demands during listening; however, this manipulation did not directly change the cognitive effort that listeners might exert to meet increased auditory demands during listening. A more direct approach would have been to use a task that manipulated speech intelligibility (e.g., varying signal-to-noise ratio) as a means to evaluate the effort exerted by listeners in response to different levels of demand due to signal factors. Second, the low spatial resolution and penetration depth of fNIRS limited our ability to measure cognitive effort in anatomically precise regions of the cortical surface (e.g., Broca's area; Wild et al., 2012) or in deeper regions thought to be involved in effortful listening (e.g., the anterior cingulate cortex; Peelle, 2018), respectively.

Moreover, several factors may have limited our ability to find an effect of hearing aid use on listening effort, including the fact that the study was likely underpowered to detect medium-sized effects. There may have also been positive effects of hearing aids that we were unable to detect, such as benefits to speech comprehension rather than performance on the n-back task. This could not be robustly assessed, given that our only measure of speech comprehension was a brief questionnaire completed before the start of the n-back task. Finally, we were also limited in our ability to determine the factors that influenced the ability of hearing aids to reduce

listening effort. For instance, participants wore their own hearing aids at settings of their choosing, which may have contributed to the variability in hearing aid benefit for listening effort. Our attempt to standardize the unaided audibility of stimuli may have also contributed to this variability, as it is likely that, due to their differing hearing profiles, not all participants (even those with the same PTA) understood the stimuli equally well when presented at 35 dB SL above PTA. We also lacked the participant information needed (e.g., a measure of verbal processing speed) to adequately investigate the mechanisms underlying the correlations of hearing aid benefit for listening effort in the L-LPFC with age and PTA.

### Conclusion

In sum, this study was motivated by the absence of an accepted brain-based measure of listening effort, particularly for clinical use. The results suggested that, for both the visual and auditory n-back tasks, fNIRS was able to measure an increase in PFC oxygenation corresponding to increased task demand in older adults who use hearing aids. This supported the potential use of fNIRS to measure listening effort in this population. However, PFC oxygenation was weakly or nonsignificantly correlated with other dependent measures, which supports a multidimensional view of listening effort. Overall, the use of hearing aids did not decrease any measure of listening effort or improve performance on the n-back task. While this could be due to the methods that we used, it could also relate to the negative effects of hearing aids on sound quality. Nonetheless, some participants did experience a reduction in listening effort when wearing hearing aids, and the extent to which hearing aids reduced oxygenation in the L-LPFC in particular was greater in participants who were older and had greater hearing loss. This could be accounted for by their slower verbal processing or reduced ability to perceive the artifacts produced by hearing aids, respectively. The ability of fNIRS to serve as a reliable and convenient measure of listening effort in older adults who use hearing aids supports its eventual application to a clinical setting, in which it could be used, perhaps in conjunction with other measures of listening effort, to help clinicians identify hearing difficulties or assess hearing instruments. To this end, future fNIRS studies should seek to measure listening effort in older adults with hearing aids during more ecologically valid tasks, such as a speech-in-noise task, and should determine how fNIRS and other measures of listening effort could be used in practice to inform clinical decisions.

## Acknowledgments

The authors would like to thank Rebecca Nurgitz for her contributions to an fNIRS pilot study with normal-hearing participants that informed the protocol implemented in this study. The authors are also indebted to Fran Copelli for her contributions to the figures as well as the coding of the experiment.

## Data Accessibility Statement

The authors will make the data collected in this study available upon reasonable request.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant awarded to the last author from the Natural Sciences and Engineering Research Council of Canada [RGPIN-2017-06969] as well as a Research Chair sponsored by Sonova AG.

## ORCID iD

Huiwen Goy  <https://orcid.org/0000-0003-4767-7863>

## References

- Agnew, J. (1998). The causes and effects of distortion and internal noise in hearing aids. *Trends in Amplification*, 3, 82–118. doi:10.1177/108471389800300302
- Ahlstrom, J. B., Horwitz, A. R., & Dubno, J. R. (2014). Spatial separation benefit for unaided and aided listening. *Ear and Hearing*, 35, 72–85. doi:10.1097/aud.000000000000068
- Alhanbali, S., Dawes, P., Millman, R. E., & Munro, K. J. (2019). Measures of listening effort are multidimensional. *Ear and Hearing*, 40(5), 1084–1097. doi:10.1097/aud.0000000000000697
- Anderson, A. A., Parsa, K., Geiger, S., Zaragoza, R., Kermanian, R., Miguel, H., . . . Gandjbakhche, A. H. (2018). Exploring the role of task performance and learning style on prefrontal hemodynamics during a working memory task. *PLoS One*, 13(6), 1–16. doi:10.1371/journal.pone.0198257
- Aranyi, G., Charles, F., & Cavazza, M. (2015). Anger-based BCI using fNIRS neurofeedback. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 511–521). New York, NY: ACM. doi:10.1145/2807442.2807447
- Ayaz, H. (2010). *Functional near infrared spectroscopy based brain computer interface* (Unpublished doctoral dissertation). Drexel University, Philadelphia, PA.
- Ayaz, H., Izzetoglu, M., Shewokis, P. A., & Onaral, B. (2010). Sliding-window motion artifact rejection for functional near-infrared spectroscopy. *Conference proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010, 6567–6570. doi:10.1109/IEMBS.2010.5627113
- Ayaz, H., & Onaral, B. (2005). Analytical software and stimulus-presentation platform to utilize, visualize and analyze near-infrared spectroscopy measures (Unpublished master's thesis). Drexel University, Philadelphia, PA.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *NeuroImage*, 59, 36–47. doi:10.1016/j.neuroimage.2011.06.023
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412. doi:10.1016/j.jml.2007.12.005
- Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times. *Journal of Rehabilitation Research and Development*, 30, 49–72. Retrieved from <https://pubs.asha.org/journal/jslhr>
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, 8, 1–13. doi:10.3389/fpsyg.2017.00456
- Beissner, F., Meissner, K., Bär, K. J., & Napadow, V. (2013). The autonomic brain: An activation likelihood estimation meta-analysis for central processing of autonomic function. *Journal of Neuroscience*, 33, 10503–10511. doi:10.1523/JNEUROSCI.1103-13.2013
- Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*, 5, 49–62. doi:10.1006/nimg.1996.0247
- Buxton, R. B., Uludağ, K., Dubowitz, D. J., & Liu, T. T. (2004). Modeling the hemodynamic response to brain activation. *NeuroImage*, 23, S220–S233. doi:10.1016/j.neuroimage.2004.07.013
- Callicott, J. H., Mattay, V., Bertolino, A., Finn, K., Coppola, R., Frank, J., . . . Weinberger, D. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex*, 9, 20–26. doi:10.1093/cercor/9.1.20
- Campbell, T., Kerlin, J. R., Bishop, C. W., & Miller, L. M. (2012). Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. *Ear and Hearing*, 33, 144–150. doi:10.1097/AUD.0b013e3182280353.
- Carson, N., Leach, L., & Murphy, K. J. (2017). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *International Journal of Geriatric Psychiatry*, 33, 379–388. doi:10.1002/gps.4756
- Cavazza M., Aranyi G., Charles F. (2015) Brain-computer interfacing to heuristic search: First results. In J. Ferrández Vicente, J. Álvarez-Sánchez, F. de la Paz López, F. Toledo-Moreo, & H. Adeli (Eds.), *Lecture Notes in Computer Science: Vol. 9107. Artificial Computation in Biology and Medicine* (pp. 312–321). Cham, Switzerland: Springer. doi:10.1007/978-3-319-18914-7\_33
- Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P. (2008). Attentional modulation of word recognition by

- children in a dual-task paradigm. *Journal of Speech Language and Hearing Research*, 51, 1042–1055. doi:10.1044/1092-4388(2008/076)
- Cox, R. M., & Alexander, G. C. (1999). Measuring satisfaction with amplification in daily life: The SADL scale. *Ear and Hearing*, 20, 306–320. doi:10.1097/00003446-199908000-00004
- Crottaz-Herbette, S., Anagnoson, R., & Menon, V. (2004). Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. *NeuroImage*, 21, 340–351. doi:10.1016/j.neuroimage.2003.09.019
- Dawes, P., Munro, K. J., Kalluri, S., & Edwards, B. (2014). Acclimatization to hearing aids. *Ear and Hearing*, 35, 203–212. doi:10.1097/aud.0b013e3182a8eda4
- Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47, 189–193. doi:10.1044/jshd.4702.189
- Eckert, M. A. (2010). Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*, 4, 1–14. doi:10.3389/neuro.09.010.2010
- Eckert, M. A., Teubner-Rhodes, S., & Vaden, K. I. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear and Hearing*, 37, 101S–110S. doi:10.1097/aud.0000000000000300
- Ferreri, L., Bigand, E., Perrey, S., & Bugaiska, A. (2014). The promise of Near-Infrared Spectroscopy (NIRS) for psychological research: A brief review. *L'Année Psychologique*, 114, 537–569. doi:10.4074/s0003503314003054
- Feuerstein, J. F. (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear and Hearing*, 13, 80–86. doi:10.1097/00003446-199204000-00003
- Fishburn, F. A., Norr, M. E., Medvedev, A. V., & Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Frontiers in Human Neuroscience*, 8, 1–11. doi:10.3389/fnhum.2014.00076
- Gagné, J., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm. *Trends in Hearing*, 21, 1–25. doi:10.1177/2331216516687287
- Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, 24, 63–68. doi:10.3109/03005369009077843
- Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, 22, 133–139. doi:10.1016/j.nec.2010.11.001
- Harrivel, A. R., Weissman, D. H., Noll, D. C., & Peltier, S. J. (2013). Monitoring attentional state with fNIRS. *Frontiers in Human Neuroscience*, 7, 1–10. doi:10.3389/fnhum.2013.00861
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology Human Mental Workload*, 52, 139–183. doi:10.1016/s0166-4115(08)62386-9
- Héту, R., Riverin, L., Lalande, N., Getty, L., & St-Cyr, C. (1988). Qualitative analysis of the handicap associated with occupational hearing loss. *British Journal of Audiology*, 22, 251–264. doi:10.3109/03005368809076462
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99. doi:10.1016/j.cognition.2003.10.011
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402. doi:10.1038/nrn2113
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70. Retrieved from <https://onlinelibrary.wiley.com>
- Homan, R. W., Herman, J., & Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, 66, 376–382. doi:10.1016/0013-4694(87)90206-9
- Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, 34, 523–534. doi:10.1097/aud.0b013e31828003d8
- Houben, R., Doorn-Bierman, M. V., & Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology*, 52, 753–761. doi:10.3109/14992027.2013.832415
- Irwin-Chase, H., & Burns, B. (2000). Developmental changes in children's abilities to share and allocate attention in a dual task. *Journal of Experimental Child Psychology*, 77, 61–85. doi:10.1006/jecp.1999.2557
- Izzetoglu, M. (2012). Functional optical brain imaging. In H. Liang, J. D. Bronzino, & D. R. Peterson (Eds.), *Biosignal processing: Principles and practices* (pp. 1–34). Boca Raton, FL: CRC Press.
- Izzetoglu, M., Bunce, S., Izzetoglu, K., Onaral, B., & Pourrezaei, K. (2007). Functional brain imaging using near-infrared technology. *IEEE Engineering in Medicine and Biology Magazine*, 26, 38–46. doi:10.1109/memb.2007.384094
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall, Inc.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York City, NY: Cambridge University Press.
- Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T.,... von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *International Journal of Audiology*, 42, 92–101. doi:10.3109/14992020309074650
- Kim, K., Punte, A. K., Mertens, G., Van de Heyning, P., Park, K. J., Choi, H.,... Song, J. J. (2015). A novel method for device-related electroencephalography artifact suppression to explore cochlear implant-related cortical changes in single-sided deafness. *Journal of Neuroscience Methods*, 255, 22–28. doi:10.1016/j.jneumeth.2015.07.020
- Kim, S. (2015) ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22, 665–674. doi:10.5351/CSAM.2015.22.6.665
- Kocsis, L., Herman, P., & Eke, A. (2006). The modified Beer-Lambert law revisited. *Physics in Medicine and Biology*, 51, N91–N98. doi:10.1088/0031-9155/51/5/n02



- Koelwijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and Hearing, 33*, 291–300. doi:10.1097/aud.0b013e3182310019
- Korczak, P. A., Kurtzberg, D., & Stapells, D. R. (2005). Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. *Ear and Hearing, 26*, 165–185. doi:10.1097/00003446-200504000-00005
- Kuruville, M. S., Green, J. R., Ayaz, H., & Murman, D. L. (2013). Neural correlates of cognitive decline in ALS: An fNIRS study of the prefrontal cortex. *Cognitive Neuroscience, 4*, 115–121. doi:10.1080/17588928.2013.797889
- Lawrence, R. J., Wiggins, I. M., Anderson, C. A., Davies-Thompson, J., & Hartley, D. E. (2018). Cortical correlates of speech intelligibility measured using functional near-infrared spectroscopy (fNIRS). *Hearing Research, 370*, 53–64. doi:10.1016/j.heares.2018.09.005
- Lemke, U. & Besser, J. (2016). Cognitive load and listening effort: Concepts and age-related considerations. *Ear and Hearing, 37*, 77S–84S. doi:10.1097/aud.0000000000000304
- León-Carrión, J., & León-Domínguez, U. (2012). Functional near-infrared spectroscopy (fNIRS): Principles and neuroscientific applications. In P. Bright (Ed.), *Neuroimaging – Methods* (pp. 47–74). doi:10.5772/23146
- Li, X., Nie, K., Karp, F., Tremblay, K. L., & Rubinstein, J. T. (2010). Characteristics of stimulus artifacts in EEG recordings induced by electrical stimulation of cochlear implants. *Proceedings of the 3rd International Conference on Biomedical Engineering and Informatics* (pp. 700–803). Piscataway, NJ: IEEE. doi:10.1109/bmei.2010.5639942
- Liang, L., Shewokis, P. A., & Getchell, N. (2016). Brain activation in the prefrontal cortex during motor and cognitive tasks in adults. *Journal of Behavioral and Brain Science, 6*, 463–474. doi:10.4236/jbbs.2016.612042
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q., Harris, T. B., Purchase-Helzner, E., ... Simonsick, E. M. (2013). Hearing loss and cognitive decline among older adults. *JAMA Internal Medicine, 173*(4), 1–14. doi:10.1001/jamainternmed.2013.1868
- Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Scientific Reports, 7*(1), 1–13. doi:10.1038/srep43293
- Mackersie, C. L., & Calderon-Moultrie, N. (2016). Autonomic nervous system reactivity during speech repetition tasks. *Ear and Hearing, 37*, 118S–125S. doi:10.1097/aud.0000000000000305
- Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology, 22*, 113–122. doi:10.3766/jaaa.22.2.6
- Mackersie, C. L., Macphree, I. X., & Heldt, E. W. (2015). Effects of hearing loss on heart rate variability and skin conductance measured during sentence recognition in noise. *Ear and Hearing, 36*, 145–154. doi:10.1097/aud.0000000000000091
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion, 23*, 209–237. doi:10.1080/02699930802204677
- McCormack, A., & Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *International Journal of Audiology, 52*, 360–368. doi:10.3109/14992027.2013.769066
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *The Quarterly Journal of Experimental Psychology Section A, 58*, 22–33. doi:10.1080/02724980443000151
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. *International Journal of Audiology, 53*, 433–445. doi:10.3109/14992027.2014.890296
- McKendrick, R., Parasuraman, R., & Ayaz, H. (2015). Wearable functional near infrared spectroscopy (fNIRS) and transcranial direct current stimulation (tDCS): Expanding vistas for neurocognitive augmentation. *Frontiers in Systems Neuroscience, 9*, 1–14. doi:10.3389/fnsys.2015.00027
- Montgomery, C., Fisk, J. E., & Roberts, C. A. (2017). Updating of working memory in ecstasy polydrug users: Findings from fNIRS. *Human Psychopharmacology: Clinical and Experimental, 32*(3), 1–12. doi:10.1002/hup.2609
- Moore, T. M., & Picou, E. M. (2018). A potential bias in subjective ratings of mental effort. *Journal of Speech Language and Hearing Research, 61*(9), 1–17. doi:10.1044/2018\_jslhrh-17-0451
- Mull, B. R., & Seyal, M. (2001). Transcranial magnetic stimulation of left prefrontal cortex impairs working memory. *Clinical Neurophysiology, 112*, 1672–1675. doi:10.1016/s1388-2457(01)00606-x
- Nyberg, L., Dahlin, E., Neely, A., & Bäckman, L. (2009). Neural correlates of variable working memory load across adult age and skill: Dissociative patterns within the frontoparietal network. *Scandinavian Journal of Psychology, 50*, 41–46. doi:10.1111/j.1467-9450.2008.00678.x
- Obleser, J., & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *NeuroImage, 55*, 713–723. doi:10.1016/j.neuroimage.2010.12.020
- Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience, 32*, 12376–12383. doi:10.1523/jneurosci.4908-11.2012
- Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A., ... Kramer, S. E. (2017). Effects of hearing impairment and hearing aid amplification on listening effort. *Ear and Hearing, 38*, 267–281. doi:10.1097/aud.0000000000000396
- Olds, C., Pollonini, L., Abaya, H., Larky, J., Loy, M., Bortfeld, H., ... Oghalai, J. S. (2016). Cortical activation patterns correlate with speech understanding after cochlear



- implantation. *Ear and Hearing*, 37, e160–e172. doi:10.1097/aud.0000000000000258
- Ong, C., Hutch, M., & Smirnakis, S. (2018). The effect of ambient light conditions on quantitative pupillometry. *Neurocritical Care*, 30(2), 316–321. doi:10.1007/s12028-018-0607-8
- Orihuela-Espina, F., Leff, D. R., James, D. R., Darzi, A. W., & Yang, G. Z. (2010). Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation. *Physics in Medicine and Biology*, 55, 3701–3724. doi:10.1088/0031-9155/55/13/009
- Owen, A. M., Mcmillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25, 46–59. doi:10.1002/hbm.20131
- Peelle, J. E. (2018). Listening Effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing*, 39, 204–214. doi:10.1097/aud.0000000000000494
- Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology*, 42, 26–32. doi:10.3109/14992020309074641
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear & Hearing*, 37, 5S–27S. doi:10.1097/AUD.0000000000000312
- Pichora-Fuller, M. K., Mick, P., & Reed, M. (2015). Hearing, cognition, and healthy aging: Social and public health implications of the links between age-related declines in hearing and cognition. *Seminars in Hearing*, 36, 122–139. doi:10.1055/s-0035-1555116
- Picou, E. M., Ricketts, T. A., & Hornsby, B. W. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear and Hearing*, 34, E52–E64. doi:10.1097/aud.0b013e31827f0431
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D.; R Core Team. (2018). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-137. Retrieved from <https://CRAN.R-project.org/package=nlme>
- Propper, R. E., Patel, N., Christman, S. D., & Carlei, C. (2017). Superior episodic memory in inconsistent-handers: A replication and extension using fNIRS. *Memory*, 25, 1390–1395. doi:10.1080/09658211.2017.1309051
- Quaresima, V., & Ferrari, M. (2019). Functional near-infrared spectroscopy (fNIRS) for assessing cerebral cortex function during human behavior in natural/social situations. *Organizational Research Methods*, 22, 46–68. doi:10.1177/1094428116658959
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ragland, J. D., Turetsky, B. I., Gur, R. C., Gunning-Dixon, F., Turner, T., Schroeder, L., ... Gur, R. E. (2002). Working memory for complex figures: An fMRI comparison of letter and fractal n-back tasks. *Neuropsychology*, 16, 370–379. doi:10.1037//0894-4105.16.3.370
- Rodriguez-Jimenez, R., Avila, C., Garcia-Navarro, C., Bagney, A., Aragon, A. M. de, Ventura Campos, N., ... Palomo, T. (2009). Differential dorsolateral prefrontal cortex activation during a verbal n-back task according to sensory modality. *Behavioural Brain Research*, 205, 299–302. doi:10.1016/j.bbr.2009.08.022
- Rönnerberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, 42, 68–76. doi:10.3109/14992020309074626
- Rönnerberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, 58, 247–261. doi:10.1080/14992027.2018.1551631
- Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech Language and Hearing Research*, 52, 1230–1240. doi:10.1044/1092-4388(2009/08-0111)
- Scarapicchia, V., Brown, C., Mayo, C., & Gawryluk, J. R. (2017). Functional magnetic resonance imaging and functional near-infrared spectroscopy: Insights from combined recording studies. *Frontiers in Human Neuroscience*, 11, 1–12. doi:10.3389/fnhum.2017.00419
- Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, 85, 6–27. doi:10.1016/j.neuroimage.2013.05.004
- Schumacher, E. H., Lauber, E., Awh, E., Jonides, J., Smith, E. E., & Koeppe, R. A. (1996). PET evidence for an amodal verbal working memory system. *NeuroImage*, 3, 79–88. doi:10.1006/nimg.1996.0009
- Signoret, C., Johnsrude, I., Classon, E., & Rudner, M. (2018). Combined effects of form- and meaning-based predictability on perceived clarity of speech. *Journal of Experimental Psychology: Human Perception and Performance*, 44, 277–285. doi:10.1037/xhp0000442
- Smith, E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, 6, 11–20. doi:10.1093/cercor/6.1.11
- Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., & Smith, J. (2018). Measuring listening effort: Convergent validity, sensitivity, and links with cognitive and personality measures. *Journal of Speech Language and Hearing Research*, 61(6), 1–24. doi:10.1044/2018\_jslhr-h-17-0257
- Van de Rijt, L. P. H., Opstal, A. J., Mylanus, E. A., Straatman, L. V., Hu, H. Y., Snik, A. F., & Wanrooij, M. M. (2016). Temporal cortex activation to audiovisual speech in normal-hearing and cochlear implant users measured with functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, 10, 1–14. doi:10.3389/fnhum.2016.00048
- Van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, 25(6), 2005–2015. doi:10.3758/s13423-018-1432-y
- Vermeij, A., van Beek, A. H. E. A., Olde Rikkert, M. G. M., Claassen, J. A. H. R., & Kessels, R. P. C. (2012). Effects of aging on cerebral oxygenation during working-memory

- performance: A functional near-infrared spectroscopy study. *PLoS One*, 7(9), 1–11. doi:10.1371/journal.pone.0046210
- Ward, C. M., Rogers, C. S., Van Engen, K. J., & Peelle, J. E. (2015). Effects of age, acoustic challenge, and verbal working memory on recall of narrative speech. *Experimental Aging Research*, 42, 97–111. doi:10.1080/0361073x.2016.1108785
- Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, 58, 475–482. doi:10.1007/bf03395630
- Wijayasiri, P., Hartley, D. E., & Wiggins, I. M. (2017). Brain activity underlying the recovery of meaning from degraded speech: A functional near-infrared spectroscopy (fNIRS) study. *Hearing Research*, 351, 55–67. doi:10.1016/j.heares.2017.05.010
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience*, 32, 14010–14021. doi:10.1523/jneurosci.1528-12.2012
- Wong, P. C. M., Ettliger, M., Sheppard, J. P., Gunasekera, G. M., & Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear and Hearing*, 31, 471–479. doi:10.1097/aud.0b013e3181d709c2