What is the Optimum Way to Measure the Concept of Listening Effort in Children with Hearing Impairment?

A thesis submitted to the University of Manchester for the degree of Master of Philosophy in the Faculty of Biology, Medicine and Health

2023

Callum A. Shields

School of Biological Sciences

Contents

Word Count 32095

List of figures, tables, and appendices included in subsequent chapters

List of abbreviations

Abbreviation	Meaning
BVP	Blood Volume Pulse
CICS	Cochlear Implant in Children's Support group
CIS	Checklist of Individual Strengths
CPY	Children and Young People
DIN	Digits in Noise
EEG	Electroencephalography
EMG	Electromyography
ENT	Ear, Nose and Throat
f-MRI	Functional Magnetic Resonance Imaging
FNIRS	Functional Near Infra-red Spectroscopy
FUEL	Framework for Understanding Effortful Listening
GRADE	Grading of Recommendations, Assessment, Development, and Evaluations
HRA	Health Research Authority
HRV	Heart Rate Variability
ICF	Informed Consent Form
IQR	Interquartile Range
ITU-T	International Telecommunication Union
LE	Listening Effort
MFT	Manchester Foundation Trust
MOSS	Manchester Online Speech-perception Suite
NASA TXL	National Aeronautics and Space Administration Task Load Index
NHS	National Health Service

NIHR	National Institute of Health and Care Research
OP	Outpatient
PICOS	Population, Intervention, Control, Outcome, Study design
PIS	Participant Information Sheet
PoMS	Profile of Mood States
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSS	Perceived Stress Scale
REC	Research Ethics Committee
RMCH	Royal Manchester Children's Hospital
RoB	Risk of Bias
RSES	Rosenberg Self-esteem Scale
SD	Standard Deviation
SHQ	Spatial Hearing Questionnaire
SNR	Signal to Noise Ratio
SSQ	Speech, Spatial Qualities
VAS	Visual Analogue Scale
VAS-F	Visual Analogue Scale - Fatigue
VFS-C	Vanderbilt Fatigue Scale - Children

Abstract

Background

Listening effort (LE) can be described as the effort required to understand and interpret an auditory stimulus. It is a phenomenon experienced by everyone but is most acutely felt by those with a hearing impairment. Currently, measuring LE is not part of the routine clinical assessment of children and young people with a hearing impairment. This is despite growing evidence to suggest that it may have a detrimental effect on quality-of-life factors and educational attainment. The importance of considering LE as an outcome measure is further emphasized when discussing those individuals with a mild or asymmetrical hearing impairment. In this group, traditional outcome measures such as pure tone audiometry may not adequately capture the burden associated with their hearing difficulties. One of the main barriers to incorporating LE into clinical practice is the lack of consensus on how to measure the concept. Many proposed tools exist, but there is a lack of evidence over which is the optimum instrument. This thesis aims to explore this issue by outlining the current challenges associated with measuring listening effort, exploring this further by conducting a systematic review of the correlations between different LE tools and combining this with normative research to look at potential alternative options for measuring LE.

Methods

Initially, an editorial letter was composed to outline the key challenges preventing LE from being used routinely. Subsequently, a systematic review was undertaken to explore the different tools that measure LE. Papers that compared two or more specific tools via correlational analysis were included. The outline of the search process is presented in a PRISMA diagram. The results of the data extraction were presented as a narrative synthesis due to the heterogeneity of included studies. Risk of bias was assessed using the GRADE and ROBINs I tools. Additionally, a protocol for exploratory research was developed to determine the relationship between LE and other related factors such as fatigue,

confidence, and stress. A cohort of 100 participants with a range of hearing abilities will be recruited. Each participant will be sent a series of online questionnaires which measured the aforementioned quality of life factors.

Results

The editorial letter highlighted the overarching barrier associated with LE as an outcome measure relates to the lack of consensus on the best way to measure the concept. This finding is enshrined in the multidimensional model of LE explaining that each measuring instrument taps into a different component of LE and therefore making it difficult to select one all-encompassing tool. This formed the basis of the systematic review of this thesis which looked to see whether this lack of correlation applied across the literature. The systematic review results show that current LE measures are weakly and insignificantly correlated with one another. Even when significant correlations did occur, they were often dependent on the experimental conditions of each study and, therefore, may not translate into real-world clinical practice. Together this supports the idea the no single tool encompasses the complex concept of LE and thus supports the multidimensional mode. Therefore, an alternative approach to capturing LE would be beneficial. An exploratory study investigating the use of downstream consequences as a proxy measure of LE is proposed in chapter 4.

Conclusion

Current measures of LE do not correlate well with one another and therefore appear to capture different elements of the concept. The work within this thesis has suggested the potential of a relationship between LE and fatigue, stress, and self-confidence. The protocol contained within this thesis details a study whereby the validity of using these downstream effects as a proxy measure for LE can be assessed. Further work is needed to undertake this study and subsequently investigate other potential sequelae of prolonged exposure to effortful listening. This approach may represent a more real-world embodiment of the challenges associated with LE and may prove more clinically acceptable due to their comparatively lower resource intensity.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or another institute of learning.

Copyright Statement

Acknowledgments

I would like to thank all the staff members within the ENT department at Royal Manchester Children's Hospital. They have provided me with expert guidance and direction throughout my time working alongside them. Special thanks to Professor Iain Bruce and Mrs Jaya Nichani, my co-supervisors, who have proved instrumental in the development of the studies contained within this thesis. I would also like to thank Mark Sladen for committing countless hours helping to screen and extract data for the systematic review. Again, special thanks to Dr. Karolina Kluk de-Kort for co-supervising and assisting with ethical approval for my thesis.

Outside of the direct clinical team, I would like to thank Tasha Ghouri, Global Ambassador for Deafness, who helped with the promotion of the exploratory research contained within this thesis. This gave us a tremendous platform to disseminate our research to a wider audience.

Finally, I would like to thank the team at Cochlear Limited who helped to fund this project and facilitated me to present this research at international conferences.

Chapter 1: Introduction

Effort applied to listening

The conceptualisation of the human experience of ‘effort’ is a relatively new area of psychological research. First formally characterised by Kahneman, effort was described as an objective process, distinct from physical strain and inextricably linked to attention (Kahneman 1973).

Although Kahneman focused on the subject of effort broadly, iterations of his models have been applied specifically to different branches of psychology to provide a more nuanced explanation to certain observations. Within the world of audiology and hearing science, the first working definition of the effort evoked specifically from a listening task, or “listening effort”, is provided by Hick and Tharpe as

"*Listening effort refers to the attention and cognitive resources required to understand speech"* (Hick and Tharpe 2002)

Listening effort is a phenomenon experienced by all individuals, but it most acutely affects those with a degree of hearing impairment. Many studies have supported this observation by contrasting the effort experienced by individuals with hearing impairment with their normal-hearing counterparts (Ohlenforst et al. 2017).

Since Hick and Tharpe first characterised listening effort, research within this domain has accelerated. Much of our contemporary understanding of listening effort stems from the Framework of Understanding Effortful Listening (FUEL) workshop, organised in 2016 by Pichora-Fuller and colleagues (Pichora-Fuller et al. 2016). The FUEL workshop expanded on the above definition to reflect that listening effort applies more generally than just understanding speech. Here the concept was defined as

"*The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when tasks involve listening*" (Pichora-Fuller et al. 2016)

Within the FUEL workshop, reference was also given to additional determinants that may affect the amount of effort experienced for a given listening situation. These include signal characteristics, the listener's hearing ability, the context of the conversation, and the speaker's vocabulary (Pichora-Fuller et al. 2016). These factors represent targets for interventions that aim to alleviate some of the detrimental effects of effortful listening. The impact of prolonged exposure to effortful listening is far-reaching (discussed in subsequent chapters of this thesis); hence it is vital to explore how we can further our understanding of listening effort to enact positive changes for those most affected by the construct.

Measuring listening effort

Currently, the measurement of LE is not incorporated into routine clinical practice. One of the barriers preventing LE's application is the lack of consensus on the best way to measure the construct (Alhanbali et al. 2019; McGarrigle et al. 2014; Pichora-Fuller et al. 2016). No single instrument has been validated as a reliable outcome measure. Paradoxically, many proposed measures have been used to sample LE in a research context. These tools are classified into three overarching categories – physiological, behavioural, and self-reported measures. Within each category exists a cornucopia of specific tools that tap into a particular component of effort perception.

Physiological measures include functional imaging modalities such as infra-red spectroscopy and magnetic resonance studies (Rovetti et al. 2019; White and Langdon 2021; Zekveld et al. 2014); hormonal tests such as cortisol and chromogranin A (Kramer et al. 2016; Dwyer et al. 2019); and electrophysiological measures such as EEG and pupillometry (Kramer et al. 2016; Lau et al. 2019; Magudilu Srishyla Kumar 2020; Key et al. 2017; Bertoli and Bodmer 2014).

Behavioural measures focus on the individual's performance to a specific task to determine their degree of effort. They work under the premise of the cognitive spare capacity model, in that they rely on the idea that each individual has a finite amount of mental resources to devote to a particular task, once these have been extinguished, effort ensues (Degeest et al. 2016; Keidser et al. 2015). This is commonly achieved through a dual-task paradigm where the primary task is designed to consume the individual's mental resources, thereby reducing the performance of the secondary task (Strand et al. 2018; Zekveld et al. 2020; Picou et al. 2017; Picou and Ricketts 2018; Hornsby 2013). This method requires the primary task to be sufficiently difficult to bankrupt cognitive resources, which may not always be feasible (McGarrigle et al. 2014). In addition, it may be difficult to administer this type of test to particular population groups, such as children and those with cognitive impairment (Choi et al. 2008).

Self-reported measures are subject and most often take the form of questionnaires; however, no specific listening effort questionnaire has been validated. As a result, these often take the form of generic effort questionnaires, such as the NASA task load index (Alhanbali et al. 2019; Rovetti et al. 2019; Seeman and Sims 2015), or by adopting relevant sections of existing hearing questionnaires as is done with the Speech, Spatial and Qualities (SSQ) questionnaire (Megha and Maruthy 2019; Russo et al. 2020). Self-reported instruments rely on the perceived effort being congruent to the intrinsic effort, which conflicts Kahneman's view that effort is an objective experience (Bruya and Tang 2018). They are also inherently susceptible to recall bias, which must be considered when interpreting the results (Choi and Pak 2004).

The motivation for exploring listening effort

Several studies have shown the potential of including listening effort within an audiological test battery (Strand et al. 2020; Winn and Teece 2021). Those with mild and asymmetrical hearing losses may represent the population group who stand to benefit most from wider adoption of LE. This reasoning mainly relates to detecting differences that traditional outcome measures may not identify. Take for example an individual with a mild hearing loss, they may perform adequately for outcomes such as speech recognition and sound localisation but these measures fail to appreciate the added effort they experience compared with others (Bess and Hornsby 2014; Hick and Tharpe 2002). Indeed, it may be the case that they exert more mental effort on a day-to-day basis than someone with a more severe hearing loss who benefits from a given audiological intervention (Picou et al. 2013; Downs 1982; Abdel-Latif and Meister 2022). The pitfalls associated with relying solely on sound perception in the assessment of hearing loss has been demonstrated in research. One study showed that intelligibility scores remain constant across different masking conditions, whereas listening effort, measured by pupillometry, increased as the masking became more difficult (Winn and Teece 2021). This finding resonates with another study that investigated the impact of various hearing device features (Wendt et al. 2017). Here, when intelligibility was set to 95%, there was no discernable difference in speech recognition scores obtained with or without the noise reduction feature enabled. However, noise reduction did lead to an improvement in listening effort scores again measured by pupillometry (Wendt et al. 2017). Therefore, the clinical evaluation of LE may provide useful information to guide the advice and recommendations given to those with milder losses whom may have historically been managed more conservatively.

From a real-world perspective, the detrimental impact of LE can be explored by considering academic and social implications. Research has shown that hearing loss can lead to adverse effects on academic performance for children and young people (CYP) (Bess et al. 1998; Khairi Md Daud et al. 2010; Traxler 2000). As a result, the modern education system responding by incorporating a range of strategies such as preferential seating, teaching assistants, microphones and loop recorders, laptops and other technology-based aids, and the use of multimedia style teaching (ADCET 2023). The aim of these interventions is to reduce the educational disparity between those with normal hearing and those with a hearing impairment. As a result, more and more CYP with a hearing impairment are being educated in mainstream schools (Reed et al. 2008). The recurring issue that stems from the current approach relates to those CYP who have mild and asymmetrical loses. As mentioned above, this cohort may go under the radar and be deemed to cope sufficiently without the added support. Currently, little attention is paid to the extra mental resources spent by these CYP in the pursuit of educational attainment. Furthermore, focus group research has shown that the prospect of listening effort can deter individuals with a hearing impairment from partaking in social situations, thus leading to isolation (Hughes et al. 2018). The desire to interact with others improved following cochlear implant surgery, which was thought to be due to the improvement in listening effort elicited by the device (Hughes et al. 2018).

Together these studies demonstrate that individuals with hearing impairment are experiencing an unmet burden due to LE. To remedy this situation, we must broaden our understanding of this complex topic and strive toward enshrining LE measurement into routine clinical practice.

Aims of thesis

This thesis aims to determine whether LE can be measured accurately using currently available instruments. It is predicted that it will not be possible to identify one specific measure which can capture LE in its entirety. As such, this thesis will also investigate potential alternative approaches to incorporate LE into routine clinical practice.

Outline of thesis

This thesis is divided into several chapters, each representing a different manuscript which helps to address the aim presented above. An outline of each chapter, including the rationale for its inclusion, key findings, and student contributions, is shown below.

Chapter 2: Listening Effort: WHAT is it, HOW is it measured and WHY is it important? (Shields et al. 2021) Editorial letter, published in Cochlear Implants International.

Rationale: This editorial letter was written to provide an overview of LE, including some key challenges associated with the topic. It discusses the current instruments used to measure LE, the clinical need for LE as an outcome measure, and the multi-dimensional model of LE.

Key findings: This editorial letter highlights the lack of congruency between outcome measures for LE found within individual studies. It also indicates the potential application of LE in borderline candidacy.

Student contribution: Write up of editorial letter.

Chapter 3: Exploring the Correlations Between Measures of Listening Effort in adults and children: A Systematic Review with Narrative Synthesis. (C. Shields et al. 2023), published in Trends in Hearing.

Rationale: To provide a comprehensive evaluation of the current literature basis for the correlations between measures of LE. To assess the quality of the literature and suggest ways to plan future research to alleviate any deficiencies.

Key findings: Correlational analyses between measures of LE most often yield non-significant results with a fair strength of association. Significant and more substantial correlations can occur but tend to be specific to the experimental conditions, thus limiting generalisability. These findings support the notion of the multi-dimensional model of LE.

Student contribution: Protocol development, literature searching, screening, data extraction, risk of bias, statistical analysis, data visualisation, and write-up.

Chapter 4: Listening effort and downstream effects due to hearing loss in children and young people: an online quantitative questionnaire-based observational study. (C. A. Shields et al. 2023), published in BMJ Open.

Rationale: To develop a peer-reviewed and transparent protocol for our subsequent study. This ensures the methodological rigor of the normative research contained in the following chapter.

Key findings: Full outline of the methodological process used for chapter 5, including study design, consent process, statistical analysis plan, data management plan, and dissemination plan. It also presents favourable opinions from industry sponsors, study sponsors, and ethical approval.

Student contribution: Co-investigator, protocol development/write-up, grant acquisition, statistical plan.

Following these chapters, this thesis will present a general discussion section that summarises the key findings across all manuscripts, provides an answer to the research questions posited within the thesis aim section above, addresses potential limitations of the research contained herewith and outlines the next steps which should be considered to explore this subject further. Figures and tables will be included at the end of each chapter; an overarching list of all figures and tables contained within this thesis can be found on page 3.

References

Abdel-Latif, K.H.A. and Meister, H. (2022). Speech Recognition and Listening Effort in Cochlear Implant Recipients and Normal-Hearing Listeners. *Frontiers in Neuroscience*, 15. [online]. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2021.725412 [Accessed May 19, 2023].

ADCET. (2023). Inclusive Teaching: Deaf and Hard of Hearing. *ADCET*. [online]. Available from: https://www.adcet.edu.au/inclusive-teaching/specific-disabilities/deaf-hearing-impaired [Accessed May 19, 2023].

Alhanbali, S. et al. (2019). Measures of Listening Effort Are Multidimensional. *Ear and Hearing*, 40(5), pp.1084–1097.

Bertoli, S. and Bodmer, D. (2014). Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 125(5), pp.1030–1041.

Bess, F., Dodd-Murphey, J. and Parker, R. (1998). Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status - PubMed. [online]. Available from: https://pubmed.ncbi.nlm.nih.gov/9796643/ [Accessed December 15, 2021].

Bess, F.H. and Hornsby, B.W.Y. (2014). Commentary: listening can be exhausting--fatigue in children and adults with hearing loss. *Ear and Hearing*, 35(6), pp.592–599.

Bruya, B. and Tang, Y.-Y. (2018). Is Attention Really Effort? Revisiting Daniel Kahneman's Influential 1973 Book Attention and Effort. *Frontiers in Psychology*, 9, p.1133.

Choi, B.C.K. and Pak, A.W.P. (2004). A Catalog of Biases in Questionnaires. *Preventing Chronic Disease*, 2(1), p.A13.

Choi, S. et al. (2008). Attentional Modulation of Word Recognition by Children in a Dual-Task Paradigm. *Journal of Speech, Language, and Hearing Research*, 51(4), pp.1042–1054.

Degeest, S., Keppler, H. and Corthals, P. (2016). The Effect of Tinnitus on Listening Effort in Normal-Hearing Young Adults: A Preliminary Study. *Journal of Speech Language and Hearing Research*, 60.

Downs, D.W. (1982). Effects of hearing and use on speech discrimination and listening effort. *The Journal of Speech and Hearing Disorders*, 47(2), pp.189–193.

Dwyer, R.T. et al. (2019). Diurnal Cortisol Levels and Subjective Ratings of Effort and Fatigue in Adult Cochlear Implant Users: A Pilot Study. *American Journal of Audiology*, 28(3), pp.686–696.

Hick, C.B. and Tharpe, A.M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of speech, language, and hearing research: JSLHR*, 45(3), pp.573–584.

Hornsby, B.W.Y. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, 34(5), pp.523–534.

Hughes, S.E. et al. (2018). Social Connectedness and Perceived Listening Effort in Adult Cochlear Implant Users: A Grounded Theory to Establish Content Validity for a New Patient-Reported Outcome Measure. *Ear and Hearing*, 39(5), pp.922–934.

Kahneman, D. (1973). *Attention and Effort*. 1st ed. Prentice-Hall.

Keidser, G. et al. (2015). Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations. *Frontiers in Psychology*, 6, p.597.

Key, A.P. et al. (2017). Speech-Processing Fatigue in Children: Auditory Event-Related Potential and Behavioral Measures. *Journal of Speech, Language, and Hearing Research : JSLHR*, 60(7), pp.2090–2104.

Khairi Md Daud, M. et al. (2010). The effect of mild hearing loss on academic performance in primary school children. *International Journal of Pediatric Otorhinolaryngology*, 74(1), pp.67–70.

Kramer, S.E., Teunissen, C.E. and Zekveld, A.A. (2016). Cortisol, Chromogranin A, and Pupillary Responses Evoked by Speech Recognition Tasks in Normally Hearing and Hard-of-Hearing Listeners: A Pilot Study. *Ear and Hearing*, 37 Suppl 1, pp.126S–35S.

Lau, M.K. et al. (2019). Effect of Auditory Task Type on Physiological and Subjective Measures of Listening Effort in Individuals With Normal Hearing. *Journal of speech, language, and hearing research: JSLHR*, 62(5), pp.1549–1560.

Magudilu Srishyla Kumar, L. (2020). *Measuring listening effort using physiological, behavioral and subjective methods in normal hearing subjects: Effect of signal to noise ratio and presentation level*. James Madison University. [online]. Available from: https://commons.lib.jmu.edu/diss202029/7.

McGarrigle, R. et al. (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(7), pp.433–440.

Megha and Maruthy, S. (2019). Auditory and Cognitive Attributes of Hearing Aid Acclimatization in Individuals With Sensorineural Hearing Loss. *American Journal of Audiology*, 28(2S), pp.460–470.

Ohlenforst, B. et al. (2017). Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: A Systematic Review. *Ear and Hearing*, 38(3), pp.267–281.

Pichora-Fuller, M.K. et al. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37 Suppl 1, pp.5S-27S.

Picou, E.M., Moore, T.M. and Ricketts, T.A. (2017). The Effects of Directional Processing on Objective and Subjective Listening Effort. *Journal of speech, language, and hearing research: JSLHR*, 60(1), pp.199–211.

Picou, E.M. and Ricketts, T.A. (2018). The relationship between speech recognition, behavioural listening effort, and subjective ratings. *International Journal of Audiology*, 57(6), pp.457–467.

Picou, E.M., Ricketts, T.A. and Hornsby, B.W.Y. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear and Hearing*, 34(5), pp.e52-64.

Reed, S., Antia, S.D. and Kreimeyer, K.H. (2008). Academic Status of Deaf and Hard-of-Hearing Students in Public Schools: Student, Home, and Service Facilitators and Detractors. *The Journal of Deaf Studies and Deaf Education*, 13(4), pp.485–502.

Rovetti, J. et al. (2019). Functional Near-Infrared Spectroscopy as a Measure of Listening Effort in Older Adults Who Use Hearing Aids. *Trends in Hearing*, 23, p.2331216519886722.

Russo, F.Y. et al. (2020). Pupillometry Assessment of Speech Recognition and Listening Experience in Adult Cochlear Implant Patients. *Frontiers in Neuroscience*, 14, p.556675.

Seeman, S. and Sims, R. (2015). Comparison of Psychophysiological and Dual-Task Measures of Listening Effort. *Journal of speech, language, and hearing research: JSLHR*, 58(6), pp.1781–1792.

Shields, C. et al. (2023). Exploring the Correlations Between Measures of Listening Effort in Adults and Children: A Systematic Review with Narrative Synthesis. *Trends in Hearing*, 27, p.23312165221137116.

Shields, C. et al. (2021). Listening effort: WHAT is it, HOW is it measured and WHY is it important? *Cochlear Implants International*, 0(0), pp.1–4.

Shields, C.A. et al. (2023). Listening effort and downstream effects due to hearing loss in children and young people: an online quantitative questionnaire-based observational study. *BMJ Open*, 13(2), p.e069719.

Strand, J.F. et al. (2018). Measuring Listening Effort: Convergent Validity, Sensitivity, and Links With Cognitive and Personality Measures. *Journal of speech, language, and hearing research: JSLHR*, 61(6), pp.1463–1486.

Strand, J.F. et al. (2020). Understanding Speech amid the Jingle and Jangle: Recommendations for Improving Measurement Practices in Listening Effort Research. *Auditory Perception & Cognition*, 3(4), pp.169–188.

Traxler, C.B. (2000). The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students. *The Journal of Deaf Studies and Deaf Education*, 5(4), pp.337–348.

Wendt, D., Hietkamp, R.K. and Lunner, T. (2017). Impact of Noise and Noise Reduction on Processing Effort: A Pupillometry Study. *Ear and Hearing*, 38(6), pp.690–700.

White, B.E. and Langdon, C. (2021). The cortical organization of listening effort: New insight from functional near-infrared spectroscopy. *NeuroImage*, 240, p.118324.

Winn, M.B. and Teece, K.H. (2021). Listening Effort Is Not the Same as Speech Intelligibility Score. *Trends in Hearing*, 25, p.23312165211027690.

Zekveld, A. et al. (2020). The Influence of Hearing Loss on Cognitive Control in an Auditory Conflict Task: Behavioral and Pupillometry Findings. *Journal of Speech, Language, and Hearing Research*, 63, pp.1–10.

Zekveld, A.A. et al. (2014). The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*, 101, pp.76–86.

Chapter 2: Listening Effort: *WHAT* is it, *HOW* is it measured and *WHY* is it important?

Published editorial: C. Shields[1,] H. Willis[2], J. Nichani[1], M. Sladen[1], and K. Kluk-de Kort[3] (2021). Listening Effort: *WHAT* is it, *HOW* is it measured and *WHY* is it important?. Cochlear Implants International VOL. 23, NO. 2, 114-117. DOI: https://doi.org/10.1080/14670100.2021.1992941

[1] Department of Paediatric Otolaryngology, Royal Manchester Children's Hospital, Manchester Foundation Trust, Manchester, UK

[2] Independent Stress Management Consultant and CI User, Reading, UK

[3] Division of Human Communication, Development & Hearing, University of Manchester, Manchester, UK

To start with a question, how do you connect the following group of measurements: EEG spikes, Cortisol levels, Chromogranin A, Reaction times, Infra-red spectroscopy, Driving simulators, Questionnaires, Pupil diameter, Response times, Fatigue scales, Functional MRIs, and Skin conductance? Although seemingly unrelated, these tools have all been proposed as a proxy measure for listening effort (LE). The variety and abundance of research approaches emphasise the lack of a clear consensus in the literature regarding the best way to measure LE. Even defining the topic of LE has its challenges, perhaps the most structured definition stems from the work of Pichora-Fuller and colleagues via the Framework for Understanding Effortful Listening (FUEL):

“The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when tasks involve listening” (Pichora-Fuller et al., 2016)

This definition highlights some of the key influencing factors which contribute to the level of LE experienced by individuals, namely task difficulty; listening ability; and motivation to succeed (Ayasse & Wingfield, 2018; Bess & Hornsby, 2014; Brookhart et al., 2006). It seems prudent within this paper to further explore each of these components to construct a broader picture of LE.

First, task difficulty is enshrined within Pichora-Fuller's definition through the phrase 'overcome obstacles', which creates an image of the participant needing to actively strive to match the demands of the task. At face value, this seems straightforward: the more difficult the task, the more effort required. Indeed, this notion is reflected if the effort necessary to understand a conversation in a busy restaurant is compared with the effort necessitated by a peaceful environment such as a library. Furthermore, this underlying principle is mirrored by many LE studies (Bernarding et al., 2010; McGarrigle et al., 2019). However, an interesting caveat to this need to strive is the existence of a "tipping point", whereby above a certain level of difficulty, the individual may disengage from the task, giving up due to a belief that the task is impossible thus leading to a diminution of effort (Ayasse & Wingfield, 2018).

Closely aligned with task difficulty is motivation to succeed, or 'goal pursuit' as delineated in the FUEL definition. It is not unreasonable to surmise that the more desire an individual has 'to overcome the obstacle', the more resources they would be willing to allocate to achieve this goal. Educational models use this premise to accentuate the importance of bolstering students' motivation to improve their performance in the classroom (Kusurkar et al., 2013). This represents a deficiency within current lab-based studies as it may be difficult to foster motivation for an arbitrary task, therefore creating the potential for skewed reports of effort. This is evidenced by the work of Zekveld et al. who examined the effect of participant feedback on listening task performance: they found that individuals who received feedback on their performance were more likely to improve in subsequent tasks (Zekveld et al., 2019). A well-established rationale for this observed effect is the understanding that feedback can be an important driver for motivation (Egeth & Kahneman, 1975). This supports shifting the paradigm of research towards the "real-world" setting within LE studies, although a novel approach (such as those adopted by Zekveld) may help remedy this potential bias.

The final and most overt factor affecting LE relates to the participants' listening ability. This is the most widely reported facet of LE within the literature. Many studies have shown that people with hearing impairment appear to experience increased LE compared to normal hearing counterparts, regardless of the outcome measure used (Bess & Hornsby, 2014). This finding resonates well with our current understanding of hearing impairment. Furthermore, measurements from before and after interventional studies have demonstrated that certain therapies such as hearing aids, cochlear implants, and other devices can reduce LE scores amongst hearing impaired individuals (Holman et al., 2020). Since, at present, those with a more severe hearing impairment are more likely to qualify for the aforementioned interventions, a paradoxical situation is created whereby those with mild to moderate impairment may in fact experience more LE than those with severe impairment. This incongruency emanates from persevering with long-established traditional criteria in terms of what we choose to consider in decisions regarding eligibility for implants and other devices. For instance, speech and language development remains the paramount criterion when determining the need for and benefit of hearing interventions. This is evidenced by the plethora of papers which use speech and language as the primary, or even sole, outcome of interest when investigating implantable devices (Geers, 2004; Iwasaki et al., 2012; May-Mederake, 2012; Nikolopoulos et al., 2004). This paper does not wish to discredit the importance of speech and language development as there are pertinent connotations involved from both personal and educational perspectives. However, there is the concern that an over-dependence on speech and language outcomes may inadvertently create a false dichotomy where other important factors, such as LE, are overshadowed. For instance, when examining the evidence of benefit arising from bilateral cochlear implantation, it appears limited to a modest improvement in language development and sound localisation (Balkany et al., 2008; Lammers et al., 2014). This seemingly modest benefit may not be due to an implicit inadequacy of the underlying research strategy but rather from discounting other valid markers of success. This scenario reinforces the potentially critical role that a validated measure of LE may have in the audiological test battery, thereby addressing this hidden burden.

Notwithstanding the challenges involved in researching the abstract concept of LE, several issues result from the way we measure this construct. As aforementioned in the list outlined in the first paragraph, there is an array of potential methods available to capture the mental load of effortful listening. To simplify and compartmentalise this wealth of tools, researchers often index these measures into three broad categories: physiological; behavioural; and self-reported measures. These groupings are shown in table 2.1.

To further confound LE measurement, studies have demonstrated inconsistent correlations between each overarching category of LE measure. Whilst several papers have shown significant correlations between physiological and self-reported measures (Bernarding et al., 2014, 2017; Dimitrijevic et al., 2019), other studies demonstrate much weaker correlations, highlighting the poor reproducibility of these findings (Holube et al., 2016; Zekveld et al., 2011). This issue of inconsistency is echoed across the literature with a similar data trend arising no matter which methods are compared. Not only this, the test-retest validity of specific measures tends to be poorly reported as well, leaving little guidance for future studies as to the consistency of each measure. Instead of emphasising the weaknesses of each tool, Alhanbali and colleagues offer an alternative hypothesis whereby each measure may encompass a different dimension of LE. This thus suggests that LE is too complex to be addressed with a single tool approach (Alhanbali et al., 2019). A simplified depiction of this multidimensional theory of LE is illustrated in figure 2.1.

Building upon this multidimensional model, the next step would be to quantify the degree of statistical correlation between themes more accurately. Indeed, it may not be possible to identify exactly which LE dimension is measured by each specific tool, but it would still be invaluable knowledge to clarify whether the various LE measures reflect similar or completely different determinants of effort. This is an important step in the move towards a validated clinical tool for LE as it combines a high degree of test sensitivity with clinical acceptability, by minimising the number of tests which need to be performed.

Another issue to address is the alignment between population group and proposed tool. The vast majority of the literature currently focuses on the adult demographic which inherently benefits from a higher willingness to cooperate during testing procedures. Indeed, it may not be practical to attempt to coax a young child into sitting long enough to gain an accurate pupil diameter recording. However, children with hearing impairment may actually benefit the most from strategies to overcome LE. There is mounting evidence to suggest that children with mild to moderate or unilateral hearing loss may display poorer educational outcomes compared with normal hearing students (Bess et al., 1998; Niedzielski et al., 2006; Purcell et al., 2016). As such, the traditional belief that the impact of mild hearing loss on childhood development is negligible or insignificant is no longer generalisable, thus necessitating an individualised approach founded upon a clear understanding of potential hindrances such as LE. There is also growing concern about LE being a stressor in its own right, thus increasing the risk of inducing chronic stress and the associated deleterious consequences on physical and mental health (Kramer et al., 2006; Mattys et al., 2012; Pichora-Fuller, 2016; Pichora-Fuller et al., 2016; Sandi & Haller, 2015; Schneiderman et al., 2005). Indeed, neuroendocrine research conducted in deaf children has revealed elevations in the cortisol awakening response at the beginning of the school day (Bess et al., 2016). These increases in cortisol have been previously associated with unusual stress and even burnout (Bess & Hornsby, 2014; Kumari et al., 2009; Schlotz et al., 2004). The link between LE and fatigue/stress may become more important following the shift in research to focus more on patient-centred outcomes rather than purely clinical metrics. It is thought that research, which recognises patient outcomes may help to bolster engagement in clinical decision making, thereby leading to an overall more positive experience for the patient (Hill-Feltham et al., 2021; Kirwan et al., 2007; Oliver & Greenberg, 2009). Therefore, it is crucial that any measure of LE shown to be applicable in adults is not only appropriately adapted and trialled in the paediatric population but that it also captures the down-stream effects of LE such as burnout, fatigue and stress.

Considering the methodological challenges involved in quantifying LE, some may believe that future research is too problematic to even attempt. However, it is important to be cognisant of the potential positive outcomes which may arise from deepening our understanding of the topic. Ultimately, this may lead to a tangible way of identifying an unmet burden of hearing impairment and allow us to act readily and holistically to improve the quality of life of affected individuals. Likewise, it may also enable the "real-life" benefit of hearing devices and implants to be measured in a truly meaningful manner. Failure to explore the impact of devices upon LE risks missing (or even under-estimating) benefit and result in restriction in access to hearing interventions.

References

Alhanbali, S., Dawes, P., Millman, R. E., & Munro, K. J. (2019). Measures of Listening Effort Are Multidimensional. *Ear and Hearing*, *40*(5), 1084–1097. https://doi.org/10.1097/AUD.0000000000000697

Ayasse, N. D., & Wingfield, A. (2018). A Tipping Point in Listening Effort: Effects of Linguistic Complexity and Age-Related Hearing Loss on Sentence Comprehension. *Trends in Hearing*, *22*, 1–14. https://doi.org/10.1177/2331216518790907

Balkany, T., Hodges, A., Telischi, F., Hoffman, R., Madell, J., Parisier, S., Gantz, B., Tyler, R., Peters, R., & Litovsky, R. (2008). William House Cochlear Implant Study Group: Position statement on bilateral cochlear implantation. *Otology and Neurotology*, *29*(2), 107–108. https://doi.org/10.1097/mao.0b013e318163d2ea

Bernarding, C., Corona-Strauss, F. I., Latzel, M., & Strauss, D. J. (2010). Auditory streaming and listening effort: An event related potential study. *2010 Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society, EMBC'10, 6817–6820. https://doi.org/10.1109/IEMBS.2010.5625957

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2014). Objective assessment of listening effort in the oscillatory EEG: Comparison of different hearing aid configurations. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2653–2656. https://doi.org/10.1109/EMBC.2014.6944168

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2017). Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort. *Cognitive Neurodynamics*, *11*(3), 203–215. https://doi.org/10.1007/s11571-017-9425-5

Bess, F., Dodd-Murphy, J., & Parker, R. (1998). Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status. *Ear and Hearing*, *19*(5), 339–354. https://doi.org/10.1097/00003446-199810000-00001

Bess, F., & Hornsby, B. (2014). Commentary: listening can be exhausting--fatigue in children and adults with hearing loss. *Ear and Hearing*, *35*(6), 592–599. https://doi.org/10.1097/AUD.0000000000000099

Brookhart, S. M., Walsh, J. M., & Zientarski, W. A. (2006). The dynamics of motivation and effort for classroom assessments in middle school science and social studies. *Applied Measurement in Education*, *19*(2), 151–184. https://doi.org/10.1207/s15324818ame1902_5

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. *Scientific Reports*, *9*(1), 11278. https://doi.org/10.1038/s41598-019-47643-1

Egeth, H., & Kahneman, D. (1975). Attention and Effort. In *The American Journal of Psychology* (Vol. 88, Issue 2). https://doi.org/10.2307/1421603

Geers, A. E. (2004). Speech, Language, and Reading Skills after Early Cochlear Implantation. *Archives of Otolaryngology - Head and Neck Surgery*, *130*(5), 634–638. https://doi.org/10.1001/archotol.130.5.634

Hill-Feltham, P. R., Johansson, M. L., Hodgetts, W. E., Ostevik, A. V., McKinnon, B. J., Monksfield, P., Sockalingam, R., Wright, T., & Tysome, J. R. (2021). Hearing outcome measures for conductive and mixed hearing loss treatment in adults: a scoping review. *International Journal of Audiology*, *60*(4), 239–245. https://doi.org/10.1080/14992027.2020.1820087

Holman, J. A., Drummond, A., & Naylor, G. (2020). The Effect of Hearing Loss and Hearing Device Fitting on Fatigue in Adults: A Systematic Review. *Ear and Hearing*, *42*(1), 1–11. https://doi.org/10.1097/AUD.0000000000000909

Holube, I., Haeder, K., Imbery, C., & Weber, R. (2016). Subjective Listening Effort and Electrodermal Activity in Listening Situations with Reverberation and Noise. *Trends in Hearing*, *20*, 1–15. https://doi.org/10.1177/2331216516667734

Iwasaki, S., Nishio, S., Moteki, H., Takumi, Y., Fukushima, K., Kasai, N., & Usami, S. ichi. (2012). Language development in Japanese children who receive cochlear implant and/or hearing aid. *International Journal of Pediatric Otorhinolaryngology*, *76*(3), 433–438. https://doi.org/10.1016/j.ijporl.2011.12.027

Kirwan, J. R., Minnock, P., Adebajo, A., Bresnihan, B., Choy, E., De Wit, M., Hazes, M., Richards, P., Saag, K., Suarez-Almazor, M., Wells, G., & Hewlett, S. (2007). Patient perspective: Fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *Journal of Rheumatology*, *34*(5), 1174–1177. https://doi.org/10.1002/art.24270

Kramer, S. E., Kapteyn, T. S., & Houtgast, T. (2006). Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *International Journal of Audiology*, *45*(9), 503–512. https://doi.org/10.1080/14992020600754583

Kumari, M., Badrick, E., Chandola, T., Adam, E. K., Stafford, M., Marmot, M. G., Kirschbaum, C., & Kivimaki, M. (2009). Cortisol secretion and fatigue: Associations in a community based cohort. *Psychoneuroendocrinology*, *34*(10), 1476–1485. https://doi.org/10.1016/j.psyneuen.2009.05.001

Kusurkar, R. A., Ten Cate, T. J., Vos, C. M. P., Westers, P., & Croiset, G. (2013). How motivation affects academic performance: A structural equation modelling analysis. *Advances in Health Sciences Education*, *18*(1), 57–69. https://doi.org/10.1007/s10459-012-9354-3

Lammers, M. J. W., Van Der Heijden, G. J. M. G., Pourier, V. E. C., & Grolman, W. (2014). Bilateral cochlear implantation in children: A systematic review and best-evidence synthesis. *Laryngoscope*, *124*(7), 1694–1699. https://doi.org/10.1002/lary.24582

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

May-Mederake, B. (2012). Early intervention and assessment of speech and language development in young children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, *76*(7), 939–946. https://doi.org/10.1016/j.ijporl.2012.02.051

McGarrigle, R., Gustafson, S. J., Hornsby, B. W. Y., & Bess, F. H. (2019). Behavioral Measures of Listening Effort in School-Age Children: Examining the Effects of Signal-to-Noise Ratio, Hearing Loss, and Amplification. *Ear and Hearing*, *40*(2), 381–392. https://doi.org/10.1097/AUD.0000000000000623

Niedzielski, A., Humeniuk, E., Błaziak, P., & Gwizda, G. (2006). Intellectual efficiency of children with unilateral hearing loss. *International Journal of Pediatric Otorhinolaryngology*, *70*(9), 1529–1532. https://doi.org/10.1016/j.ijporl.2006.02.011

Nikolopoulos, T. P., Dyar, D., Archbold, S., & O'Donoghue, G. M. (2004). Development of Spoken Language Grammar Following Cochlear Implantation in Prelingually Deaf Children. *Archives of*

Otolaryngology - Head and Neck Surgery, *130*(5), 629–633. https://doi.org/10.1001/archotol.130.5.629

Oliver, A., & Greenberg, C. C. (2009). Measuring Outcomes in Oncology Treatment: The Importance of Patient-Centered Outcomes. *Surgical Clinics of North America*, *89*(1), 17–25. https://doi.org/10.1016/j.suc.2008.09.015

Pichora-Fuller, M. K. (2016). How social psychological factors may modulate auditory and cognitive functioning during listening. *Ear and Hearing*, *37*, 92S-100S. https://doi.org/10.1097/AUD.0000000000000323

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

Purcell, P., Shinn, J., Davis, G., & Sie, K. (2016). Children with unilateral hearing loss may have lower intelligence quotient scores: A meta-analysis. *The Laryngoscope*, *126*(3), 746–754. https://doi.org/10.1002/LARY.25524

Sandi, C., & Haller, J. (2015). Stress and the social brain: Behavioural effects and neurobiological mechanisms. *Nature Reviews Neuroscience*, *16*(5), 290–304. https://doi.org/10.1038/nrn3918

Schlotz, W., Hellhammer, J., Schulz, P., & Stone, A. A. (2004). Perceived Work Overload and Chronic Worrying Predict Weekend-Weekday Differences in the Cortisol Awakening Response. *Psychosomatic Medicine*, *66*(2), 207–214. https://doi.org/10.1097/01.psy.0000116715.78238.56

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology*, *1*(Lacey 1967), 607–628. https://doi.org/10.1146/annurev.clinpsy.1.102803.144141

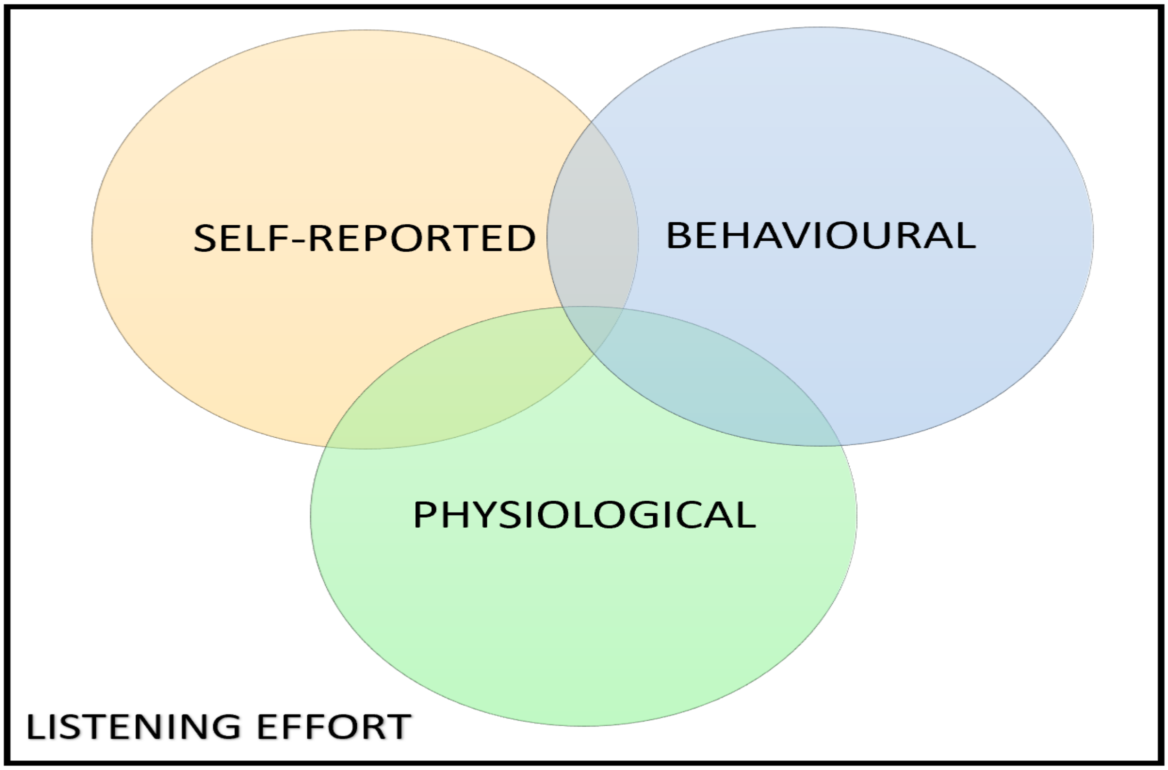
Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, *32*(4), 498–510. https://doi.org/10.1097/AUD.0b013e31820512bb

Zekveld, A. A., van Scheepen, J. A. M., Versfeld, N. J., Veerman, E. C. I., & Kramer, S. E. (2019). Please try harder! The influence of hearing status and evaluative feedback during listening on the pupil dilation response, saliva-cortisol and saliva alpha-amylase levels. *Hearing Research*, *381*, 107768. https://doi.org/10.1016/j.heares.2019.07.005

Table 2.1 Showing the different measures of LE and their associated category

Physiological	Behavioural	Self-reported
EEG spikes	Reaction times	Effort based questionnaires
Pupil size	Response times	Fatigue based questionnaires
Hormonal measures	Correct v incorrect response	
Infra-red Spectroscopy	Driving simulators	
Functional MRI imaging		

Figure 2.1 Visual representation of multidimensional theory of LE. Demonstrating the overlap between self-reported/questionnaires, behavioural and physiological measures in how they capture LE.



Chapter 3: Exploring the Correlations Between Measures of Listening Effort in adults and children: A Systematic Review with Narrative Synthesis

Published systematic review: C. Shields[1], M. Sladen[1], I. A. Bruce[1], K. Kluk-de Kort[2], and J Nichani[1], (2022). Exploring the Correlations Between Measures of Listening Effort in adults and children: A Systematic Review with Narrative Synthesis (2022). Trends in Hearing, Volume 26: 1–20 DOI: 10.1177/23312165221137116.

[1] Department of Paediatric Otolaryngology, Royal Manchester Children's Hospital, Manchester Foundation Trust, Manchester, UK

[2] Division of Human Communication, Development & Hearing, University of Manchester, Manchester, UK

Abstract

Listening effort (LE) describes the cognitive resources needed to process an auditory message. Our understanding of this notion remains in its infancy, hindering our ability to appreciate how it impacts individuals with hearing impairment effectively. Despite the myriad of proposed measurement tools, a validated method remains elusive. This is complicated by the seeming lack of association between tools demonstrated via correlational analyses. This review aims to systematically review the literature relating to the correlational analyses between different measures of LE. Five databases were used– PubMed, Cochrane, EMBASE, PsychINFO, and CINAHL. The quality of the evidence was assessed using the GRADE criteria and risk of bias with ROBINS-I/GRADE tools. Each statistically significant analysis was classified using an approved system for medical correlations. The final analyses included 48 papers, equating to 274 correlational analyses, of which 99 reached statistical significance (36.1%). Within these results, the most prevalent classifications were poor or fair. Moreover, when moderate or very strong correlations were observed, they tended to be dependent on experimental conditions. The quality of evidence was graded as very low. These results show that measures of LE are poorly correlated and supports the multi-dimensional concept of LE. The lack of association may be explained by considering where each measure operates along the effort perception pathway. Moreover, the fragility of significant correlations to specific conditions further diminishes the hope of finding an all-encompassing tool. Therefore, it may be prudent to focus on capturing the consequences of LE rather than the notion itself.

Keywords: Hearing-related mental fatigue, Quality of Life, Outcome measures, Listening effort, Mental effort

Introduction

The concept of listening effort (LE) is still yet to be fully understood, and as a result, it is not formally integrated into routine clinical practice. Many attempts have been made to define LE, but perhaps the most contemporary definition stems from the Framework of Understanding Effortful Listening (FUEL) workshop.

"The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when tasks involve listening" (Pichora-Fuller et al. 2016, page 10)

This definition and the resulting model poised by FUEL builds upon the effort perception theories first developed by Kahneman to better apply to the world of hearing (Kahneman 1973). It relays that the experience of LE is a complex interaction between factors such as working memory, attention, motivation, task difficulty, and cognitive capacity (Kahneman 1973; Pichora-Fuller et al. 2016; Shields et al. 2021). Despite these recent developments, our understanding of the effort perception pathways involved with LE is limited.

The ability to accurately measure LE has wide-reaching potential ramifications for individuals with hearing impairment. Perhaps this is most acutely demonstrated by considering those with mild to moderate or asymmetrical hearing loss. These patients are often overlooked with regards to interventions, especially surgical devices, and may therefore represent an unmet disease burden. While this degree of hearing loss may not be severe enough to interfere with traditional outcomes such as speech and language development, it cannot be said to have no negative connotations for quality of life. Further, a workshop conducted via the Nottingham Hearing Biomedical Research Unit identified "What adverse effects are associated with not treating mild-to-moderate hearing loss in adults?" as the number one research priority for adults with mild to moderate hearing loss (Henshaw et al. 2015). In the paediatric population, there is growing research to suggest that overexposure to LE can negatively affect educational outcomes (Bess, Dodd-Murphey, and Parker 1998; Downs and Crum 1978; Hsu, Vanpoucke, and van Wieringen 2017). This is particularly concerning when 11.3% of the general school population is thought to have some degree of hearing-impairment (Bess et al. 1998). Since children with mild-to-moderate hearing loss are less frequently identified compared to those with more severe impairments, they may be more likely to endure LE without the benefit of interventions and thus further hindering their educational attainment (Archbold et al. 2015). With this considered, research must include the needs of the younger generation, especially as they are perhaps more susceptible to the sequelae of LE.

Outside of educational domains, other recognized consequences accrue following prolonged periods of effortful listening. Some of the most established outputs include fatigue and stress, which, in the long run, can have a deleterious impact on mental wellbeing (Bess et al. 2016; Bess and Hornsby 2014; Kramer, Kapteyn, and Houtgast 2006; Kumari et al. 2009; Mattys et al. 2012; Pichora-Fuller 2016; Pichora-Fuller et al. 2016; Schlotz et al. 2004; Schneiderman, Ironson, and Siegel 2005). Together these educational and psychological corollaries emphasise the importance of broadening our understanding of LE and the need to continue exploring this relatively uncharted territory to discern other possible consequences of LE. It also provides an alternative cache for developing outcome measures of LE by using these downstream effects as proxy measures for LE itself.

There are many barriers to incorporating LE as an outcome measure within clinical practice, paramount of which relates to the lack of consensus regarding the best way to measure LE (Pichora-Fuller et al. 2016). Three broad categories of LE measures are used across the literature: self-reported questionnaires, behavioural measures, and physiological measures (McGarrigle et al. 2014). Within each category exists a cornucopia of different tools postulated to capture LE, each has its advocates and dissidents, yet the ‘gold standard’ measure remains frustratingly elusive. Moreover, it is not uncommon for tools that aim to represent the consequences of LE, such as fatigue questionnaires, to be used as a proxy measure of LE (Alhanbali et al. 2018; Dwyer et al. 2019). A brief outline of these measures and their advantages and disadvantages are presented below.

Self-reported measures: Questionnaires that assess the individual's perceived effort levels are subjective. These questionnaires often include a numerical rating scale to quantify the degree of effort. The advantages of self-reported measures include that they are quick to administer and can be used to trend LE over a period of time or assess for benefit before and after an intervention (McGarrigle et al. 2014). Potential pitfalls to questionnaires relate mainly to their subjective nature and the subsequent effect of user bias. Indeed, Kahneman (1973) distinguishes subjective and objective effort as two separate entities, inferring that mental effort is solely an objective construct, independent of perceived effort (Bruya and Tang 2018). Examples of self-reported measures include the NASA task load index or more superficial LE visual analogue scales (Alhanbali et al. 2019; Hart 2006). Alongside effort questionnaires, fatigue questionnaires operate similarly but focus on the tiredness evoked via effortful listening. This is an example of how the sequelae of LE can be utilised in the assessment of hearing-related effort (McGarrigle et al. 2014). They have the same advantages and disadvantages as effort questionnaires.

Behavioural measures: Centre around an individual's performance within a particular task and use their performance as an objective measure of effort. This is often through a single or dual-task paradigm and thus resonates with the capacity model of mental effort discussed above (Pichora-Fuller et al. 2016). A benefit to mapping performance is that it can resemble real-world challenges that listeners face, such as classroom or workplace scenarios, and thus has more ecological value (McGarrigle et al. 2014). The contrary stance is that laboratory-based testing frequently subverts the impact of motivation on capacity since it may be challenging to foster participant motivation for an arbitrary task (Shields et al. 2021). Additionally, their reliance on the capacity model to demonstrate effort requires the task to stretch the individual to their cognitive limit, which may not be feasible in all experiments (Pichora-Fuller et al. 2016). Behavioural measures include response times and reaction times within a dual-task (Rovetti et al. 2019; Visentin and Prodi 2021).

Physiological measures: This type of tool uses a myriad of instruments to determine effort, relaying the autonomic reaction of body systems to a stressful testing stimulus (McGarrigle et al. 2014). Their interpretation of neural and endocrine signals leaves little room for subjectivity, indicating that these measures reflect Kahneman's view of mental effort more closely than the other measures (Bruya and Tang 2018; Kahneman 1973). Unfortunately, their objective nature does not eliminate all bias; they are prone to confounding bias in that many other factors could amplify or dampen somebody's physiological responses (Ayres et al. 2021). For instance, concurrent comorbidities (neurological disorders, endocrine disorders, ophthalmic conditions, or cardiovascular pathology), medication use, caffeine or alcohol intake, or even sleep hygiene (Ayres et al. 2021; Kemp et al. 2014; Pichora-Fuller et al. 2016). Therefore, physiological measures may represent the gold standard index of effort in ideal experimental conditions, but this is unlikely to translate into real-world environments. Examples of physiological measures include pupillometry, EEG oscillation, cortisol level, functional imaging, heart rate variability, and skin conductance (Francis et al. 2021; Kramer, Teunissen, and Zekveld 2016; Rovetti et al. 2019; Zekveld et al. 2019).

To add further complexity to the matter, individual studies have shown that these measures are poorly correlated, diminishing the hope of a validated clinical tool (Alhanbali et al. 2019). This observation has given rise to the multidimensional model of LE and suggests that no one tool can fully encompass LE. If multiple measures are required to assess LE accurately, the clinical acceptability of LE as an outcome measure will be seriously hindered (Mosadeghrad 2012). Given the vast array of tools and experimental conditions that can be applied to investigate LE, it remains problematic to generalise results from individual studies across the literature. Therefore, it is prudent to systematically review these individual studies to map out correlational data between measures across all relevant studies to understand this concern fully. If the results of the individual studies harmonise the lack of correlation between measures of LE, then it would be important to consider alternative strategies to embrace the concept as an outcome.

Aims

The primary aim of this paper is to perform a systematic review of the literature pertaining to the correlations between methods of measuring LE in adults and children to determine how related they are to one another. If the multidimensional model is accurate, this review hypothesises that weak and non-significant correlations will be observed across the literature.

A secondary aim is to compare the correlations between various LE instruments to measures of fatigue to investigate the association between the two concepts. If fatigue is a downstream consequence of LE, this paper predicts that the strength of correlations for effort-effort comparisons will be similar to effort-fatigue comparisons.

Methods

A protocol outlining the methodology of this review was published on PROSPERO prior to the commencement of the literature search. Access to the full PROSPERO protocol can be obtained via the following link: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=253010.

Literature search

The literature search for this review was conducted using five databases – PubMed, Cochrane Central, EMBASE, PsycINFO, and CINAHL. An appropriate search strategy was used that utilised medical subject heading (MeSH) and operators was used. The two core themes of the search were – listening effort/fatigue and hearing impairment – a detailed overview of this strategy can be found in the PROSPERO registration. A date limit of 2000-present was applied, given the relative novelty of the term listening effort. The strategy was validated by a medical librarian, independent of the study. Once the final set of papers had been retrieved, they were uploaded to the systematic review software EPPI-reviewer to be screened (Thomas, Brunton, and Graziosi 2021).

Study selection

The initial screening stage involved removing duplicates; this was done initially by the tool located within EPPI-reviewer and subsequently verified manually by one reviewer. Once duplicates had been removed, screening on title and abstract was conducted by two reviewers independently with blinding applied to the other's decisions to minimise any bias effects. Inclusion and Exclusion criteria were generated using a modified version of the Population, Intervention, Comparison, Outcome and Study design (PICOS) framework (Armstrong 1999; Ebell 1999). In this review, the domain 'intervention' was replaced with 'exposure' to resonate with the study objective. The following inclusion and exclusion criteria were applied:

Inclusion:

• Participants – The study population includes the following categories: Normal hearing, Hearing-impaired with a device, Hearing-impaired without a device, Adults and Children. Adults and children have been included in this review to allow for age-specific comparisons of the correlations between instruments. Children may be more prone to the negative consequences of LE in terms of educational attainment and therefore represent an important population to consider. However, there is only a small number of published papers on paediatric individuals; therefore, this review included adults to allow for a complete narrative.

• Exposure – The paper includes any tool used to capture LE or fatigue related to a listening task.

• Comparison – The paper uses statistical correlational methods to compare the results obtained for the different measures of LE

• Outcome – The paper produces correlation coefficients between measures and associated p-values

• Study design – Primary research, study must be available in English

Exclusion:

• Participants – Any documented comorbidities that may influence listening effort results or fatigue assessment. This may include chronic health conditions, fatigue-causing conditions, or dual sensory impairment.

• Exposure – Any tool which does not primarily measure listening effort or fatigue related to a listening task. This included general quality of life measures.

• Comparison – No analysis provided to compare different tools.

• Outcome – The paper does not comment on the relationship between two measures of LE derived from the results of the correlational analysis.

• Study design – Secondary research, editorials, interviews, book chapters or case reports, Studies not available in English

Any disagreement between reviewers during the coding stage was referred to a panel constituting an ENT consultant, audiologist and another ENT doctor for reconciliation. This process was then repeated similarly for full-text screening. After obtaining an initial set of included results, forward citation chaining was conducted to further search the literature for relevant papers. Forward citation was chosen as it had the added benefit of finding papers published after the initial search and papers contained within the grey literature and, therefore, may not have been identified (Lefebvre et al. 2021). An outline of this process is shown in the PRISMA flowchart contained within the results section of this paper.

Data extraction and data analysis

Data extraction was performed via a google form document by two reviewers. Key information captured for each study included:

• Details of the study – Type of paper, date published, authors, country of publication

• Study design – Observational vs. interventional

• Study population – Cohorts, sample size, and whether a power analysis had been performed

• Setting of the study – Listening task and listening conditions

• Measure of LE – What exact tools were used, and which overarching category do they fall under (Behavioural/physiological/self-reported)

• Type of comparison method used for the analysis of LE measures (Spearman's/Pearson's)

• Strength of correlation.

Quality assessment and risk of bias

To assess the quality of evidence and potential risk of bias, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria was applied to the pool of included studies (Schünemann et al. 2013). This approach has the benefit of assessing the quality of evidence in its entirety rather than focusing on individual studies. This was deemed particularly appropriate for this review, given the heterogenous designs of the studies included. The GRADE criteria contain several domains that assess the overall quality of evidence (Schünemann et al. 2013).

First, 'inconsistency' refers to how homologous the evidence is, where studies performed in similar conditions on comparable participants. 'Imprecision' describes the statistical power of the studies in relation to their sample sizes. 'Indirectness' is how applicable the included studies are when examining the research question of this review. Determinates of indirectness include the population under investigation and the outcome measures used within the included studies. Each item can be graded as 'not serious', 'serious', or 'very serious' depending on the proportion of studies that satisfy each criterion (Schünemann et al. 2013). The effect of publication bias was also investigated within the overall GRADE assessment. This was determined by examining whether studies were consistent between their methods and reported results.

In addition to the above markers for study quality, the risk of bias was also assessed by applying an appropriate checklist to each study which varied based on individual study design. The ROBINS-I checklist was used for interventional studies to determine study limitations (Sterne et al. 2016). The current review classified a study as interventional if it attempted to measure the change in LE within study participants upon application of an accepted therapy such as hearing aid settings. This tool was chosen as it was specifically designed for non-randomised interventional studies, which aligns well with the subset of included studies in this review. Additionally, it is endorsed by the Cochrane Handbook for systematic reviews (Sterne et al. 2021). ROBINS-I assesses several domains, including – Confounding bias due to differences in baseline characteristics; selection bias emanating from differences in participants between study groups; classification bias due to misclassifying participants based on intervention status; deviation bias when study participants deviate from intended interventions and outcome bias arising from differences in the methods used to measure outcomes between participants (Sterne et al. 2016). Following this assessment, the overall risk of bias is formulated as – 'low', 'moderate', 'serious', or 'critical' depending on how the study is assessed in the domains mentioned above (Sterne et al. 2016). An example of how the ROBINS-I checklist was applied within this study can be found in the supplementary information.

For observational studies, the GRADE checklist was utilised, which examines factors including - whether a control group was present, how the study confirmed the exposure of interest, how consistently was the outcome measure applied to participants, did the study account for possible confounders, and finally did the study measure at multiple time points (Schünemann et al. 2013). For this review, studies were classed as observational if they measured LE within a cohort and did not attempt to discern the effect of a particular intervention such as hearing aid settings on the participant's LE. The overall risk of bias for the observational studies can be judged as 'low' whereby most information from studies is deemed to have a small risk of bias; 'unclear' if the information from studies indicates a small or uncertain risk of bias; or 'high' in which the information from studies suggests that the risk of bias is sufficient to affect the interpretation of the results (Schünemann et al., 2013). An example of how the GRADE checklist was considered in the current study can be found in the supplementary information.

Data synthesis

It was not appropriate to conduct a meta-analysis of the included studies as the pooled sample of papers was heterogenous in design. Factors that contributed to the lack of congruency between papers included measurement tools, experimental design, and study population. Therefore, a narrative synthesis of the result was produced to answer the research question posed by this review.

To overcome the challenge associated with directly comparing correlations amongst a heterogeneous study pool, this paper classified each r-value of statistically significant results according to the taxonomy of correlations outlined by Chan (Chan 2003). This reporting system has been specifically designed for application to medical papers. Chan reported the following grading system, a perfect correlation (coefficient +1 or -1), very strong (coefficient 0.8 to 1 or -0.8 to -1), moderate (coefficient 0.6 to 0.8 or -0.6 to -0.8), fair (coefficient 0.3 to 0.6 or -0.3 to -0.6), poor (coefficient 0 to 0.3 or 0 to -0.3) and none (coefficient 0) (Chan 2003). This review evaluated the performance of each measurement theme against these criteria to provide an overall understanding of the degree of association between effort questionnaires, behavioural measures, physiological measures, and fatigue questionnaires in relation to capturing LE. A Chi-squared test was used to determine if there was a statistically significant difference between the observed classifications across all significant results. Additionally, a Fisher's exact test was used to examine for observed differences between the results for the adult and child population groups. Fisher's exact test was chosen due to the relatively small number of children within the included papers.

Results

Upon completion of both title and abstract, full-text screening, and forward citation searching, 48 papers were included based on the above PICOS criteria. The PRISMA diagram below summaries the results of this process (Page, McKenzie, et al. 2021). The full details of each included paper, with reference to the experimental design, can be found in the supplementary information.

The adult population across the papers comprised 775 hearing-impaired participants and 815 individuals with normal hearing, totalling 1687 participants. Of the 775 patients with hearing impairment, 124 were listed as hearing aid users, whereas 117 had cochlear implants. For the paediatric population, 34 had a hearing impairment, and 179 had normal hearing, totalling 213 paediatric participants. The total sample size across all studies, ages, and hearing level equated to 1900.

Summary of the strength of correlations

The total number of correlational analyses across all included papers totalled 274. These 274 were stratified across ten different ‘subgroups’ each of which provided a comparison between two categories of instruments (EQ-EQ, EQ-B, EQ-P, EQ-FQ, B-B, B-P, B-FQ, P-P, P-FQ, and FQ-FQ – EQ = Effort questionnaire, B = Behavioural measures, P = Physiological measure and FQ = Fatigue questionnaire).

This review defined the results as either statistically significant or non-significant based on the p-values reported by the individual papers. A reported p-value of less than 0.05 was considered statistically significant in line with the traditional cut-off (Andrade 2019). Across all subgroups, 99 analyses reached statistical significance representing 36.1% of the total analyses. The range of r-values has been plotted using a violin chart and stratified according to significance – figure 2. Only those analyses that reported the r-value, despite non-statistically significant results were included in the figure (133/175).

Each statistically significant trial was classified using the criteria outlined by Chan, which provides an overview of how related each measure is to one another (Chan 2003). This is summarised in table 1, which is overlaid with a heat map to illustrate the density of the results. The most common classification identified across all significant correlations was fair (n = 62). The observed difference between the frequency of each classification was shown to be statistically significant through Chi-squared testing ($p < 0.01$).

Comparison between adults and children

A total of 253/274 correlational analyses exclusively contained an adult population, whereas only 21/274 correlational analyses contained children. The number of significant trials amongst the adult population was 89 (35.2%). In contrast, 10 (47.6%) of the analyses in children reached significance. Fisher's exact test showed no significant difference between the frequency of each type of classification between the adult and child population groups ($p = 0.84$).

Conditional and unconditional correlations

This review recognises two distinct types of correlations: conditional and unconditional. The latter refers to correlations in which the paper has provided a general comparison between two measures, collapsing for any changes to experimental conditions or study populations. Conditional correlations exist when the paper has published the strength of association between two measures of LE for a specific study condition. This review identified three categories of experimental conditions discussed below. The conditions for each analysis can be found in the supplementary information.

The first relates to the listening task used to evoke effort. This may include noise, light, signal quality, and task difficulty. This review noted that, in general, more challenging tasks tended to produce stronger correlations. For example, in one study, the r-value between a Visual analogue scale/Peak pupil diameter was 0.6 ($p < 0.01$) within the SRT50% condition. When this was altered to quiet conditions, the r-value fell to 0.008 ($p > 0.05$) (Kramer et al. 2016). In another study, an analysis comparing Peak pupil diameter/Need for recovery questionnaire revealed an r-value of - 0.35 ($p < 0.05$) within the light experimental conditions. When this was changed to dark conditions, the r-value was found to be - 0.18 ($p > 0.05$) (Wang et al. 2018).

Next is the time point at which LE is measured. Examples of this condition include pre-test, during-test, post-test, pre-post-test difference (Gustafson et al. 2018; Key et al. 2017; Kramer et al. 2016; Zekveld et al. 2019), baseline period, retention period, speech period (Alhanbali et al. 2019; Russo et al. 2020), before response cue, and after response cue (Giuliani 2021). Again, similarly to the listening task, this review found that correlations were stronger when LE was sampled immediately following the listening task. When lapse of attention was compared to a fatigue scale, the pre-test r value changed from -0.5 ($p < 0.05$) to -0.702 ($p < 0.05$) in the post-test condition (Key et al. 2017). This was again shown during a comparison of the Speech, Spatial, and Qualities (SSQ) questionnaire, and pupillometry. Here, the baseline r value was not documented due to a lack of significance, but the retention interval findings reached significance with a coefficient of -0.78 (Russo et al. 2020).

The third is the hearing ability of the population used within the experiment. Normal hearing, hearing-impaired, and hearing device users constitute this condition. While it is accepted that those with a hearing impairment suffer from higher levels of effort, there was no apparent variation in the strength of correlations based on population conditions (Alhanbali et al. 2017; Holube et al. 2016).

It is rational to surmise that conditional correlations confer less evidence towards an association than unconditional correlations as their reproducibility would be limited to the exact experimental conditions used to derive the finding.

Strength of correlation stratified by subgroup

To further explore the correlations reported, this review presents both a within subgroup and between subgroup comparisons below. Each section compares two subgroups of measures against one another to provide an insight into whether each theme captures similar or different components of LE. A summary of the number of significant analyses for each subgroup comparison is shown in table 2.

Effort questionnaires v Other Effort questionnaires – A within subgroup comparison

This review identified three papers that compared different effort-based questionnaires (Perreau et al., 2017; Picou et al., 2017; Picou & Ricketts, 2018). This resulted in five correlational analyses, all of which reached statistical significance. All five analyses were unconditional and therefore not dependent on changes to experimental conditions. Within this subgroup, the strongest association was observed between single entity questions that aimed to elicit 'loss of control' and questions that focused on the 'work required' to complete a task. This had an r-value of 0.82. The weakest correlation was noted between 'work required' questions and questions which explored an individual's desire to 'give up' with an r-value of 0.34. Three of the correlations were classified as fair, one as moderate, and one as very strong.

Effort questionnaires v Behavioural measures – A between subgroup comparison

Out of the 48 included papers, 13 of them compared effort questionnaires to behavioural measures of LE (Desjardins, 2016; Johnson et al., 2015; Megha & Maruthy, 2019; Perreau et al., 2017; Picou et al., 2017; Picou & Ricketts, 2018; Rovetti et al., 2019; Seeman & Sims, 2015; Strand et al., 2018; Visentin et al., 2019, 2021; Visentin & Prodi, 2021; White & Langdon, 2021). These papers produced 28 distinct correlational analyses. Only nine (32.1%) of these correlations reached statistical significance. For these nine correlations, six were unconditional, and three were conditional. The strongest association was demonstrated between control-themed questions and reaction times, with an r-value of - 0.59. This was a conditional correlation found only when the stimulus was presented in noise. The weakest correlation existed between a five-point LE scale/response times with an r-value of 0.16. Regarding the classification of correlations, one was poor, and eight were fair. The remaining 19 results were not statistically significant.

Effort questionnaires v Physiological measures – A between subgroup comparison

The comparisons between effort questionnaires and physiological measures of LE represent the most investigated subgroup within this review, with results stemming from 23 papers (Alhanbali et al., 2019, 2021; Bernarding et al., 2017; Dimitrijevic et al., 2019; Dwyer et al., 2019; Finke et al., 2016; Francis et al., 2021; Holube et al., 2016; Kramer et al., 2016; Lau et al., 2019; Mackersie et al., 2015; Mackersie & Cones, 2011; Magudilu Srishyla Kumar, 2020; McGarrigle et al., 2020; Rovetti et al., 2019; Russo et al., 2020; Schafer et al., 2015; Seeman & Sims, 2015; Strand et al., 2018; Visentin et al., 2021; White & Langdon, 2021; Wisniewski et al., 2015; A. A. Zekveld et al., 2011). This culminated in a total of 73 correlational analyses. Within these 73 analyses, 21 (28.8%) reached significance, and 52 failed to reach significance. Focusing on the 21 significant results, four were unconditional, and 17 were conditional. The joint strongest association compared National Aeronautics Space Administration (NASA) task load index/EEG oscillation and a five-point LE scale/EEG frontal midline theta power, both showing r values of 0.97. This represents a near-perfect correlation; similar results were also observed between a seven-point LE scale/EEG oscillation (r-value 0.94). These very strong analyses stemmed from three separate papers. The weakest association was between NASA task load index/Functional infra-red spectroscopy with an r-value of 0.13. In terms of classification, five were poor, eight were fair, four were moderate, and four were very strong.

Effort questionnaires v Fatigue questionnaires – A between subgroup comparison

Within the category contrasting effort questionnaires to fatigue questionnaires, a total of eight papers published comparisons (Alhanbali et al., 2017, 2018, 2019, 2021; Dwyer et al., 2019; McGarrigle et al., 2020; Picou et al., 2017; Picou & Ricketts, 2018). This resulted in 24 separate correlation analyses; 19 (79.2%) reached significance, while five did not. Among the significant findings, 11 were unconditional, and eight were condition dependant. The largest effect size was demonstrated by Amsterdam Checklist for Hearing and Work/Profile of Mood States - frequency of tiredness questions with an r-value of 0.80. The weakest association was between an Effort assessment scale/Fatigue assessment scale resulting in an r-value of 0.30. The significant correlations obtained the following classifications of strength, 14 were fair, four were moderate, and one was very strong.

Behavioural measures v Other Behavioural measures – A within subgroup comparison

Two papers compared different behavioural measures of LE and generated 25 distinct correlational analyses (Strand et al., 2018; Wu et al., 2014). Ten (66.7%) of these analyses were significant, and 15 were non-significant. Of the ten significant results, one was unconditional, and nine were conditional. The largest r value existed for listening span/cognitive spare capacity test, and this was a conditional correlation found within the high predictability/updating conditions with an r-value of 0.53. The weakest correlation was observed between complex reaction times/cognitive spare capacity tests, again conditional for the 2-word recall condition; here, the r value was 0.38. All the results in this section were classified as fair.

Behavioural measures v Physiological measures – A between subgroup comparison

A total of nine papers compared behavioural and physiological measures of LE (Bertoli & Bodmer, 2014; Gustafson et al., 2018; Key et al., 2017; Rovetti et al., 2019; Strand et al., 2018; Visentin et al., 2021; White & Langdon, 2021; A. Zekveld et al., 2020; Zhao et al., 2019). They conducted 26 correlational analyses, with only seven (26.9%) reaching statistical significance. The seven analyses which reached significance were all conditional. The largest effect size was between hit rate (correct response within a certain time frame)/pupillometry - pupil dilation response, r = - 0.65. This result was a conditional correlation found within the young cohort (age 18-35 years) within the 15-20 second time period after the stimulus. The joint weakest correlation was between reaction times/pupillometry - mean pupil dilation in the compatible condition and reaction times/pupillometry - peak pupil dilation in the compatible/conflict condition with r values of 0.28 and - 0.28 respectively. Two analyses were poor, four were fair, and one was moderate.

Behavioural measures v Fatigue questionnaires – A between subgroup comparison

Six papers reported correlations between behavioural measures of LE and questionnaires designed to measure fatigue, resulting in 19 correlational analyses (Athey, 2021; Gustafson et al., 2018; Hornsby, 2013; Key et al., 2017; Picou et al., 2017; Picou & Ricketts, 2018). Within this category, nine results reached significance; seven of these significant results were conditional, and all were dependent on the time point used to measure LE. The largest coefficient was shown between lapses of attention/fatigue scale within the post-test time condition with an r value - 0.70. The weakest correlation was between reaction times/tiredness questions, with an r-value of 0.22. These results are classified as follows, one poor, seven fair, and one moderate.

Physiological measures v Other Physiological measures – A within subgroup comparison

Eight of the 48 papers included in this study compared physiological measures of LE with other physiological measures (Alhanbali et al., 2019; Francis et al., 2021; Giuliani, 2021; Kramer et al., 2016; McMahon et al., 2016; Miles et al., 2017; Seifi Ala et al., 2020; A. A. Zekveld et al., 2019). Altogether, this translated into 43 distinct correlational analyses. For these results, seven (16.3%) reached significance, and 36 failed to reach significance. When examining the significant correlations, this review found that two out of the seven were unconditional correlations, and the remaining five were conditional. The strongest association existed between skin conductance response quantity/skin conductance response amplitude, with an r-value of 0.89. This was a conditional correlation demonstrated only in the before response cue period. The weakest association was between skin conductance/pupillometry - peak pupil diameter with an r-value of 0.20. These results can be stratified according to Chan's classification as follows, three poor, two fair, one moderate, and one very strong.

Physiological measures v Fatigue questionnaires – A between subgroup comparison

Eight papers compare physiological measures against questionnaires related to fatigue (Alhanbali et al., 2019, 2021; Dwyer et al., 2019; Gustafson et al., 2018; Key et al., 2017; McGarrigle et al., 2020; Wang, Kramer, et al., 2018; Wang, Naylor, et al., 2018). These papers produced 30 correlational analyses, 11 (36.7%) reached significance. Eight of the 11 significant results were unconditional correlations. The strongest association existed between pupillometry - task-evoked pupillary response/fatigue scale 0-100. The r-value for this result was - 0.48. The weakest correlation was between pupillometry - peak pupil diameter/visual analogue scale - tiredness with an r-value of - 0.19. The classifications for the significant results were six poor and five fair.

Fatigue questionnaires v Other Fatigue questionnaires – A within subgroup comparison

The final correlation comparison between fatigue questionnaires only yielded one trial (Wang, Naylor, et al., 2018). This result demonstrated a significant unconditional correlation between Need for Recovery questionnaire/Checklist Individual Strengths questionnaire. The r-value for this analysis was 0.57, resulting in a fair classification.

Quality assessment and risk of bias

The GRADE approach was applied to the 48 included studies to provide an overview of the quality of the evidence contained within this review. This review utilised either the GRADE criteria or the ROBINS-I. In total, 38 papers were classified as observational and nine as interventional. Overall, the risk of bias was judged to be high/serious for both types of study design. At an individual study level, 38/48 were deemed to have a serious/high risk of bias, 9/48 a moderate risk of bias, and 5/48 a low risk of bias. The rationale behind this assessment was due to several factors. First, 31 out of the 48 studies only included one cohort; 17 of these studies only had a normal hearing group, and 14 only contained individuals with hearing impairment. Inconsistency was also noted between the exclusion criteria applied by each study with regard to possible confounding comorbidities. Some papers accounted for a range of possible pre-existing factors which may influence the measure of LE obtained. Examples include neurological disorders, learning disabilities, psychiatric conditions, chronic fatigue, medication use, smoking history, alcohol intake, and caffeine use. The antithesis was found whereby some papers only partially attempted to adjust for this potential risk of bias, introducing a high degree of confounder bias. Finally, most papers, 44/48, only measured LE at a single time point and therefore had no follow-up. Generally, the literature mitigated the effects of exposure bias – by ensuring normal hearing participants were screened via audiological testing prior to inclusion - and outcome bias – by keeping missing data to a minimum.

The included studies displayed a large amount of heterogeneity in terms of study design; therefore, 'inconsistency' was judged to affect the quality of evidence seriously. Specifically, a large variation existed between studies regarding which tool was utilised to measure LE. Indeed, variations existed between specific types of tools, for example, differences in the numerical ranges contained within self-reported LE scales. The setting of the study experiments was also inconsistent in terms of listening task and listening conditions. Finally, the study population differed significantly between papers with reference to hearing level and age. 'Imprecision', most studies did not report a sample size calculation, 30/48, an additional 3/48 performed a power calculation but subsequently did not have sufficient sample size, one study sample fulfilled the power calculation for some but not all outcomes, and only 14/48 studies satisfied the required sample size. Therefore, the effect of 'Imprecision' on the quality of evidence was graded as serious. This review applied the criterion of 'indirectness' by considering three factors – Did the study contain hearing-impaired participants? Did the study contain both an adult and paediatric population? Did the study use an accepted measure for LE? Overall, no study satisfied all three factors; 17 met two, 30 met one, and one failed to meet any of the above factors. This finding resulted in 'indirectness' achieving a rating of moderate. Publication bias moderately impacted the quality of evidence due to the fact 42/274 correlational analysis did not report the r values of non-significant results (Page, Higgins, and Sterne 2021). Combining each of the GRADE criteria outlined above translates into an overall quality of evidence rating of very low, which has important implications for the conclusions drawn by this review. This aligns with previous systematic reviews investigating the literature pertaining to LE and reflects one of the inherent challenges to validating a clinical tool to measure this phenomenon (Holman, Drummond, and Naylor 2021; Ohlenforst et al. 2017). A summary of the quality of evidence assessment for the included papers is shown in table 3 below.

Discussion

This review aimed to determine the association between different instruments used to capture LE by evaluating the correlational analyses performed across the literature. Additionally, the correlations between effort and fatigue instruments were contrasted to investigate the relationship between the two concepts of LE and fatigue.

Multi-dimensional model of listening effort

Spatial stratification

This review notes that only 36.1% of the 274 correlation analyses reached statistical significance, and correlation strength was most often classified as fair. The distribution of classifications towards fair was found to be statistically significant. This lack of association lobbies the notion that different measures encompass different components of LE. Previous studies have explained this phenomenon by referring to a multi-dimensional model of LE; therefore, this idea is reaffirmed by the findings of this review. (Alhanbali et al. 2019; Francis and Love 2020).

This model outlines that behavioural, physiological, and self-reported measures capture different components of LE (Alhanbali et al. 2019; Francis and Love 2020; Peelle 2018). McMahon et al (2016) hypothesize that this can be rationalised by considering that different measures align with the different cognitive processes involved with sound perception (McMahon et al. 2016). Only when we understand these cognitive processes will we be able to fully interpret information gathered from measures of LE. Hughes et al. (2018) have attempted to categorise some of these cognitive pathways to help clarify this topic. They outlined three distinct cognitive stages which can give rise to LE. These were described as the effort evoked by ‘attending to’, ‘processing’, and ‘adapting to’ an auditory message. The measures of LE, therefore, may resonate more closely with one particular part of the overall process

(Hughes et al. 2018). From a physiological perspective, this theory is reflected by examining functional imaging studies highlighting the myriad of cortical areas involved with understanding degraded speech (Peelle 2018).

Further to these intrinsic processes involved in speech perception, individual factors may also contribute to the effort experienced. Francis and Love (2020) delineated the differences between exerted and assessed effort in that the tangible cognitive resources required to match the demands of a task may not be congruent to the perceived effort required as interpreted by the individual listener. This theory also accounts for the interrelated effects of motivation and pleasure on the perception of effort (Pichora-Fuller et al. 2016). Therefore, a combination of exerted effort and assessed effort contribute to the overall LE experienced (Francis and Love 2020). From this depiction, it is clear to see the complexities involved with LE mechanisms and the challenges associated with its measurement.

Together, the lack of association between subgroups demonstrated across the literature and the cognitive mechanisms discussed above suggest that measures of LE are spatially stratified according to which components of the cognitive pathways they encapsulate. This means that behavioural, physiological, and self-reported measures have the potential to encode different aspects of the neural and cognitive routes involved with effort perception.

Effect of experimental conditions

To further build upon the aforementioned model, it is important to consider the effect that experimental conditions have on the degree of association between measures of LE. As stated, significant correlations between measures can and do occur, a finding which would seemingly refute the premise of the multi-dimensional nature of LE. To help explain this incongruency, this review highlights the fragility of significant correlations by referencing the influence of experimental conditions on the strength of correlations. This is reflected by the fact that 59/99 (56%) of the significant correlation were related to a specific experimental condition. Changing either the listening task or the timeframe LE was measured has been shown to alter the strength of correlation between measures of LE. This inconsistency may reflect the sensitivity of different instruments and their reliance on the cognitive capacity model (Kahneman 1973; Pichora-Fuller et al. 2016). This model posits that more difficult scenarios require individuals to draw upon more cognitive reserves. The more of these limited resources the individual uses, the more effort they experience. The ability of a measure to capture LE is, therefore, dependent on the relationship between its sensitivity and the task difficulty (Hick and Tharpe 2002). Differences in the sensitivity of each measure make it challenging to demonstrate associations in easier listening conditions hence explaining why the strength of correlations appears to change based upon the timeframe LE was sampled and the task used to evoke the effort. This highlights how measures of LE are context dependant and thus are appropriately termed ‘markers’ rather than an ‘invariant’ as a one to one relationship between instrument and outcome cannot be demonstrated (Richter and Slade 2017). Additionally,cautions against generalising the results of one paper across the entire literature. One possible way to mitigate this effect is to conduct more ecological studies outside the laboratory to help encapsulate real-world conditions more representationally (Bruya and Tang 2018).

To analyse this finding, and in keeping with the astronomical theme of multi-dimensions, this paper denotes a model of LE which accounts for the conditional effect on the association between measures, termed the solar system model of LE. In this model, LE is depicted as the sun, and the different measurement tools take the role of the orbiting planets. When the various planets align (Experiments with difficult listening situations), they absorb the same solar outputs (LE dimensions). This mirrors the notion that when the conditions are right, the measures capture the same dimensions of LE. As the experimental conditions become easier, the sensitivity of each measure pulls the planets out of alignment, and the solar outputs they capture are different. This suggests that LE measures depend on the experimental conditions within which they are applied. A diagrammatic representation of this model is shown in figure 3 below. This model only applies to a minority of LE measures; the remainder, as shown by this review, encompass different dimensions of LE no matter what experimental conditions are enacted.

Temporal stratification

The final observation noted from this review is the lack of association between different measures within the same subgroup. This can be evidenced by considering two factors:

1. How many 'within subgroup' comparisons reached statistical significance? (See table 2)

2. What strength did the statistically significant results reach? (See table 1)

The comparisons between EQ - EQ, B - B, and FQ - FQ subgroups satisfy the first point above, with the percentage of trials reaching significance equating to 100%, 66.7%, and 100%, respectively. Despite this, the modal classification for each remained as fair (coefficient 0.3 to 0.6 or -0.3 to -0.6). Moreover, the P - P subgroup failed to meet either point. Indeed, the number of significant results totalled 16.3%, representing the lowest from any subgroup comparison, and the modal classification was poor (coefficient 0 to 0.3 or 0 to -0.3). Interestingly, six of the seven significant results within the P - P subgroup comparison involved measures that utilise the same underlying mechanism to sample effort. For instance, Skin conductance/EEG alpha were significantly correlated, and both measure the electrophysiological response to the effort. A significant correlation was also noted between Cortisol/Chromogranin A with both being hormonal measures. Given this disparity amongst physiological measures, it may be more accurate to subclassify them according to their mechanism of action. Logical groupings include functional imaging, electrophysiological measures, and hormonal measures.

Together, these results suggest that measures of LE may also be temporally stratified along a timeline which begins at the onset of an effortful stimulus. For example, functional imaging measures such as infra-red spectroscopy may reflect the more real-time cerebral processes that result from LE.

Electrophysiological measures, including pupillometry, capture the subsequent neuronal response to these processes. Followed by the cascade of endocrine mechanisms recruited in response to effort, which would be sampled using hormonal measures of LE such as cortisol. Subsequently, behavioural measures would then encode for the immediate reactive response of the individual exposed to effort via their reaction times. Finally, both effort and fatigue questionnaires would capture the downstream consequences of LE, which manifests much further along the effort perception pathway. The decision to group the subtypes of questionnaires together can be rationalised by examining that 79.2% of trials that compared EQ - FQ reached significance satisfying the first point above. A representation of the temporal stratification concept is found in figure 4 below.

The spatial and temporal stratification of LE measures explains why the multi-dimensional model exists and why specific tools are weakly correlated. Moreover, the fragility of significant correlations has been explained by examining the effect of experimental conditions on the strength of association between two tools. Accepting these findings clarifies why the literature has not been able to pin down a single optimum measure for LE and advocates exploring different approaches to incorporate LE into clinical assessment.

An alternative approach to measuring listening effort

Given the seemingly perpetual challenge of selecting one measure of LE to clinically assess individuals with hearing impairment, it is prudent to consider other possible alternatives to help document the abstract notion. As previously discussed, there are several recognised consequences associated with LE, which may represent a more suitable target for screening tools. Furthermore, this review has shown that fatigue questionnaires perform similarly to effort indices in the classification matrix of subgroups. Therefore, shifting attitudes towards capturing these downstream effects creates an opportunity to design a composite measure comprised of previously validated tools.

First, to start with fatigue, many of the papers included in this review have referred to fatigue, and some have already adopted the approach of measuring this metric alongside effort in hearing-impaired individuals (Alhanbali et al. 2018, 2019; Dwyer et al. 2019; Key et al. 2017; Picou, Moore, and Ricketts 2017; Wang et al. 2018). Many of the proposed measures for LE demonstrate considerable overlap with fatigue. Take, for instance, the use of cortisol, performance over time or effort/fatigue assessment scales (Bess and Hornsby 2014). The evidence for the connection between effort and fatigue is becoming increasingly formidable and is now widely accepted as a consequence of prolonged intensive listening. Similarly to LE, fatigue displays multi-dimensional characteristics making it challenging to measure (Bess and Hornsby 2014; Gawron 2016). As a result, many self-reported tools have been proposed for the adult population. Gawron (2016) provides a detailed overview of these fatigue questionnaires and provides information on their possible optimal application (Gawron 2016). Examples that may be relevant to LE include the Checklist for individuals Strength's questionnaire due to its application to chronic fatigue and visual analogue scales for fatigue which have been widely used for hearing-impaired individuals already (Alhanbali et al. 2021; Gawron 2016). Regarding individuals with hearing loss, recent work by Hornsby et al. has validated a fatigue scale that aims to capture hearing-related fatigue specifically – The Vanderbilt Fatigue Scale. This instrument has now been validated in adults and children (Hornsby et al. 2021, 2022).

The next consequence to consider is the stress generated from effortful listening. Again, several studies included in this review have acknowledged stress as a corollary of LE and attempted to represent this within the testing strategy (Dwyer et al. 2019; Kramer et al. 2016; Mackersie, MacPhee, and Heldt 2015; Zekveld et al. 2019). Stress accumulates due to over-activation of autonomic processes and disruption of the endocrine system's cortisol axis (Sharma 2021). Given this, physiological measures (Cortisol, Skin conductance, Heart rate variability) which directly sample these changes may provide the most accurate insight. Their use in clinical practice is hindered by the ability to test repeatedly to establish a longitudinal trend of stress levels. An alternative to these objective measures is the Perceived Stress Scale (PSS), first proposed by Cohen in the 1980s (Cohen, Kamarck, and Mermelstein 1983). The PSS is the most widely used self-reported measure for stress and has been shown to apply to adults, children, and across ethnic groups (Baik et al. 2019; Huang et al. 2020; White 2014).

Outside of fatigue and stress, which are well-documented consequences of LE, there are more subtle effects that this paper will now explore. Individuals with hearing impairment have been shown to suffer lower levels of self-assurance and confidence compared with normal hearing counterparts (Bat-Chava 1993; Theunissen et al. 2014). Whilst this is generally accepted, an interesting observation was noted by Warner-Czyz et al, whereby children who had received a hearing intervention (hearing aid or cochlear implant) rated their self-esteem higher than normal-hearing children (Warner-Czyz et al. 2015). This highlights the reversibility of the detrimental impact hearing loss has on self-worth and the potential for confidence to be used as an indicator of benefit in before and after studies. The relationship between hearing ability and confidence is likely a multifactorial process influenced by factors such as the ability to communicate, social connectedness, physical appearance, and personal circumstances. Notwithstanding these many facets, qualitative data from Hughes et al. (2018) did elicit a relationship between low confidence and the polarity between listening effort and reward through focus group interviews (Hughes et al. 2018). Further to these internal determinants of confidence, educational outcomes are closely linked to an individual's sense of esteem (Rubin, Dorle, and Sandidge 1977). As elaborated previously, many studies have observed a trend for poorer school-based outcomes amongst those with hearing impairment (Bess et al. 1998; Niedzielski et al. 2006; Purcell et al. 2016). It is therefore plausible to suggest that LE may adversely affect self-esteem levels through its connection to the aforementioned factors. The Rosenberg scale (1965) is perhaps the most widely utilised measure for self-esteem (Rosenberg 1965). It has undergone rigorous evaluation and consistently yielded impressive reliability indices across age groups, ethnicities, languages, and genders (Dukes and Martinez 1994). Within the UK population, the Rosenberg scale was demonstrated to have a Cronbach alpha level of 0.9, indicating strong internal consistency (Schmitt and Allik 2005). Moreover, it has already been used within the hearing-impaired population (Theunissen et al. 2014; Warner-Czyz et al. 2015). Therefore, the Rosenberg scale represents a strong candidate for a composite tool to measure the possible confidence sequalae emanating from LE.

The final downstream effects of LE identified through this review is an individual's desire to take control of a situation or to give up on a listening task. This was enshrined within the work of Picou and colleagues in the effect of hearing aid settings on subjective ratings of LE (Picou et al. 2017; Picou and Ricketts 2018). The idea is that at increasing levels of task difficulty, individuals would be more likely to act to change something about the situation (move to a quieter room) or give up on the task altogether (Picou and Ricketts 2018). As opposed to the other downstream effects discussed in this section, these reflect more acute consequences of LE and will depend heavily on the task's difficulty. The decision to take control and give up on a task represents reactive tipping points for action in response to effortful listening. They used specific questions to assess these phenomena – 'How likely would you be to try to do something else to improve the situation (e.g. move to a quiet room, ask the speaker to speak louder)?' and 'How likely would you be to give up or just stop trying?' (Picou and Ricketts 2018).

Justification for a questionnaire-based approach

This review has provided evidence for the various downstream effects of LE in the form of – fatigue, stress, low self-assurance, and increased desire to take control and give up on a task. They represent some of the acute and chronic consequences faced by individuals exposed to prolonged periods of intense listening. Given the widespread use of fatigue scales in LE studies and the difficulties associated with measuring LE directly, this paper advocates for a shift in approaches to focus on capturing these negative sequelae. As discussed, several self-reported tools are already validated for each domain, creating the possibility of amalgamating the most relevant parts of each tool into a single measure which provides a holistic insight into the quality-of-life detriments stemming from LE.

Additionally, questionnaire-based instruments are less resource-intensive than more complex physiological and behaviour measures of LE. This method allows for longitudinal sampling of LE, which helps to trend how the individual is affected over time (Farrington 1991). This approach reflects the dynamic nature of LE. It would also allow for 'real-world' data collection which may help alleviate some of the issues that stem from the influence that experimental conditions have on LE instruments discussed previously (Bruya and Tang 2018; Katkade, Sanders, and Zou 2018). Finally, since questionnaires can be administered outside the hospital/laboratory setting, this approach would align with the post-COVID shift to try and move towards a remote consultation model wherever possible (Gupta, Gkiousias, and Bhutta 2021; Hutchings 2020; Samarrai et al. 2021).

Review limitations

Several factors limit the results obtained through this systematic review. The most important relates to the quality of included studies. Many papers were graded as having a serious risk of bias and low overall quality, which has overt connotations for the above inferences. The low sample size of many of the included papers may have limited ability to demonstrate statistical significance, creating the possibility for the results to appear poorly correlated when in fact, this may be due to poor study design. Linked to study design, publication bias my have artificially influenced the observed results due to a tendency for those articles demonstrating significant results being looked upon more favourably in publication process (Nair 2019). Thus, even weaker correlations may have been noted from unpublished research. Furthermore, the high degree of heterogeneity between included studies prevented the use of meta-analytical methods for data synthesis. Another inherent challenge with reviewing correlational data from a heterogenous study pool, is the difficulty in using quantitative methods to compare individual analyses. Although this review did not find a difference in the strength of correlations observed for adults and children, only three papers out of the 48 included studies contained a paediatric population. This small sample limits the ability to draw inferences and should be addressed in future work. Finally, a date limit of 2000 was placed on the search strategy of this review. While the specific terminology of "listening effort" may not have been used commonly prior to this date, there is still potential for some relevant articles which explore the effort related to hearing being missed. For example, Feuerstein correlated "perceived ease of listening" to word recognition in 1992, and Kramer correlated "Self-rated handicap" to pupil dilation in 1996 (Feuerstein 1992; Kramer et al. 1997).

Future work

Future work should focus on more robust studies to compare different measures of LE, including hearing-impaired individuals alongside normal hearing controls, measuring LE at multiple timepoints, ensuring sample size meets power requirements and the inclusion of paediatric participants. This will help address some of the fundamental limitations of this review and provide further clarification regarding the multi-dimensional model of LE. This paper has probed into some of the downstream effects of LE; however, the list created above is not exhaustive, and further research should aim to identify other potential consequences. Finally, exploratory research should be conducted to assess whether a composite tool, proposed within the discussion section, may be feasible as a proxy measure for LE.

Conclusion

The results of this review have shown that the LE measures are poorly correlated, further supporting the notion of the multi-dimensional model of LE. Taken further, this review has identified some of the potential dimensions that may impact LE, namely the cognitive processes involved with LE (Spatial dimension), the experimental conditions used to elicit LE and the time lapse between stimulus and dimension along the effort pathway measured (temporal dimension). This may indicate that no one measure will be adequate to capture LE in its entirety. To overcome this problematic finding, this review discusses the possibility of shifting focus from measuring LE directly to measuring the consequences of prolonged exposure to effortful listening. These include, but are not limited to, fatigue, stress, low confidence, and increased desire to take control of a situation or give up on a task. A composite tool that comprises components of previously validated measures for these effects may provide a better insight into the burden of LE.

Funding statement: This research was funded through a grant provided by Cochlear Research & Development Ltd under reference number IIR-2186

References

Included studies

Alhanbali, S., Dawes, P., Lloyd, S., & Munro, K. J. (2017). Self-Reported Listening-Related Effort and Fatigue in Hearing-Impaired Adults. Ear and Hearing, 38(1), e39. https://doi.org/10.1097/AUD.0000000000000361

Alhanbali, S., Dawes, P., Lloyd, S., & Munro, K. J. (2018). Hearing Handicap and Speech Recognition Correlate With Self-Reported Listening Effort and Fatigue. Ear and Hearing, 39(3), 470–474. https://doi.org/10.1097/AUD.0000000000000515

Alhanbali, S., Dawes, P., Millman, R. E., & Munro, K. J. (2019). Measures of Listening Effort Are Multi-dimensional. Ear and Hearing, 40(5), 1084–1097. https://doi.org/10.1097/AUD.0000000000000697

Alhanbali, S., Munro, K. J., Dawes, P., Carolan, P. J., & Millman, R. E. (2021). Dimensions of self-reported listening effort and fatigue on a digits-in-noise task, and association with baseline pupil size and performance accuracy. International Journal of Audiology, 60(10), 762–772. https://doi.org/10.1080/14992027.2020.1853262

Athey, H. (n.d.). 'The effect of auditory fatigue on reaction time in normal hearing list' by Haley Athey. Retrieved 19 December 2021, from https://commons.lib.jmu.edu/diss201019/108/

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2017). Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort. Cognitive Neurodynamics, 11(3), 203–215. https://doi.org/10.1007/s11571-017-9425-5

Bertoli, S., & Bodmer, D. (2014). Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 125(5), 1030–1041. https://doi.org/10.1016/j.clinph.2013.09.045

Desjardins, J. L. (2016). The Effects of Hearing Aid Directional Microphone and Noise Reduction Processing on Listening Effort in Older Adults with Hearing Loss. Journal of the American Academy of Audiology, 27(1), 29–41. https://doi.org/10.3766/jaaa.15030

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. Scientific Reports, 9(1), 11278. https://doi.org/10.1038/s41598-019-47643-1

Dwyer, R. T., Gifford, R. H., Bess, F. H., Dorman, M., Spahr, A., & Hornsby, B. W. Y. (2019). Diurnal Cortisol Levels and Subjective Ratings of Effort and Fatigue in Adult Cochlear Implant Users: A Pilot Study. American Journal of Audiology, 28(3), 686–696. https://doi.org/10.1044/2019_AJA-19-0009

Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., & Sandmann, P. (2016). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. Neuropsychologia, 87, 169–181. https://doi.org/10.1016/j.neuropsychologia.2016.05.019

Francis, A. L., Bent, T., Schumaker, J., Love, J., & Silbert, N. (2021). Listener characteristics differentially affect self-reported and physiological measures of effort associated with two challenging listening conditions. Attention, Perception & Psychophysics, 83(4), 1818–1841. https://doi.org/10.3758/s13414-020-02195-9

Giuliani, N. P. (n.d.). Comparisons of physiologic and psychophysical measures of listening effort in normal-hearing adults [University of Iowa]. Retrieved 20 December 2021, from https://iro.uiowa.edu/esploro/outputs/doctoral/Comparisons-of-physiologic-and-psychophysical-measures/9983776858302771

Gustafson, S. J., Key, A. P., Hornsby, B. W. Y., & Bess, F. H. (2018). Fatigue Related to Speech Processing in Children With Hearing Loss: Behavioral, Subjective, and Electrophysiological Measures. Journal of Speech, Language, and Hearing Research: JSLHR, 61(4), 1000–1011. https://doi.org/10.1044/2018_JSLHR-H-17-0314

Holube, I., Haeder, K., Imbery, C., & Weber, R. (2016). Subjective Listening Effort and Electrodermal Activity in Listening Situations with Reverberation and Noise. Trends in Hearing, 20, 2331216516667734. https://doi.org/10.1177/2331216516667734

Hornsby, B. W. Y. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear and Hearing, 34(5), 523–534. https://doi.org/10.1097/AUD.0b013e31828003d8

Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A Comparison of Two Methods for Measuring Listening Effort As Part of an Audiologic Test Battery. American Journal of Audiology, 24(3), 419–431. https://doi.org/10.1044/2015_AJA-14-0058

Key, A. P., Gustafson, S. J., Rentmeester, L., Hornsby, B. W. Y., & Bess, F. H. (2017). Speech-Processing Fatigue in Children: Auditory Event-Related Potential and Behavioral Measures. Journal of Speech, Language, and Hearing Research : JSLHR, 60(7), 2090–2104. https://doi.org/10.1044/2016_JSLHR-H-16-0052

Kramer, S. E., Teunissen, C. E., & Zekveld, A. A. (2016). Cortisol, Chromogranin A, and Pupillary Responses Evoked by Speech Recognition Tasks in Normally Hearing and Hard-of-Hearing Listeners:

A Pilot Study. Ear and Hearing, 37 Suppl 1, 126S-35S. https://doi.org/10.1097/AUD.0000000000000311

Lau, M. K., Hicks, C., Kroll, T., & Zupancic, S. (2019). Effect of Auditory Task Type on Physiological and Subjective Measures of Listening Effort in Individuals With Normal Hearing. Journal of Speech, Language, and Hearing Research: JSLHR, 62(5), 1549–1560. https://doi.org/10.1044/2018_JSLHR-H-17-0473

Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indices of listening effort in a competing-talker task. Journal of the American Academy of Audiology, 22(2), 113–122. https://doi.org/10.3766/jaaa.22.2.6

Mackersie, C. L., MacPhee, 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(1), 145–154. https://doi.org/10.1097/AUD.0000000000000091

Magudilu Srishyla Kumar, L. (2020). Measuring listening effort using physiological, behavioral and subjective methods in normal hearing subjects: Effect of signal to noise ratio and presentation level [James Madison University]. https://commons.lib.jmu.edu/diss202029/7

McGarrigle, R., Rakusen, L., & Mattys, S. (2020). Effortful listening under the microscope: Examining relations between pupillometric and subjective markers of effort and tiredness from listening. https://doi.org/10.31234/osf.io/kych5

McMahon, C. M., Boisvert, I., de Lissa, P., Granger, L., Ibrahim, R., Lo, C. Y., Miles, K., & Graham, P. L. (2016). Monitoring Alpha Oscillations and Pupil Dilation across a Performance-Intensity Function. Frontiers in Psychology, 7, 745. https://doi.org/10.3389/fpsyg.2016.00745

Megha, null, & Maruthy, S. (2019). Auditory and Cognitive Attributes of Hearing Aid Acclimatization in Individuals With Sensorineural Hearing Loss. American Journal of Audiology, 28(2S), 460–470. https://doi.org/10.1044/2018_AJA-IND50-18-0100

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., de Lissa, P., Graham, P., & Lyxell, B. (2017). Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG. Trends in Hearing, 21, 2331216517706396. https://doi.org/10.1177/2331216517706396

Perreau, A. E., Wu, Y.-H., Tatge, B., Irwin, D., & Corts, D. (2017). Listening Effort Measured in Adults with Normal Hearing and Cochlear Implants. Journal of the American Academy of Audiology, 28(8), 685–697. https://doi.org/10.3766/jaaa.16014

Picou, E. M., Moore, T. M., & Ricketts, T. A. (2017). The Effects of Directional Processing on Objective and Subjective Listening Effort. Journal of Speech, Language, and Hearing Research: JSLHR, 60(1), 199–211. https://doi.org/10.1044/2016_JSLHR-H-15-0416

Picou, E. M., & Ricketts, T. A. (2018). The relationship between speech recognition, behavioural listening effort, and subjective ratings. International Journal of Audiology, 57(6), 457–467. https://doi.org/10.1080/14992027.2018.1431696

Rovetti, J., Goy, H., Pichora-Fuller, M. K., & Russo, F. A. (2019). Functional Near-Infrared Spectroscopy as a Measure of Listening Effort in Older Adults Who Use Hearing Aids. Trends in Hearing, 23, 2331216519886722. https://doi.org/10.1177/2331216519886722

Russo, F. Y., Hoen, M., Karoui, C., Demarcy, T., Ardoint, M., Tuset, M.-P., De Seta, D., Sterkers, O., Lahlou, G., & Mosnier, I. (2020). Pupillometry Assessment of Speech Recognition and Listening Experience in Adult Cochlear Implant Patients. Frontiers in Neuroscience, 14, 556675. https://doi.org/10.3389/fnins.2020.556675

Schafer, P. J., Serman, M., Arnold, M., Corona-Strauss, F. I., Strauss, D. J., Seidler-Fallbohmer, B., & Seidler, H. (2015). Evaluation of an objective listening effort measure in a selective, multi-speaker listening task using different hearing aid settings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2015, 4647–4650. https://doi.org/10.1109/EMBC.2015.7319430

Seeman, S., & Sims, R. (2015). Comparison of Psychophysiological and Dual-Task Measures of Listening Effort. Journal of Speech, Language, and Hearing Research: JSLHR, 58(6), 1781–1792. https://doi.org/10.1044/2015_JSLHR-H-14-0180

Seifi Ala, T., Graversen, C., Wendt, D., Alickovic, E., Whitmer, W. M., & Lunner, T. (2020). An exploratory Study of EEG Alpha Oscillation and Pupil Dilation in Hearing-Aid Users During Effortful listening to Continuous Speech. PloS One, 15(7), e0235782. https://doi.org/10.1371/journal.pone.0235782

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: JSLHR, 61(6), 1463–1486. https://doi.org/10.1044/2018_JSLHR-H-17-0257

Visentin, C., & Prodi, N. (2021). How Reliable are 11- to 13-Year-Olds' Self-Ratings of Effort in Noisy Conditions? Frontiers in Built Environment, 7, 108. https://doi.org/10.3389/fbuil.2021.688016

Visentin, C., Prodi, N., Cappelletti, F., Torresin, S., & Gasparella, A. (2019). Speech intelligibility and listening effort in university classrooms for native and non-native Italian listeners. Building Acoustics, 26, 1351010X1988231. https://doi.org/10.1177/1351010X19882314

Visentin, C., Valzolgher, C., Pellegatti, M., Potente, P., Pavani, F., & Prodi, N. (2021). A comparison of simultaneously-obtained measures of listening effort: Pupil dilation, verbal response time and self-rating. International Journal of Audiology, 1–13. https://doi.org/10.1080/14992027.2021.1921290

Wang, Y., Kramer, S. E., Wendt, D., Naylor, G., Lunner, T., & Zekveld, A. A. (2018). The Pupil Dilation Response During Speech Perception in Dark and Light: The Involvement of the Parasympathetic Nervous System in Listening Effort. Trends in Hearing, 22, 2331216518816603. https://doi.org/10.1177/2331216518816603

Wang, Y., Naylor, G., Kramer, S. E., Zekveld, A. A., Wendt, D., Ohlenforst, B., & Lunner, T. (2018). Relations Between Self-Reported Daily-Life Fatigue, Hearing Status, and Pupil Dilation During a Speech Perception in Noise Task. Ear and Hearing, 39(3), 573–582. https://doi.org/10.1097/AUD.0000000000000512

White, B. E., & Langdon, C. (2021). The cortical organization of listening effort: New insight from functional near-infrared spectroscopy. NeuroImage, 240, 118324. https://doi.org/10.1016/j.neuroimage.2021.118324

Wisniewski, M. G., Thompson, E. R., Iyer, N., Estepp, J. R., Goder-Reiser, M. N., & Sullivan, S. C. (2015). Frontal midline θ power as an index of listening effort. Neuroreport, 26(2), 94–99. https://doi.org/10.1097/WNR.0000000000000306

Wu, Y.-H., Aksan, N., Rizzo, M., Stangl, E., Zhang, X., & Bentler, R. (2014). Measuring listening effort: Driving simulator versus simple dual-task paradigm. Ear and Hearing, 35(6), 623–632. https://doi.org/10.1097/AUD.0000000000000079

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. Ear and Hearing, 32(4), 498–510. https://doi.org/10.1097/AUD.0b013e31820512bb

Zekveld, A. A., van Scheepen, J. A. M., Versfeld, N. J., Veerman, E. C. I., & Kramer, S. E. (2019). Please try harder! The influence of hearing status and evaluative feedback during listening on the pupil dilation response, saliva-cortisol and saliva alpha-amylase levels. Hearing Research, 381, 107768. https://doi.org/10.1016/j.heares.2019.07.005

Zekveld, A., Scheepen, J., Versfeld, N., Kramer, S., & Steenbergen, H. (2020). The Influence of Hearing Loss on Cognitive Control in an Auditory Conflict Task: Behavioral and Pupillometry Findings. Journal of Speech, Language, and Hearing Research, 63, 1–10. https://doi.org/10.1044/2020_JSLHR-20-00107

Zhao, S., Bury, G., Milne, A., & Chait, M. (2019). Pupillometry as an Objective Measure of Sustained Attention in Young and Older Listeners. Trends in Hearing, 23, 2331216519887815. https://doi.org/10.1177/2331216519887815

General references

Alhanbali, Sara, Piers Dawes, Simon Lloyd, and Kevin J. Munro. 2017. 'Self-Reported Listening-Related Effort and Fatigue in Hearing-Impaired Adults'. Ear and Hearing 38(1):e39. doi: 10.1097/AUD.0000000000000361.

Alhanbali, Sara, Piers Dawes, Simon Lloyd, and Kevin J. Munro. 2018. 'Hearing Handicap and Speech Recognition Correlate With Self-Reported Listening Effort and Fatigue'. Ear and Hearing 39(3):470–74. doi: 10.1097/AUD.0000000000000515.

Alhanbali, Sara, Piers Dawes, Rebecca E. Millman, and Kevin J. Munro. 2019. 'Measures of Listening Effort Are Multidimensional'. Ear and Hearing 40(5):1084–97. doi: 10.1097/AUD.0000000000000697.

Alhanbali, Sara, Kevin J. Munro, Piers Dawes, Peter J. Carolan, and Rebecca E. Millman. 2021. 'Dimensions of Self-Reported Listening Effort and Fatigue on a Digits-in-Noise Task, and Association with Baseline Pupil Size and Performance Accuracy'. International Journal of Audiology 60(10):762–72. doi: 10.1080/14992027.2020.1853262.

Andrade, Chittaranjan. 2019. 'The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives'. Indian Journal of Psychological Medicine 41(3):210–15. doi: 10.4103/IJPSYM.IJPSYM_193_19.

Archbold, Sue, Zheng Yen Ng, Suzanne Harrigan, Sue Gregory, Lynda Holland B. Ed, and Imran Mulla. 2015. 'Views of Parents and Teachers'. 48.

Armstrong, E. C. 1999. 'The Well-Built Clinical Question: The Key to Finding the Best Evidence Efficiently'. WMJ: Official Publication of the State Medical Society of Wisconsin 98(2):25–28.

Ayres, Paul, Joy Yeonjoo Lee, Fred Paas, and Jeroen J. G. van Merriënboer. 2021. 'The Validity of Physiological Measures to Identify Differences in Intrinsic Cognitive Load'. Frontiers in Psychology 12:3905. doi: 10.3389/fpsyg.2021.702538.

Baik, Sharon H., Rina S. Fox, Sarah D. Mills, Scott C. Roesch, Georgia Robins Sadler, Elizabeth A. Klonoff, and Vanessa L. Malcarne. 2019. 'Reliability and Validity of the Perceived Stress Scale-10 in Hispanic Americans with English or Spanish Language Preference'. Journal of Health Psychology 24(5):628–39. doi: 10.1177/1359105316684938.

Bat-Chava, Yael. 1993. 'Antecedents of Self-Esteem in Deaf People: A Meta-Analytic Review'. Rehabilitation Psychology 38(4):221–34. doi: 10.1037/h0080303.

Bess, Fred, Jeanne Dodd-Murphey, and Robert Parker. 1998. 'Children with Minimal Sensorineural Hearing Loss: Prevalence, Educational Performance, and Functional Status - PubMed'. Retrieved 15 December 2021 (https://pubmed.ncbi.nlm.nih.gov/9796643/).

Bess, Fred H., Samantha J. Gustafson, Blythe A. Corbett, E. Warren Lambert, Stephen M. Camarata, and Benjamin W. Y. Hornsby. 2016. 'Salivary Cortisol Profiles of Children with Hearing Loss'. Ear and Hearing 37(3):334–44. doi: 10.1097/AUD.0000000000000256.

Bess, Fred H., and Benjamin W. Y. Hornsby. 2014. 'Commentary: Listening Can Be Exhausting--Fatigue in Children and Adults with Hearing Loss'. Ear and Hearing 35(6):592–99. doi: 10.1097/AUD.0000000000000099.

Bruya, Brian, and Yi-Yuan Tang. 2018. ‘Is Attention Really Effort? Revisiting Daniel Kahneman’s Influential 1973 Book Attention and Effort’. Frontiers in Psychology 9:1133. doi: 10.3389/fpsyg.2018.01133.

Chan, Y. H. 2003. ‘Biostatistics 104: Correlational Analysis’. Singapore Medical Journal 44(12):614–19.

Cohen, S., T. Kamarck, and R. Mermelstein. 1983. ‘A Global Measure of Perceived Stress’. Journal of Health and Social Behavior 24(4):385–96.

Downs, D. W., and M. A. Crum. 1978. ‘Processing Demands during Auditory Learning under Degraded Listening Conditions’. Journal of Speech and Hearing Research 21(4):702–14. doi: 10.1044/jshr.2104.702.

Dukes, Richard L., and Rubén Martinez. 1994. ‘The Impact of Ethgender on Self-Esteem among Adolescents’. Adolescence 29(113):105–15.

Dwyer, Robert T., René H. Gifford, Fred H. Bess, Michael Dorman, Anthony Spahr, and Benjamin W. Y. Hornsby. 2019. ‘Diurnal Cortisol Levels and Subjective Ratings of Effort and Fatigue in Adult Cochlear Implant Users: A Pilot Study’. American Journal of Audiology 28(3):686–96. doi: 10.1044/2019_AJA-19-0009.

Ebell, M. 1999. ‘Information at the Point of Care: Answering Clinical Questions’. The Journal of the American Board of Family Practice 12(3):225–35. doi: 10.3122/jabfm.12.3.225.

Farrington, DAVID P. 1991. ‘Longitudinal Research Strategies: Advantages, Problems, and Prospects’. Journal of the American Academy of Child & Adolescent Psychiatry 30(3):369–74. doi: 10.1097/00004583-199105000-00003.

Feuerstein, J. F. 1992. ‘Monaural versus Binaural Hearing: Ease of Listening, Word Recognition, and Attentional Effort’. Ear and Hearing 13(2):80–86.

Francis, Alexander L., Tessa Bent, Jennifer Schumaker, Jordan Love, and Noah Silbert. 2021. 'Listener Characteristics Differentially Affect Self-Reported and Physiological Measures of Effort Associated with Two Challenging Listening Conditions'. Attention, Perception & Psychophysics 83(4):1818–41. doi: 10.3758/s13414-020-02195-9.

Francis, Alexander L., and Jordan Love. 2020. 'Listening Effort: Are We Measuring Cognition or Affect, or Both?' WIREs Cognitive Science 11(1):e1514. doi: 10.1002/wcs.1514.

Gawron, Valerie J. 2016. 'Overview of Self-Reported Measures of Fatigue'. The International Journal of Aviation Psychology 26(3–4):120–31. doi: 10.1080/10508414.2017.1329627.

Giuliani, Nicholas Patrick. 2021. 'Comparisons of Physiologic and Psychophysical Measures of Listening Effort in Normal-Hearing Adults'. University of Iowa.

Gupta, Tanya, Vasileios Gkiousias, and Mahmood F. Bhutta. 2021. 'A Systematic Review of Outcomes of Remote Consultation in ENT'. Clinical Otolaryngology 46(4):699–719. doi: 10.1111/coa.13768.

Gustafson, Samantha J., Alexandra P. Key, Benjamin W. Y. Hornsby, and Fred H. Bess. 2018. 'Fatigue Related to Speech Processing in Children With Hearing Loss: Behavioral, Subjective, and Electrophysiological Measures'. Journal of Speech, Language, and Hearing Research: JSLHR 61(4):1000–1011. doi: 10.1044/2018_JSLHR-H-17-0314.

Hart, Sandra G. 2006. 'Nasa-Task Load Index (NASA-TLX); 20 Years Later'. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50(9):904–8. doi: 10.1177/154193120605000909.

Henshaw, Helen, Linda Sharkey, David Crowe, and Melanie Ferguson. 2015. 'Research Priorities for Mild-to-Moderate Hearing Loss in Adults'. The Lancet 386(10009):2140–41. doi: 10.1016/S0140-6736(15)01048-X.

Hick, Candace Bourland, and Anne Marie Tharpe. 2002. 'Listening Effort and Fatigue in School-Age Children with and without Hearing Loss'. Journal of Speech, Language, and Hearing Research: JSLHR 45(3):573–84. doi: 10.1044/1092-4388(2002/046).

Holman, Jack A., Avril Drummond, and Graham Naylor. 2021. 'The Effect of Hearing Loss and Hearing Device Fitting on Fatigue in Adults: A Systematic Review'. Ear and Hearing 42(1):1–11. doi: 10.1097/AUD.0000000000000909.

Holube, Inga, Kristina Haeder, Christina Imbery, and Reinhard Weber. 2016. 'Subjective Listening Effort and Electrodermal Activity in Listening Situations with Reverberation and Noise'. Trends in Hearing 20:2331216516667734. doi: 10.1177/2331216516667734.

Hornsby, Benjamin W. Y., Stephen Camarata, Sun-Joo Cho, Hilary Davis, Ronan McGarrigle, and Fred H. Bess. 2021. 'Development and Validation of the Vanderbilt Fatigue Scale for Adults (VFS-A)'. Psychological Assessment 33(8):777–88. doi: 10.1037/pas0001021.

Hornsby, Benjamin W. Y., Stephen Camarata, Sun-Joo Cho, Hilary Davis, Ronan McGarrigle, and Fred H. Bess. 2022. 'Development and Evaluation of Pediatric Versions of the Vanderbilt Fatigue Scale for Children With Hearing Loss'. Journal of Speech, Language, and Hearing Research 65(6):2343–63. doi: 10.1044/2022_JSLHR-22-00051.

Hsu, Benson Cheng-Lin, Filiep Vanpoucke, and Astrid van Wieringen. 2017. 'Listening Effort Through Depth of Processing in School-Age Children'. Ear and Hearing 38(5):568–76. doi: 10.1097/AUD.0000000000000436.

Huang, Feifei, Huijun Wang, Zhihong Wang, Jiguo Zhang, Wenwen Du, Chang Su, Xiaofang Jia, Yifei Ouyang, Yun Wang, Li Li, Hongru Jiang, and Bing Zhang. 2020. 'Psychometric Properties of the Perceived Stress Scale in a Community Sample of Chinese'. BMC Psychiatry 20(1):130. doi: 10.1186/s12888-020-02520-4.

Hughes, Sarah E., Hayley A. Hutchings, Frances L. Rapport, Catherine M. McMahon, and Isabelle Boisvert. 2018. 'Social Connectedness and Perceived Listening Effort in Adult Cochlear Implant Users: A Grounded Theory to Establish Content Validity for a New Patient-Reported Outcome Measure'. Ear and Hearing 39(5):922–34. doi: 10.1097/AUD.0000000000000553.

Hutchings, Rachel. 2020. 'The Impact of Covid-19 on the Use of Digital Technology in the NHS'. Nuffieldtrust 23.

Kahneman, Daniel. 1973. Attention and Effort. 1st ed. Prentice-Hall.

Katkade, Vaibhav B., Kafi N. Sanders, and Kelly H. Zou. 2018. 'Real World Data: An Opportunity to Supplement Existing Evidence for the Use of Long-Established Medicines in Health Care Decision Making'. Journal of Multidisciplinary Healthcare 11:295–304. doi: 10.2147/JMDH.S160029.

Kemp, Andrew H., Andre R. Brunoni, Itamar S. Santos, Maria A. Nunes, Eduardo M. Dantas, Roberta Carvalho de Figueiredo, Alexandre C. Pereira, Antonio L. P. Ribeiro, José G. Mill, Rodrigo V. Andreão, Julian F. Thayer, Isabela M. Benseñor, and Paulo A. Lotufo. 2014. 'Effects of Depression, Anxiety, Comorbidity, and Antidepressants on Resting-State Heart Rate and Its Variability: An ELSA-Brasil Cohort Baseline Study'. American Journal of Psychiatry 171(12):1328–34. doi: 10.1176/appi.ajp.2014.13121605.

Key, Alexandra P., Samantha J. Gustafson, Lindsey Rentmeester, Benjamin W. Y. Hornsby, and Fred H. Bess. 2017. 'Speech-Processing Fatigue in Children: Auditory Event-Related Potential and Behavioral Measures'. Journal of Speech, Language, and Hearing Research : JSLHR 60(7):2090–2104. doi: 10.1044/2016_JSLHR-H-16-0052.

Kramer, Sophia E., Theo S. Kapteyn, Joost M. Festen, and Dirk J. Kuik. 1997. 'Assessing Aspects of Auditory Handicap by Means of Pupil Dilatation'. Audiology 36(3):155–64. doi: 10.3109/00206099709071969.

Kramer, Sophia E., Theo S. Kapteyn, and Tammo Houtgast. 2006. ‘Occupational Performance: Comparing Normally-Hearing and Hearing-Impaired Employees Using the Amsterdam Checklist for Hearing and Work’. International Journal of Audiology 45(9):503–12. doi: 10.1080/14992020600754583.

Kramer, Sophia E., Charlotte E. Teunissen, and Adriana A. Zekveld. 2016. ‘Cortisol, Chromogranin A, and Pupillary Responses Evoked by Speech Recognition Tasks in Normally Hearing and Hard-of-Hearing Listeners: A Pilot Study’. Ear and Hearing 37 Suppl 1:126S-35S. doi: 10.1097/AUD.0000000000000311.

Kumari, Meena, Ellena Badrick, Tarani Chandola, Emma K. Adam, Mai Stafford, Michael G. Marmot, Clemens Kirschbaum, and Mika Kivimaki. 2009. ‘Cortisol Secretion and Fatigue: Associations in a Community Based Cohort’. Psychoneuroendocrinology 34(10):1476–85. doi: 10.1016/j.psyneuen.2009.05.001.

Lefebvre, Carol, Julie Glanville, Simon Briscoe, Anne Littlewood, Chris Marshall, Maria-Inti Metzendorf, Anna Noel-Storr, Tamara Rader, Farhad Shokraneh, James Thomas, and Susan Wieland. 2021. ‘Chapter 4: Searching for and Selecting Studies’. Retrieved 22 November 2021 (https://training.cochrane.org/handbook/current/chapter-04).

Mackersie, Carol L., Imola X. MacPhee, and Emily W. Heldt. 2015. ‘Effects of Hearing Loss on Heart-Rate Variability and Skin Conductance Measured During Sentence Recognition in Noise’. Ear and Hearing 36(1):145–54. doi: 10.1097/AUD.0000000000000091.

Mattys, Sven L., Matthew H. Davis, Ann R. Bradlow, and Sophie K. Scott. 2012. ‘Speech Recognition in Adverse Conditions: A Review’. Language and Cognitive Processes 27(7–8):953–78. doi: 10.1080/01690965.2012.705006.

McGarrigle, Ronan, Kevin J. Munro, Piers Dawes, Andrew J. Stewart, David R. Moore, Johanna G. Barry, and Sygal Amitay. 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(7):433–40. doi: 10.3109/14992027.2014.890296.

McMahon, Catherine M., Isabelle Boisvert, Peter de Lissa, Louise Granger, Ronny Ibrahim, Chi Yhun Lo, Kelly Miles, and Petra L. Graham. 2016. 'Monitoring Alpha Oscillations and Pupil Dilation across a Performance-Intensity Function'. Frontiers in Psychology 7:745. doi: 10.3389/fpsyg.2016.00745.

Mosadeghrad, Ali Mohammad. 2012. 'A Conceptual Framework for Quality of Care'. Materia Socio-Medica 24(4):251–61. doi: 10.5455/msm.2012.24.251-261.

Nair, Abhijit S. 2019. 'Publication Bias - Importance of Studies with Negative Results!' Indian Journal of Anaesthesia 63(6):505–7. doi: 10.4103/ija.IJA_142_19.

Niedzielski, Antoni, Ewa Humeniuk, Przemysław Błaziak, and Grazyna Gwizda. 2006. 'Intellectual Efficiency of Children with Unilateral Hearing Loss'. International Journal of Pediatric Otorhinolaryngology 70(9):1529–32. doi: 10.1016/j.ijporl.2006.02.011.

Ohlenforst, Barbara, Adriana A. Zekveld, Elise P. Jansma, Yang Wang, Graham Naylor, Artur Lorens, Thomas Lunner, and Sophia E. Kramer. 2017. 'Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: A Systematic Review'. Ear and Hearing 38(3):267–81. doi: 10.1097/AUD.0000000000000396.

Page, Matthew J., Julian Higgins, and Jonathan Sterne. 2021. 'Chapter 13: Assessing Risk of Bias Due to Missing Results in a Synthesis'. Retrieved 6 February 2022 (https://training.cochrane.org/handbook/current/chapter-13).

Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. 2021. 'The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews'. Systematic Reviews 10(1):89. doi: 10.1186/s13643-021-01626-4.

Peelle, Jonathan E. 2018. 'Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior'. Ear and Hearing 39(2):204–14. doi: 10.1097/AUD.0000000000000494.

Pichora-Fuller, M. Kathleen. 2016. 'How Social Psychological Factors May Modulate Auditory and Cognitive Functioning During Listening'. Ear and Hearing 37 Suppl 1:92S-100S. doi: 10.1097/AUD.0000000000000323.

Pichora-Fuller, M. Kathleen, Sophia E. Kramer, Mark A. Eckert, Brent Edwards, Benjamin W. Y. Hornsby, Larry E. Humes, Ulrike Lemke, Thomas Lunner, Mohan Matthen, Carol L. Mackersie, Graham Naylor, Natalie A. Phillips, Michael Richter, Mary Rudner, Mitchell S. Sommers, Kelly L. Tremblay, and Arthur Wingfield. 2016a. 'Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL)'. Ear and Hearing 37 Suppl 1:5S-27S. doi: 10.1097/AUD.0000000000000312.

Pichora-Fuller, M. Kathleen, Sophia E. Kramer, Mark A. Eckert, Brent Edwards, Benjamin W. Y. Hornsby, Larry E. Humes, Ulrike Lemke, Thomas Lunner, Mohan Matthen, Carol L. Mackersie, Graham Naylor, Natalie A. Phillips, Michael Richter, Mary Rudner, Mitchell S. Sommers, Kelly L. Tremblay, and Arthur Wingfield. 2016b. 'Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL)'. Ear and Hearing 37 Suppl 1:5S-27S. doi: 10.1097/AUD.0000000000000312.

Picou, Erin M., Travis M. Moore, and Todd A. Ricketts. 2017. 'The Effects of Directional Processing on Objective and Subjective Listening Effort'. Journal of Speech, Language, and Hearing Research: JSLHR 60(1):199–211. doi: 10.1044/2016_JSLHR-H-15-0416.

Picou, Erin M., and Todd A. Ricketts. 2018. ‘The Relationship between Speech Recognition, Behavioural Listening Effort, and Subjective Ratings’. International Journal of Audiology 57(6):457–67. doi: 10.1080/14992027.2018.1431696.

Purcell, Patricia L., Justin R. Shinn, Greg E. Davis, and Kathleen C. Y. Sie. 2016. ‘Children with Unilateral Hearing Loss May Have Lower Intelligence Quotient Scores: A Meta-Analysis’. The Laryngoscope 126(3):746–54. doi: 10.1002/lary.25524.

Richter, Michael, and Kate Slade. 2017. ‘Interpretation of Physiological Indicators of Motivation: Caveats and Recommendations’. International Journal of Psychophysiology 119:4–10. doi: 10.1016/j.ijpsycho.2017.04.007.

Rosenberg, Morris. 1965. Society and the Adolescent Child. Princeton, NJ.

Rovetti, Joseph, Huiwen Goy, M. Kathleen Pichora-Fuller, and Frank A. Russo. 2019. ‘Functional Near-Infrared Spectroscopy as a Measure of Listening Effort in Older Adults Who Use Hearing Aids’. Trends in Hearing 23:2331216519886722. doi: 10.1177/2331216519886722.

Rubin, Rosalyn A., Jeanne Dorle, and Susanne Sandidge. 1977. ‘Self-Esteem and School Performance’. Psychology in the Schools 14(4):503–7. doi: 10.1002/1520-6807(197710)14:4<503::AID-PITS2310140423>3.0.CO;2-Z.

Russo, Francesca Yoshie, Michel Hoen, Chadlia Karoui, Thomas Demarcy, Marine Ardoint, Maria-Pia Tuset, Daniele De Seta, Olivier Sterkers, Ghizlène Lahlou, and Isabelle Mosnier. 2020. ‘Pupillometry Assessment of Speech Recognition and Listening Experience in Adult Cochlear Implant Patients’. Frontiers in Neuroscience 14:556675. doi: 10.3389/fnins.2020.556675.

Samarrai, Ruwaa, Aaliyah C. Riccardi, Belachew Tessema, Michael Setzen, and Seth M. Brown. 2021. ‘Continuation of Telemedicine in Otolaryngology Post-COVID-19: Applications by Subspecialty’. American Journal of Otolaryngology 42(3):102928. doi: 10.1016/j.amjoto.2021.102928.

Schlotz, Wolff, Juliane Hellhammer, Peter Schulz, and Arthur A. Stone. 2004. 'Perceived Work Overload and Chronic Worrying Predict Weekend-Weekday Differences in the Cortisol Awakening Response'. Psychosomatic Medicine 66(2):207–14. doi: 10.1097/01.psy.0000116715.78238.56.

Schmitt, David P., and Jüri Allik. 2005. 'Simultaneous Administration of the Rosenberg Self-Esteem Scale in 53 Nations: Exploring the Universal and Culture-Specific Features of Global Self-Esteem'. Journal of Personality and Social Psychology 89(4):623–42. doi: 10.1037/0022-3514.89.4.623.

Schneiderman, Neil, Gail Ironson, and Scott D. Siegel. 2005. 'STRESS AND HEALTH: Psychological, Behavioral, and Biological Determinants'. Annual Review of Clinical Psychology 1:607–28. doi: 10.1146/annurev.clinpsy.1.102803.144141.

Schünemann, H., J. Brożek, G. Guyatt, and A. Oxman. 2013. 'GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations'. Retrieved 22 November 2021 (https://gradepro.org/terms/cite).

Sharma, Dushyant. 2021. 'Physiology of Stress and Its Management by Dushyant Kumar Sharma'. 2018, 1: 001:01–05.

Shields, C., H. Willis, J. Nichani, M. Sladen, and K. Kluk-de Kort. 2021. 'Listening Effort: WHAT Is It, HOW Is It Measured and WHY Is It Important?' Cochlear Implants International 0(0):1–4. doi: 10.1080/14670100.2021.1992941.

Sterne, Jonathan AC, Miguel A. Hernán, Barnaby C. Reeves, Jelena Savović, Nancy D. Berkman, Meera Viswanathan, David Henry, Douglas G. Altman, Mohammed T. Ansari, Isabelle Boutron, James R. Carpenter, An-Wen Chan, Rachel Churchill, Jonathan J. Deeks, Asbjørn Hróbjartsson, Jamie Kirkham, Peter Jüni, Yoon K. Loke, Theresa D. Pigott, Craig R. Ramsay, Deborah Regidor, Hannah R. Rothstein, Lakhbir Sandhu, Pasqualina L. Santaguida, Holger J. Schünemann, Beverly Shea, Ian Shrier, Peter Tugwell, Lucy Turner, Jeffrey C. Valentine, Hugh Waddington, Elizabeth Waters, George

A. Wells, Penny F. Whiting, and Julian PT Higgins. 2016. 'ROBINS-I: A Tool for Assessing Risk of Bias in Non-Randomised Studies of Interventions'. BMJ 355:i4919. doi: 10.1136/bmj.i4919.

Sterne, Jonathan, Miguel Hernán, Alexandra McAleenan, Barnaby Reeves, and Julian Higgins. 2021. 'Chapter 25: Assessing Risk of Bias in a Non-Randomized Study'. Retrieved 22 November 2021 (https://training.cochrane.org/handbook/current/chapter-25).

Theunissen, Stephanie C. P. M., Carolien Rieffe, Anouk P. Netten, Jeroen J. Briaire, Wim Soede, Maartje Kouwenberg, and Johan H. M. Frijns. 2014. 'Self-Esteem in Hearing-Impaired Children: The Influence of Communication, Education, and Audiological Characteristics'. PLOS ONE 9(4):e94521. doi: 10.1371/journal.pone.0094521.

Thomas, J., J. Brunton, and S. Graziosi. 2021. 'Eppi-Reviewer 4.0: Software for Research Synthesis'.

Visentin, Chiara, and Nicola Prodi. 2021. 'How Reliable Are 11- to 13-Year-Olds' Self-Ratings of Effort in Noisy Conditions?' Frontiers in Built Environment 7:108. doi: 10.3389/fbuil.2021.688016.

Wang, Yang, Sophia E. Kramer, Dorothea Wendt, Graham Naylor, Thomas Lunner, and Adriana A. Zekveld. 2018. 'The Pupil Dilation Response During Speech Perception in Dark and Light: The Involvement of the Parasympathetic Nervous System in Listening Effort'. Trends in Hearing 22:2331216518816603. doi: 10.1177/2331216518816603.

Warner-Czyz, Andrea D., Betty A. Loy, Christine Evans, Ashton Wetsel, and Emily A. Tobey. 2015. 'Self-Esteem in Children and Adolescents With Hearing Loss'. Trends in Hearing 19:2331216515572615. doi: 10.1177/2331216515572615.

White, Barbara. 2014. 'The Perceived Stress Scale for Children: A Pilot Study in a Sample of 153 Children'. International Journal of Pediatrics and Child Health 2:45–52. doi: 10.12974/2311-8687.2014.02.02.4.

Zekveld, Adriana A., Johanna A. M. van Scheepen, Niek J. Versfeld, Enno C. I. Veerman, and Sophia E. Kramer. 2019. 'Please Try Harder! The Influence of Hearing Status and Evaluative Feedback during Listening on the Pupil Dilation Response, Saliva-Cortisol and Saliva Alpha-Amylase Levels'. Hearing Research 381:107768. doi: 10.1016/j.heares.2019.07.005.

Figure 3.1 PRISMA diagram illustrating literature search and screening process.

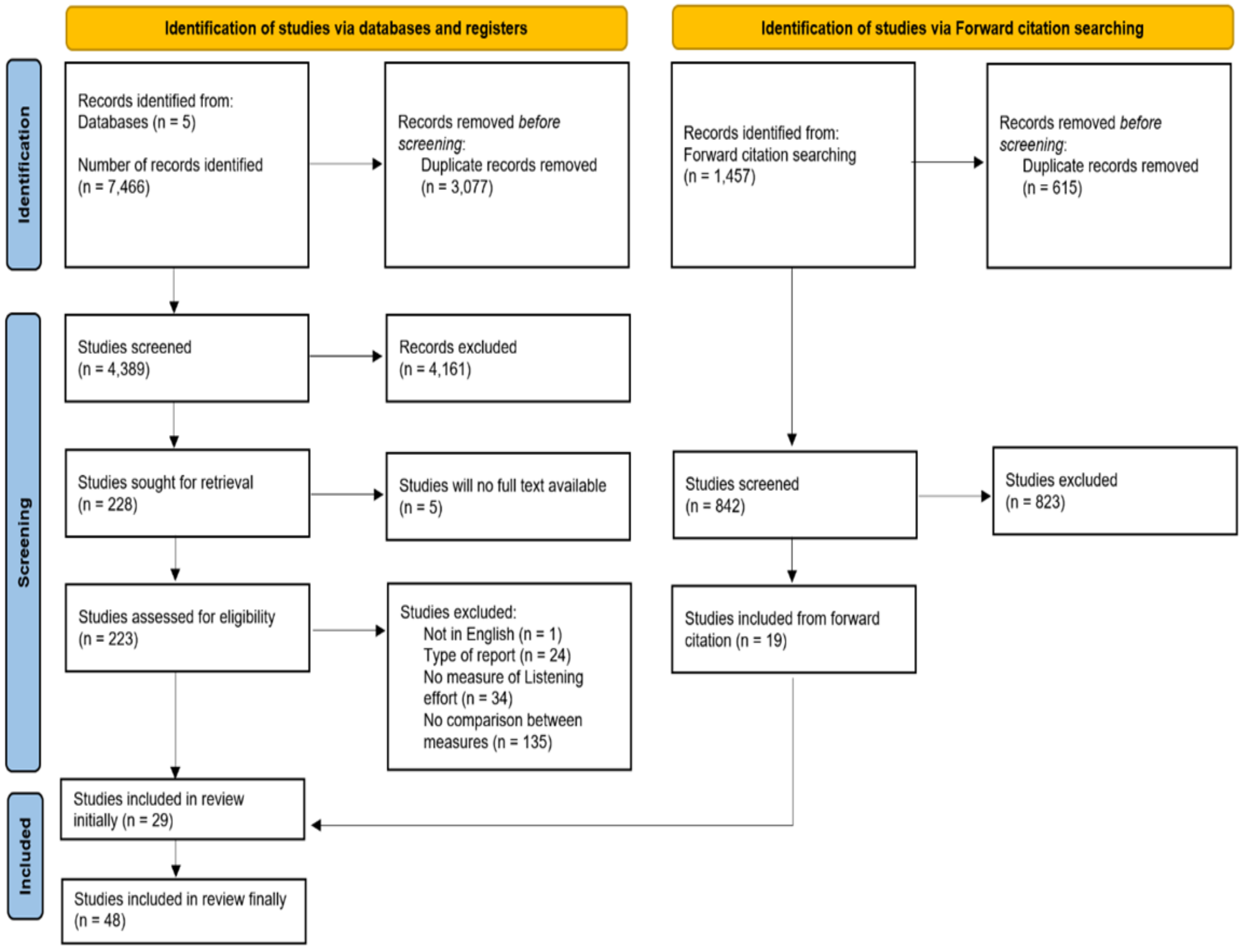


Figure 3.2 Violin plots showing the range of r-values (x-axis) against each subgroup comparison (y-axis) for both the significant (top graph) and non-significant (bottom graph) correlational analyses. Colour is matched to subgroup. (EQ = Effort Questionnaire, B = Behavioural Measures, P = Physiological Measures and FQ = Fatigue Questionnaire).

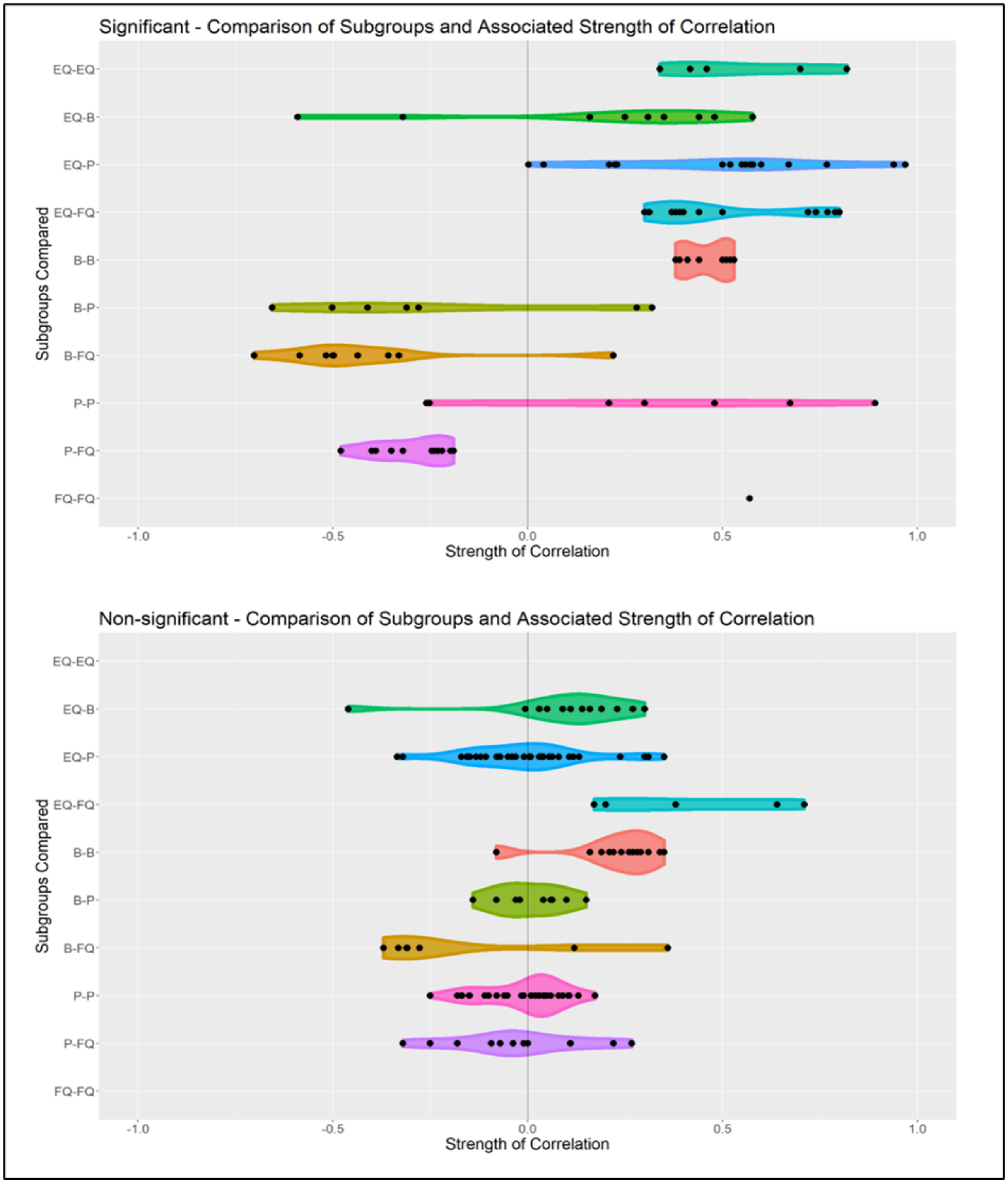


Table 3.1 Showing the number (%) of each classification type against each subgroup comparison. Overlaid with a heat map to illustrate the concentration of results. The model value for each subgroup comparison is indicated by bold and italicised text. (EQ = Effort Questionnaire, B = Behavioural Measures, P = Physiological Measures and FQ = Fatigue Questionnaire).

Strength / Subgroup	Poor (0 to 0.3 or 0 to -0.3)	Fair (0.3 to 0.6 or -0.3 to -0.6)	Moderate (0.6 to 0.8 or -0.6 to -0.8)	Very strong (0.8 to 1 or -0.8 to -1)	Total
EQ-EQ	0	***3 (60.0%)***	1 (20.0%)	1 (20.0%)	5
EQ-B	1 (11.1%)	***8 (88.9%)***	0	0	9
EQ-P	5 (23.8%)	***8 (38.0%)***	4 (19.1%)	4 (19.1%)	21
EQ-FQ	0	***14 (73.7%)***	4 (21.0%)	1 (5.3%)	19
B-B	0	***10 (100%)***	0	0	10
B-P	2 (28.6%)	***4 (57.1%)***	1 (14.3%)	0	7
B-FQ	1 (11.1%)	***7 (77.8%)***	1 (11.1%)	0	9
P-P	***3 (42.8%)***	2 (28.6%)	1 (14.3%)	1 (14.3%)	7
P-FQ	***6 (54.5%)***	5 (45.5%)	0	0	11
FQ-FQ	0	***1 (100%)***	0	0	1

Table 3.2 Summary of the number of significant correlational analyses across each comparison of subgroups.

Subgroup	Total number of trials	Number of significant results	% Of trials reaching significance
EQ - EQ	5	5	100
EQ - B	28	9	32.1
EQ - P	73	21	28.8
EQ - FQ	24	19	79.2
B - B	15	10	66.7
B - P	26	7	26.9
B - FQ	19	9	47.4
P - P	43	7	16.3
P - FQ	30	11	36.7
FQ - FQ	1	1	100

Table 3.3 Summary of risk of bias assessment

Risk of bias tool	*Risk of bias*	*Inconsistency*	*Imprecision*	*Indirectness*	*Publication bias*	*Quality*	*Importance*
GRADE (n=39)	High	Serious	Serious	Moderate	Moderate	Very low	Important
ROBINS-I (n=9)	Serious						

Figure 3.3 Illustration of the effects experimental conditions have when measuring LE. The component of LE captured is represented by the solar rays.

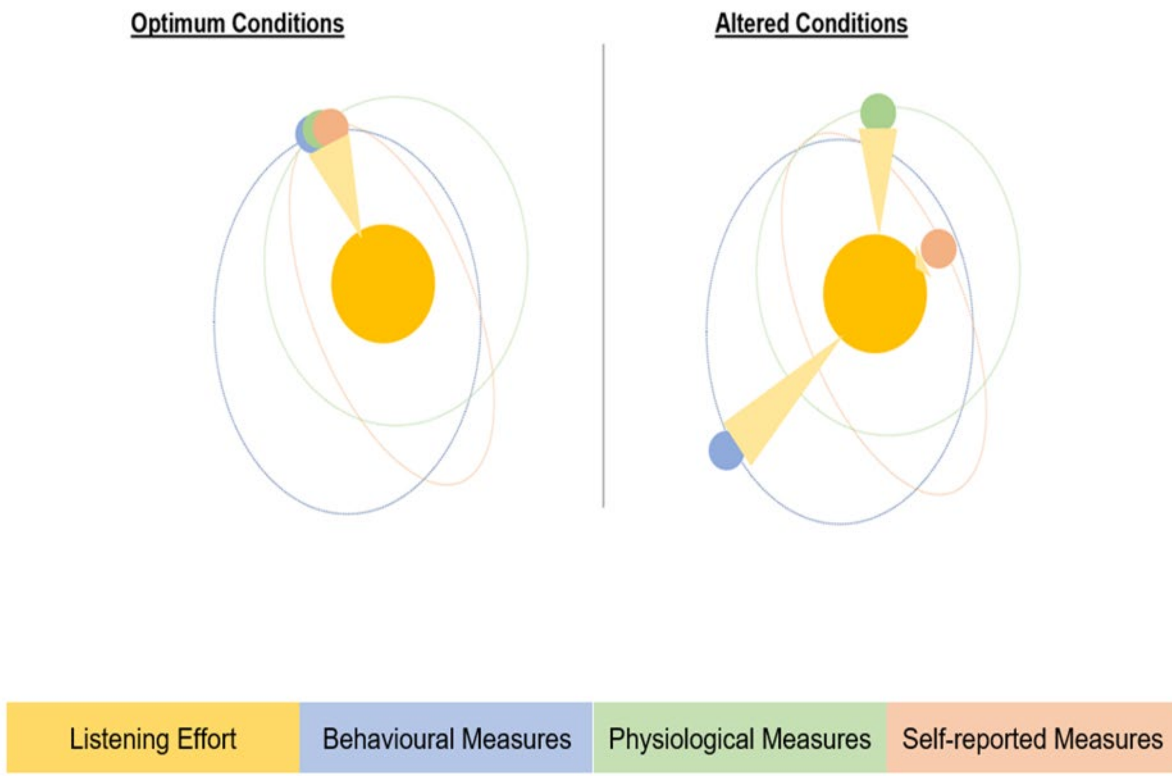


Figure 3.4 Metro map model indicating the time point each tool captures LE. (BVP = Blood volume pulse, EMG = Electromyogram, HRV = Heart rate variability, EEG = Electroencephalogram, FNIRS = Functional near infra-red spectroscopy, f-MRI = Functional magnetic resonance imaging, NASA TLX = National Aeronautics and Space Administration task load index, SSQ = Speech, spatial and qualities questionnaire, VAS – E = Visual analogue scale – effort, SHQ = Spatial hearing questionnaire, CIS = Checklist Individual Strengths, PoMS = Profile of mood states, ITU-T = International telecommunications union – tiredness, VAS – F = Visual analogue scale - fatigue

Appendix

Appendix 3.1: PRISMA checklist

		Reporting Item	Page Number
Title			
Title	#1	Identify the report as a systematic review	1
Abstract			
Abstract	#2	Report an abstract addressing each item in the PRISMA 2020 for Abstracts checklist	2
Introduction			
Background/rationale	#3	Describe the rationale for the review in the context of existing knowledge	3
Objectives	#4	Provide an explicit statement of the objective(s) or question(s) the review addresses	6
Methods			
Eligibility criteria	#5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses	7
Information sources	#6	Specify all databases, registers, websites, organisations, reference lists, and other sources searched or consulted to identify studies. Specify	Supplementary information

		the date when each source was last searched or consulted	
Search strategy	#7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used	6
Selection process	#8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and, if applicable, details of automation tools used in the process	7
Data collection process	#9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and, if applicable, details of automation tools used in the process	8
Data items	#10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (for example, for all measures, time points, analyses), and, if not, the methods used to decide which results to collect	8

Study risk of bias assessment	#11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and, if applicable, details of automation tools used in the process	9
Effect measures	#12	Specify for each outcome the effect measure(s) (such as risk ratio, mean difference) used in the synthesis or presentation of results	10
Synthesis methods	#13a	Describe the processes used to decide which studies were eligible for each synthesis (such as tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5))	NA
Synthesis methods	#13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics or data conversions	10
Synthesis methods	#13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses	NA
Synthesis methods	#13d	Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of	10

		statistical heterogeneity, and software package(s) used	
Synthesis methods	#13e	Describe any methods used to explore possible causes of heterogeneity among study results (such as subgroup analysis, meta-regression)	10
Synthesis methods	#13f	Describe any sensitivity analyses conducted to assess robustness of the synthesised results	NA
Reporting bias assessment	#14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases)	9
Certainty assessment	#15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome	9
Data items	#10b	List and define all other variables for which data were sought (such as participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information	8
Results			
Study selection	#16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram	11

		(http://www.prisma-statement.org/PRISMAStatement/FlowDiagram)	
Study selection	#16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded	11
Study characteristics	#17	Cite each included study and present its characteristics	Supplementary information
Risk of bias in studies	#18	Present assessments of risk of bias for each included study	Supplementary information
Results of individual studies	#19	For all outcomes, present for each study (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (such as confidence/credible interval), ideally using structured tables or plots	Supplementary information
Results of syntheses	#20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies	16
Results of syntheses	#20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (such as confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect	11

Results of syntheses	#20c	Present results of all investigations of possible causes of heterogeneity among study results	NA
Results of syntheses	#20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesised results	NA
Risk of reporting biases in syntheses	#21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed	16
Certainty of evidence	#22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed	16
Discussion			
Results in context	#23a	Provide a general interpretation of the results in the context of other evidence	18
Limitations of included studies	#23b	Discuss any limitations of the evidence included in the review	25
Limitations of the review methods	#23c	Discuss any limitations of the review processes used	25
Implications	#23d	Discuss implications of the results for practice, policy, and future research	22
Other information			

Registration and protocol	#24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered	6
Registration and protocol	#24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared	6
Registration and protocol	#24c	Describe and explain any amendments to information provided at registration or in the protocol	NA
Support	#25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review	26
Competing interests	#26	Declare any competing interests of review authors	NA
Availability of data, code, and other materials	#27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review	Supplementary information

Chapter 4: Listening effort and downstream effects due to hearing loss in children and young people: an online quantitative questionnaire-based observational study

Published Protocol: C. Shields[1], M. Sladen[1], A. Rajai[2], H Guest[3] I. A. Bruce[1], K. Kluk-de Kort[3], and J Nichani[1] (2023). Listening effort and downstream effects due to hearing loss in children and young people: an online quantitative questionnaire-based observational study. BMJ Open. DOI 10.1136/bmjopen-2022-069719

[1] Department of Paediatric Otolaryngology, Royal Manchester Children's Hospital, Manchester Foundation Trust, Manchester, UK

[2] Medical Statistician, Manchester Foundation Trust, Manchester, UK

[3] Division of Human Communication, Development & Hearing, University of Manchester, Manchester, UK

Abstract

Introduction: The clinical application of listening effort (LE) is challenging due to the lack of consensus regarding measuring the concept. Correlational analysis between different measuring instruments often shows conditional and weak relationships, indicating they capture different dimensions of LE. Current research has suggested possible links between LE and downstream consequences such as fatigue, stress, and confidence. One way to clinically measure LE would be to focus on its corollaries. Further research is needed to explore whether the tools used to measure these downstream effects can be applied to capture LE. This study explores using existing questionnaire-based outcome measurement instruments to evaluate LE and its associated consequences in children and young people (CYP), with and without hearing loss.

Methods & Analysis: One hundred children & young people (CYP) aged 12-17 years old with normal hearing and a range of hearing loss levels will be invited to complete a series of online questionnaires (Speech, Spatial and Qualities, Vanderbilt Fatigue Scale – Child, Perceived Stress Scale, and Rosenberg Self-Esteem Scale) and a hearing test (Digits in Noise). They will complete the questionnaires at two time points 1: at the end of a REST day; 2: at the end of a WORK day. Standard demographic and hearing health information will be collected. The sample size was determined pragmatically due to a lack of comparable published data to power the study. Tests are exploratory and for generating hypotheses; therefore, the standard criterium of $p < 0.05$ will be used.

Article summary

Strengths and limitations of this study

Strengths:

• Records data at two time points to allow within-subject comparison.

• Relatively large sample size (n = 100).

• Objectively and subjectively assess for hearing impairment.

• Entirely remote to avoid unnecessary expense/time commitments for the participants.

Limitations:

• Questionnaires only capture subjective elements of their domain – no objective instruments are used.

• Questionnaires administered remotely may be misinterpreted or risk bias from a third party influencing the participant's responses.

Study protocol

Introduction

Listening effort (LE) remains an elusive concept and, as such, is yet to be integrated as an outcome within clinical audiology. Perhaps the most contemporary and inclusive definition of LE concept stems from the Framework of Understanding Effortful Listening (FUEL).

"The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a listening based task" [1].

The importance of deepening our insight into LE can be illustrated by considering the possible unmet burden experienced within hearing-impaired populations. The impact conditions such as mild-moderate and unilateral hearing loss have on individuals has consistently been underreported [2]. However, recent studies have documented the negative educational, professional, and behavioural ramifications of mild hearing loss, especially amongst children and young people (CYP) [3–5]. This observation is particularly concerning given that hearing loss affects around 11.3% of the general school population [3]. This lack of awareness may have contributed to a paradox whereby, in certain circumstances, children with severe/profound hearing loss perform better educationally than those with mild hearing loss [2]. Moreover, a clinically acceptable measure of LE may also allow us to discern the real-life benefit of particular interventions, for example, bilateral cochlear implantation, which may be missed through over-reliance on traditional audiological outcomes [6].

One of the significant challenges to the clinical application of LE is the lack of consensus on measuring the abstract notion. Many instruments have been postulated, ranging from self-reported measures to physiological markers [7]. Nevertheless, despite the myriad of proposed tools, the evidence to support one optimum measure is nominal [1,8].

This issue is further complicated because correlational analysis between different measures often shows conditional and weak relationships supporting the idea that each tool captures a different component of LE [9]. This finding was corroborated in a recent systematic review within our department which is in press with Trends in Hearing.

An additional challenge pertains to our understanding of cognitive processes involved with effort perception. The interplay between working memory, processing speed, attention, motivation, task difficulty, arousal, and cognitive capacity is yet to be fully mapped [1]. Although these theories have been developed since Kahneman first posited effort = attention, these mechanistic deficiencies hinder our ability to explain the inconsistencies noted within the literature [10,11].

With the potential benefit from broadening our application of LE to clinical practice harshly stifled by these barriers, it may be more prudent to consider a different approach to LE. One such way would be to focus on its corollaries rather than the intrinsic concept. The association between fatigue and LE has been well documented in the literature [8,12–14]. However, research remains in its infancy regarding other possible consequences of prolonged exposure to effortful listening. Manifestations in the form of stress, low self-esteem, low motivation, and low mood are yet to be fully explored. If an association can be shown, then there is a potential to utilise pre-existing and pre-validated instruments for these downstream consequences as a clinically informative marker for LE.

Study Aim

This study explores the use of existing questionnaire-based outcome measurement instruments in evaluating LE and its associated consequences in CYP with different hearing abilities.

Primary Objective:

Explore the relationship between LE and measures of fatigue, stress, and confidence in CYP and the impact of hearing loss using existing questionnaires.

Secondary Objective:

Using existing questionnaires, explore the relationship between LE and fatigue, stress, and confidence measures at different time points (working day vs. rest day) in CYP with normal hearing and hearing loss.

Methods and analysis

This is an exploratory study of the relationship between LE and its downstream effects on fatigue, stress, and confidence in a range of CYP with normal hearing and varying degrees of hearing loss; therefore, there will be no initial sampling. This study will take the form of a quantitative questionnaire-based design. The questionnaires will be delivered entirely online using the secure platform REDCap [15,16]. CYP with normal hearing and a range hearing loss level will be recruited (n=min 100). Once 50 responses have been collected, the distribution of hearing loss, age, and gender of the participants will be evaluated. If there is a bias towards either normal hearing or hearing loss or a cluster of ages and gender, efforts will be made to purposively sample to optimise recruitment with an equal distribution of normal hearing and hearing loss, age, and gender.

Participants will be recruited through three avenues:

1. Subjects will be selected from our database (Auditbase) of paediatric patients who have undergone cochlear implant and/or hearing aid fitting at the Royal Manchester Children's Hospital (RMCH). This will be done by a clinical audiologist who is normally involved with the care of this subgroup of patients.

Potential participants will be sent an email invitation to sign up for the study. This will be via a link to the eligibility questionnaire on REDCap. This will be the source for the hearing-impaired participants.

2. An invitation link will also be sent to hearing charities/social media groups for them to share. This will link to the eligibility questionnaire on REDCap. This will rely on the participants self-referring to the study as they will not be directly contacted via this avenue.

3. Recruitment will also take place in the Fracture clinic within the outpatient department (OP) of RMCH with posters. Fracture clinic has been specifically chosen as this population is less likely to have chronic health conditions, which may confound the questionnaire results. The participants will self-refer by following the link on the posters to the eligibility questionnaire on REDCap. Periodically, study team members will be in the waiting room of the fracture clinic to highlight the chance to participate in the study to patients before they are seen in the clinic. We will not use any database to target patients within this cohort specifically.

Once an individual has declared interest in participating in the study, they will follow a link to the eligibility questionnaire on REDCap. This process will ensure they meet the criteria set out in table 1. If they are not eligible for the study, REDCap will inform them, and no details from their questionnaire will be recorded.

The relevant participant information sheet (PIS) will be available via the eligibility link before confirming interest. Those eligible for the study who register their interest via the initial link will then be emailed a copy of the PIS immediately after registration. The participant will then receive a link to the consent process.

Following eligibility screening and consent gathering, participants will be invited to complete an online battery of questionnaires. This battery consists of four separate sessions, with a total completion time of approximately 1 hour 30 minutes over two weeks. The sessions are outlined below:

1. Participant demographic information: Basic information relating to age, sex, and hearing device use will be collected (see table 2 for the full list of questions). Perceived hearing ability will also be measured using the Speech, Spatial, and Qualities questionnaire. First created by Gatehouse and noble and then later validated in children 11 years and over by Galvin and noble [17,18]. This session will take approximately 10 minutes to complete.

2. Digits in Noise (DIN): Next, participants will be asked to complete a DIN hearing test (see below for full details). This test will be done using the University of Manchester Software MOSS. The DIN may be done on the same day as the demographic questionnaire or on a different day if the participant wishes. The DIN will allow us to establish an objective spectrum of hearing abilities across the participants. The test will take approximately 20 minutes to complete.

3. REST day questionnaire session: The participant will be asked to complete a series of four validated questionnaires (see below for full details). These questionnaires aim to measure LE, fatigue, stress, and confidence. This session will occur at the end of a REST day when the participant has not been to school/college/work. This session will take approximately 30 minutes to complete.

4. WORK day questionnaire session: Finally, the participant will be asked to complete the above outcome questionnaires again. This session will occur at the end of a day when the participant has been in school/college/work. This session will take approximately 30 minutes to complete.

Audiological Assessment

Each participant will complete an online hearing test, Digits in Noise test (DIN), with instructions on how to complete it. The test will be completed at home in quiet using the participant's device (must be a computer or laptop and not a phone or tablet, due to software compatibility) and head/earphones. If a participant uses a hearing device such as a hearing aid or cochlear implant, they should remove this prior to completing the test. The DIN test will take 10-20 minutes.

The DIN is an online speech in noise test. The participant is presented with a series of 3 randomised digits within a steady-state speech-shaped background noise. The participant must select (from a number pad within the platform) the digits they heard. They are then presented with further digit triplets until the end of the test. They are given a decibel signal-to-noise ratio (dB SNR), defined as the difference between target digits and background noise, in dB, that is required to correctly recognise 50% of the digit-triplets (SNR50). This measure was chosen to provide a relatively quick and better approximate hearing function in a "real-life" environment that can be completed at home.

Listening Effort, Fatigue, Confidence, and Stress, Questionnaires

For sessions 3 and 4, each participant will be presented with four questionnaires, each measuring a different domain (see below). Parents/guardians can ask the participant the items provided in the questionnaire and encourage the participant to answer the item as detailed as possible, or the participant can complete it by themself. The questionnaires will take approximately 30 minutes to complete.

The participant will be asked to complete the questionnaires twice. The two time points are defined as part 1: the evening of a rest day (rest day is defined as not attending school/college/work; part 2: the evening of a workday (workday is defined as attending school/college/work). Each standardised questionnaire will be checked and validated by a research team member. If a participant does not complete a questionnaire session fully and there is missing data, we will send an email requesting the complete the entire session again.

The questionnaires are as follows:

Listening Effort: Speech, Spatial and Qualities of Hearing Scale for children (SSQ) Section C Qualities of Hearing. This questionnaire was first created by Gatehouse and Noble, then later validated in children 11 years and over by Galvin and Noble [17,18]. Several studies have used a subscale of the qualities section of the SSQ, which contains three items on a sliding scale, to subjectively measure

LE [13,19]. The sliding scale corresponds to a numerical value (1-10). The sum of each scale is determined, with lower scores representing higher effort levels.

Fatigue: Vanderbilt Fatigue Scale – Child (VFS-C). First created by Hornsby et al. for use in adults and recently validated for use in children aged 6-17 years [20,21]. VFS is specifically designed to measure fatigue related to hearing. VFS-C is a 10-item, five-point Likert scale with numerical weighting (0-4). The sum of the ten items equates to the level of fatigue, with higher scores representing higher fatigue levels.

Stress: Perceived Stress Scale (PSS). First created by Cohen and later shown to apply to children aged 5-18 by White [22,23]. The PSS is a 10-item, five-point Likert scale with a numerical weighting (0-4). The sum of the ten items equates to the stress level, with higher scores representing higher stress levels. Some questions are reversed scored.

Confidence: Rosenberg Self-Esteem Scale (RSES). First created by Rosenberg and has shown to apply to children aged 12-17 by Bagley & Mallick [24,25]. The RSES is a 10-item, four-point Likert scale with a numerical weighting (0-3). The sum of the ten items equates to the confidence level, with higher scores representing higher confidence levels. Some questions are reversed scored.

A flow diagram outlining the process a participant takes through this study is represented in figure 1.

Statistical Analysis Plan

Without previous information to power the study, the sample size was determined pragmatically. Tests are exploratory and for hypothesis-generating purposes - the standard criterium of $p < 0.05$ highlights areas with potential significance.

An appropriate descriptive analysis of participants' demographics and baselines will be provided. The number (%) of participants who completed each part of the tests will be reported. Depending on the distribution, numeric outcome measures will be summarised using mean/median (SD/IQR) for each test condition. Binary outcomes will be summarized by count (%).

Listening effort will be correlated to other outcome measures using appropriate methods depending on the distribution of outcome measures. A sample of 100 will have 80% power to detect a Pearson correlation of 0.3 or larger. Univariable regression models may be used to explore the effect of different factors such as day of assessment, age, and gender on the correlations. The relationship between outcome measures with demographics will be explored.

Missing data: If there is missing data or clarification is needed the participant will be contacted to try and ascertain missing data. If they are not able to do this, the quantity of missing data on each variable per individual will be assessed. For individuals with less than 20% missing items per questionnaire, the missing data will be predicted using an appropriate multiple imputation method, such as delta adjusted. This has been chosen to reduce the risk of selection bias associated with complete case sampling (de Goeij et al., 2013; Mackinnon, 2010). A complete case analysis will be conducted alongside multiple imputation.

The data generated by the study will be analysed at Manchester University NHS Foundation Trust, and the analysis will be performed by researchers employed by Manchester University NHS Foundation Trust.

Data Statement

Data collected as part of this study will be pseudonymised and labelled with a unique personal study participant code rather than names. The study ID is assigned to the participant when they sign the ICF. The contact email address and study ID are downloaded from REDCap and stored in a password-protected file within the MFT secure server drive.

There is also an intention to share an anonymous final data set with external researchers; however, this will only occur if the correct data-sharing agreement is in place.

Ethics and dissemination

This study has been reviewed within the funding organisation (Cochlear Research and Development Limited) by an independent and relevant peer reviewer/committee. Favourable comments and approval of the protocol were given during the peer review process. A favourable opinion has also been granted by the Health Research Authority (HRA) and NHS Research Ethics Committee (REC) for the study and all the supporting documents, including this protocol, information sheets, informed consent forms, and other relevant documents (22/WS/0134)

Study results will be disseminated via the Cochlear Implanted Children's Support Group (CICS), e.g., newsletters, and social media, as deemed appropriate by the charity and mailing lists. A summary of this study's results will be provided to all participants who have consented to be contacted. Aim for publication of the final study results in medical literature and presentation at medical conferences.

References

1 Pichora-Fuller MK, Kramer SE, Eckert MA, et al. Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). Ear Hear 2016;37 Suppl 1:5S-27S. doi:10.1097/AUD.0000000000000312

2 Archbold S, Ng ZY, Harrigan S, et al. Views of parents and teachers. 2015;:48.

3 Bess F, Dodd-Murphey J, Parker R. Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status - PubMed. 1998.https://pubmed.ncbi.nlm.nih.gov/9796643/ (accessed 15 Dec 2021).

4 Cone BK, Wake M, Tobin S, et al. Slight-mild sensorineural hearing loss in children: audiometric, clinical, and risk factor profiles. Ear Hear 2010;31:202–12. doi:10.1097/AUD.0b013e3181c62263

5 Khairi Md Daud M, Noor RM, Rahman NA, et al. The effect of mild hearing loss on academic performance in primary school children. Int J Pediatr Otorhinolaryngol 2010;74:67–70. doi:10.1016/j.ijporl.2009.10.013

6 Lammers MJW, van der Heijden GJMG, Pourier VEC, et al. Bilateral cochlear implantation in children: a systematic review and best-evidence synthesis. The Laryngoscope 2014;124:1694–9. doi:10.1002/lary.24582

7 Shields C, Willis H, Nichani J, et al. Listening effort: WHAT is it, HOW is it measured and WHY is it important? Cochlear Implants Int 2021;0:1–4. doi:10.1080/14670100.2021.1992941

8 McGarrigle R, Munro KJ, Dawes P, et al. Listening effort and fatigue: what exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group "white paper." Int J Audiol 2014;53:433–40. doi:10.3109/14992027.2014.890296

9 Alhanbali S, Dawes P, Millman RE, et al. Measures of Listening Effort Are Multidimensional. Ear Hear 2019;40:1084–97. doi:10.1097/AUD.0000000000000697

10 Bruya B, Tang Y-Y. Is Attention Really Effort? Revisiting Daniel Kahneman's Influential 1973 Book Attention and Effort. Front Psychol 2018;9:1133. doi:10.3389/fpsyg.2018.01133

11 Kahneman D. Attention and Effort. 1st ed. Prentice-Hall 1973.

12 Bess FH, Hornsby BWY. Commentary: listening can be exhausting--fatigue in children and adults with hearing loss. Ear Hear 2014;35:592–9. doi:10.1097/AUD.0000000000000099

13 Alhanbali S, Dawes P, Lloyd S, et al. Self-Reported Listening-Related Effort and Fatigue in Hearing-Impaired Adults. Ear Hear 2017;38:e39. doi:10.1097/AUD.0000000000000361

14 Athey H. "The effect of auditory fatigue on reaction time in normal hearing list" by Haley Athey. 2021.https://commons.lib.jmu.edu/diss201019/108/ (accessed 19 Dec 2021).

15 Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81. doi:10.1016/j.jbi.2008.08.010

16 Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019;95:103208. doi:10.1016/j.jbi.2019.103208

17 Galvin KL, Noble W. Adaptation of the speech, spatial, and qualities of hearing scale for use with children, parents, and teachers. Cochlear Implants Int 2013;14:135–41. doi:10.1179/1754762812Y.0000000014

18 Gatehouse S, Noble W. The Speech, Spatial and Qualities of Hearing Scale (SSQ). Int J Audiol 2004;43:85–99.

19 Dawes P, Munro KJ, Kalluri S, et al. Acclimatization to Hearing Aids. Ear Hear 2014;35:203–12. doi:10.1097/AUD.0b013e3182a8eda4

20 Hornsby BWY, Camarata S, Cho S-J, et al. Development and Validation of the Vanderbilt Fatigue Scale for Adults (VFS-A). Psychol Assess 2021;33:777–88. doi:10.1037/pas0001021

21 Hornsby BWY, Camarata S, Cho S-J, et al. Development and Evaluation of Pediatric Versions of the Vanderbilt Fatigue Scale for Children With Hearing Loss. J Speech Lang Hear Res 2022;65:2343–63. doi:10.1044/2022_JSLHR-22-00051

22 Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav 1983;24:385–96. doi:10.2307/2136404

23 White B. The Perceived Stress Scale for Children: A Pilot Study in a Sample of 153 Children. Int J Pediatr Child Health 2014;2:45–52. doi:10.12974/2311-8687.2014.02.02.4

24 Bagley C, Mallick K. Normative Data and Mental Health Construct Validity for the Rosenberg Self-Esteem Scale in British Adolescents. Int J Adolesc Youth 2001;9:117–26. doi:10.1080/02673843.2001.9747871

25 Rosenberg M. Society and the adolescent child. Princeton, NJ 1965. https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft01038-000 (accessed 26 Jan 2022).

Authors' statement

IAB, JN, KDK, CS, AR, and MS conceived the study concept and design. MS and CS will conduct screening and data collection. AR will perform the analysis. IAB, JN, CS, and MS prepared the first draft of the manuscript. All authors provided edits and critiqued the manuscript for intellectual content.

Funding

This research is funded by [Cochlear Research and Development Limited] grant number [IIR-2186] and supported by the NIHR Manchester Biomedical Research Centre.

Conflict of interest

None declared

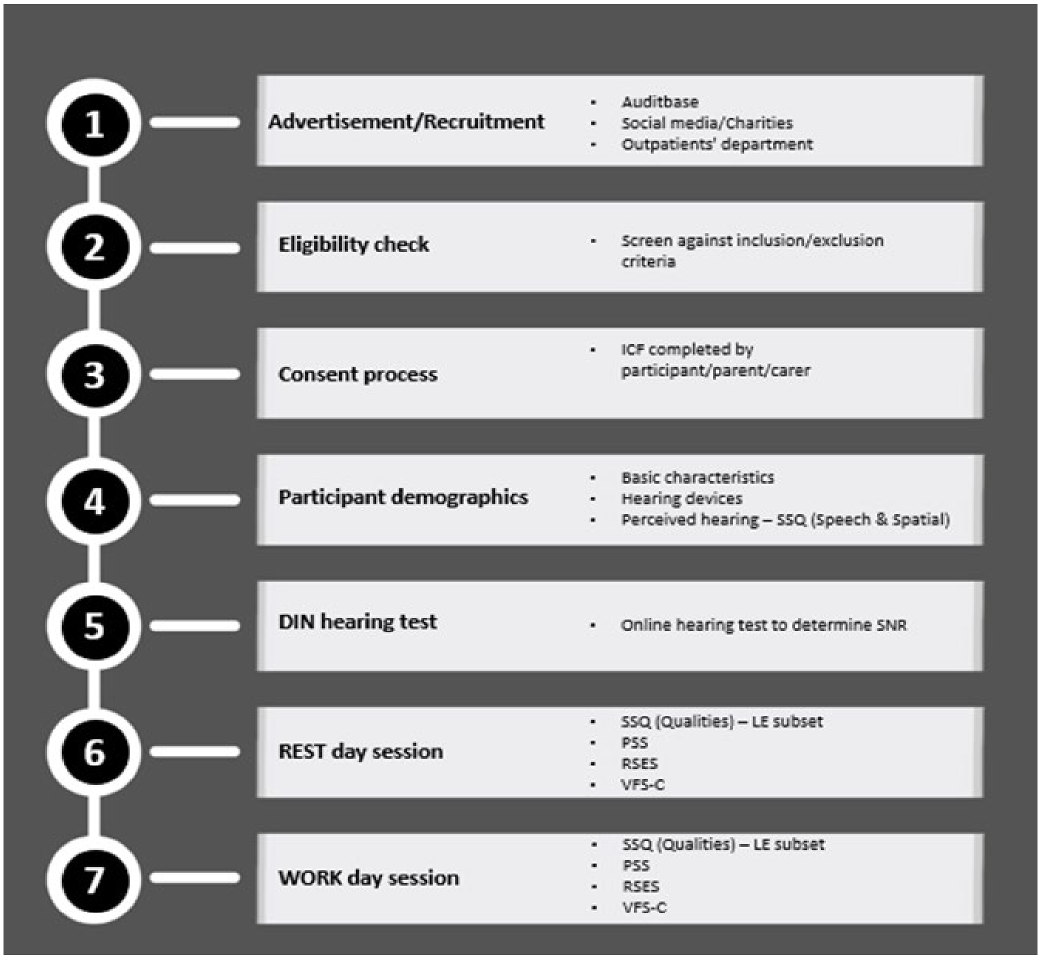
Table 4.1 Outline of the inclusion and exclusion screening criteria

Inclusion criteria	• Aged 12 to 17 years old. • Normal hearing or hearing loss of any aetiology. • No other co-morbidity or chronic health condition associated with fatigue. • No dual sensory impairment (uncorrected vision loss). • Adequate English language understanding. • Access to the internet.
Exclusion criteria	• Aged younger than 12 years old or older than 17 years old. • Presence of any other co-morbidity or chronic health condition associated with fatigue. • Dual sensory impairment. • Inadequate English language understanding. • No access to the internet.

Table 4.2 Demographic Information Form

Initial question	Follow-up question(s)
Age of participant	NA
Gender of participant	NA
Does the participant use a hearing device	1. What device do they wear 2. Which side do they wear the device 3. How long have they used the device
Does the participant use any other hearing equipment	1. What equipment do they use
Has the participant been diagnosed with any medical conditions	1. Please indicate the conditions
SSQ sections A (Speech) and B (Spatial), 22 items to assess for perceived hearing ability	NA

Figure 4.1 Timeline of participant's journey through this study



<u>Appendix</u>

Appendix 4.1: Outline of audiological assessment

Please find the details regarding the DIN testing procedure:

Stimulus

- Structure: A carrier phrase ("The digits...") followed by three digits
- Timing: The digits are separated by silent gaps of 180-250 ms duration (varied randomly)
- Included digits: 0-9 are included
- Digit exemplars: 6 exemplars of each digit are included
- Talker: Female British
- Digit selection: Digits are selected at random, with the constraint that a digit cannot be repeated within a trial
- Exemplar selection: Exemplars are selected at random
- Masker: Speech-spectrum-shaped Gaussian noise
- Frequency content: Spans 120 to 8000 Hz (so that the upper-frequency limit of listeners' headphones/earphones does not introduce unwanted variability)

Presentation level

- Approach: Total stimulus level (target + masker) is held constant throughout the experiment
- The rationale for the above approach: Ensures that stimuli do not become uncomfortably loud and reduce the risk of the target falling below the threshold of audibility

- Calibration method: At the start of the experiment, the listener is presented with a "loud" and a "quiet" calibration phrase, separated in RMS by 25 dB. They adjust their volume control until the "quiet" phrase is clear, and the "loud" phrase is not uncomfortably loud.
- Stimulus presentation level: 5 dB below the level of the "loud" calibration phrase

Basic adaptive procedure

- Scoring criterion: 2/3 or 3/3 digits must be entered correctly for a trial to count as correct
- Stepping rule: 2-down 1-up (i.e., two correct trials in a row causes a step down in SNR, one incorrect trial causes a step up in SNR)
- Starting SNR: 2 dB
- Block 1 (practice)
- Reversals: 2
- Step size: 6 dB
- Block 2 (real)
- Number of phases: 2 ("initial" and "measurement")
- Reversals in the initial phase: 2
- Step size in the initial phase: 6 dB
- Reversals in measurement phase: 6
- Step size in measurement phase: 2 dB
- Threshold calculation: Mean of SNRs at final 6 reversals

Listener feedback

- Post-trial feedback: Feedback displayed on the correctness of the response to each trial
- Difficulty level: Information displayed on current and past "difficulty level" (a linear transform of SNR) to preserve interest and motivation
- End-of-block feedback: Feedback displayed on the lowest SNR at which the listener scored 2/3 or 3/3 correct

Chapter 5: Discussion

This thesis aims to determine whether an optimum measure for LE exists and can be incorporated into the routine clinical practice of children with hearing loss. It was hypothesised that no single measure can capture LE in its entirety. This aim has been investigated through a systematic review of the literature and the design of exploratory research in the form of an online quantitative questionnaire study. These papers have been supported by an editorial letter to outline the current challenges associated with measuring LE and a protocol study to ensure a high level of methodological rigor was applied to the normative research element of this thesis.

Together these papers support the hypothesis stated above by providing further evidence for the multi-dimensional model of LE. This was achieved by demonstrating that the correlations between measures of LE are most often fair and non-significant across the literature. This thesis has developed this model by elaborating on the effects of experimental conditions on the strength of association between two instruments. As a result, this thesis has postulated an alternative approach to incorporating LE into routine clinical practice, focusing on the downstream consequences of prolonged exposure to effortful listening rather than the construct of LE itself. This approach will be investigated and the protocol published in the previous chapter provides an outline for an exploratory study to achieve this purpose.

Limitations of evidence

The evidence base for instruments to measure LE is fraught with issues that limit the interpretation of findings from the collective literature. This section will discuss some fundamental limitations to help guide future work.

The sample size of studies that examine LE is usually small, especially when involving children. This is demonstrated by the results of chapter 3, where just 34 out of 1900 total participants represented children with a hearing loss. An incongruous notion, given that evidence suggests this population may be most acutely affected by the burden of LE (Archbold et al. 2015; Downs and Crum 1978; Hsu et al. 2017). Furthermore, 30 out of the 48 studies did not perform sample size calculations, a finding that is echoed across medical research (Vasileiou et al. 2018). Consistently underpowered sample sizes may have affected the ability of some studies to demonstrate significant correlational analyses between two measures of LE and hence skewed the findings of the systematic review (Faber and Fonseca 2014). One possible way to remedy this situation is for studies to undertake purposive sampling techniques during participant recruitment (Etikan 2016). Although all non-probability sampling methods are also subject to bias, homogeneous sampling for hearing-impaired participants may improve the overall quality of evidence concerning LE (Andrade 2021; Etikan 2016).

Another issue identified from the literature is that most papers only measure LE at a single time point. This observation was reflected in the results of chapter 3, where only two out of the 48 included papers contained a follow-up measurement. By only recording LE cross-sectionally, it assumes that the construct is static when it is more likely to be a dynamic process that changes over time (Peelle 2018). Moreover, it limits the ability of instruments to demonstrate test-retest reliability, a critical stage in validating an outcome measure (Matheson 2019; Hornsby et al. 2022; Hornsby et al. 2021). Therefore, longitudinal measures of LE are a crucial next step for future research.

Studies pertaining to LE often lack ecological validity in that very few take place in a "real-world" setting. This is particularly problematic given that the correlations between outcome measures have been shown to depend on the study's experimental design – see chapter 3. This creates uncertainty regarding the ability of these measures to capture LE. Is it due to an inherent flaw of the outcome measure or because the environment, stimuli, or test did not evoke a meaningful experience of effort? (Schmuckler 2001) Moving research out of the laboratory and into "real-world" settings such as the workplace/classroom is a valuable step in developing the evidence base of LE.

Outside of limitations in study design, our understanding of the cognitive processes involved within the effort perception pathway also hinders the clinical use of LE as an outcome. Even the notion of 'effort' remains elusive in terms of an agreed definition (Steele 2020). Although the subsequent iterations of Kahneman's model of effort and attention have helped to conceptualise the topic, the influence of factors such as motivation, arousal, fatigue, working memory, speed of processing, and perception remain to be fully charted (Kahneman 1973; Pichora-Fuller et al. 2016). This deficiency in knowledge prevents us from being able to explain certain inconsistencies in studies and thus serves as another barrier to the clinical application of LE.

Limitations of the current thesis

Casting a critical eye to the methodology of the papers within the thesis raises important points for discussion. These issues will now be explored to help guide future work.

Within chapter 3, the systematic review placed a date limit of 2000 during the literature search. The justification was to temper the search's sensitivity against maintaining efficiency and reproducibility (Unger 2019). Additionally, because "listening effort" is a relatively modern term, it was thought that very few, if any, relevant papers would be found by widening the search dates (Cooper et al. 2018; Lefebvre et al. 2021). Upon reflection, it is possible that some articles which did not specifically reference "listening effort" but contained insightful findings were missed. Subsequently, the review did identify two papers published before 2000, which may have proved relevant (Feuerstein 1992; Kramer et al. 1997). Although these two papers would have been excluded based upon other factors, it cautions against applying definitive criteria such as date limits. This stance is in line with Cochrane guidance and therefore represents best practice when conducting systematic reviews (Lefebvre et al. 2021).

The explorative research proposed within this thesis – chapters 4 – utilised questionnaires as the sole archetype of outcome measure. The rationale behind this decision stems from the fact that these papers aimed to identify correlates between LE and several downstream effects. This research was exploratory by nature, so the use of numerous invasive/resource-heavy measures such as hormonal or electrophysiological instruments was deemed unwarranted (Hinkin 1998). Notwithstanding this logic, many issues are inherent with questionnaires and introduce bias which will now be discussed (Choi and Pak 2004).

Questionnaire fatigue: The questionnaire sessions proposed within chapter 4 will generally take 20-30 minutes for the participant to complete. This is towards the upper limit of what is considered acceptable for self-administered tools (Hartge and Cahill 1998). By exceeding this time period, there is a chance that participants will become disinterested and began to answer questions indifferently (Choi and Pak 2004; Aday and Cornelius 2006). Perhaps, a more concerning consequence related to lengthy questionnaires is that it may induce a level of fatigue in the participant, which could unduly skew the results of the fatigue questionnaire (Choi and Pak 2004).

Proxy responses: The delivery of the questionnaires proposed in chapter 4 is entirely remote. This may led to parents/caregivers aiding participants with their responses (Choi and Pak 2004). This has obvious implications for the results of the questionnaires and may be mitigated in future work by conducting virtual/telephone/in-person sessions.

Selection bias: One of the inclusion criteria set forward in chapter 4 was that the participant must have access to an internet-enabled device to complete the questionnaires. This may prevent those from lower socioeconomic backgrounds from participating in the study and thus decreased the diversity of the study population (Choi and Pak 2004). Using postal questionnaires alongside online versions may help eliminate this barrier and should be considered in future work (Menon and Muraleedharan 2020).

A final issue associated with using questionnaires as an outcome measure is their subjective nature. If Kahneman's theory regarding effort and attention still holds, then effort must be viewed as an objective construct and, therefore, cannot be captured using tools such as questionnaires (Kahneman 1973). Empirical research by Robinson and Morsella would seemingly refute Kahneman's claim regarding the objectivity of effort (Robinson and Morsella 2014). Here it was shown that perceived ratings of effort demonstrate significant linear associations with task difficulty indicating the congruency between subjective and objective effort (Robinson and Morsella 2014). Bruya and Tang have criticised Kahneman's stance, remarking that the topic was not adequately covered in his book "Attention and Effort" and therefore requires further study before disregarding the role of self-reported feelings in effort pathways (Bruya and Tang 2018). Given this uncertainty, a pragmatic solution may be to use a mix of subjective and objective measures of effort and its associated consequences in further research.

The original intention was to include the results of the exploratory study within this thesis. Unfortunately, due to several delays with study sponsorship and therefore subsequent ethical applications this was not possible. Every attempt was made to try and mitigate for this delay but the decision to postpone the start date of recruitment was made to prioritize the quality of the research generated from the study above all else. At the time of writing of this thesis, the study is in an active recruitment phase and has received ethical approval for NHS and non-NHS based participants.

Planning future research

Alongside the suggestions made in previous sections of this chapter, this thesis recommends several areas that should be investigated further in subsequent research.

The results of chapter 4 set out a way to investigate the correlations between effort and fatigue/stress/confidence as measured using self-reported instruments. The logical next step would be to therefore conduct this study. Indeed, with ethical approval now in place for both NHS and non-NHS domains, recruitment for this study is active at the time of writing of this thesis. Through this study the validity of using proxy measures of LE will be assessed. Thereafter, if correlations are established, a useful next step would be to conduct a before and after study using a particular hearing device as the intervention. This will allow for within-subject comparisons to be performed and would identify how LE and the associated downstream consequences change following auditory rehabilitation. If LE and fatigue/stress/confidence are related, then it is expected that the correlation coefficient would remain stable while the raw scores for each domain would decrease following intervention.

To address the issue of subjective versus objective effort, it is prudent to consider potential objective measures of LE, fatigue, stress, and confidence when planning future research. An instrument that shows promise in this area is automated facial expression recognition. This tool uses a camera to detect facial expressions during a particular task (Dinges et al. 2005). Machine learning algorithms then analyses the pattern of expressions to identify the emotional state of the participant (Dinges et al. 2005). The benefit of this measure is that it would be suitable to capture a wide range of downstream consequences. This subverts the need for multiple instruments to capture each domain of interest. Interestingly, studies have already shown the ability of automated facial expression recognition to capture fatigue, stress, and effort (Gao et al. 2014; Kawamura et al. 2015; Dinges et al. 2005; Venkitakrishnan et al. 2020). Moreover, this tool has been effective in both adult and child populations and those with a hearing impairment (Venkitakrishnan et al. 2020; Littlewort et al. 2011). With this considered, facial expression recognition constitutes an exciting and promising area of research that overcomes several limitations presented within this thesis.

References

Aday, L. and Cornelius, L. (2006). *Designing and Conducting Health Surveys: A Comprehensive Guide*. 3rd ed. USA: Jossey-Bass. [online]. Available from: https://books.google.co.uk/books?hl=en&lr=&id=Xr-OCQAAQBAJ&oi=fnd&pg=PA15&ots=VquN3vIYIF&sig=OKkI_bHaoiNVwlLFsedMDwO8ij0&redir_esc=y#v=onepage&q&f=false [Accessed July 13, 2022].

Andrade, C. (2021). The Inconvenient Truth About Convenience and Purposive Samples. *Indian Journal of Psychological Medicine*, 43(1), pp.86–88.

Archbold, S. et al. (2015). Views of parents and teachers. , p.48.

Bruya, B. and Tang, Y.-Y. (2018). Is Attention Really Effort? Revisiting Daniel Kahneman's Influential 1973 Book Attention and Effort. *Frontiers in Psychology*, 9, p.1133.

Choi, B.C.K. and Pak, A.W.P. (2004). A Catalog of Biases in Questionnaires. *Preventing Chronic Disease*, 2(1), p.A13.

Cooper, C. et al. (2018). Revisiting the need for a literature search narrative: A brief methodological note. *Research Synthesis Methods*, 9(3), pp.361–365.

Dinges, D.F. et al. (2005). Optical computer recognition of facial expressions associated with stress induced by performance demands. *Aviation, Space, and Environmental Medicine*, 76(6 Suppl), pp.B172-182.

Downs, D.W. and Crum, M.A. (1978). Processing demands during auditory learning under degraded listening conditions. *Journal of Speech and Hearing Research*, 21(4), pp.702–714.

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), p.1.

Faber, J. and Fonseca, L.M. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics*, 19(4), pp.27–29.

Feuerstein, J.F. (1992). Monaural versus binaural hearing: ease of listening, word recognition, and attentional effort. *Ear and Hearing*, 13(2), pp.80–86.

Gao, H., Yüce, A. and Thiran, J.-P. (2014). Detecting emotional stress from facial expressions for driving safety. *2014 IEEE International Conference on Image Processing (ICIP)*, pp.5961–5965.

Hartge, P. and Cahill, J. (1998). Field methods in epidemiology. In *Modern Epidemiology*. Philadelphia: Lippincott-Raven, pp. 67–68.

Hinkin, T.R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1(1), pp.104–121.

Hornsby, B.W.Y. et al. (2022). Development and Evaluation of Pediatric Versions of the Vanderbilt Fatigue Scale for Children With Hearing Loss. *Journal of Speech, Language, and Hearing Research*, 65(6), pp.2343–2363.

Hornsby, B.W.Y. et al. (2021). Development and Validation of the Vanderbilt Fatigue Scale for Adults (VFS-A). *Psychological assessment*, 33(8), pp.777–788.

Hsu, B.C.-L., Vanpoucke, F. and van Wieringen, A. (2017). Listening Effort Through Depth of Processing in School-Age Children. *Ear and Hearing*, 38(5), pp.568–576.

Kahneman, D. (1973). *Attention and Effort*. 1st ed. Prentice-Hall.

Kawamura, R., Takemura, N. and Sato, K. (2015). Mental fatigue estimation based on facial expressions during speech. *2015 IEEE/SICE International Symposium on System Integration (SII)*, pp.223–228.

Kramer, S.E. et al. (1997). Assessing Aspects of Auditory Handicap by Means of Pupil Dilatation. *Audiology*, 36(3), pp.155–164.

Lefebvre, C. et al. (2021). Chapter 4: Searching for and selecting studies. [online]. Available from: https://training.cochrane.org/handbook/current/chapter-04 [Accessed November 22, 2021].

Littlewort, G.C. et al. (2011). Automated measurement of children's facial expressions during problem solving tasks. *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, pp.30–35.

Matheson, G.J. (2019). We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *PeerJ*, 7, p.e6918.

Menon, V. and Muraleedharan, A. (2020). Internet-based surveys: relevance, methodological considerations and troubleshooting strategies. *General Psychiatry*, 33(5), p.e100264.

Peelle, J.E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing*, 39(2), pp.204–214.

Pichora-Fuller, M.K. et al. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37 Suppl 1, pp.5S-27S.

Robinson, M.M. and Morsella, E. (2014). The subjective effort of everyday mental tasks: Attending, assessing, and choosing. *Motivation and Emotion*, 38(6), pp.832–843.

Schmuckler, M.A. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, 2(4), pp.419–436.

Steele, J. (2020). What is (perception of) effort? Objective and subjective effort during task performance. *PsyArXiv*. [online]. Available from: https://pure.solent.ac.uk/en/publications/what-is-perception-of-effort-objective-and-subjective-effort-duri [Accessed July 12, 2022].

Unger, J.M. (2019). Reproducible Findings in Systematic Reviews and Meta-analyses in Oncology: Verify, Then Trust. *JAMA Oncology*, 5(11), pp.1545–1546.

Vasileiou, K. et al. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, 18(1), p.148.

Venkitakrishnan, S., Giullani, N. and Yu-Hsiang, W. (2020). Can Facial Expressions Measure Listening Difficulty? [online]. Available from: https://haar.lab.uiowa.edu/sites/haar.lab.uiowa.edu/files/wysiwyg_uploads/venkitakrishnan_giuliani_and_wu_2020_aaa.pdf.

Chapter 7: Conclusion

This thesis aimed to investigate the existence of an optimum measure for listening effort with the ultimate goal of incorporation into routine clinical practice. The narrative from the literature thus far suggests that due to the complexity of the concept one single measure may not be versatile enough to capture all aspects of LE. This notion created the idea of the multidimensional model of listening effort and through a systematic review, this thesis has provided evidence for this model. Furthermore, this thesis has developed this model by elaborating on the effects of experimental conditions on the strength of association between two instruments. In conjunction, these finding highlight the futility of advocating for one sole measure and provides justification for exploring alternative options. Following this rationale, this thesis has postulated an alternative approach to incorporating LE into routine clinical practice, focusing on the long-term effects of continuous effortful listening rather than the LE concept itself. Although using this method has certain inherent drawbacks, such as the subjective nature of surveys and the disregard for the cognitive and physiological processes involved in listening effort, its potential to reflect real-world burden warrants additional investigation.

To explore further, the thesis has designed a protocol for an online questionnaire based study to investigate the correlation between listening effort and fatigue, stress and low confidence. The study outlined in this protocol is currently open for recruitment and the results will be written up for publication. If association can be demonstrated future research should be undertaken to determine the validity of this approach. This may take the form of before and after intervention studies or longitudinal studies over a longer period of time.

Overall, this thesis has emphasised the difficulties with measuring listening effort in its entirety. Despite this, from a clinical perspective an holistic self reporting method may provide the most robust technique to discern meaningful information from children and young people with a hearing impairment.

Appendix

NHS ethical approval

WoSRES
West of Scotland Research Ethics Service

Dear Dr Shields

Study title:	**Listening effort and downstream effects due to hearing loss in children and young people**
REC reference:	**22/WS/0134**
IRAS project ID:	**314726**

Thank you for your letter responding to the Research Ethics Committee's (REC) request for further information on the above research and submitting revised documentation.

The further information was considered in correspondence by a Sub-Committee of the REC. A list of the Sub-Committee members is attached.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Good practice principles and responsibilities

The UK Policy Framework for Health and Social Care Research sets out principles of good practice in the management and conduct of health and social care research. It also outlines the responsibilities of individuals and organisations, including those related to the four elements of research transparency:

- registering research studies
- reporting results
- informing participants
- sharing study data and tissue

Dr Callum Shields Research Fellow
Manchester University NHS Foundation Trust
Royal Manchester Children's Hospital
Oxford Road
Manchester M13 9WL

West of Scotland REC 1
West of Scotland Research Ethics Service
Ward 11
Dykebar Hospital
Grahamston Road
Paisley PA2 7DE
www.nhsggc.org.uk

Date 12 October 2022
Direct line 0141-314-0212
e-mail WosRec1@ggc.scot.nhs.uk

<u>Please note</u>: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Number	Condition	Response from Applicant
1	Please add into the PISs, that hearing aids will need to be removed or switched off for the DIN test.	

You should notify the REC once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. Revised documents should be submitted to the REC electronically from IRAS. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which you can make available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Confirmation of Capacity and Capability (in England, Northern Ireland and Wales) or NHS management permission (in Scotland) should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations

Registration of Clinical Trials

All research should be registered in a publicly accessible database and we expect all researchers, research sponsors and others to meet this fundamental best practice standard.

It is a condition of the REC favourable opinion that **all clinical trials are registered** on a publicly accessible database within six weeks of recruiting the first research participant. For this purpose, ‘clinical trials’ are defined as:

- clinical trial of an investigational medicinal product
- clinical investigation or other study of a medical device
- combined trial of an investigational medicinal product and an investigational medical device
- other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice.

Failure to register a clinical trial is a breach of these approval conditions, unless a deferral has been agreed by the HRA (for more information on registration and requesting a deferral see: Research registration and research project identifiers).

If you have not already included registration details in your IRAS application form you should notify the REC of the registration details as soon as possible.

Publication of Your Research Summary

We will publish your research summary for the above study on the research summaries section of our website, together with your contact details, no earlier than three months from the date of this favourable opinion letter.

Should you wish to provide a substitute contact point, make a request to defer, or require further information, please visit:
https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/

N.B. If your study is related to COVID-19 we will aim to publish your research summary within 3 days rather than three months.

During this public health emergency, it is vital that everyone can promptly identify all relevant research related to COVID-19 that is taking place globally. If you haven't already done so, please register your study on a public registry as soon as possible and provide the REC with the registration detail, which will be posted alongside other information relating to your project. We are also asking sponsors not to request deferral of publication of research summary for any projects relating to COVID-19. In addition, to facilitate finding and extracting studies related to COVID-19 from public databases, please enter the WHO official acronym for the coronavirus disease (COVID-19) in the full title of your study. Approved COVID-19 studies can be found at: https://www.hra.nhs.uk/covid-19-research/approved-covid-19-research/

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

After ethical review: Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study, including early termination of the study
- Final report
- Reporting results

The latest guidance on these topics can be found at https://www.hra.nhs.uk/approvals-amendments/managing-your-approval/.

Ethical review of research sites

NHS/HSC sites

The favourable opinion applies to all NHS/HSC sites taking part in the study, subject to confirmation of Capacity and Capability (in England, Northern Ireland and Wales) or management permission (in Scotland) being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS/HSC sites

I am pleased to confirm that the favourable opinion applies to any non-NHS/HSC sites listed in the application, subject to site management permission being obtained prior to the start of the study at the site.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
IRAS Application Form [IRAS_Form_08082022]		08 August 2022
Letters of invitation to participant [Invite/Advertisement]	1.0	27 June 2022
Letters of invitation to participant [Emails to participants for the Listening Effort Study]	1.0	22 July 2022
Letters of invitation to participant [RAS314726 LE study Recruitment and eligibility]	1.0	22 June 2022
Non-validated questionnaire [Demographics]	1.0	22 July 2022
Non-validated questionnaire [Demographics]	2	26 September 2022
Other [Student Research Declaration (Form A)]		
Participant consent form [Consent_16+]	1.0	20 July 2022
Participant consent form [Consent confirmation]	1.0	20 July 2022
Participant consent form [Consent_parent]	2	26 September 2022
Participant information sheet (PIS) [PIS_parents]	2	26 September 2022
Participant information sheet (PIS) [PIS_16+]	2	26 September 2022
Participant information sheet (PIS) [PIS_12-15]	2	26 September 2022
Participant information sheet (PIS) [PIS finding of an unexpected hearing loss]	1	26 September 2022
Referee's report or other scientific critique report		10 December 2021
Research protocol or project proposal [LE Protocol V2 260922]	2	26 September 2022
Response to Request for Further Information [Response to REC comments]		
Summary CV for Chief Investigator (CI) [CV (Professor Iain Bruce)]		16 April 2021
Summary CV for student [Dr Callum Shields]		28 September 2022
Validated questionnaire [Rest day questionnaires]		
Validated questionnaire [Work day questionnaires]		

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at:
https://www.hra.nhs.uk/planning-and-improving-research/learning/

IRAS project ID: 314726   Please quote this number on all correspondence

With the Committee's best wishes for the success of this project.

Yours sincerely

On behalf of
Dr Katriona Brooksbank
Chair

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments

After ethical review guidance for sponsors and investigators – Non CTIMP Standard Conditions of Approval]

Copy to: Ms Danielle Marrochia
Professor Iain Alexander Bruce, Royal Manchester Children's Hospital
Lead Nation

West of Scotland REC 1
Attendance at Sub-Committee of the REC meeting on 11 October 2022

Committee Members:

Name	*Profession*	*Present*	*Notes*
Dr Patricia Roxburgh	Medical Oncologist (REC Vice Chair)	Yes	Chair of Meeting
Miss Santa Walker	Student	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Mrs Kirsty Burt	Senior Co-ordinator

University of Manchester ethical approval

Human Communication, Development & Hearing Division Panel

Division of Human Communication, Development & Hearing

The University of Manchester

Manchester

M13 9PL

Email: HCDHethics@manchester.ac.uk; keith.jensen@manchester.ac.uk

Ref: 2022-15274-25361

20/09/2022

Dear Mr Callum Shields, , Dr Karolina Kluk-De kort, Prof Iain Bruce

Study Title: Listening effort and downstream effects due to hearing loss in children & young people

Human Communication, Development & Hearing Division Panel

I write to thank you for submitting the final version of your documents for your project to the Committee on 19/09/2022 15:17 . I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation as submitted and approved by the Committee.

Two small comments:

Comments were adequately addressed; in future applications, it would help to highlight the changes in the modified documents and to include only these documents in the resubmission.

On the DMP, you say "Manchester University". That university is in the USA. You should correct this to the University of Manchester
On the DMP, you say "Manchester University". That university is in the USA. You should correct this to the University of Manchester.

Please note that your approved data collection window expires on 30/11/2022 , as per the information you provided in section D of your application. If you wish to extend this you must submit a formal amendment before this date or a new ethics application may be required. The maximum window for data collection the Committee is able to approve is 5 years from the date ethics approval is granted (5 years from 20/09/2022). If you wish to collect data beyond 5 years a new ethics application will be required.

Please ensure you review the Research Ethics website throughout the duration of your project to keep up to date on current UoM guidance and best practice.

Please see below for a table of the titles, version numbers and dates of all the final approved documents for your project:

Document Type	File Name	Date	Version
Additional docs	IRAS314726 LE IRAS V1 280722 CS	28/07/2022	1
Additional docs	LE Study Advertisment V1 CS 090822	09/08/2022	1
Additional docs	LE study PIS parents and carers V1 CS 090822	09/08/2022	1
Additional docs	LE study PIS 16 yo and over V1 CS 090822	09/08/2022	1
Additional docs	LE Study recruitment and eligibility REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study ICF (16 years and over) REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study ICF and Assent (12-15 years) REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study participant demographics REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study REST day questionnaire session REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study WORK day questionnaire session REDCap version 1 CS 090822	09/08/2022	1
Additional docs	LE Study Emails to participants V1 CS 090822	09/08/2022	1
Additional docs	LE study Social Media Post V1 CS 090822	09/08/2022	1
Additional docs	Revisions to Ethics Applications - CS - 100922	10/09/2022	1
Additional docs	LE study PIS 12-15 yo V1 CS 170922	17/09/2022	2
Additional docs	IIR-2186 LE Protocol Paper V1 CS 170922	17/09/2022	2
Data Management Plan	DMP UoM V1 170922 CS	17/09/2022	2

If you wish to propose any changes to the methodology or any other specifics within the project, including the dates of data collection, an application to seek an amendment must be submitted for review. Failure to do so could invalidate the insurance and constitute research misconduct.

You are reminded that, in accordance with University policy, any data carrying personal identifiers must be encrypted when not held on a secure university computer or kept securely as a hard copy in a location which is accessible only to those involved with the research.

For those undertaking research requiring a DBS Certificate: As you have now completed your ethical application if required a colleague at the University of Manchester will be in touch for you to undertake a DBS check. Please note that you do not have DBS approval until you have received a DBS Certificate completed by the University of Manchester, or you are an MA Teach First student who holds a DBS certificate for your current teaching role.

Reporting Requirements:

You are required to report to us the following:

1. Amendments: Guidance on what constitutes an amendment
2. Amendments: How to submit an amendment in the ERM system
3. Ethics Breaches and adverse events
4. Data breaches

We wish you every success with the research.

Yours sincerely,

Human Communication, Development & Hearing Division Panel

HCDHethics@manchester.ac.uk; keith.jensen@manchester.ac.uk