Effects of Remote Microphone Hearing Aids (RMHAs) on Listening-in-Noise, Attention and Memory in School-Aged Children with Auditory Processing Disorder (APD)

Georgios Stavrinos

A thesis submitted for the degree of Doctor of Philosophy

University College London

Faculty of Brain Sciences

Ear Institute

University College London

September 2018

DECLARATION

I, Georgios Stavrinos, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGMENTS

This has been one of the most challenging long-term projects I ever took on and it would not have been completed without the contribution, help, guidance, discussions and support from all the wonderful people mentioned here. I would first like to give warm thanks to the children, their parents and teachers who took part in my study. They were all very willing to contribute to the study and their patience and dedication is much appreciated. Of course, this project would not have been realised without the contribution from the funders of the project, GN ReSound and Action Medical Research. I am therefore grateful for their support.

His support was constant for thirty-something years and without him I would not have been in the position to even start this PhD project. A lot of gratitude and love goes out to my father, Panagiotis Stavrinos, whom I had the misfortune to lose before he saw me complete this journey of which he was so supportive.

The supervision and guidance of my supervisor, Professor Doris Bamiou, throughout this project have been nothing short of excellent. Her style of supervision fit exactly my style of work and this meant I was working in an environment where I felt motivated and confident. She has always been there to give the right advice, answer questions, listen, support and guide me whenever required. I have developed as a professional and as a person through her mentorship. I thank her for everything and feel extremely happy to be submitting this work having her name on it as my supervisor.

The help of Professor Maria Chait, who served as my secondary supervisor, was invaluable and it shaped a large part of this work and my way of thinking, as well. She is scientist I look up to and her approach to research has taught me to be a better researcher myself and has definitely improved the quality of work in this thesis. I am grateful for her input and I thank her for everything I have learned from her.

There are a number of people who practically contributed in various ways and varying degrees to the PhD project. They either discussed ideas, addressed concerns, gave me practical advice, read

parts of my work or offered tips on where to find answers. Debi Vickers, Ulrich Pomper, Ed Sohoglou, Vivian Iliadou, Irini Symeonidou, Orestis Georgiou, Nehzat Koohi, Hannah Cooper, Lindsey Edwards, Rajeev Mathew, Theo Petsas, Anahita Mehta, and Deborah Ridout, a thousand thanks! I also wish to give special thanks to Zelia Gregoriou who inspired me through her lectures during my undergraduate and played a big part in my love and interest for research and academia.

And of course, last but not least, I am giving love and thanks to friends and relatives for the immense support and love who provided unconditionally to me throughout this PhD adventure. This PhD has taken me to many peaks and many valleys, but I am grateful to have had the following group of beautiful people be with me to help me up or cheer with me in all these moments, sometimes from distance (thank you skype, too!); Orestis, the other Orestis, Ianthi and Rallou, Giannis, Christiana and Teddy, my sister Constantia, Alexis, my nephews Nikolas and Ifigenia, huge thanks to Claudia who helped me immensely, Silvia, Andreas and Alexia, Irini and Christos (and Donald), Panos, Xenios and Eugenia, Giulia, Gyorgyi and Melinda, George and Maggie, Steph and Tom, Agata, Marta, Marios and Stella and Yiannakis Heracleous, and the rest of my family in London, Cyprus or Canada who have loved me, fed me and worried about me throughout this journey.

ABSTRACT

Auditory Processing Disorder (APD) is characterised by poor Auditory Processing (AP) in the presence of normal hearing. Children with APD have poorer Speech-in-Noise (SIN) skills than their peers and are often reported to have poor attention and Auditory Working Memory (AWM). Management approaches include auditory training and Remote Microphone Hearing Aids (RMHAs). Research on the effects of RMHAs on children with APD is scarce. Some studies claim improvements in SIN performance following RMHA intervention, others suggest that AWM might not be assisted, while the system's effect on attention remains unexplored. Also, the relationship between types of attention and AP skills has not been adequately explored and APD diagnostic criteria do not routinely include attention measures despite evidence linking the two. This thesis aims to address these gaps in the field through three studies.

The first two studies examine the effects of RMHAs on measures of SIN and AWM but also on the novel measures of attention and spatial listening in children with APD using randomised controlled trials. The third study is the first to use correlation analyses to examine the relationship between AP skills and attention in a sample of only APD-diagnosed children. Furthermore, this study looked at a newly proposed APD diagnostic process which included measures of attention with the aim to better inform APD management.

Overall, findings from the clinical trials did not find improvements in any of the attention tests following RMHA use for 6 months, while spatial listening was not negatively affected. In the final study, a strong relationship between divided attention and the dichotic digits test (an AP test) was revealed. Finally, the proposed diagnostic procedure provides useful information to the clinician in terms of management, which may potentially better address children's needs. All findings are discussed in the context of the studies' limitations.

IMPACT STATEMENT

The results of the two clinical trials showed improved listening when using the Remote Microphone Hearing Aids (RMHAs) in some challenging acoustic situations in the classroom. We did not find any significant long-term improvement in attention, working memory, or listening-in-noise tests, and no adverse effect on spatial listening skills as documented on the listening-in-spatialised-noise test (for up to 6 months of use). These findings can inform clinical management of Auditory Processing Disorder (APD). For example, families who are recommended to use RMHAs with their child should be informed that while the RMHAs will make listening easier for the child, they will not have a direct impact on children's attention, working memory, and listening-in-noise ability when not wearing the hearing aids. Therefore, additional interventions should be implemented to address such skills. At the same time, children who are diagnosed with Spatial Processing Disorder may use the RMHAs for at least up to 6 months (if the other auditory processing tests warrant a use of the device), as it will not negatively influence their spatial listening skills.

The findings from the correlational study showed that divided auditory attention has a partial correlation with the dichotic digits test (a test already used to diagnose APD) in a sample of only APD-diagnosed children. This is a novel finding and it may have a direct impact on the APD diagnostic procedure. It is suggested that the two tests should be administered in combination with each other during APD assessments. This will help clinicians have a more complete profile of the patient's dichotic listening skills and abilities in divided attention tasks.

Furthermore, both the findings from the two clinical trials and the correlational study provide with a basis for future studies to build on. Firstly, future clinical trials should test RMHAs under longer intervention periods (of at least up to 12 months) to test long-term (unaided) benefits in attention, working memory, and listening-in-noise. Additionally, aided conditions (i.e. when RMHAs are used) should also be included in the research design to test whether they provide with an improvement in these outcome measures, compared to the unaided conditions. Based on limitations that were identified in the two clinical trials, it is proposed that future work using

RMHAs with APD-diagnosed children should follow these guidelines: not using other interventions alongside RMHAs unless part of the research protocol, verifying functionality of RMHAs, monitoring and reporting the total time the system was used during the study period, and administering a children’s questionnaire on levels of satisfaction in using the RMHA to be used as a control factor in the analysis. These guidelines aim to minimise bias, control confounding factors, while they also ensure uniformity across RMHAs trials on APD studies. Finally, based on the findings from the correlational study research should further explore the relationship between divided attention and the dichotic digits test in children with APD.

PUBLICATIONS AND PRESENTATIONS

Published article resulting from this work

1. Original research article: **Stavrinos, G.**, Iliadou, V.-M., Edwards, L., Sirimanna, T., & Bamiou, D.-E. (2018). The Relationship Between Types of Attention and Auditory Processing Skills: Reconsidering Auditory Processing Disorder Diagnosis. Frontiers in Psychology, 9(January), 1–13. doi:10.3389/fpsyg.2018.00034.

Platform presentations resulting from this work

1. FM systems for APD, Auditory Processing Disorder Masterclass at the Ear Institute, UCL, London, 10th February 2016.
2. The effects of Personal Remote Microphone Hearing Aids on listening skills and auditory attention in children, aged 7-12, with Auditory Processing Disorder, GN ReSound Complex Needs Educational Events,
    a. Bristol, 19th May 2017
    b. London, 24th May 2017
    c. Edinburgh, 25th May 2017

Published article not linked to this work

1. Case report article: Murphy, C. F. B., **Stavrinos, G.**, Chong, K., Sirimanna, T., & Bamiou, D.-E. (2017). Auditory Processing after Early Left Hemisphere Injury: A Case Report. Frontiers in Neurology, 8. doi:10.3389/fneur.2017.00226.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAA	American Academy of Audiology
AD	Attention Deficit
ADHD	Attention Deficit Hyperactivity Disorder
Adv.	Advantage
AFG	Auditory Figure Ground
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ANSD	Auditory Neuropathy Spectrum Disorder
AP	Auditory Processing
APD	Auditory Processing Disorder
ASD	Autism Spectrum Disorders
ASHA	American Speech-Language-Hearing Association
AT	Auditory Training
AWM	Auditory Working Memory
AWMA	Automated Working Memory Assessment
B	Borderline
BSA	British Society of Audiology
CANS	Central Auditory Nervous System
CAPD	Central Auditory Processing Disorder
CCC-2	Children’s Communication Checklist
CHAPS	Children’s Auditory Performance Scale
CI	Confidence Interval
D	Disordered
DDT	Dichotic Digits Test
Div-AA	Divided Auditory Attention
Div-AVA	Divided Auditory-Visual Attention
DLD	Developmental Language Disorder
DSM	Diagnostic and Statistical Manual
DT	Dual Task
EEG	Electroencephalogram
ERP	Event Related Potential
FM	Frequency Modulation
FPT	Frequency Pattern Test
GCC	General Communication Composite
GIN	Gaps-in-Noise
GOSH	Great Ormond Street Hospital

GP	General Practitioner
HHL	Hidden Hearing Loss
HINT	Hearing in Noise Test
HL	Hearing Level
i-APD	Inattentive subtype of APD
ICD-10-CM	International statistical Classification of Diseases and related health problems – 10th revision – Clinical Modifications
IID	Interaural Intensity Difference
IQ	Intelligence Quotient
ISTS	International Speech Test Signal
ITD	Interaural Time Difference
ITT	Intention-to-Treat
L	Left
LIFE	Listening Inventory for Education
LIFE-R	Listening Inventory for Education – Revised
LiSN-S	Listening in Spatialised Noise – Sentences
MMN	Mismatch Negativity
N	Normal
NAL	National Acoustic Laboratories
NSL	Non-Standard Language
PAC	Primary Auditory Cortex
PAS	Peripheral Auditory System
PLI	Pragmatic Language Impairment
PTA	Pure Tone Audiometry
Q	Question
R	Right
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
REM	Real Ear Measurement
RMHA	Remote Microphone Hearing Aid
SD	Standard Deviation
SE	Standard Error
Sel-AA	Selective Auditory Attention
Sel-VA	Selective Visual Attention
SENCO	Special Educational Needs Co-ordinators
SIDC	Social Interaction Deviance Composite

SIFTER	Screening Instrument for Targeting Educational Risk
SIN	Speech-in-Noise
SL	Standard Language
SNR	Signal-to-Noise Ratio
SPD	Spatial Processing Disorder
SPL	Sound Pressure Level
SRT	Speech Reception Threshold
Sus-AA	Sustained Auditory Attention
Sus-VA	Sustained Visual Attention
TEACh	Test of Everyday Attention for Children
UCL	University College London
w/o	Without
WISC	Wechsler Intelligence Scale for Children
WNV	Wechsler NonVerbal

CHAPTER 1 – INTRODUCTION

1.1 Overview

1.1.1 Brief background overview

Developmental Auditory Processing Disorder (APD) is characterised by difficulties in processing sound or speech, especially in background noise, in the presence of normal peripheral hearing. It is a controversial disorder in terms of its definition, diagnosis and treatment. The main controversies revolve around the debate on the effect top-down functions might have on Auditory Processing (AP), the co-occurrence of APD with other developmental and language disorders and the fact that there are no standard diagnostic APD criteria. In mixed samples of children suspected of or diagnosed with APD associations between AP tests and cognitive measures such as attention or memory have been revealed by several studies (Ahmmed et al., 2014; Gyldenkærne, Dillon, Sharma, & Purdy, 2014; Lotfi, Moossavi, Abdollahi, Bakhshi, & Sadjedi, 2016; Martin, Jerger, & Mehta, 2007; Moore, Ferguson, Edmondson-Jones, Ratib, & Riley, 2010; Sharma, Purdy, & Kelly, 2009; Tomlin, Dillon, Sharma, & Rance, 2015), although direction of causality has not been determined. Therefore, children suspected of or diagnosed with APD often present with poor listening in noise, attention or memory skills and these may also affect their academic performance (Martin et al., 2007). In order to assist learning and development of children with APD, evidence on the effectiveness of specific management strategies should be examined through Randomised Control Trials (RCTs).

Proposed management and treatment approaches for APD include bottom-up and top-down auditory training as well as the use of personal Remote Microphone Hearing Aids (RMHAs). This latter strategy focuses on enhancing the Signal-to-Noise Ratio (SNR) for the child in the classroom. There are only a handful of studies examining the effects of RMHAs on specific aspects in children with APD (e.g. language, speech perception in noise and memory), with their findings being inconclusive (Johnston, John, Kreisman, Hall III, & Crandell, 2009; Sharma, Purdy, & Kelly, 2012; Umat, Mukari, Ezan, & Din, 2011). Most of these trials did not include a large enough sample size or they demonstrated other methodological shortcomings, such as use of non-randomised control groups. Previous research showed improved questionnaire scores or self-report scores measuring attention in children with APD following the use of low-gain hearing aids equipped with a directional microphone and noise reduction (Kuk, Jackson, Keenan, & Lau, 2008). However, other behavioural attention tests have not been used in RMHA intervention studies on children with APD. The studies of the thesis aim to contribute to the field by assessing whether the use of RMHAs influences attention, memory and speech-in-noise perception, in children diagnosed with APD. This is the first work to look at the effects of RMHAs on behavioural measures of attention in this population.

In addition, there have been studies demonstrating correlations between sustained attention and AP skills in samples of children suspected of APD (Gyldenkærne et al., 2014; Sharma et al., 2009). Moreover, composite scores (auditory plus visual) of selective and sustained attention have been found to be worse in a group of children diagnosed with APD compared to a group suspected of but not diagnosed with the disorder (Allen & Allan, 2014). This, though, may not be specific to children with APD, as there have been studies on preschool children and adults with dyslexia that found poorer visual spatial attention and attentional modulation of visual processing in the clinical groups compared to typically-developing controls (Facoetti, Corradi, Ruffino, Gori, & Zorzi, 2010; Roach & Hogben, 2004). At the same time, diagnosis of APD has no universal standards and does not routinely include attention tests, even though attention is linked to APD. Therefore, the thesis also aims to further examine the relationship between types of attention and AP skills in a sample of children diagnosed with APD and to assess other diagnostic procedures that include attention measures in order to inform APD management.

1.1.2 Structure of Chapter 1

In this chapter the structure and functions of the Peripheral Auditory System (PAS) and of the Central Auditory Nervous System (CANS) will first be briefly described. This will lead to a discussion of the controversies in the definition of APD, the diagnosis of APD, the co-occurring disorders, and a presentation of APD symptoms and prevalence. Then, the relationship between

APD and speech-in-noise, attention and memory will be outlined. This will be followed by a discussion on APD management which will mainly focus on studies using RMHAs to test their effects on children with APD. Justification of how RMHAs in the studies are expected to bring changes are then detailed, leading up to a brief presentation of the hypotheses; the detailed hypotheses are discussed later under each study chapter.

1.2 Anatomy of the auditory system

1.2.1 Peripheral auditory system

The PAS consists of the outer, middle and inner ear and auditory nerve, all working together to introduce auditory stimuli further up the auditory system. Both the outer and middle ear comprise the conductive system, as their chief function is to transfer/ conduct the sound coming from the environment to the inner ear (Gelfand, 2009). The sound entering the ear canal (enhanced due to the canal's shape) hits the tympanic membrane and causes it to vibrate (T. Wright, 2001). Inside the middle ear there are three ossicles (the malleus, incus and stapes), which conduct the vibrations from the tympanic membrane to the oval window of the inner ear through a lever-like, mechanical effect (Gelfand, 2009). In turn, these vibrations initiate changes in pressure inside the inner ear.

In the inner ear the main structure is the cochlea, which resembles a snail-like shell (T. Wright, 2001). The cochlea itself analyses high frequencies around its base and lower frequencies toward the apex, a function known as tonotopic organisation (Gelfand, 2009). A thin bony structure divides the cochlea into two fluid compartments, the scala vestibuli and the scala tympani (see Figure 1.1). The area of the scala vestibuli and scala tympani communicate through an opening at the apex of the cochlea and thus both share the same fluid, the perilymph (Hayes, Ding, Salvi, & Allman, 2013).

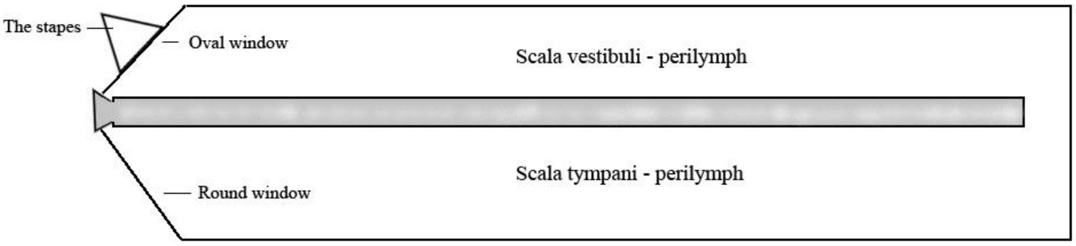

Figure 1.1 – Schematic representation of an uncoiled cochlea
The areas of scala vestibuli and scala tympani communicate through the opening at the apex of the cochlea and share the same fluid, the perilymph.

The scala media is also running through the cochlea and contains the fluid endolymph, while the basilar membrane creates a division between the scala media and the scala tympani (see cross-sectional view of this in Figure 1.2). The organ of Corti is found within the basilar membrane and just like the scala media, it also runs along the cochlea (Hayes et al., 2013).

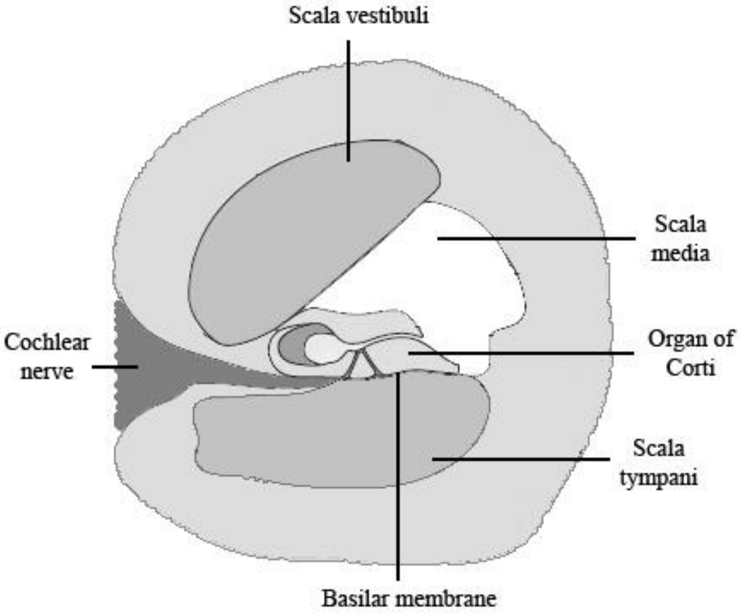

Figure 1.2 – Cross-sectional view of the cochlea
Cross-sectional view of the cochlea indicating the areas of scala vestibuli, scala tympani, scala media, the organ of Corti, the basilar membrane and the cochlear nerve.

The changes in pressure initiated at the oval window are transferred through changes in pressure in the perilymph and endolymph. This causes the basilar membrane to vibrate and create movement in the hair cells. Hair cells, which are the sensory structures of the inner ear, reside in the organ of Corti and their function is to convert the sound coming in the middle ear (in the form of mechanical energy) into electricity so that it can be then transmitted to the brain (Hayes et al., 2013). These cells are divided into inner and outer hair cells and both have nerve fibres attached to them, with the former having more fibres than the latter (T. Wright, 2001). These fibres, along with other fibres originating from the inner ear, form the auditory nerve which carries on further up the CANS (T. Wright, 2001).

1.2.2 Central auditory pathways

The role of the CANS is to transform the information it receives from the cochlea into meaningful auditory objects and this is achieved through the processes of the different auditory pathways of the structure (Brugge, 2013). Figure 1.3 presents a highly schematic diagram of the CANS, as described in the next sections. This is a simplified diagram and does not intend to present all pathways and connections of the CANS.

Auditory nerve and cochlear nucleus

The auditory nerve leaves the inner ear, reaches the brainstem and terminates at the cochlear nucleus (Gelfand, 2009). The tonotopic organisation observed in the cochlea is also reproduced in the auditory nerve (Clopton, Winfield, & Flammino, 1974; Gelfand, 2009). In addition, the auditory nerve is capable of accurately representing the information it receives from the inner ear to the CANS and thus a dysfunction in the auditory nerve affects people’s ability to further break down the information they hear (Bellis, 2011).

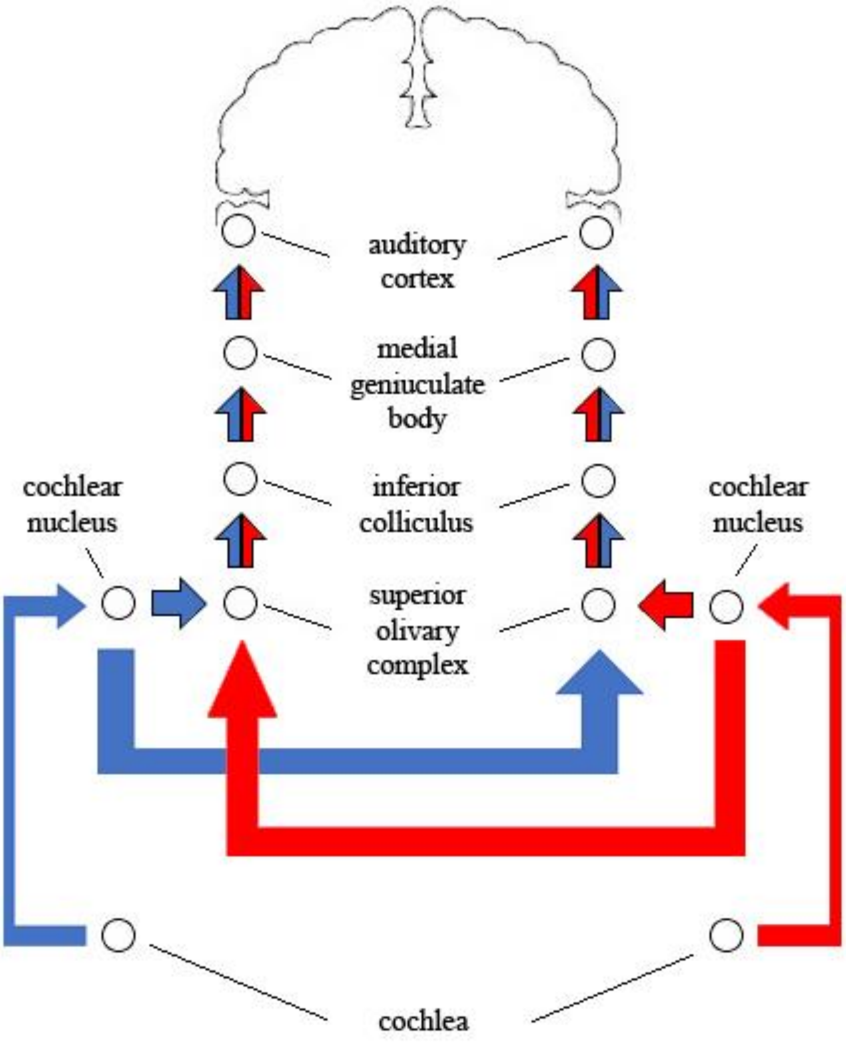

Figure 1.3 – Highly schematic representation of the CANS
A highly schematic diagram of the CANS showing the main pathways and connections. This is a simplified representation and does not intend to present all pathways and connections of the CANS. The red and blue lines are used to demonstrate the contralateral and ipsilateral projections to each structure. Red lines represent the projections coming from the right side, blue lines show projections coming from the left side. CANS: Central Auditory Nervous System.

The auditory nerve, terminating at the cochlear nucleus, synapses with the nerve cells of this level. The cochlear nucleus' main function appears to be the enrichment of features in the neural signal, while the dorsal cochlear nucleus may play a role in the amplification of signals in noise (Bellis, 2011; Pickles, 2015). The cochlear nucleus is the first level of the central auditory system where auditory fibres decussate and project contralaterally in addition to projecting ipsilaterally, thus dysfunctions at the cochlear nucleus and further up usually translate into bilateral and contralateral test abnormalities (Bellis, 2011; Hackett, 2009).

Superior olivary complex and inferior colliculus

Some neurons leaving the cochlear nucleus will go to the superior olivary complex on the same side (ipsilateral), while the majority will cross contralaterally (Gelfand, 2009). This function of decussation, along with patterns of divergence and convergence of neurons may help the superior olivary complex achieve processing of the binaural input (Bellis, 2011; Brugge, 2013). The superior olivary complex uses information cues related to timing and loudness of the signal in each ear and this could subserve critical processes for listening in background noise, localisation and lateralisation of the auditory signal (Brugge, 2013; Wada, 2007). Therefore, abnormalities in the superior olivary complex might be linked to temporal processing problems and reduced ability to listen in background noise.

The inferior colliculus is the structure where streams coming from the lower levels converge (Gelfand, 2009; Pickles, 2015). Binaural information is said to be processed at this level and it is suggested that simple acoustic characteristics are transformed into more meaningful auditory objects in the inferior colliculus (Brugge, 2013; Pickles, 2015).

Thalamus and the medial geniculate body

The thalamus receives both ipsilateral and contralateral projections from the inferior colliculus, which terminate at the medial geniculate body of the thalamus (Bellis, 2011). Additionally, it is supported that the medial geniculate body contributes to the further refinement of the neural coding and to the processing of complex signal processing which are then communicated to the cortex (Bellis, 2011; Pickles, 2015). Moreover, it is thought that vowels and slowly changing consonants are well encoded in the thalamus at a level which may preserve this information even when the cortex is damaged (Bellis, 2011).

Primary auditory cortex and the corpus callosum

Auditory stimuli activate neurons in the Primary Auditory Cortex (PAC), as well as in several other regions of the cerebrum (Hackett, 2015). The PAC receives its main input from the medial geniculate body (Hackett, 2015). The frequency modulation that is encoded by the neurons of the

PAC could serve as an important cue for speech perception, as their modulation range falls within the envelope of running speech, while the encoding of rapid acoustic speech stimuli that are crucial for discriminating consonants may also take place in the PAC (Brugge, 2013; Nourski & Brugge, 2011). When the PAC is damaged, speech-sound discrimination is directly affected, with stop consonants becoming more difficult to identify than vowels (Bellis, 2011). Adding to that, the PAC appears to play a role in the perception of sound localisation (Phillips, 2014). Even though the PAC is the area where auditory stimuli are being formed into sounds or auditory objects, other regions of the cerebrum can also respond to auditory stimuli (Brugge, 2013) revealing the complexity of the activity network.

The two auditory cortical areas in the left and right hemisphere communicate through the corpus callosum. This interhemispheric transfer may assist dichotic listening, as split brain patients have been found to perform poorly in some dichotic listening tasks (Chermak & Musiek, 2014). Moreover, the corpus callosum is associated with a number of other functions, from basic sensory perceptual functions to higher-order cognitive functions, such as selective, divided and sustained attention or auditory-verbal learning and memory (Bellis, 2011; Hutchinson, Mathias, & Banich, 2008; van der Knaap & van der Ham, 2011). Therefore, dysfunctions in the corpus callosum may interrupt any of these processes.

1.3 Auditory processing disorder

1.3.1 Bottom-up and top-down auditory processing

Before talking about APD it is important to briefly describe the bottom-up and top-down processes and their contribution to audition. The initial processing of acoustic stimuli occurs pre-consciously, meaning that some auditory processing takes place automatically even without the listener’s directed or conscious attention to the auditory signal. This constitutes the bottom-up process (Alain, Arnott, & Picton, 2001; Bellis, 2011). Top-down processing, a more detailed processing of the stimulus using prior knowledge, context and memory, helps extract the relevant information (Alain et al., 2001). In short, the bottom-up process groups sounds based on physical characteristics, whereas the top-down process makes sense of the auditory stimulus by using cognitive functions and previous knowledge (Alain et al., 2001). Both these functions contribute to the successful processing of auditory information and it is argued that a disordered auditory system could be helped by supporting bottom-up processing and training top-down processing (Bellis, 2011; Chermak & Musiek, 1997; Pimentel & Inglebret, 2014). For instance, the use of bottom-up training (e.g. temporal processing training) or signal amplification may provide support to CANS functions or to peripheral hearing and help process or amplify the incoming auditory signal (Chermak & Musiek, 2002). In turn, further processing of auditory information could be supported by training language or cognition (i.e. memory, attention), as the trained individual may be able to use the new knowledge or strategies to make sense of or further break down the incoming information (Bellis & Bellis, 2015). These bottom-up and top-down management approaches are further discussed later under Section 1.7.

1.3.2 Definitions and theoretical models

APD categories

Auditory processing is the perceptual processing of auditory input within the CANS that is driven by the neurobiological activity of the system, and is reflected in electrophysiological auditory potentials (American Speech-Language-Hearing Associeation [ASHA], 2005b). In other words, AP is the processing that represents how efficient and effective the CANS is in using the auditory information it receives (ASHA, 2005b). Auditory processing disorder is characterised by poor speech and non-speech perception often in the presence of normal peripheral hearing (British Society of Audiology [BSA], 2018). The BSA proposes three categories of APD (BSA, 2018):

- Developmental APD: This category includes children that present with listening difficulties in the presence of normal peripheral hearing. They may have a family history of language or developmental disorders but not any other known risk factors.
- Acquired APD: In this category individuals acquire APD either later in life (i.e. with ageing) or after a specific event (such as brain or neurological trauma).
- Secondary APD: The individuals in this category present with hearing impairment, which may be causing their APD symptoms or may coexist alongside APD.

The present work will focus on children with Developmental APD (and thus normal peripheral hearing). Therefore, this thesis is specific to Developmental APD and from this point on the term APD will be used for brevity.

Controversies in defining APD

Auditory processing disorder is considered a controversial disorder in terms of its definition and diagnosis. One of the controversies is reflected in the terms used by the ASHA (2005a) and by the BSA (2018) to refer to AP deficits, with the former using the term ‘Central Auditory Processing Disorder’ (CAPD), and the latter using ‘Auditory Processing Disorder’. The addition of ‘Central’ in ASHA’s term, signifies that the disorder is explained by a dysfunction in the central auditory pathway alone (Bellis, 2011). It is hence purported that CAPD is not caused by higher-level language or cognitive deficits – i.e. top-down effects – but that it is primarily the product of a dysfunction or dysfunctions in the neural processing of auditory signals (ASHA, 2005a; Cowan, Rosen, & Moore, 2009). It is nonetheless made clear that other disorders may coexist alongside APD, but not resulting in or explaining the disorder (American Academy of Audiology [AAA], 2010; ASHA, 2005a). To add to the discussion, a recent ‘European consensus’ perspective on APD has been proposed by a number of leading researchers in the field (Iliadou et al., 2017). This argues that defining the disorder as APD rather than CAPD may include cases that have Hidden Hearing Loss (HHL; i.e. a dysfunction that occurs between the inner ear hair cells and the auditory nerve fibres) or even Auditory Neuropathy Spectrum Disorder (ANSD; i.e. auditory nerve dysfunction and desynchronised auditory brainstem response; Iliadou et al., 2017; Rance & Starr, 2016). The ‘European consensus’ paper states that the use of CAPD is a narrower disorder category which excludes the cases mentioned above, whereas the term APD is a broader category (Iliadou et al., 2017). Thus, the American definition of the disorder is based on the assumption that bottom-up processing deficits are the primary cause of the clinical presentation of APD rather than cognitive or language related deficits. The ‘European consensus’ group does not take an absolute stance on which term to be used, but does mention that the majority of the authors favour the term APD.

The BSA's (2018, 2011b) position statement on APD posits that poor perception characterises the disorder. Perception in this case is said to incorporate and use both bottom-up (sensory activation) and top-down processes (i.e. attention and memory; BSA, 2018). Consequently, according to the British definition, a dysfunction in either pathway may explain the presence of APD (BSA, 2011a). In the BSA's position statement the important role of top-down processes is highlighted and in some occasions dysfunctions in top-down processes are thought to generate the presentation of APD symptoms. This view is largely based on the findings of a population study by Moore et al. (2010), in which 1469 UK children, aged 6-11, were tested on measures of AP (i.e. frequency discrimination, simultaneous and backward masking, and speech-in-noise) and extrinsic and intrinsic attention (i.e. phasic auditory and visual alertness). The study also calculated derived measures of APD - frequency discrimination and temporal resolution - ensuring that variability produced by the procedural and cognitive demands of the task were controlled and thus the tests reflected sensory deficits (Moore et al., 2010). Results demonstrated that the two derived AP measures were related to poor extrinsic visual alertness in children who performed more poorly on AP measures, but they were not linked to auditory alertness. For intrinsic auditory attention, the poor AP performers had a more variable performance in the attention task than typical AP performers. The authors concluded that attention (along with other top-down, cognitive functions such as memory) might be contributing to the poorer presentation of some children on the AP measures and could explain APD (Moore et al., 2010). The different definitions are important because they have implications for intervention. For instance, ASHA (2005a) notes that for managing APD equal emphasis should be put on both bottom-up and top-down strategies. It maintains that with top-down management strategies, individuals with APD may support their disordered AP skills that were not fully treated through bottom-up training. The 'European consensus' group also highlights the contribution of possible bottom-up and/ or top-down dysfunctions in explaining APD and suggests that a multidisciplinary management approach designed for the individual should be used making use of environmental modifications, bottom-up, and top-down strategies (Iliadou et al., 2017).

Just last year, the disorder was included, for the first time, in the US version of the International statistical Classification of Diseases and related health problems – 10th revision – Clinical Modifications (ICD-10-CM) with the code H93.25 (ICD-10-CM, 2018). The definition given by the ICD-10-CM is similar to the ASHA definition, as it characterises the disorder as a dysfunction solely in the auditory pathway. The ICD-10-CM mentions that possible causes may be delays in brain maturation or other brain traumas (ICD-10-CM, 2018). It has therefore been argued that the ICD-10-CM definition is presented more from an audiological perspective (Wilson, 2018). Conversely, APD was not included in the 2017 update of the Diagnostic and Statistical Manual (DSM) of Mental Disorders – Fifth edition (American Psychiatric Association, 2013, 2017). Therefore, the classification of APD as a disorder is found only in the US version of the ICD-10, and the fact that it is not included in the DSM-5 indicates that the disorder is not universally recognised.

APD subtypes – Spatial processing disorder

In addition, to the terminological controversies, there is also no consensus regarding the clinical presentation of APD. Some researchers have suggested that APD can be classified into different subtypes. For instance, the Buffalo Model (Katz, 1992) describes four subtypes of APD: the Decoding subtype (problems in phonemic processing), the Tolerance-Fading Memory subtype (speech-in-noise difficulties and poor short-term memory), the Organisational subtype (problems in organisation of auditory information), and the Integration subtype (difficulties in integrating information from both the auditory and visual modality; Katz, 1992). Similarly, Bellis and Ferre (1999) proposed three APD subtypes: Auditory Decoding Deficit (difficulties in bilateral integration and monaural separation and closure, and poor speech and spelling), Prosodic Deficit (difficulties using prosodic and pragmatic features to extract meaning), and Integration Deficit (poor abilities in tasks requiring interhemispheric transfer – binaural integration/ separation; Bellis & Ferre, 1999). However, these APD categorisations were drawn from the authors' clinical impressions (Dawes & Bishop, 2009), were not based on data from peer-reviewed sources (Jutras et al., 2007), and were not widely accepted (ASHA, 2005b; BSA, 2011b). In addition, there is lack of further research to support these two models and they are not acknowledged by the recent

BSA position statement (BSA, 2018) nor by the AAA and the ‘European consensus’ group (AAA, 2010; Iliadou et al., 2017).

There is also the distinction between spatial processing disorder and other forms of APD. Spatial Processing Disorder (SPD) is considered a subtype of APD, with specific test deficits that arise in a significant proportion of affected individuals in the presence of chronic otitis media (Cameron, Dillon, Glyde, Kanthan, & Kania, 2014). Spatial listening is the ability to localise sound in space by using the auditory information from both ears (Kitterick, Lovett, Goman, & Summerfield, 2011). Sound localisation is achieved by the CANS by comparing differences in time and intensity arriving at the two ears (Cameron & Dillon, 2008). Interaural Time Differences (ITDs) are created because of the different distance the sound requires to travel to reach each ear and this constitutes the time cue the CANS uses (Cameron & Dillon, 2008). Head shadowing, causing a decrease in intensity in the ear furthest from the target and thus difference in the intensity reaching each ear, produces the Interaural Intensity Differences (IIDs), which constitute the intensity cue the CANS uses (Cameron & Dillon, 2008). Both these cues assist the CANS in localising the target in space (Litovsky, 2005). When noise or distractors are emanating from the same source as the target, the ability to recognise the target is decreased but when the distractor or noise is spatially separated from the target, target intelligibility is assisted (Hawley, Litovsky, & Culling, 2004; Litovsky, 2005). Studies have shown that children suspected of APD perform worse in tasks with spatial separation compared to typically developing age-matched controls (Cameron, Dillon, & Newall, 2006; Cameron & Dillon, 2008; Lotfi, Moossavi, et al., 2016). This decreased ability to use binaural cues to spatially separate the target from distractors has been termed SPD and is considered a subtype of APD (Cameron et al., 2014). Therefore, individuals with SPD cannot efficiently use ITDs or IIDs to achieve spatial release from masking (Cameron, Glyde, & Dillon, 2012).

Additionally, Cameron and Dillon (2008) found that all children who failed the spatial listening conditions of the Listening in Spatialised Noise – Sentences (LiSN-S) test had normal non-spatial speech-in-noise scores (i.e. normal score in the Low-cue Speech Reception Threshold [SRT] condition, where target and distractor come from the same location). This further strengthens the argument for the presentation of the SPD subtype in some children with auditory processing deficits (Moore, Rosen, Bamiou, Campbell, & Sirimanna, 2013). Nonetheless, a previous study by the same group of researchers, which included children with an APD diagnosis (and not suspected of APD as in the 2008 study), found that children with APD performed significantly worse than controls even in the Low-cue SRT condition (Cameron et al., 2006). The authors indicated that in the 2006 study a subgroup of the sample may have had undiagnosed deficits (perhaps memory problems) that influenced their performance in the Low-cue SRT test (Cameron & Dillon, 2008). Another problem with the 2006 trial is that its design possibly introduced bias, as children had to give details of the stories they heard (instead of repeating back fixed words or sentences) meaning their responses were subjective (Cameron & Dillon, 2008). At the same time, the argument has been made that SPD may be influenced by language processing and that this relationship should be further investigated (Moore et al., 2013).

There is some evidence to suggest that SPD can be ameliorated following deficit-specific training (Cameron & Dillon, 2011; Lotfi, Moossavi, et al., 2016). Cameron and Dillon (2011) found that nine children, aged 7-11, diagnosed with SPD (via the LiSN-S) significantly improved their scores in the spatial listening conditions of the LiSN-S after 3-months of training on the LiSN & Learn software, which specifically targets spatial processing deficits. The Low-cue SRT condition (where target and distractors come from the same source) remained unchanged. The scores in the spatial listening conditions maintained the same improved levels even 3 months after the end of the study; a period during which training was stopped (Cameron & Dillon, 2011). This indicates that the effect can be sustained for at least 3 months following this deficit-specific intervention. Apart from the small sample used, the study did not use an RCT design and had children acting as their own controls. The improvement may have either been the result of maturation, or due to a placebo/ Hawthorne effect. Furthermore, the trial only included children suspected of and not diagnosed with APD.

The authors conducted a similar study with an improved design where they compared the effectiveness of the LiSN & Learn training programme against Earobics using a single-blind (i.e. participants were blinded) RCT design (Cameron et al., 2012). Earobics is an auditory training programme aimed at teaching auditory and phonological awareness skills and is presented in a non-spatial format (unlike the LiSN & Learn). The authors randomly assigned five children, aged 6-10, with SPD in each programme, with LiSN & Learn serving as the intervention programme (because of the deficit-specific training) and the Earobics as the control programme (because the training was presented in a non-spatialised format). After three months of 15 minutes daily training, the treatment group significantly improved its scores in the LiSN-S spatial conditions while the control group did not (Cameron et al., 2012). These improvements were also mirrored in questionnaires (completed by parents, teachers and children themselves) assessing speech-in-noise performance, with the LiSN & Learn group being the only group that had significantly improved questionnaire scores (Cameron et al., 2012). The authors concluded that non-spatial listening training did not improve spatial listening-in-noise and highlighted the importance of providing a deficit-specific intervention.

Another group of researchers randomly assigned 30 children suspected of APD into two groups (control and intervention with mean ages of 9 years; Lotfi, Moossavi, et al., 2016). The controls did not follow any training, whereas the intervention group used an auditory training program. The training was targeting auditory lateralisation listening skills using ITDs, a process important for spatial listening. After a 6-week intervention, only the training group recorded significant improvements in the spatial listening tests (Lotfi, Moossavi, et al., 2016). Conclusions were similar to those drawn previously by Cameron et al. (2012), but the interesting finding here was that significant improvements were also observed in a non-spatial speech-in-noise test (Lotfi, Moossavi, et al., 2016). Pairing the findings from the two studies together it can be argued that non-spatial listening-in-noise training would not be expected to improve spatial listening, but a spatial SIN training may improve non-spatial listening-in-noise ability, as well as spatial listening ability.

1.3.3 Symptoms and co-occurrence with other disorders

Speech-in-noise difficulties

The AP of speech signals appears to be at a disadvantage when listening conditions are suboptimal, such as when listening in the presence of background noise or in reverberant rooms (W. J. Keith & Purdy, 2014). Reverberation time and distance from the speaker could introduce a detriment to the listening ability of individuals with APD (W. J. Keith & Purdy, 2014; Moore et al., 2013); however, this has not been scientifically examined but rather extrapolated from findings of worse speech-in-noise performance by children with APD compared to typically developing controls (the study is further discussed at the end of this paragraph; Lagacé, Jutras, Giguère, & Gagné, 2011). Reverberation time is the reflection of sound waves off of surfaces in rooms, before arriving at the listener's ears (Toe, 2009). Rooms high in reverberation time contribute to the degradation of speech perception (Toe, 2009), as the reflected speech signal overlaps and masks acoustic characteristics of the direct speech signal (Flexer, 2014). A difficulty in Speech-in-Noise (SIN) perception is reported as one of the main symptoms of children with APD (BSA, 2011b), also reflected on parental reports and questionnaires (Ferguson, Hall, & Moore, 2011; Wilson et al., 2011). Understanding speech in background noise is a challenge even for normal hearing children (Schafer et al., 2014; Wong, Uppunda, Parrish, & Dhar, 2008), but studies reveal that for children with APD this is more of a problem than for their peers. A study showed that a group of 10 children, aged 9-12, with APD had significantly poorer scores in a sentences-in-noise test at the +3 dB SNR and 0 dB SNR conditions compared to a group of 10 age-matched typically developing children (Lagacé et al., 2011). Another study found that the group performance of 25 children with APD, aged 6-12, was worse than the group performance of 25 typically developing age-matched controls even at +20 dB SNR in a monosyllabic words in noise test (Rocha-Muniz et al., 2014).

Other findings, though, do not support the view that children with APD have poor SIN abilities when compared to non-APD peers (Ferguson et al., 2011). Ferguson et al. (2011) did not find statistically significant differences in sentences-in-noise and non-words-in-noise test scores between 29 children labelled as having APD, aged 6-13, and 55 randomly selected children from mainstream schools aged 6-12, despite parental reports prior to the study suggesting that SIN difficulty was the most common symptom in the APD group. The researchers however, did not exclude any children from the control group, even though some were suspected of developmental disorders (Ferguson et al., 2011). This decision might have influenced the findings as potential developmental disorders in a subgroup of the controls may have brought down the mean scores in the two SIN tests. Adding to this, Ferguson et al. (2011) did not run any AP tests on the children labelled as having APD and these children were only said to have APD based solely on parental reports of APD symptoms. It could be therefore argued that evidence up to now suggests that children with APD have been performing worse in SIN tasks than typically developing peers.

Other APD symptoms

The rest of the observed and reported symptoms of children with APD are varied. As well as difficulties in understanding speech and auditory stimuli in non-ideal listening conditions (W. J. Keith & Purdy, 2014; Lagacé et al., 2011), other characteristics often include (but are not limited to) the following: mishearing words or entire sentences, delayed responses to requests/ questions, poor attention and specifically sustained attention, distractibility, need for repetitions, poor ability to follow directions, sound localisation issues, memory and other learning problems, weak social skills, low self-esteem and frustration (AAA, 2010; ASHA, 2005b). These behaviours might vary among children with APD depending on the severity (Martin & Keith, 2009) and/ or the type of the disorder, such as typical APD or spatial processing disorder as mentioned earlier (Cameron et al., 2014).

Co-occurrence with other disorders

It is often the case that APD is discussed in the context of other developmental disorders with similar symptoms and is even found to co-occur with these disorders (Bamiou, Musiek, & Luxon, 2001; Chermak, Somers, & Seikel, 1998; Dawes & Bishop, 2010). Co-occurrence takes place when two disorders coexist but are distinct in the same person (Cacace & McFarland, 2006). Co-occurrence of APD with other disorders may be due to: (1) shared neuroanatomic areas where the physiologic dysfunctions are located; (2) common symptoms and aetiologies; (3) coincidental, unrelated co-occurrence (Chermak, Bellis, & Musiek, 2014); or (4) atypical neurodevelopment leading to clustering of shared symptoms (i.e. without the need for a shared neurological/aetiological origin). However, given the high co-occurrence of APD with other developmental disorders coincidental co-occurrence is more unlikely. Regarding the second point, common aetiologies may also include shared genetic risk factors between APD and other co-occurring disorders. For instance, both Developmental APD and Developmental Language Disorder (DLD) are said to be linked to a family history of developmental language and communication disorders (BSA, 2018; Tomblin, Smith, & Zhang, 1997). Other models that attempt to explain the co-occurrence of APD and DLD are discussed later below.

One of the most common disorders that co-occurs with APD is Attention Deficit Hyperactivity Disorder (ADHD; Cacace & McFarland, 2013). Cacace and McFarland (2006) discuss the possibility that the co-occurrence between these two disorders might be due to mistaken diagnosis. They argued that many of the tests for APD require attention, while conversely the AP abilities of children suspected of ADHD might not be assessed during ADHD assessment (Cacace & McFarland, 2006). Despite the overlap in some symptoms between ADHD and APD, other authors have argued that the two are distinguishable (Bamiou et al., 2001; Chermak et al., 1998). This is because on the one hand, the diagnosis of ADHD is usually made based on symptom reports[1] and it usually comes from professionals working in the field of psychology, while on the other hand diagnosis of APD relies on multidisciplinary assessments that include tests of hearing and AP (Bamiou et al., 2001; Chermak et al., 1998). However, given the shared symptomatology of the two disorders it can be argued that the two disorders are a single disorder (Bamiou et al., 2001). It can also be supported that the overlap in symptoms indicates a shortcoming in the accuracy of diagnosis (Cacace & McFarland, 2006). Nevertheless, Chermak et al. (1998) found

---

[1] There are contradicting views between the DSM-5 and the ICD-10 on how categorisation of ADHD subtypes is made (National Collaborating Centre for Mental Health, 2009). Without going into details regarding that, both the DSM-5 and ICD-10 require at least six symptoms out of nine in each of the two groups of symptoms (i.e. inattentive, and hyperactive-impulsive) for any ADHD diagnosis to be given. These symptoms are drawn from rating scales and questionnaires completed by both teachers and parents, and there is no gold standard as to which questionnaires are to be administered (National Collaborating Centre for Mental Health, 2009).

that clinicians making the diagnoses for APD and ADHD identified, through the use of behavioural checklists, different behaviours and characteristics for each of the disorders. Specifically, clinicians diagnosing each disorder were given a checklist of behaviours and were asked to rank them based on frequency of appearance in their clinical group. The behaviours that were 1 Standard Deviation (SD) above the grand mean in each group were reported. For children with APD the five most frequently observed behaviours were SIN difficulties, difficulties following oral instructions, poor listening skills, academic struggles and poor auditory association skills. For children with ADHD the five most frequently observed behaviours were inattention, distractibility, hyperactivity, restlessness and impulsivity. Even though the five most frequently observed behaviours were different between the two groups, inattentiveness and distractibility were recorded often enough in children with APD to be 1 SD above the mean (but observed less often than the top five APD behaviours; Chermak et al., 1998). Therefore, it is yet unclear how these two entities interact and more research is required to determine their relationship (Cacace & McFarland, 2006).

Other disorders commonly reported to co-occur with APD are DLD, dyslexia, Autism Spectrum Disorder (ASD) and reading disorders (Bamiou et al., 2001; Dawes & Bishop, 2010; Sharma et al., 2009). Dawes and Bishop (2010) compared children with a diagnosis of APD[2] (mean age 10.4) and children with a diagnosis of dyslexia[3] (mean age 10.1) in order to examine whether they have distinct psychometric profiles. They found that 52% of children with a diagnosis of APD also met criteria for dyslexia, DLD[4] or both, while almost 20% of children with dyslexia fit a diagnosis of DLD. The percentage of children who met DLD criteria was not significantly different between the two studied groups (Dawes & Bishop, 2010). Moreover, Ferguson et al.

---

[2] Details of APD diagnosis are given in Appendix A. In short, children had to have a score at least -2 SD from the mean on the SCAN-C or SCAN-A and failure of at least one or more AP test (i.e. Pitch Patterns test, Duration Patterns test, Random Gap Detection test; but no mention of cut-offs). Normal non-verbal Intelligence Quotient (IQ), above 80 standard score, was also required.

[3] In Dawes & Bishop (2010) dyslexia was diagnosed when a reading or spelling test was below 85 standard score plus a normal non-verbal IQ (above 80 standard score).

[4] In Dawes & Bishop (2010) DLD was diagnosed when non-verbal IQ (Wechsler, 1991) was above 80 and two of the six language tests (i.e. sentence repetition, non-word repetition, storytelling reception and recall, mean length of utterance, story comprehension) were below -1 SD from the mean (Bishop, 2004, 2005; Korkman, Kirk, & Kemp, 1997).

(2011) found that on the parental Children's Communication Checklist – 2 (CCC-2) questionnaire (Bishop, 2003), which measures language and communication skills in children, two clinical groups of children labelled as having APD and DLD did not significantly differ in any of the 10 subscales except one[5]: speech, which assesses children's articulation/pronunciation and fluency. Another study also found a co-occurrence of language impairment[6] in children diagnosed with APD (Sharma et al., 2009). Specifically, 7 out of 49 children with APD (7%) met the criteria for language impairment, a further 7% met the criteria for reading disorder, while 32 (65%) met the criteria for both language impairment and reading disorder[7] in addition to their APD.

These studies reveal a high co-occurrence of APD and language problems (i.e. DLD, dyslexia), while DLD appears both in children with APD and with dyslexia. In trying to explain the high co-occurrence of APD and DLD, it has been suggested that symptoms of listening deficits may be a consequence of language disorders. While causal models have argued that APD could be the cause of language disorders (Goswami, 2011; Tallal, 2004), there is evidence for the support of consequence models that propose that auditory deficits may be a consequence rather than a cause of language disorders (Bishop, Hardiman, & Barry, 2012). In addition, Ferguson et al. (2011) suggested that the children in their two groups of APD and DLD were given different diagnoses based on the differential referral route they followed and not because of the actual differences in their behaviours. They considered the possibility that the two disorders may identify the same problems in children by giving them two separate clinical labels (Ferguson et al., 2011). However,

---

[5] In Ferguson et al. (2010) children labelled as having APD did not differ from children with DLD on the following CCC-2 subscales: syntax, semantics, coherence, inappropriate initiation, stereotyped language, use of context, nonverbal communication, social relations, interests (Bishop, 2013). They only differed in the speech subscale.

[6] In Sharma et al. (2009) language impairment was identified if children had receptive and/ or expressive language standard scores (on the Clinical Evaluation of Language Fundamentals - 4) below 80, and/ or receptive and expressive language standard scores that were above 80 but where either one of those scores was more than 1 SD below non-verbal IQ, and/ or both receptive and expressive language scores were within normal range but the receptive language score was more than 1.25 SDs below the expressive language score (Brown, Sherbenou, & Johnson, 1997; Semel, Wiig, & Secord, 2003).

[7] Reading disorder in Sharma et al. (2009) was identified if children scored at least -1 SD below the normative mean in reading accuracy and fluency, *and* scored at least -1 SD below the mean on at least two of the Queensland University Inventory of Literacy phonological awareness tasks (i.e. nonword reading, rhyming, phoneme detection/ segmentation/ manipulation and syllable identification; Dodd, Holm, Orelemans, & McCormick, 1996; Madelaine & Wheldall, 2002).

Ferguson et al. (2011) labelled children as diagnosed with APD based on parental reports of symptoms of APD (see Appendix A for details) without administration of behavioural AP tests.

To date only one study has assessed the overlap between APD and ASD symptoms in children. In Dawes & Bishop's (2010) study they also analysed scores on a subset of their initial samples using a questionnaire completed by parents that measured autistic spectrum features. Parents of children with APD rated their children higher on these features than parents of children with dyslexia. Based on the recommended cut-off score that identifies cases with ASD, 33% of children with APD met ASD criteria compared to none from the dyslexia group. Nonetheless, comparison of these two proportions was marginally non-significant (Dawes & Bishop, 2010). In conclusion, there are no primary symptoms that can undoubtedly be assigned to APD and this appears to be a problem that is universally recognised (Iliadou et al., 2017). This means that it is difficult to have an APD diagnosis without other co-occurring disorders, as evidenced previously. The relationship between APD and other disorders is still not fully understood and the use of multidisciplinary assessments to identify primary and secondary deficits may provide information to help understand why these co-occurrences happen (Chermak & Bellis, 2014).

1.4 APD diagnosis and diagnostic criteria

1.4.1 Diagnostic procedure and controversies

Assessment for APD should be performed by audiologists, even though referral may come from other professionals (e.g. General Practitioner [GP], Speech and Language Therapists; AAA, 2010; ASHA, 2005b; BSA, 2018). However, the BSA (2018) indicates that not all audiologists may be aware that listening problems may originate from other deficits (e.g. language impairments). Additionally, the professionals who refer children to APD clinics may or may not refer children depending on their awareness of APD and might even make a differential referral, again based on their knowledge and experience. This highlights the importance for standardised referral procedures for APD assessment, as the level of APD awareness by referring professionals can have an impact on APD prevalence, diagnosis and management. The position statements by the

ASHA, AAA and BSA all highlight that children need to be at least seven years old before they are tested for APD (AAA, 2010; ASHA, 2005b; BSA, 2011b). This is because younger children often show high variability in performance and present difficulties comprehending the tasks (AAA, 2010; ASHA, 200b). For children younger than seven, if any AP tests are administered they should be interpreted cautiously (AAA, 2010). For these young children behavioural checklists and screening tools can be used instead of AP behavioural tests. They also need to be followed-up regularly until they are able to perform behavioural tasks so that they are diagnosed as early as possible (AAA, 2010).

As a first step in an APD assessment, self-reports or parental reports in the form of questionnaires are employed as an initial screening tool (Hind, 2006) but are not part of the diagnostic tests. The case history of patients informs the clinician of the type of disorder or disorders the child might have (Bamiou & Iliadou, 2014). Case history should include information on genetic history (e.g. family history of ear or other problems), medical history (e.g. pregnancy, premature birth), language and communication abilities, educational, social and psychological aspects (ASHA, 2005b). The ASHA (2005b) and the BSA (2018; 2011b) position statements agree that reports/ questionnaires and case history are essential parts of the initial screening. However, there is no universal agreement on which specific screening tools should be used during screening, while the BSA (2018) maintains that there is the research need to develop a structured case history for children referred for APD assessment.

Baseline audiological tests comprising Pure Tone Audiometry (PTA), tympanogram (to test middle ear function) and otoacoustic emissions (to test inner ear function) are essential as they are required to verify normal peripheral hearing (Bamiou, Campbell, & Sirimanna, 2006). Furthermore, the auditory brainstem response and/ or acoustic reflexes help assess auditory nerve function (BSA, 2011b; Martin & Keith, 2009), while other electrophysiologic tests (e.g. the Mismatch Negativity [MMN] response) may be more time-consuming with parametric outcomes less investigated on specific diagnostic entities (Liasis et al., 2003; Roggia & Colares, 2008). There is an infrequent association of APD with some structural brain lesions (Iliadou et al., 2017) but in general the evidence that specific electrophysiologic or imaging tests could routinely identify APD is weak (BSA, 2018). As part of the general APD assessment, audiologists should consult reports and assessments from a multidisciplinary team including speech and language therapists, educational psychologists and special education teachers (Bamiou et al., 2006). These tests and reports, which cover areas not examined by audiologists, can better inform the diagnosis of APD and help characterise the nature of the child's deficits (Bamiou et al., 2006).

Following, are tests that make up the APD diagnostic test battery. These include temporal processing, frequency discrimination, dichotic listening, gaps-in-noise, binaural integration and reduced redundancy speech tests (ASHA, 2005b; Martin & Keith, 2009) and listening in spatialised noise tests (Cameron & Dillon, 2011). It should be noted that both the ASHA (2005b) and the BSA (2011a) recommend diagnoses based on normative scores on these AP tests. Nevertheless, there is no universally accepted diagnostic APD procedure and the debate on APD diagnosis is ongoing (BSA, 2018; Iliadou et al., 2017). The ASHA (2005b) suggests that a diagnosis of APD should be given when performance on at least two AP tests (described previously) falls -2 SDs from the mean or when one test falls -3 SDs from the mean or when one test is at -2 SDs from the mean and there is significant functional difficulty in auditory behaviours during test administration. While the first two conditions are reliant on measurable poor performance on standardised tests, the third condition allows subjectivity from the clinician. The AAA (2010) agrees on the first condition (i.e. APD diagnosis based on performance of 2 SD below the mean in at least two AP tests) but does not use the other two conditions described by the ASHA. In the BSA position statements (2011b; 2018) there is no mention of the specific cut-offs and minimum number of failed tests in order for APD diagnosis to be given. The 'European consensus' group on APD (Iliadou et al., 2017) though, which has representation from the UK, agrees with the AAA's (2010) set of cut-offs and criteria. The 'European consensus' group also highlights that poor performance on only one AP test or performance that falls below 3 SD from the mean are not considered adequate criteria to give an APD diagnosis (Iliadou et al., 2017). The disagreement on what criteria and cut-offs should be used to diagnose APD unavoidably impacts the diagnosis rate and accuracy, prevalence and management of the disorder and further emphasises the need for universally agreed APD criteria.

Moreover, some of the tests used to diagnose APD have been found to correlate with cognitive measures, such as attention and memory tests in samples of children suspected of APD, and mixed samples of children suspected of APD and typically developing children (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). These studies will be explored in detail at a later section but in short, their findings suggest that APD diagnostic criteria may require re-evaluation. A recent paper discusses ways in which audiologists can limit linguistic and cognitive influence during the APD diagnostic procedure (Chermak, Bamiou, Iliadou, & Musiek, 2017). The authors suggest the selection of AP tests that minimise the effect of these factors. They recommend that the cognitive tests to be administered by the multidisciplinary team should be discussed by the whole team, so that clinicians better understand the influence of cognition and language on the AP performance of patients (Chermak et al., 2017). Moreover, even though both the ASHA (2005b) and the BSA (2018) advocate for the involvement of a multidisciplinary team in APD assessments, the role and nature of the multidisciplinary team should be clearly defined and accepted universally.

A recent systematic review evaluated the APD clinical practice guidelines from different organisations, including the ones from the BSA and ASHA (Heine & O’Halloran, 2015). Based on the Appraisal of Guidelines and Research and Evaluation Tool, it was concluded that the BSA position statement on APD received the highest scores among five practice guidelines (i.e. the other four were the AAA Clinical Practice Guidelines, ASHA Technical Report, Canadian Guidelines on APD in Children and Adults, and the Colorado Department of Education-APD). The authors, though, stressed the need for changes in all five position statements, including the one from the BSA. Recommendations were given to all organisations to better define APD, to provide management recommendations based on studies with improved level of evidence and to reconsider the involvement of stakeholders (Heine & O’Halloran, 2015). The need for a universal and more clearly defined position statement on APD remains. Research should focus on illuminating the relationship between APD and attention and on assessing the usefulness of clinical diagnostic procedures. To that end, the thesis aims to examine whether a differential diagnostic APD test battery which includes measures of attention can provide useful information in terms of APD management. The research aims of this study will be presented in detail in Chapter 5.

1.4.2 Diagnostic criteria in research studies

Research studies and research groups also use differential diagnostic criteria, reflecting the fact that there are no universally accepted APD criteria. The table in Appendix A presents the criteria that each research study reviewed in this thesis has used to identify children as having APD. These studies have either included children with APD in their entire samples or have used APD criteria to identify a subset of APD-diagnosable children from their initial mixed samples. It can be observed that there is no consistency in the criteria that APD studies use. The following points show the inconsistencies in the APD diagnostic criteria used across APD studies:

- Some studies did not mention the specific AP tests they used or which cut-off criteria they followed. They relied on only reporting that children were diagnosed with APD from the APD clinics.
- There is inconsistency in the number of AP tests used between studies, as well as inconsistency in the AP tests per se that were administered.
- Some studies did not clarify whether they used normative data to identified children as having APD.
- Some studies required native language speakers (of the country the study was carried out) in order to be eligible for inclusion, but not all studies set that as a criterion.
- Not all studies tested all octave frequencies between 250-8000 Hz (i.e. some checked octave frequencies between 500-4000 Hz), while a few studies did not refer to middle ear function check or otoscopic tests.
- Poor performance on non-verbal IQ tests was not mentioned as an exclusion criterion in all studies.

- Children with co-occurring disorders were not excluded in some studies. In addition, there was inconsistency between studies on which co-occurring disorders would lead to exclusion from an APD diagnosis.

The use of different criteria to diagnose APD by research studies is problematic, as this has an impact on the findings and interpretations. Different APD criteria may produce different results, which in turn will be interpreted differently depending on the chosen criteria. As a result, it appears that each research or clinical institution draws its own conclusions on how APD should be characterised and managed. Therefore, this impedes the development of a universally accepted understanding of APD and its relationship to other cognitive factors and other disorders.

The current work uses APD criteria used at the Great Ormond Street Hospital (GOSH) in London. These criteria closely follow the ASHA (2005b) position statement but are not identical. They are described in detail in Section 2.2.1 (under Inclusion and exclusion criteria). This means that findings from the study cannot be generalised to APD populations and would only apply to other clinical populations who have used these exact criteria to diagnose APD in children.

1.4.3 APD prevalence

An estimated prevalence of APD in children falls between 2-3% (Chermak & Musiek, 1997), but this has been an outdated estimation, which was based on the two authors' clinical experience and on prevalence information from other co-occurring conditions. Most recent estimations are still using the above percentage to report prevalence in children (Bamiou et al., 2001; Bellis & Bellis, 2015), highlighting the need for a better estimation. Another prevalence rate for children with APD often reported is between 0.5-1.0%, while child referrals to audiological services who obtained normal hearing thresholds is reported to be at 5% (Hind et al., 2011). The difference in the estimated numbers may be due to the differential views on APD and to the fact that there are no universally accepted diagnostic criteria (as discussed in the previous sections). Thus, without standardised criteria for assessment it would not be possible to draw an accurate estimate.

1.5 Attention and APD

Attention is defined as the selection of information from a number of available sensory information, which are then further processed (Broadbent, 1958; Hahn et al., 2008). Sustained attention is the vigilant focus on (often unpredictably occurring) stimuli (Sarter, Givens, & Bruno, 2001). Sustained attention is considered a basic attentional function that influences how efficiently other higher aspects of attention function (e.g. selective, divided attention; Sarter et al., 2001). Selective attention is the process where the focus is on a specific, task-relevant input over task-irrelevant distractors (Parasuraman, 1998). On the other hand, during divided attention the individual rapidly splits or shifts focus between different stimuli that are presented simultaneously (Parasuraman, 1998).

1.5.1 Sustained auditory attention

Children diagnosed with APD (aged 7-17) have been shown to perform more poorly in tasks of Sustained Auditory Attention (Sus-AA) compared to children suspected of but not meeting APD criteria (Allen & Allan, 2014). In this study, Sus-AA was measured through the use of the Test of Everyday Attention for Children (TEACh; Manly, Robertson, Anderson, & Nimmo-Smith, 1999). The authors used a composite score for Sus-AA (Allen & Allan, 2014), as the TEACh has three different subtests to measure this type of attention (for a detailed description of the TEACh tasks please see Section 2.2.3 under Attention in Chapter 2). However, the use of a composite score may not reflect true Sus-AA deficits, as two of the three Sus-AA subtests were found to correlate with cognitive tests (e.g. the Wechsler Intelligence Scale for Children [WISC] III, the Block Design test [assessing spatial visualisation and motor skills], and the Object Assembly test [assessing visual-motor and problem-solving skills]; Manly et al., 1999). This means that children with APD may face cognitive deficits (along with Sus-AA deficits) compared to children without APD.

Moreover, in a study examining the relationship between Sus-AA and AP tests, of the 58 children that had been identified with APD (aged 7-12), 20 (34%) had Sus-AA scores 2 SDs below the normative mean (Gyldenkærne et al., 2014). Of the entire sample of 119 children suspected of

APD, 51 (43%) had a deficit in Sus-AA. Additionally, the researchers used a correlation analysis on the entire sample of 119 children suspected of APD, finding statistically significant correlations between Sus-AA and the Dichotic Digits Test (DDT) and Frequency Pattern Test (FPT)[8], both part of a standard APD diagnostic test battery (Gyldenkærne et al., 2014). Tomlin et al. (2015) also ran a correlation analysis on a sample of 105 children suspected of APD, aged 7-12, examining the relationship between the same sustained attention test and AP tests. Nevertheless, instead of using the Sus-AA score on its own (as was done in Gyldenkaerne et al., 2014) they used the composite scores from both the Sus-AA and Sustained Visual Attention (Sus-VA) tests, finding similar significant correlations[9] (Tomlin et al., 2015). Because of the weak to moderate correlations found in both studies, the authors concluded that Sus-AA and APD are not completely independent from each other and that the former could influence but not solely determine AP deficits (Gyldenkærne et al., 2014; Tomlin et al., 2015).

On the other hand, Sharma et al. (2009) interpreted their findings of overlap between APD and Attention Deficits (ADs) differently. Their sample of 68 children suspected of APD, aged 7-12, confirmed that 49 (72%) met criteria for an APD diagnosis (see Appendix A for details), 26 (30%) met criteria for APD diagnosis without ADs and 9 children (8%) had ADs without APD diagnosis (Sharma et al., 2009). Of the 49 children diagnosed with APD, 30 (59%) also had an AD. In addition, 36 of the 68 children (53%) had a Sus-AA deficit, a difference of 10% compared to the 43% found in Gyldenkærne et al. (2014). However, Sharma et al. (2009) mentioned that children were classified as having ADs when the attention score was at least 1 SD below the mean, whereas in Gyldenkærne et al.'s (2014) study the score had to be at least 2 SDs below the mean for an AD to be identified. Moreover, the researchers ran a partial correlation analysis, controlling for age and memory, finding significant associations between the same Sus-AA the two previous studies used and the DDT (averaged across the two ears) and between Sus-AA and the FPT[10] (Sharma et

---

[8] $r = .29$ between Sus-AA and DDT left ear, $r = .25$ between Sus-AA and DDT right ear and $r = .33$ between Sus-AA and FPT.
[9] $r = .22$ between attention score and DDT left, $r = .22$ between attention score and DDT right and $r = .27$ between attention score and FPT.
[10] $r = .30$ between Sus-AA and DDT and $r = .44$ between Sus-AA and FPT.

al., 2009). These are slightly stronger correlations compared to the ones from the two previous studies perhaps because of the controlled covariates of age and memory used here. The same conclusions though were drawn; Sus-AA only explained a small amount of variance and it is possible that APD and ADs co-occurred independently in this subgroup (Sharma et al., 2009).

1.5.2 Selective auditory attention

There is scarcity of studies examining Selective Auditory Attention (Sel-AA) in children with APD. The only study that uses a Sel-AA test is one including a sample of 10 children with dichotic listening deficits, aged 8-13, also suspected of (but not diagnosed with) APD (Martin et al., 2007). Behavioural performance and electrophysiologic measures from this group were compared against those from a control group of 10 aged-matched typically developing children. The tasks used were a Sel-AA and a divided auditory attention task, both using a ‘same-different’ judgement paradigm. The Sel-AA task required the subjects to attend only to one of the ears in a dichotic listening task, while in the divided attention task they had to attend and respond to both dichotic stimuli. The electrophysiologic outcome measures were the N400 component, aimed to reflect semantic analysis of the ‘same-different’ judgement, and a late-positive component (between 500-1100 msec) expected to represent cognitive aspects of stimulus analysis (Martin et al., 2007).

Results showed that the late-positive component was significantly smaller in amplitude and more delayed for the experimental group as compared to the control group in the divided attention tasks, whereas when children were instructed to direct attention to one ear (i.e. Sel-AA) these differences disappeared (Martin et al., 2007). On the other hand, significant group differences in the N400 were observed in both tasks, with the dichotic deficits group having larger negative Event Related Potential (ERP) amplitudes than the controls. As a control diotic task (also employed) did not show group differences in the N400, authors argued that the N400 component could reflect attentional resource allocation problems during the Sel-AA task in the dichotic deficits group. However, it is important to note that there were no indications that this group of children had APD, other than the authors’ suspicion. Parents had concerns about their child’s academic performance and attention, and children were tested to ensure normal peripheral hearing and were assessed on a dichotic listening test (Martin et al., 2007). These tests are only a few of the tests in an APD test battery with key AP tests missing that are essential for diagnosing APD. Thus, there is the need for trials to study Sel-AA in samples of children with APD diagnosis so that the relationship of Sel-AA and APD is understood.

The reason Sel-AA is important to investigate in children with APD is the possible link it has to SIN perception. There is a series of studies that demonstrate how Sel-AA influences SIN ability (Mesgarani & Chang, 2012; O'Sullivan et al., 2015; Puvvada & Simon, 2017; Zion Golumbic et al., 2013). These studies are discussed in more detail in Section 1.8.1, but their premise is that being able to selectively attend to the target stimulus while supressing distractors (an action constituting Sel-AA) enhances SIN performance.

1.5.3 Divided attention

Chermak, Tucker, and Seikel (2002) asked audiologists diagnosing APD which were the most common behavioural symptoms they observed in children with APD. Divided Auditory Attention (Div-AA) was ranked as one of the most common behaviours (Chermak et al., 2002). On a scale from 1 (never observed) to 5 (always observed), Div-AA was rated an average of 3.76, which was more than 1 SD above the mean (Chermak et al., 2002). However, this survey did not incorporate behavioural measures and just as for Sel-AA, research on behavioural tests of Div-AA in children with APD is limited.

The only study examining Div-AA in a group of children with dichotic deficits (and suspected of APD), was the one mentioned earlier by Martin et al. (2007). As previously noted, in the Div-AA task the dichotic deficits group had significantly smaller amplitudes and more delayed ERPs (i.e. 800-1000 msec) in the late-positive component compared to a typically developing control group (who had 600-800 msec). Moreover, significantly larger N400 amplitudes were observed in the experimental group during this task compared to the control group. This difference is said to reflect an additional cognitive effort to allocate resources in the dichotic deficits group (Martin et al., 2007). The experimental group also had behaviourally poorer accuracy in the Div-AA task compared to the controls but only for the left ear, a finding consistent with the right-ear advantage

(Martin et al., 2007). Nonetheless, the issue that was highlighted previously remains; children in the experimental group did not go through any formal APD diagnostic testing and were only suspected of APD. Thus, the findings cannot be generalised to the APD population.

In terms of Divided Auditory-Visual Attention (Div-AVA) in children with APD, again not much has been researched up to this point. What is known for normal listening adults is that adding a visual task alongside an auditory one, thus creating a bimodal paradigm, reduces performance in the auditory task compared to when performing the auditory task alone (Mattys & Palmer, 2015). But there have been no studies to examine the performance of children with APD in bimodal tasks or to compare their performance to that of typically developing children. One study looked at the ability of 18 different measures (including a Div-AVA measure) to predict APD (Lam & Sanchez, 2007). The rest of the measures in that study comprised a competing sentences test, an auditory figure ground test, an auditory performance questionnaire and other attention tests. These measures were compared with results from the APD diagnostic tests (that were used to diagnose APD in the study) to examine their sensitivity and specificity as APD screening tools. The Div-AVA measure was not found to be predictive of APD and only the competing sentences (left ear) measure was significantly predicting APD (Lam & Sanchez, 2007). However, the study recruited children (aged 7-10) that also had an ADHD diagnosis. In addition, the threshold for the PTA test the authors had used was set at 25 dB and not at 20 dB which is considered the standard (BSA, 2012). These issues make the study's results difficult to generalise to the APD population. Therefore, divided attention and its relationship to APD has not been adequately studied.

1.5.4 Visual attention

Children diagnosed with APD were found to have poorer scores in tasks of Selective Visual Attention (Sel-VA) as opposed to children suspected of but not meeting diagnostic APD criteria (Allen & Allan, 2014). Two subtests of the TEACh test that measure Sel-VA were combined to create a composite Sel-VA score (Allen & Allan, 2014; Manly et al., 1999). But the use of a composite score may not accurately indicate Sel-VA deficits, as one of the TEACh Sel-VA subtests was found to correlate with the Object Assembly test (assessing visual-motor and problem-solving skills; Manly et al., 1999). Therefore, it is not clear if the observed deficits in the composite Sel-VA score that children with APD have is due to genuine Sel-VA deficits or if it is influenced by deficits in visual-motor or problem-solving skills.

Other studies have used another type of visual attention (i.e. Sus-VA) to study its relationship to AP skills in children suspected of APD (Gyldenkærne et al., 2014; Sharma et al., 2009). In Gyldenkærne et al.'s (2014) study, from the entire sample of 119 children suspected of APD, 25 (21%) had a deficit (defined as -2 SDs from the mean) in Sus-VA. From the 58 children in the study identified with APD, 39 (67%) had sustained attention deficits (visual and/ or auditory) as well, but only 5 (9%) had a deficit in Sus-VA alone (i.e. without a Sus-AA deficit). Sharma et al. (2009) found almost a doubled proportion of Sus-VA deficits in their sample of 68 children suspected of APD. However, as discussed previously this is probably due to their classification of an attentional problem when the score was at least -1 SD from the mean. Specifically, they found that 27 of the 68 children (40%) had a Sus-VA deficit.

Gyldenkærne et al. (2014) also examined the correlation between Sus-VA and AP measures. They found weak, albeit significant, correlations between Sus-VA and the DDT and between Sus-VA and the FPT[11]. They also looked at the correlation between Sus-AA and Sus-VA, revealing a significant moderate correlation[12] which they interpreted as an indication of a common factor shared between the auditory and visual modality. They supported that this may explain the existence of the (weak) correlations they found between visual (i.e. Sus-VA) and auditory tasks (i.e. DDT and FPT). The fact that other possible covariates (e.g. IQ performance or memory) were not controlled in the analysis could also explain the significant association between Sus-VA and AP measures.

1.5.5 Suspicion of APD

Most studies used in this review on attention and APD had mixed samples of children suspected of and diagnosed with APD. Recruitment of children in these studies was usually based on

---

[11] r = .25 between Sus-VA and DDT left, r = .27 between Sus-VA and DDT right and r = .28 between Sus-VA and FPT.

[12] r = .55 between Sus-AA and Sus-VA.

parental reports of suspected listening difficulties and/ or referrals based on suspicion of APD, while a few of the studies' samples included children with ADHD (Gyldenkærne et al., 2014; Lam & Sanchez, 2007; Martin et al., 2007; Sharma et al., 2009; Tomlin et al., 2015). This is a potential shortcoming when interpreting their results, as the correlation and regression analyses are making use of this mixed sample, while attempting to draw conclusions about APD from APD-suspected findings (Iliadou, Sirimanna, & Bamiou, 2016). It has been shown that the relationship between auditory skills and language skills varies between different study samples (Grube, Cooper, Kumar, Kelly, & Griffiths, 2014; Kuppen & Goswami, 2016) and one can argue that the relationship between auditory and attention skills might also vary depending on the studied population. It is therefore essential for trials to examine the relationship of attention and AP skills in APD-diagnosed samples. Another issue to be considered is the lack of a common APD diagnostic procedure used in these studies (see Section 1.4.2 earlier and Appendix A for more details). For instance, in their sample suspected of APD Gyldenkærne et al. (2014) identified children as having APD by using at least one AP test with -2 SDs from the mean. Sharma et al. (2009) on the other hand, identified children as having APD in their sample when they either had at least two AP test scores at -2 SDs from the mean or when they had one AP test score -3 SDs from the mean. Thus, the choice of the diagnostic criteria can influence the proportion of children diagnosed with APD and in consequence affect the findings and interpretation. This is a broader problem in the field of APD which calls for a universal APD diagnostic procedure.

1.5.6 Summary

Summing up, there is evidence for deficits in Sus-AA in children diagnosed with APD (Allen & Allan, 2014), but thus far studies indicate that Sus-AA explains only a small amount of the variance of APD (see previous Section 1.5.1 for details; Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). The relationships between measures of other types of auditory attention (Sel-AA, Div-AA, Div-AVA) and APD have not been sufficiently examined yet, highlighting the need for future research on the field. While Sus-VA is shown to have weak correlations with the DDT and FPT (Gyldenkærne et al., 2014), its contribution to APD remains unclear. At the same time, Sel-VA deficits may be present in children with APD (Allen & Allan, 2014). Further studies on visual attention and APD could provide evidence as to whether children with APD have global attentional difficulties. To address the limitations discussed here, the study in Chapter 5 will include in its design different measures of attention (auditory and visual) to examine their relationship to AP tests. The research questions for that study are broadly discussed in Section 1.9.1 and are presented in detail in Chapter 5.

1.6 Memory and APD

Working memory is a cognitive system that temporarily stores information and also temporarily keeps this information accessible to be used by a mental task (e.g. immediate recall, language comprehension, reasoning or problem-solving; Baddeley, 1992; N. Cowan, 1998). There are a number of studies showing worse Auditory Working Memory (AWM) scores in children with APD compared to typically developing peers (Lotfi, Mehrkian, Moossavi, Zadeh, & Sadjedi, 2016; Maerlender, Wallis, & Isquith, 2004; Moossavi, Mehrkian, Lotfi, Faghihzadeh, & Sadjedi, 2014; Tomlin et al., 2015). A retrospective review of medical records looked at performance in the digit span (an AWM measure) from children diagnosed with APD and children suspected of but not meeting criteria for APD diagnosis, all aged 7-15 (Maerlender et al., 2004). The authors found that children who met criteria for APD diagnosis had significantly poorer scores in the digit span forward and backward test than children who were suspected of but not diagnosed with APD. Additionally, children that met APD diagnostic criteria had a mean scaled score of 7.81 (which is close to the 16th percentile/ scaled score of 7, that identifies borderline performance[13]), whereas children that did not meet APD criteria had a mean scaled score of 9.65 (which is closer to the normative mean scaled score of 10; Maerlender et al., 2004). Even though the authors did not set an inclusion threshold for IQ scores, hence including cases with an IQ score as low as 63, they did compare the IQ performance between the two groups and found no differences. The researchers did not control for other disorders, though, such as reading impairments or ADHD

---

[13] Borderline performance is considered when the scaled score is ≤ 7 and > 2.

(Maerlender et al., 2004). Possible presentation of these disorders may have had an impact on the scores in the memory test.

Two studies used the forward and backward digit span and a non-word repetition task to measure working memory capacity in 15 children with APD, aged 9-11, and 20 typically developing age-matched controls (Lotfi, Mehrkian, et al., 2016), and in 17 children with APD and 20 typically developing controls, all aged 9-11 (Moossavi et al., 2014). Results from both studies showed significantly poorer scores in the clinical groups compared to the control groups in all three memory tasks (Lotfi, Mehrkian, et al., 2016; Moossavi et al., 2014). But the APD diagnostic criteria used in these two studies were incomplete and did not follow standard APD diagnostic protocols. Children were administered a DDT, pitch pattern and a selective attention test, while other AP tests, such as Gaps-in-Noise (GIN), SIN tests and the LiSN-S were not used. In addition, children were recruited in the studies if they failed all three of the administered tests. This contrasts all APD protocols (by the AAA, ASHA, 'European consensus' group) stated earlier in Section 1.4.1, which look for at least two failed AP tests (AAA, 2010; Iliadou et al., 2017), or one failed AP test by -3 SDs from the mean (i.e. the ASHA position statement, [ASHA, 2005b]). Therefore, it is unclear how these samples of children met standard APD criteria, while it was also not clarified whether children had any other deficits.

Another study compared 105 children suspected of APD, aged 7-13, and 50 typically developing controls, aged 7-12, in various cognitive measures, including AWM (Tomlin et al., 2015). A composite cognitive score comprising AWM, attention and nonverbal IQ was significantly worse in the clinical group in comparison to the typically developing control group. Dividing the clinical group between children who were diagnosed with APD (*n* = 36) and those who were not (*n* = 69) and comparing their scores, revealed that the former group had significantly worse composite scores compared to the non-APD group (Tomlin et al., 2015). When the authors ran a post-hoc test to examine each cognitive test individually, they failed to find significant differences between children with APD and typically developing children in the AWM test, concluding that only general cognition was impaired in the clinical group. The same study also conducted correlation analyses finding significant (but weak) correlations between the AWM task and the DDT, FPT and LiSN-S Low-cue SRT condition[14] (Tomlin et al., 2015). All three tests are used for diagnosing APD, and AWM could thus be explaining a small variance of APD.

In a task measuring short-term auditory memory in children suspected of APD, the authors recorded scores of -1 SD from the mean or worse in 59% of the total sample of 68 (Sharma et al., 2009). Even though a score of -2 SDs from the mean or worse is usually used to characterise disordered performance, percentages for this threshold were not reported in the study. Regression analyses demonstrated that age, memory and attention scores accounted for 21% of the variance in the FPT, while the same three variables explained 28.5% of the variance in the DDT (Sharma et al., 2009). Individual correlation analyses were then run, revealing a significant association only between the memory task and the FPT and not between memory and the DDT. This could mean that working memory is only influencing (or is influenced by) the FPT and not the DDT. A shortcoming of the study is that it did not exclude cases of children with dyspraxia, ADHD, ASD or reading disabilities, making it difficult to draw safe conclusions regarding the way memory interacts with AP deficits. The researchers did not exclude cases with DLD or reading disorder either, however part of their objectives was to examine the overlap between APD, DLD and reading disorder (see Section 1.3.3 for more details on this; Sharma et al., 2009).

Finally, a study was conducted in order to examine the clinical usefulness of a questionnaire with attention and memory subscales (Iliadou & Bamiou, 2012). The Children’s Auditory Performance Scale (CHAPS) questionnaire was completed by parents of three different groups of children; suspected and diagnosed with APD, suspected but not diagnosed with APD and typically developing control children, all aged between 11-13. Results showed significantly lower scores in the APD group compared to the control in all subscales including attention and memory. In addition, the APD group had lower scores in the subscales of attention and memory when compared to the group that was suspected of APD but did not meet APD criteria. Running a

---

[14] $r = .39$ between AWM and DDT left, $r = .41$ between AWM and DDT right, $r = .30$ between AWM and the average FPT and $r = .21$ between AWM and the LiSN-S Low-cue SRT.

Spearman's rho correlation analysis, the authors further examined the relationship between the CHAPS subscales and individual AP tests. They revealed a strong correlation between the memory scale and the duration pattern test (right ear), a moderate correlation between the attention scale and the duration pattern (both ears) and a moderate correlation between memory/ attention and the DDT (both ears; Iliadou & Bamiou, 2012). But as previous research cautions against the use of CHAPS as a diagnostic or stand-alone screening tool (Sharma et al., 2009; Wilson et al., 2011), findings from the CHAPS should not be generalised to APD populations.

In conclusion, measures of AWM are found to explain a small variance of APD (Sharma et al., 2009; Tomlin et al., 2015). Children with APD are shown to perform worse in memory tasks compared to their typically developing peers (Lotfi, Mehrkian, et al., 2016; Moossavi et al., 2014) and compared to children suspected of APD but not meeting APD criteria (Maerlender et al., 2004; Tomlin et al., 2015). Finally, these differences are also reflected in parental questionnaires (Iliadou & Bamiou, 2012).

1.7 Management of APD and RMHAs

1.7.1 Brief overview of APD management

Irrespective of the controversies in diagnosing APD, the fact remains that children given a diagnosis of APD struggle to listen in the classroom (Flexer, 2014) and are identified with other cognitive deficits (Allen & Allan, 2014; Moore et al., 2010). In order for children to receive help in their learning and development, research evidence of specific management approaches that improve children's performance should be found. In this section a general overview of the main APD management approaches will be first given. This will be followed by a review of evidence on the effectiveness of RMHAs as a management approach, which is the central point of this thesis (Chapters 2, 3 and 4). This discussion will lead up to the research questions and the main aims of the thesis.

Bottom-up and top-down theories drive management strategies in APD. Management of APD can be broken down into (1) environmental modifications, (2) bottom-up approaches and (3) top-down cognitive approaches (ASHA, 2005b). Environmental modifications include preferential seating, reduction in reverberation (through sound insulation techniques), teacher adaptations (to offer better access to speech stimuli), visual aids and sound-field amplification systems (ASHA, 2005b; BSA, 2011b; Toe, 2009). These modifications usually benefit the entire classroom and not just the child with listening difficulties (Toe, 2009). Bottom-up approaches focus on enhancing the auditory stimulus for the child with APD (ASHA, 2005b) and they include Auditory Training (AT) such as temporal and spectral pure-tone discrimination training (Sharma et al., 2012), as well as the use of personal assistive listening devices, such as personal RMHAs (BSA, 2011b).

On the other hand, top-down strategies include training of higher-order functions by using cognitive, metacognitive and language strategies (ASHA, 2005b). These approaches aim to improve the auditory input by teaching children how to listen, pay attention and how to use their working memory (BSA, 2011b). Top-down strategies are usually offered in the form of formal or informal AT (BSA, 2011b) and it is believed that they could give support to a dysfunctional auditory processing system (Bellis & Bellis, 2015). Administration of AT programmes should be specific to the child’s needs. As Cameron et al. (2012) revealed, spatial listening in children suspected of APD only improved using a computer training programme with listening-in-spatialised-noise tasks and not with monaural speech-in-noise tasks. There is an accumulating research interest to investigate whether AT strategies improve listening and related skills in children with APD. While the topic could be discussed extensively, the focus of this thesis is on RMHAs as an intervention for children with APD.

1.7.2 Remote microphone hearing aids

Remote microphone hearing aids are another management approach aiming at improving the acoustic signal for the child with APD. Remote microphone hearing aids are wireless listening systems, consisting of a microphone worn by the speaker (i.e. the teacher) and ear receivers worn by the user (e.g. children; BSA, 2011b). The microphone picks up the speaker’s voice and then wirelessly transmits it into the ear receivers worn by the listener (Toe, 2009). In this way, the negative effects of ambient noise, distance from the speaker and reverberation are reduced (W. J. Keith & Purdy, 2014; Kuk et al., 2008). A Mini-mic GN ReSound model is depicted in Figure 1.4. The term RMHAs is an umbrella term that is used in this thesis and it refers to all systems using wireless transmission of a speaker's voice into ear receivers, including Frequency Modulation (FM) systems.

Figure 1.4 – Mini-mic RMHA
A RMHA by GN ReSound (hearing aid manufacturer) consisting of a left-ear receiver (on the left) and a microphone transmitter (on the right). RMHA: Remote Microphone Hearing Aid.

Personal RMHAs have replaced to a large degree the sound-field FM system technology (Bamiou et al., 2006). The latter system follows the same concept as personal RMHAs, with the difference being that the child does not need to wear ear receivers, as the teacher's voice is transmitted through sound speakers placed around the classroom (Schafer & Kleineck, 2009). Not having to wear ear receivers is considered advantageous by some children (Johnston et al., 2009), but personal RMHAs offer other advantages over sound-field settings. They offer a consistent and more direct signal, they are effective in limiting reverberation time, are easily portable to different environments/ classrooms and they have individualised programming based on the child's needs (BSA, 2011b; Schafer et al., 2013).

As a management approach, personal RMHAs are recommended to children with APD who exhibit difficulties in SIN perception (BSA, 2011b). But the effects personal RMHAs have on children with APD have only been examined by a handful of studies. A systematic review on the effectiveness of RMHAs as a management approach for APD was conducted by Lemos et al. (2009). This review did not include meta-analysis hence no statistical analysis was carried out.

The review looked for evidence up until 2008 and only 19 studies matched their search criteria, all of which were of low level of evidence (i.e. descriptive studies, case studies or specialists' opinions). The authors concluded that there was a lack of higher-level evidence to support benefits of RMHA use on children with APD (Lemos et al., 2009). Since 2008, only one RCT was conducted (Sharma et al., 2012), while other intervention studies either did not use randomised designs or had other methodological problems (Johnston et al., 2009; Kuk et al., 2008; Umat et al., 2011). These studies are summarised in Table 1.1 and are discussed in detail in the upcoming sections.

Table 1.1 – Studies on APD and RMHAs/ hearing aids

An outline of all studies since 2008 that included children with APD who used RMHAs or hearing aids as interventions. ADHD: Attention Deficit Hyperactivity Disorder, APD: Auditory Processing Disorder, AT: Auditory Training, HINT: Hearing in Noise Test, RMHA: Remote Microphone Hearing Aid, SIN: Speech-in-Noise, SNR: Signal-to-Noise Ratio.

Study	Intervention group	Control group	Co-occurring disorders	Study design	Intervention	Intervention period	Results
Johnston et al. (2009)	10 children with APD (8 boys, 2 girls). Ages: 8-15 Mean age: 11 years, 8 months.	13 typically developing children (9 boys, 4 girls) Ages: 8-13 Mean age: 10 years, 6 months.	None stated.	Non-randomised controlled trial. Typically developing controls vs. children with APD.	RMHAs used by the APD group only. Controls did not receive any form of intervention.	5-month intervention. RMHAs used in classroom daily. Participants also encouraged to use the system at home. Tests on APD group at baseline and 5 months aided and unaided. Tests on controls only at baseline aided and unaided.	Statistically significant improved SNR scores in HINT test for the APD vs control group in the aided condition in noise. Statistically significant improved academic and psychosocial status for the APD group based on questionnaires from teachers and parents.
Kuk et al. (2008)	14 children with APD (8 boys, 6 girls). Ages: 7-11 Mean age: 9 years.	No control group.	5 children also had ADHD.	Single-blind, longitudinal descriptive study. Subjects served as their own controls.	Hearing aids with directional microphone and noise reduction giving a 10 dB gain.	6-month intervention. Hearing aids were used as much as possible in all environments (school, home, other settings). Tests at baseline aided and unaided, and aided only at 2 weeks, 3 months, 6 months.	Statistically significant improved SIN scores at 2 weeks, 6 months (not at 3 months). Fewer errors but not statistically significant differences in scores in the Auditory continuous performance test. No statistically significant change in parental questionnaire scores on auditory performance.
Sharma et al. (2012)	4 intervention groups. Group A: 12 children with APD (5 boys, 7 girls), Mean age: 9 years, 8 months Group B: 10 children with APD (7 boys, 3 girls), Mean age: 10 years, 9 months Group C: 12 children with APD (10 boys, 2	Group E: 11 children with APD (8 boys, 4 girls), Mean age: 8 years, 8 months.	62% both reading disorders and language impairment, 18% only reading disorders, 16% only language impairment. 4% only APD.	Randomised controlled trial.	Group A: Bottom-up AT. Group B: Bottom-up AT and RMHA. Group C: Top-down AT. Group D: Top-down AT and RMHA.	6-week intervention. AT: 15-min daily homework, weekly 1-hour sessions with therapist at clinic and 1-2 hours every week of working with parents at home. RMHA: The system was used only at school. Tests at baseline and 6 weeks unaided.	Two outcome measures (SIN and spoken word assessment) showed no statistically significant improvement between intervention groups and control group. Several outcome measures showed statistically significant improvement only in the two AT training groups. Three outcomes measures (2 phonological awareness measures and 1 language measure) showed statistically significant improvement only for the two

	girls), Mean age: 9 years, 6 months. Group D: 9 children with APD (6 boys, 3 girls), Mean age: 10 years, 1 month.				Group E: No interventions used.		AT plus RMHA groups.
Umat et al. (2011)	2 groups of intervention. Group A: 19 children with APD, Age: 7-9. Group B: 19 children with APD, Age: 7-9.	Group C: 15 children with APD, Age: 7-9.	Children with ADHD symptoms were excluded. No other disorders stated.	Longitudinal study.	Group A: Unilateral RMHA use. Group B: Bilateral RMHA use. Group C: No interventions used.	12-week intervention. RMHAs used only during classroom time 4-5 hours daily. Tests at baseline, 12 weeks and 1 year after end of study (no RMHA use during this year) aided and unaided.	No interaction between group and time mentioned for any of the three conditions. For working memory subtest of the Rey auditory verbal learning test, main effect of time was statistically improved at 12 weeks and 1 year compared to baseline. Main effect of group showed no statistical significance in scores between groups. For best learning condition of the Rey auditory verbal learning test, main effect of time was statistically improved at 1 year compared to baseline. Main effect of group showed no statistical significance in scores between groups. For retention of information subtest of the Rey auditory verbal learning test, main effect of time was statistically improved at 1 year compared to baseline. Main effect of group was not found statistically significant for this factor.

Effects of RMHAs on speech perception in noise and language skills

Auditory processing disorder is usually characterised by poor SIN perception (W. J. Keith & Purdy, 2014). It is expected that the use of a RMHA will bring an improved signal to the child's ears, bypassing the negative effects of poor room acoustics (W. J. Keith & Purdy, 2014). Results from Johnston et al.'s (2009) study confirm this, as they found a gain of 10 dB in SNR when comparing the SIN scores of children with APD, aged 8-15, in the aided condition (i.e. RMHA on) with the scores of the same children in the unaided condition (i.e. RMHA off). But the long-term effects of RMHA use in SIN tasks have not been properly looked at yet. While Johnston et al. (2009) report that an enhanced auditory system was developed after a 5-month use of RMHAs, that does not coincide with their actual findings. Examining the unaided conditions when wishing to test for long-term effects is important, as that will reveal possible lasting effects or give indications to neuroplastic changes[15]. But in Johnston et al.'s (2009) study the increase in SNR in the Hearing in Noise Test (HINT) was 0.7 dB from pre to post in the unaided conditions and this marginal change was non-significant. Moreover, comparisons were only made between post-intervention scores of the APD group and the pre-intervention scores of the age-matched control group, as the latter group was only tested at baseline and was never followed up (Johnston et al., 2009). Hence, their claims of long-term benefits on SIN performance following a 5-month RMHA intervention are not valid and the possibility that the improved scores were due to age maturation with no contribution from RMHA use cannot be ruled out.

With the study being a non-randomised trial, without having an APD control group, its findings are of low level of evidence (III; J. G. Wright, Swiontkowski, & Heckman, 2003). Adding to this, a sample size of only 10 children with APD is small, while their age range is rather varied (8-15 years old; Johnston et al., 2009). Sharma et al. (2012) used the same SIN test (i.e. the HINT) as Johnston et al. (2009), finding no improvement in any of the groups, with children aged 7-13 (including two RMHA groups), following a 6-week intervention (see Table 1.1 for more details).

---

[15] Neuroplasticity (or plasticity) describes changes in the nervous system and the brain due to internal or external stimuli and sensory experiences (Pascual-Leone, Amedi, Fregni, & Merabet, 2005; Tremblay, 2007).

These two studies give indications that after 6 weeks (Sharma et al., 2012) and after 5 months (Johnston et al., 2009) of using RMHAs, the HINT scores (in the unaided condition) did not improve in children with APD. However, Sharma et al.'s (2012) RMHA groups also used AT programmes, making it difficult to distinguish the effect of the two interventions. Sharma et al. (2012) did have AT-only groups that did not have significant improvement in the HINT scores. One can claim that since the addition of RMHA (in the AT plus RMHA groups) did not bring any improvement in the AT plus RMHA groups, that RMHAs did not add any benefit. But there is the possibility that a combination of interventions (in this case AT plus RMHA) might not be as beneficial as using only one of them separately (i.e. RMHA). It is therefore unclear from Sharma et al.'s (2012) study whether RMHA have any positive effects on the HINT test.

An earlier study by Kuk et al. (2008) on SIN perception following a 6-month use of (mild-gain) hearing aids with a directional microphone and noise reduction by children with APD, aged, 7-11, noted improvements at different time points in the aided-noise condition (see Table 1.1). They used the Northwestern University word-list as their target signal coming from the front and speech-shaped noise coming from 180°. Even though the study reports improvements in word recognition at specific time points, there are a few shortcomings in the design. In short, they did not use an RCT design, did not include a control group and the SIN test that was used had been modified, rendering it a non-standardised test (Kuk et al., 2008). Therefore, claims of improvement could have occurred due to maturation or due to practice effect. Furthermore, significant improvement in scores were found at earlier time points but not in later ones; there was significant improvement when comparing scores from baseline to 2 weeks and from baseline to 6 months but not when comparing scores from baseline to 3 months (Kuk et al., 2008). This produces an aberration left unexplained by the researchers (Kuk et al., 2008). These two studies (Johnston et al., 2009; Kuk et al., 2008) are the only ones incorporating behavioural SIN tests in their intervention designs and as noted their findings are inconclusive.

Apart from behavioural SIN tests, the studies by Kuk et al. (2008) and Johnston et al. (2009) also employed other measures such as self-reports or teacher/ parental questionnaires on listening-in-noise skills. Kuk et al. (2008) found that in the CHAPS questionnaire (measuring auditory performance in different environments) both teachers and parents considered the child with APD to be at risk in the Noise subscale at baseline. When comparing pre to post scores, these were non-significant, indicating no improvements in that subscale. Johnston et al. (2009) used the Listening Inventory for Education (LIFE) questionnaire, which is given to children to assess how well they hear their teacher in different listening conditions in the classroom. Results showed that children with APD had significantly worse scores at baseline than children in the control group and the mean score of children with APD was in the disordered range while for controls in the normal range. Comparing the results of the APD group pre- to post-fit, three of the questions showed significant improvements. These questions reflected improved ability to hear the teachers in the following situations: when teachers were talking in front of the room, when teachers had their backs turned and when other students were making babble noise (Johnston et al., 2009). It may be concluded that when children with APD were using RMHAs in the classroom they were helped in these three listening situations. However, due to the lack of an APD control group these changes may reflect maturation, placebo/ Hawthorne effects or a regression to the mean.

One study has examined linguistic skills in children with APD after using RMHAs. Sharma et al.'s (2012) RCT on 55 children with APD found an increase in phonological awareness scores after the use of RMHAs for 6 weeks. Specifically, the subtests of sentence recall, non-word spelling and syllable segmentation showed significantly better scores in the RMHA groups. The researchers included four intervention groups which received bottom-up AT only[16], top-down AT[17] only, bottom-up AT plus RMHAs and top-down AT plus RMHAs (see Table 1.1). The groups that showed improvement in the three language subtests received bottom-up AT plus RMHAs and top-down AT plus RMHAs (Sharma et al., 2012). As significant improvements were recorded only for these two groups and not for the two AT-only groups, the increase in the three

---

[16] Bottom-up AT was also called discrimination training and it involved activities of temporal and spectral discrimination of pure tones. These activities were based on suggestions drawn from case studies conducted by Musiek and Schochat (1998).

[17] Top-down AT was also called language training and it included activities to develop metacognitive and language skills (e.g. reasoning, inferencing, understanding differences in meaning, reading aloud tasks, listening for meaning; Chermak & Musiek, 2007).

outcome measures mentioned was attributed to the use of RMHAs (Sharma et al., 2012). But as the study did not include a RMHA-only group, improvements cannot be solely attributed to RMHA use, even if changes were not observed for the AT-only groups.

The fact that the RMHA intervention groups in the study did not both exhibit improvements in the same subtests presents problems to the argument that the increase was due to RMHAs (Sharma et al., 2012). The top-down AT plus RMHA group improved in sentence recall and non-word spelling, while the bottom-up AT plus RMHA group improved in syllable segmentation. One possible interpretation of these findings could be that the RMHA brings improvement in linguistic skills only when coupled with specific bottom-up or top-down AT programmes. For instance, RMHA use on its own would not have had an effect on syllable segmentation, but when coupled with bottom-up AT (and not top-down AT) it does. Similarly, it can be argued that when bottom-up AT is used on its own it does not influence syllable segmentation and needs to be combined with RMHA use to do so. There is also the question of whether the inclusion of RMHAs in combination with AT programmes interfered with possible benefits that AT would have brought in the rest of the outcome measures (as discussed in the second paragraph of this section). The study would have benefitted from the inclusion of a RMHA-only group, which would have helped answer the above questions. The study also has other limitations. While the researchers initially planned to recruit 85 subjects to reach a statistical power of 80%, they ended up recruiting 55 subjects compromising the power to 64% (Sharma et al., 2012). This of course did not prevent the study from yielding statistically significant differences in some measures, but a low statistical power may inflate the significant effects. Hence, Sharma et al.'s (2012) results need to be cross-validated by larger sample-sized studies with higher statistical power and RMHA-only intervention groups.

Finally, Johnston et al. (2009) also used the Screening Instrument for Targeting Educational Risk (SIFTER), which was completed by parents and aimed at examining learning difficulties and academic performance. At baseline, parents of children with APD rated their children significantly poorer than parents of control children in terms of their academic performance. The authors then compared the results of the APD group at baseline and 5 months after using RMHAs. Neither of the two language subscales of the SIFTER (i.e. academic performance and communication) showed significant improvement over the intervention period. However, both subscales did increase from ‘fail’ to ‘marginal’, pointing to clinical significance despite the lack of statistical significance.

Attention

There are currently no trials examining the effects of long-term use of RMHAs on attention in children with APD. However, Johnston et al. (2009) found that the attention subscale from the SIFTER did not improve and was found to remain at risk according to parent, even after 5 month’s use of RMHAs by children with APD. The rest of the four subscales in SIFTER (measuring academics, communication, class participation and school behaviour) even though they also did not have significant improvements, they did present with clinical significance since they changed from ‘fail’ to ‘marginal’. The study, though, did not employ repeated measures in an RCT design, and the small sample size (*n* = 10) cannot be used to generalise findings. The study’s shortcomings were discussed in detail earlier. Moreover, mixed findings were reported by Kuk et al. (2008) who used the attention subscale of the CHAPS questionnaire (completed by parents and teachers), while children used hearing aids with directional microphones (instead of RMHAs). At baseline, teachers rated the students marginally normal in terms of their attention, while parents rated them as being ‘at risk’. After 6 months, teachers’ scores (who viewed the children during classroom time, hence when using the hearing aids), showed slight non-significantly worse scores but enough to rate the children marginally ‘at risk’. Parents on the other hand, showed significant improvement in their ratings on their children’s attention; a rate though that remained just below the normal threshold. These findings are hard to make sense of and better research designs examining attention through attention-specific questionnaires are needed.

The same study by Kuk et al. (2008) actually did include a test measure of Sus-AA (see Table 1.1), but the test was modified to be tested both in quiet and in noise with and without the hearing aids. Thus, it cannot be considered a pure attention test, as the researchers essentially changed it into a SIN test. This means that the scores children received on the modified test were no longer reflecting the standardised scores of the tests. In non-standardised tests improvement in scores may be explained by maturation effects (unless there is an adequate control group). Moreover, the test conditions were not consistent, as the condition testing in quiet-unaided (which would be the best condition to measure sustained attention) was only measured at baseline and not post-intervention (Kuk et al., 2008). This makes it impossible to determine effectiveness of the intervention on the Sus-AA task. In addition, the sample size was small and the authors highlighted this limitation (Kuk et al., 2008). Therefore, links between use of the directional microphone hearing aids and auditory attention cannot be drawn from these findings.

With attention being a top-down cognitive function (Blake, Field, Foster, Platt, & Wertz, 1991), and the RMHA a bottom-up, passive intervention (Fey et al., 2011), one can argue that attention would not be expected to show improvements with RMHA use. However, other researchers argue that listening-in-noise ability, which is primarily found impaired in children with APD (Cameron et al., 2006; Cameron & Dillon, 2008; Lagacé et al., 2011), is intertwined with attention and memory (Wong et al., 2008). There is evidence that shows that the act of listening involves detecting, discriminating, remembering, understanding and acting on a pattern of sounds; all these being processes governed by attention and memory (Moore et al., 2013). Adding to that, Blake et al. (1991) explain that as the RMHA enhances the target signal over competing and distracting sounds, selective attention in the child is in turn assisted (i.e. target signals are easily attended versus distracting signals). In Section 1.8.1 this matter is examined in more detail and a hypothesis on how long-term RMHA use may help bring neuroplastic changes in attention in children with APD is presented.

Concluding, the lack of studies examining the effects of RMHAs on attention in children with APD make it hard to have a solid basis for discussion on this topic. Some evidence from self-reports and questionnaires are insufficient and contradicting (Johnston et al., 2009; Kuk et al., 2008). Further research is needed to determine the long-term effects of RMHA use on attention.

In Chapters 2, 3 and 4 this matter will addressed by including measures of attention in RMHA trials.

Memory

To the best of our knowledge there is only one study that tested for changes in AWM after RMHA use for 12 weeks (Table 1.1; Umat et al., 2011). The participating children in the study, aged 7-10, were also retested at 12 months (but without using the device between 12 weeks and 12 months), in order to determine the RMHA’s post-intervention efficacy (Umat et al., 2011). Statistically significant results were mentioned (claiming benefit due to RMHA use after 12 weeks and at 12 months) but this was for main effects (Umat et al., 2011). One-way Analysis of Variance (ANOVA) was used to compare scores at baseline and repeated measures ANOVA to look for score differences across time, but these two analyses are limited to only comparing differences between groups when time points are collapsed across groups (one-way ANOVA) or when groups are collapsed together to compare the different time points (repeated measures ANOVA). Even though a mixed design (split-plot) ANOVA was also used that looked at the interaction between the within subjects factor (i.e. time) and the between subjects factor (i.e. group), this was not followed up with pairwise comparisons on the separate groups to identify which groups and which time points showed significant improvement. Consequently, no specific time points were assigned to significant differences between the groups neither specific group differences were identified within time points. Moreover, while an interaction (without follow-up pairwise comparisons) was noted for the Working memory condition, no interaction was found for the Retention of information condition (Umat et al., 2011). Therefore, the data collected from the study have not been adequately investigated through statistical analysis. The authors claimed working memory improvements due to RMHA use without identifying specific significant improvements in the intervention groups over the control, while the lack of significant interaction in the Retention of information condition was not taken into account during interpretation of the findings.

Another shortcoming of the study is that despite the fact that the authors created the three groups matching their age and IQ scores, this procedure was not randomised, hence introducing bias.

Additionally, the memory test's normative data is from an American sample, while this study tested Malaysian children without mentioning whether Malaysian normative data were developed (Umat et al., 2011). Finally, in their discussion the authors indicated that the significant improved scores in two of the outcome measures at 12 months was due to the 12-week use of RMHAs. The RMHA system though was not in use by children between 12 weeks and 12 months (Umat et al., 2011). The observed improvement in scores at 12 months could have been influenced by a number of factors unrelated to the initial 12-week intervention that the study did not control for (e.g. maturation or other home or school-based interventions).

An interesting finding by Sharma et al. (2012), not reported in their results as this was not the focus of their research, is that the Digit span forward and Digit span backward measures (both working memory tests) showed no statistically significant change in any of their five study groups after the 6-week intervention period (see groups in Table 1.1). Based on this finding, it could be supported that working memory is not showing improvement after a 6-week RMHA and AT intervention. Nonetheless, in that same study Sentence recall (another measure linked to working memory) recorded significant improvement in the RMHA plus top-down AT intervention group but not in the top-down AT group (Sharma et al., 2012). This could point to a gain due to the additional use of the RMHA in the AT plus RMHA group. The fact, though, that the group using bottom-up AT (instead of top-down AT) plus RMHAs did not exhibit this gain could mean that the improvement in Sentence recall can only be achieved through combining RMHA use with top-down AT (and not RMHA with bottom-up AT).

Finally, in their directional microphone hearing aid interventional study Kuk et al. (2008) included the CHAPS questionnaire, which has an Auditory memory sequencing subscale. The questionnaire was completed by teachers and parents. Children with APD were rated as being 'at risk' at the start of the study by both their teachers and parents. After a 6-month use of the intervention at school, teachers did not demonstrate significant change in their memory subscale scores and children were still considered to be 'at risk'. Looking at the parental reports, significant improvement in the subscale was recorded, but this failed to break out of the disordered threshold, thus not exhibiting clinical significance. In conclusion, evidence on AWM improvement using RMHAs is scarce and findings from the few studies dealing with the topic suggest that working memory in children with APD remains unchanged following RMHA or hearing aid interventions. This requires further examination and in Chapter 2 an AWM test will be included to test for that.

RMHAs and other disorders

The review of the handful of studies on RMHA use on children with APD revealed lack of strong evidence, limitations in study designs, unfounded claims and small sample sizes in these studies. It is for these reasons that APD research usually turns to studies using RMHAs in children with other disorders (often co-occurring with APD) to draw conclusions on the effectiveness of the system in certain developmental aspects in children. Of course, this poses validity problems as the benefit of the intervention should be directly validated in children with APD before conclusions are drawn (BSA, 2011b). A brief overview of selected studies, which test the effectiveness of RMHAs on children with other disorders, will be presented here. The focus will be on studies that dealt with attention and neuroplasticity, which are two topics the current thesis is looking to examine.

First, Friederichs and Friederichs' (2005) study tested the effects of a one-year RMHA intervention in children with ADHD diagnosis and suspected of APD. The findings of this study are reported under this section because even though the study claimed that children were suspected of APD, none of them underwent an official diagnostic assessment for APD. The authors’ suspicion of APD was based only on a history of learning difficulties and APD symptoms. The children with ADHD (*n*=10), aged 7-14, were compared to 10 age-matched controls (also with ADHD) at 6 months and at 12 months (Friederichs & Friederichs, 2005). For the electrophysiology task, an oddball paradigm was employed with a 3:1 ratio of frequent to infrequent tones. The data were analysed using a paired t-test and demonstrated differences in the P2 component in the two groups, but without claims of statistical significance. Additionally, significant improvements in the psychoacoustic tasks of frequency discrimination and side order (i.e. high and low tone sequencing) were observed at 6 and 12 months compared to baseline in the intervention group. Other psychoacoustic outcome measures (e.g. intensity discrimination, gap detection and time order judgement) did not show significant changes over the intervention period between groups (Friederichs & Friederichs, 2005).

The statistical analysis chosen by the authors, though, might have introduced bias. This design of two independent groups (the between subjects factor) measured across three different time points (the within subjects factor) should have been analysed using a mixed design (split-plot) ANOVA, instead of a paired sample t-test. Making these comparisons using a t-test does not allow for correction of multiple comparisons, increasing the chances of a type I error; a correction made possible in a mixed design ANOVA. Adding to that, the psychoacoustic data were based on percentage of correct answers (Friederichs & Friederichs, 2005), while analysis of d-primes would have better addressed the issue of random responses of subjects due to uncertainty. Thus, it appears that the authors did not use the most appropriate methods for their statistical analysis. Other limitations in the study design include the use of a small sample size with no mention of power sample considerations and the fact that ADHD medication was continued throughout the study by some of the participants in both groups with no control over this factor. Therefore, results from this study should be interpreted cautiously, even for the population with ADHD.

Furthermore, studies using RMHAs on children with different disorders reported improved attention over time (Blake et al., 1991; Purdy, Smart, Baily, & Sharma, 2009; Schafer et al., 2013). In Blake et al.'s (1991) study children had learning disorders with reported attention problems. In their findings, observed attentive behaviours by children, aged 5-10, were improved following a 24-week RMHA intervention. Attention was measured through observation by two professional speech-language pathologists. Even though measures to control consistency across observations were taken, the observers were not blind (not possible with the use of an external hearing aid device), and the method employed for measuring attention was problematic. For instance, attention was measured through observations in the following behaviours: body-turn to signals, eyes-turn to signal and absence of unnecessary body and vocal movements (Blake et al., 1991). These are not standardised methods for measuring attention and they rely heavily on subjective observation and on quick judgment from the observer, who could have introduced observer error and in turn may have reduced validity. Schafer et al. (2013) used a design to increase reliability and validity in their observations, which appears to have successfully controlled for several factors (Schafer et al., 2013; page 36 for more details). In their results, they found reduced scores in inattention and distractibility (indicating better attentive behaviours) in children with ASD and ADHD, aged 9-12, when RMHAs were on (Schafer et al., 2013). An issue raised earlier, though, remains; observations were not made in the unaided condition, meaning that it is not possible to support neuroplastic changes in these children due to RMHA use.

Purdy et al. (2009) conducted a study looking at the effects of RMHAs on children with reading delays compared to a disordered control group in an RCT design, all children aged between 6-11. Each of the two groups had 23 children and a RMHA was given to the intervention group to use for 6 weeks. The LIFE questionnaire was administered to both teachers and children, while a standardised reading test was also included in the design. The authors found that despite teacher and children questionnaires stating improvements in reading and listening skills, this was not reflected in the reading tests (Purdy et al., 2009). While the LIFE questionnaire did not have standardised scores, the researchers had tested it in a sample of 18 typically developing controls in the school for test-retest reliability. Moreover, even though 64% of the study sample did not have English as the main language at home, this factor was added as a covariate in the analysis, revealing that children having English spoken at home had better reading test scores than their counterparts. It was concluded that the 6-week intervention period might have been adequate only to significantly improve scores in the teachers' and children's questionnaires and that for possible improvements to be reflected in the standardised reading test a longer intervention period might be required (Purdy et al., 2009).

A study looking for neuroplastic changes in phonological awareness, followed 38 children, aged 8-14, with dyslexia over a one-year RMHA intervention period (Hornickel, Zecker, Bradlow, & Kraus, 2012). Half of the participants were also identified with attention deficits, which the researchers controlled by equally dividing these children into two groups, the dyslexic intervention and the dyslexic control group. The study found that after a year's use of RMHAs, children with dyslexia in the intervention group had improved phonological awareness (reflected in a behavioural test) and also a more consistent brainstem response (in a stop consonant stimulus task) in comparison to the age and gender matched dyslexic control group and also in comparison to a third group of typically developing children (*n* = 26; Hornickel et al., 2012). The improved results from the two tasks, behavioural and electrophysiological, had a significant moderate correlation. Allocation to groups was performed using a quasi-randomisation, which could have introduced selection bias as allocation was not concealed. But the study sample of 38 with 19 children in each group tackles the common problem of small sample sizes observed in most of the previously discussed interventional studies.

In the study's conclusion it was highlighted that an "interaction of attention and enhanced signal quality" (Hornickel et al., 2012, p. 16734) is what led to these findings. A possible proposed interpretation on how this interaction takes place, according to the authors, is this: the clearer signal offered by the RMHA enhances auditory perception, promotes active engagement to the stimuli and thus alleviates the cognitive burden of constantly attending. In turn, this allows for improvement in phonological awareness which results in a more consistent brainstem response in comparison to non-RMHA users (Hornickel et al., 2012). But as this model is an empirical proposal by the authors without cross-validation from other findings, it cannot be used as an established theory that RMHAs enhance attention.

Conclusion

There is lack of reliable and valid research findings on the effects of personal RMHAs on the development of children with APD (Johnston et al., 2009). For that reason, management recommendations are usually based on empirical data (Johnston et al., 2009) or on findings from RMHA studies on children with other disorders (e.g. ADHD, ASD, dyslexia). Empirical arguments need to be backed up by research, while findings from non-APD studies need to be validated in APD studies (BSA, 2011b). This issue can only be addressed with more RCTs examining the effects of RMHAs on aspects of cognition and development in children with APD.

This present work aims to contribute to this need by looking at the effects of RMHA use in specific aspects of children with APD. The aspects of development to be examined are the ones discussed and justified in the next section; SIN perception, attention and memory.

1.8 Justification of research

The RMHA system, being one of the management approaches for children with APD, aims at improving the SNR for these children. For RMHAs to be included in legislative frameworks and promoted as a viable APD management strategy, evaluation and proof of their efficacy should first be confirmed through research (Flexer, 2014). The review on RMHA studies on children with APD revealed that firstly there is a small number of such studies and that the ones that exist are either of low level of evidence, or exhibit methodological shortcomings (Johnston et al., 2009; Kuk et al., 2008; Umat et al., 2011). This current work will contribute to the field by evaluating the efficacy of RMHAs on specific aspects of development in children with APD, in order to provide much needed evidence on the matter. The aspects of development to be examined are SIN performance, measures of auditory attention and AWM. Questionnaires with subscales measuring these skills will also be used. These aspects of development will be discussed in the following sections in relation to RMHAs and the reasons as to why they are expected to change or remain unchanged over the trial period.

1.8.1 RMHAs and changes in SIN perception

A study evaluating a model looking to explain SIN, showed that cognitive and AP factors were the most influential on SIN performance (S. Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013). Focusing on cognitive factors, there is a series of studies demonstrating how the relationship between Sel-AA and SIN performance is developed (Mesgarani & Chang, 2012; O’Sullivan et al., 2015; Zion Golumbic et al., 2013). In Zion Golumbic et al.'s (2013) study a selective attention paradigm was applied, in which an experimental ‘cocktail party’ condition was presented in surgical epilepsy patients as responses were recorded using intracranial electrocorticography. The experimental condition was compared to a single talker control condition. Results showed that in early-sensory regions (i.e. the superior temporal gyrus) there were representations of both talkers (but stronger for the attended). Moreover, the target talker was selectively represented in various higher-order areas (i.e. inferior frontal cortex, anterior and inferior temporal cortex, and inferior parietal lobule) with the distractor talker having no representation there. Based on these findings, it was suggested that while the early auditory sensory cortex maintains auditory speech representation for both attended and unattended speech, higher-order areas show a purely selective response. Additionally, the early sensory speech encoding gates sensory input to higher-order areas, depending on attentional demands and resources (Zion Golumbic et al., 2013).

Furthermore, O'Sullivan et al. (2015) used a similar 'cocktail party' paradigm with Electroencephalogram (EEG) recording and a stimulus reconstruction strategy to determine attendance to the speech input (i.e. how strong was the linear correlation of the recorded neural activity to the incoming speech signal envelope). It was shown that at late latencies (up to ~250 ms, with a peak at 218 ms) the attended stimulus had maximal response corresponding to a spatial bitemporal distribution of contributing electrode channels. The authors concluded that this later peak response was the result of selective attention, which enhanced the response to the attended versus unattended signal at the later level of semantic processing (O'Sullivan et al., 2015). This is broadly in line with the aforementioned claims of late representation of the attended speech target in higher-order brain areas due to selective attention (Zion Golumbic et al., 2013).

Furthermore, Mesgarani and Chang (2012) employed a similar experimental design to Zion Golumbic et al.'s (2013) paradigm, using a stimulus reconstruction method. They found that in the selective attention condition where subjects attended to a target speaker while ignoring another speech stream, the reconstructed spectrograms from surface neural activity differed. During trials with correct responses, these spectrograms showed a stronger correlation to the reconstructed spectrogram of the target stream. At the same time, trials where error responses were recorded did not correlate with the spectrogram of the target speech (Mesgarani & Chang, 2012). The authors concluded that attention modulates the cortical representation of speech at early levels of spectrotemporal speech feature encoding, selectively enhancing relevant-to-the-task speech and suppressing irrelevant-to-the-task speech (Mesgarani & Chang, 2012). This was also supported by O’Sullivan et al.'s (2015) similar ‘cocktail party’ paradigm using EEG described previously, where electrophysiologic and behavioural data correlated. Putting together the findings from these three studies, a schematic model is suggested in Figure 1.5 on how selective attention might be related to SIN performance.

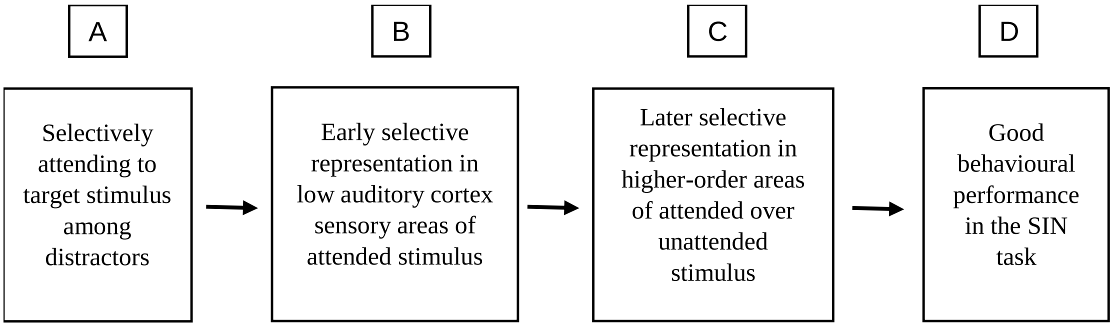

Figure 1.5 – Selective auditory attention and SIN performance
Schematic representation of (A) the effect selective attention has (B) on lower auditory cortex sensory regions, (C) on higher-order brain areas, (D) and the way that influences SIN behavioural performance (Mesgarani & Chang, 2012; O’Sullivan et al., 2015; Zion Golumbic et al., 2013). SIN: Speech-in-Noise.

In children with APD this ability to focus on a target speaker amongst competing streams (Cameron et al., 2006; Cameron & Dillon, 2008) and amongst competing speech babble (Lagacé et al., 2011) is found to be worse than in typically developing children. This difficulty exhibited by children with APD could reflect a compromised ability of the brain to represent the target stimulus in higher-order regions while supressing competing speech/ speech babble due to poorer encoding of the speech stimulus (see Figure 1.5). This could alternatively be interpreted as an extended difficulty in selective attention, meaning that children find it hard to focus on the target stream and therefore the early/ later selective representation (B and C in Figure 1.5) cannot be assisted by the compromised selective attention.

A study by Kidd, Arbogast, Mason, and Gallun (2005) on college students has shown that pre-knowledge of speech target spatial location improves behavioural task performance in comparison to a control condition (i.e. without prior knowledge of location). It was concluded that this makes it easier for the listener to selectively attend to the target and this, in turn, helps to solve the ‘cocktail party’ effect (Kidd et al., 2005). Another study on typically developing adult listeners assessing the effects of auditory object continuity on selective attention, has similarly found that selective attention is enhanced when the focus of the listener remains fixed without spatial shifts of the auditory object (Best, Ozmeral, Kopco, & Shinn-Cunningham, 2008). A task where target digits were presented in the presence of distractor digits from different sources, showed that presentation of target digits from a fixed focus location resulted in better recall, indicating better selective attention for a fixed and sustained spatial focus than for an interrupted one (Best et al., 2008). Interestingly, digit recall became better at later stages of the fixed spatial condition task, suggesting that selective attention training positively affects task performance (Best et al., 2008).

The use of RMHAs by children with APD makes it easier and less effortful for the child to selectively focus on the enhanced (in terms of acoustic fidelity and SNR) target speech versus the background distractors (Johnston et al., 2009). Therefore, it is hypothesised that the enhanced acoustic signal representation and fixed spatial location (resulting from the use of the RMHA) will make it easier for children with APD to have fixed attentional focus on the target. A sustained focus on the target is thus expected to be achieved, and based on the above-mentioned findings (Best et al., 2008; Kidd et al., 2005) this will have a positive influence on selective attention, which is expected to be assisted by the improved SNR. Subsequently, the representation of the attended stimulus should now be selectively represented in higher-order regions with simultaneous suppression of the background distractors, which will in turn have a positive impact on SIN behavioural performance (see Figure 1.5).

Extending the hypothesis and the schematic representation on Figure 1.5, it is hypothesised that by using the RMHA long-term, a lasting effect on this auditory function will be achieved. It is known that continuous interventional training can have lasting effects in the latency and amplitudes of ERPs in a number of auditory-related paradigms (Alain, Snyder, He, & Reinke, 2007). A study by Song, Skoe, Banai, and Kraus (2012) focusing on improving SIN performance, provided a 4-week training programme to the intervention group, comprising training in degraded speech, communication strategies and working memory. Significant improvements were shown in the scores of the behavioural SIN tasks and significant change in the neural representation of the acoustic signal-in-noise in the brainstem was also observed for the intervention group only. These changes were attributed to neuroplastic changes generated by the extensive use of the training programme. These plastic changes were also observed 6 months after the end of the study (Song et al., 2012). Similarly, in the case of RMHAs and SIN in children with APD, because of the continual exposure to an improved SNR via the RMHA, neuroplastic changes in the process of selectively attending and separating the target stream from the background are expected to be achieved. This will make it possible for the child to perform this separation of attended versus unattended stimuli even without the use of the assistive listening device, following long-term RMHA use, and this will be reflected behaviourally in improvements in SIN performance. To test this, behavioural SIN tests will be administered without the use of an assistive device (unaided). These tests will be described in detail in the relevant methodology sections in Chapters 2 and 3.

1.8.2 RMHAs and changes in Sus-AA and Div-AVA

It has been shown that the use of RMHAs or hearing aids with directional microphones by different groups of children (APD, ADHD and children with learning difficulties) led to improvements in their attentive behaviours as judged by their parents and teachers (Blake et al., 1991; Kuk et al., 2008; Rosenberg et al., 1999; Updike, 2006). Kuk et al.'s (2008) study, which is the only one involving children with APD, showed significant improvements in the Auditory attention span subscale of the CHAPS questionnaire, completed by parents, following a 6-month hearing aid intervention. As parents judged their child’s attention span at home, this means that their views reflected changes when the hearing aids were not used, pointing to possible neuroplastic changes. It should be taken into account that the significant improvement in scores in the Auditory attention span subscale of the CHAPS was not paired with clinical improvement as children’s mean scores after hearing aid use remained marginally ‘at risk’ (Kuk et al., 2008).

Another study by Cameron and Dillon (2011) tested omission and commission errors in an auditory attention test, following a 12-week listening-in-spatialised-noise training programme on children with SPD, aged 6-11. Commission errors significantly decreased (thus indicating improvement), whereas omission errors did not significantly change. The authors attributed this to an improvement in sustained attention, because omission errors were assessed in the first 10 minutes of the test while commission errors in the last 10 minutes. Even though the training programme did not target sustained attention (but rather listening in spatialised noise), the fact that the programme required the child to sustain its attention on the training tasks 20 minutes daily for 3 months, may have helped reduce auditory fatigue, which was reflected in the reduction of commission errors (Cameron & Dillon, 2011). Similarly, the use of a RMHA system will enable children to sustain their focus daily on the target voice (teacher's voice) for a few hours and for as long as the intervention period lasts. Therefore, it is hypothesised that similar effects on behavioural Sus-AA tasks will be reflected in the intervention group, as part of a general reduction in auditory fatigue (Cameron & Dillon, 2011). For that, a behavioural Sus-AA task will be used in the design and will be described later in the methodologies of Chapters 2 and 3.

Moreover, regarding Div-AVA, it is supported that in divided attention tasks that involve different sensory modalities (i.e. auditory-visual) there are shared attentional resources and this is dependent on the task (Wahn & König, 2017). For instance, spatial attentional tasks have this shared pool of attentional resources, while in other tasks (object-based attention tasks) there can be partially shared attentional resources (Wahn & König, 2017). Going back to the hypotheses of the thesis, it is expected that in bimodal tasks the auditory sensory modality will be assisted by the improved SNR brought by the RMHAs. This enhanced auditory signal will help reduce the attentional demand in the auditory modality and fewer resources will thus be drained from the shared pool of attentional resources, allowing more to be used by the visual sensory modality. This will result in an improved overall performance in Div-AVA tasks when using the RMHA, as the auditory modality will be aided by RMHA use and the freed-up resources could now be allocated to the visual modality. Continuous long-term use of an intervention is expected to induce a lasting effect on this process, as in the cases discussed earlier of other long-term interventional studies that brought neuroplastic changes in auditory processes (Alain et al., 2007). It is thus hypothesised that the continuous use of RMHAs at school will bring lasting changes in the process of performing bimodal (auditory-visual) tasks and these changes will be observed in behavioural Div-AVA tasks even without the use of the assistive listening device.

1.8.3 Memory

There was only one study which specifically looked at AWM measures to assess whether these improve following a one-year RMHA intervention in children suspected of APD (Umat et al., 2011). The results of the study are unclear, as highlighted earlier, because there is only mention of main effects of time and group and no reference to simple main effects, which would indicate significant improvement over time in the intervention group. Additionally, an RCT by Sharma et al. (2012), not primarily aiming to test AWM, did not find improvements in the digit span forward and backward following a combination of AT and RMHA 6-week intervention. As the post-intervention tests were administered without the use of RMHAs (unaided), this could mean that RMHAs do not generate neuroplastic changes in AWM functions after long-term RMHA use. But the fact that RMHAs were combined with AT in these intervention groups (Sharma et al., 2012) makes it difficult to attribute the lack of improvement in AWM tasks to RMHA use alone.

This argument of no improvement in AWM following long-term use of hearing aid interventions can be further supported by studies on children with hearing loss using cochlear implants (Mishra & Boddupally, 2017). In this trial, children with cochlear implants, aged 6-15, were provided with a digit backward working memory training programme for 5 weeks while another group of age-matched children with cochlear implants was given a placebo. At the end of the intervention period, the group receiving training demonstrated significant improvements in both forward and backward digit span tasks compared to the placebo group (Mishra & Boddupally, 2017). These results show that the use of a cochlear implant alone (which does provide with an enhanced signal such as RMHA do for children with APD) does not bring improvements in AWM tasks in children with hearing loss and additional interventions (in this study training on AWM tasks) are required to bring lasting improvements in AWM. The evidence thus far regarding children with APD using a RMHA point to a lack of significant improvements in working memory tasks. This assumption will be further tested and it is thus hypothesised that the scores on an AWM task will not show significant improvements in the group of children with APD using a RMHA intervention. The AWM task will be described in detail in the methodology of Chapter 2.

1.8.4 Questionnaires

In addition to the behavioural tests, questionnaires to be completed by parents, children and teachers will be used. These will be described in more detail in the methodologies of Chapters 2 and 3 but briefly, the CHAPS questionnaire (to be given to parents of children with APD) is a questionnaire with several subscales, including a Noise, Auditory attention span and Auditory memory sequencing subscales (Wilson et al., 2011). These three subscales will be analysed in relation to the relevant behavioural tests described earlier. Moreover, the Listening Inventory for Education – Revised (LIFE-R) questionnaire is a self-reported questionnaire, in which children assess the effectiveness of the intervention in the acoustic environment of the classroom (K. L. Anderson, Smaldino, & Spangler, 2011; K. L. Anderson & Smaldino, 1999). The CCC-2 will be administered to monitor children's performance on language and communication skills (Bishop, 2003). As discussed earlier, there is evidence to show that a combination of RMHAs with either bottom-up or top-down AT could bring significant improvements in sentence recall, non-word spelling and syllable segmentation tasks (Sharma et al., 2012). Additionally, the study by Hornickel et al. (2012) reviewed earlier in Section 1.7.2 (under RMHAs and other disorders), revealed that children with dyslexia using RMHAs for a year demonstrated more consistent neural representation of stop consonants which was paired behaviourally with improved reading and phonological awareness. Therefore, by administering the CCC-2 possible changes in some aspects of children's linguistic development can be examined. Specifically, conditions measuring standard language aspects (e.g. semantics and syntax) are expected to change. Conditions measuring non-standard language aspects (e.g. use of context and nonverbal communication) are expected to remain unchanged. This is because in Sharma et al.'s (2012) study children using both AT and RMHAs did not have significant changes over the 6-week intervention period in tests measuring non-standard linguistics aspects (i.e. figurative language, sarcasms, extracting meaning from context, drawing inferences). In addition, there are no studies in children's populations that indicate possible improvements in non-standard linguistic aspects due to RMHA use. The CCC-2 questionnaire was preferable to additional behavioural language or reading tests as it minimises test administration time and because examination of the effect of RMHAs on language and communication development in children with APD was a secondary objective. Finally, the SIFTER questionnaire will be given to teachers to complete and the two subscales of interest are the Attention and the Communication subscale (Wilson et al., 2011). This questionnaire can offer teachers' perspective on how children's attention and communication potentially change when RMHAs are used in the classroom.

1.9 Research questions and thesis structure

1.9.1 Research questions

There are three studies in this thesis and the detailed research questions and hypotheses of each will be described under the relevant study chapter. This section will offer a brief overview of the general research questions of the thesis. In general, the thesis aims to determine whether the improved SNR provided by the use of RMHAs (Johnston et al., 2009) has a positive long-term effect (i.e. unaided, without the use of the system) on SIN performance and on types of attention (Sus-AA, Div-AVA) in children with APD. The mechanisms expected to change and the justification behind these possible changes were described in detail in Section 1.8. Therefore, the studies in Chapters 2 and 3 will use an RCT design and will look for possible improvements in SIN performance first at 3 months and then at 6 months of RMHA use. Moreover, to date no study has looked at the effects of RMHAs on attention. The hypothesis on how RMHA use might positively influence Sus-AA and Div-AVA was described earlier in Section 1.8.2. Hence, these two types of attention will also be tested over the intervention period to monitor possible improvements in their scores following a 3-month and 6-month RMHA intervention. Furthermore, the review of studies on the effects of RMHAs on AWM indicate that perhaps the system does not have lasting benefits on memory (Sharma et al., 2012; Umat et al., 2011). Based on this evidence, AWM will be examined under the hypothesis that there will be no significant change in AWM scores after a 3- and 6-month RMHA use.

Finally, the discussion on APD diagnosis and diagnostic criteria (see Section 1.4) concluded that there is a lack of consensus on how APD should be defined and diagnosed. At the same time, measures of attention are not routinely included in APD diagnostic protocols despite findings that link the two (discussed in detail in Section 1.6; Allen & Allan, 2014; Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). Meanwhile, APD management is relying on an official APD diagnosis and for that reason the diagnostic protocol should be able to construct a more complete profile for children with APD, where their needs will be addressed through targeted interventions. To that end, differential APD diagnostic criteria including measures of attention will also be examined in the study of Chapter 5 in order to better inform management recommendations for children. This examination solely aims to provide more targeted management recommendations and does not in any way aim to propose a new diagnostic APD process. In addition, correlation analyses between measures of attention and AP tests will be conducted in Chapter 5, expecting that there will be correlations between some types of attention and AP skills (see detailed measures of attentions and AP tests and expected correlations in Section 5.1.3). Detailed hypotheses are provided in the introduction of each chapter.

1.9.2 Structure of the rest of the thesis

Chapters 2, 3, 4 and 5 will present the four studies of the thesis which aim to answer the research questions. These chapters include a brief background for each study, detailed research questions, description of the methods, presentation of the results and discussion of the findings. The final chapter (i.e. Chapter 6) brings together all the study findings, discusses their implications, identifies limitations, outlines future directions for further research on these topics and draws the final conclusions.

CHAPTER 2 – STUDY 1: EFFECTS OF RMHAs ON SIN PERCEPTION, ATTENTION AND MEMORY IN CHILDREN WITH APD

2.1 Introduction

2.1.1 APD and RMHAs

Children with APD face a number of difficulties but some of the most commonly reported are SIN difficulties, inattentiveness, auditory fatigue and poor working memory (AAA, 2010; ASHA, 2005b). Compared to typically developing peers or children suspected of but not meeting APD criteria children who meet the criteria for APD exhibit worse SIN ability and worse scores in attention tests (Allen & Allan, 2014; Cameron et al., 2006; Cameron & Dillon, 2008; Lagacé et al., 2011). This, in addition to poor classroom acoustics (noise and reverberation time), puts children with APD at a disadvantage and these difficulties listening in noise may hinder their learning (Flexer, 2014; Toe, 2009). One of the key APD management recommendations is the use of RMHAs at school (BSA, 2011b). Remote microphone hearing aids improve SNR for children with APD (Johnston et al., 2009) and help bypass the negative effects of noise, reverberation and distance from the speaker (W. J. Keith & Purdy, 2014). But it is yet unclear whether this immediate benefit in SNR has a positive long-term effect on children's functions necessary for communication or academic success. A systematic review on studies using RMHAs on trials involving children with APD concluded that only studies of low level of evidence (III) were conducted up until 2008 (Lemos et al., 2009). Since 2008, there were only three studies that used RMHAs as an intervention in children with APD examining aspects of their development, such as SIN ability, working memory, language and communication (Johnston et al., 2009; Sharma et al., 2012; Umat et al., 2011) and only one of them was an RCT (see Table 1.1 in Chapter 1 for an overview of these studies).

Johnston et al. (2009) found a gain of approximately 10 dB in a SIN test when children with APD were aided by a RMHA than when unaided at baseline. The authors' claim of this being evidence for an improved auditory system in children with APD, though, was not justified. Long-term benefits in the auditory system (i.e. even in unaided test conditions) can be claimed only when there is significant improvement in unaided conditions and in that study there was a 0.7 dB non-significant change. Results from an earlier APD study on low-gain hearing aids equipped with a directional microphone and noise reduction by Kuk et al. (2008), did not test for long-term SIN benefits of hearing aid use, as they did not include an unaided post-intervention condition. Both studies exhibited methodological shortcomings, such as small sample sizes, lack of control groups and non-randomised designs. Therefore, the hypothesis that RMHAs might bring lasting effects in SIN performance in children with APD has not been adequately researched. This study will use a SIN test to test whether its scores improve following a RMHA intervention. The way potential improvements are expected to take place long-term is described in detail in Chapter 1, Section 1.8.1.

Moreover, to date no study has examined the long-term effects of RMHAs on behavioural attention tests and only data from questionnaires exist on this topic (Kuk et al., 2008). In Kuk et al.'s (2008) study, the Auditory attention span subscale of the CHAPS questionnaire, completed by parents, did show significant improvement after 6 months of directional microphone hearing aid use, but still remained below normal levels. As parents observed their children at home (where hearing aids were not used), this indicates that the intervention may have had a lasting effect on their attentive behaviours. In the current study I will use tests of Sus-AA and Div-AVA to test whether their scores improve following a RMHA intervention. Section 1.8.2 in Chapter 1 describes in detail how possible improvements in these two types of attention are expected to occur. Two attention control conditions (i.e. Sel-VA and Div-AA) will also be included. It is not expected that the scores of these types of attention will record improvement as the RMHA does not enhance or assist divided listening or visual attention.

On working memory, only one trial employed a behavioural AWM test following a RMHA intervention on children with APD (Umat et al., 2011). Even though the authors concluded that the use of the RMHA has brought long-term improvements in AWM, in their results there is no mention of significant simple main effects of time or group and only main effects of group and time are cited. Finally, an RCT using a working memory measure did not find improvements in the two AT plus RMHA groups (Sharma et al., 2012), suggesting that AWM might not be helped long-term by RMHA interventions; at least not when coupled with AT interventions.

2.1.2 Research question and hypotheses

The study aims to examine the effects of RMHAs on SIN perception, attention and AWM in children with APD following a 3-month intervention. Questionnaires will also be given to parents and teachers to complete throughout the study. Parents will complete the CHAPS and CCC-2 questionnaires. The Noise, Multiple inputs, Auditory attention span and Auditory memory sequencing conditions from the CHAPS will be used. The CCC-2 questionnaire measures language and communication in children and composite scores of the subscales of standard language and non-standard language will be used as experimental and control conditions, respectively. Finally, the SIFTER questionnaire will be completed by teachers and the Attention and Communication subscales will be analysed.

The research hypotheses are detailed below. All tests mentioned here are described in depth later in Methodology.

1. The group of children with APD who use RMHAs will show improved scores after 3 months compared to the APD control group in:
    a. The two conditions of the SIN test (the SCAN-3 C test; unaided).
    b. The Div-AVA task and the composite score of the three Sus-AA tasks of the TEACh test (unaided).
    c. The Noise, Multiple inputs and Auditory attention span subscales of the CHAPS (completed by parents).
    d. The composite score of the four standard language conditions of the CCC-2 questionnaire (completed by parents; Speech, Syntax, Semantics, Coherence).
    e. The Attention and Communication subscales of the SIFTER questionnaire (completed by teachers).

2. The group of children with APD who use RMHAs will not show improved scores at 3 months in:
    a. The Sel-VA and Div-AA tasks of the TEACh test (unaided).
    b. The three conditions (see Table 2.4) of the AWMA (unaided).
    c. The Auditory memory sequencing subscale of the CHAPS.
    d. The composite score of the four non-standard language subscales of the CCC-2 (Inappropriate initiation, Stereotyped language, Use of context, Nonverbal communication).

If hypotheses 1a-e are confirmed, then the following additional hypotheses will also be tested:

3. Improved scores in the two SCAN-3 C conditions will correlate with improved scores in the Noise subscale of the CHAPS in the RMHA group.
4. Improved scores in the three Sus-AA tests will correlate with improved scores in the Auditory attention span subscale of the CHAPS and the Attention subscale of the SIFTER in the RMHA group.

2.2 Methodology

2.2.1 Participants

<u>Sample size</u>

The total sample size for the study had initially planned to include 32 children, aged 7-12. This sample size was decided using a power sample calculation based on an estimated 1 SD effect size, $n = 16/ f2$, 80% power at 5% significance – where n = sample size per group, f = effect size. Invitation letters were posted to 60 potential participants. Prospective participants were identified via the APD Clinic at GOSH, where they had been previously assessed and given an APD diagnosis. Diagnostic criteria used in the clinic are described in the next section. Twenty-two replies were received from which seventeen comprised the final sample size. Figure 2.1 summarises the attrition process in the form of a flowchart.

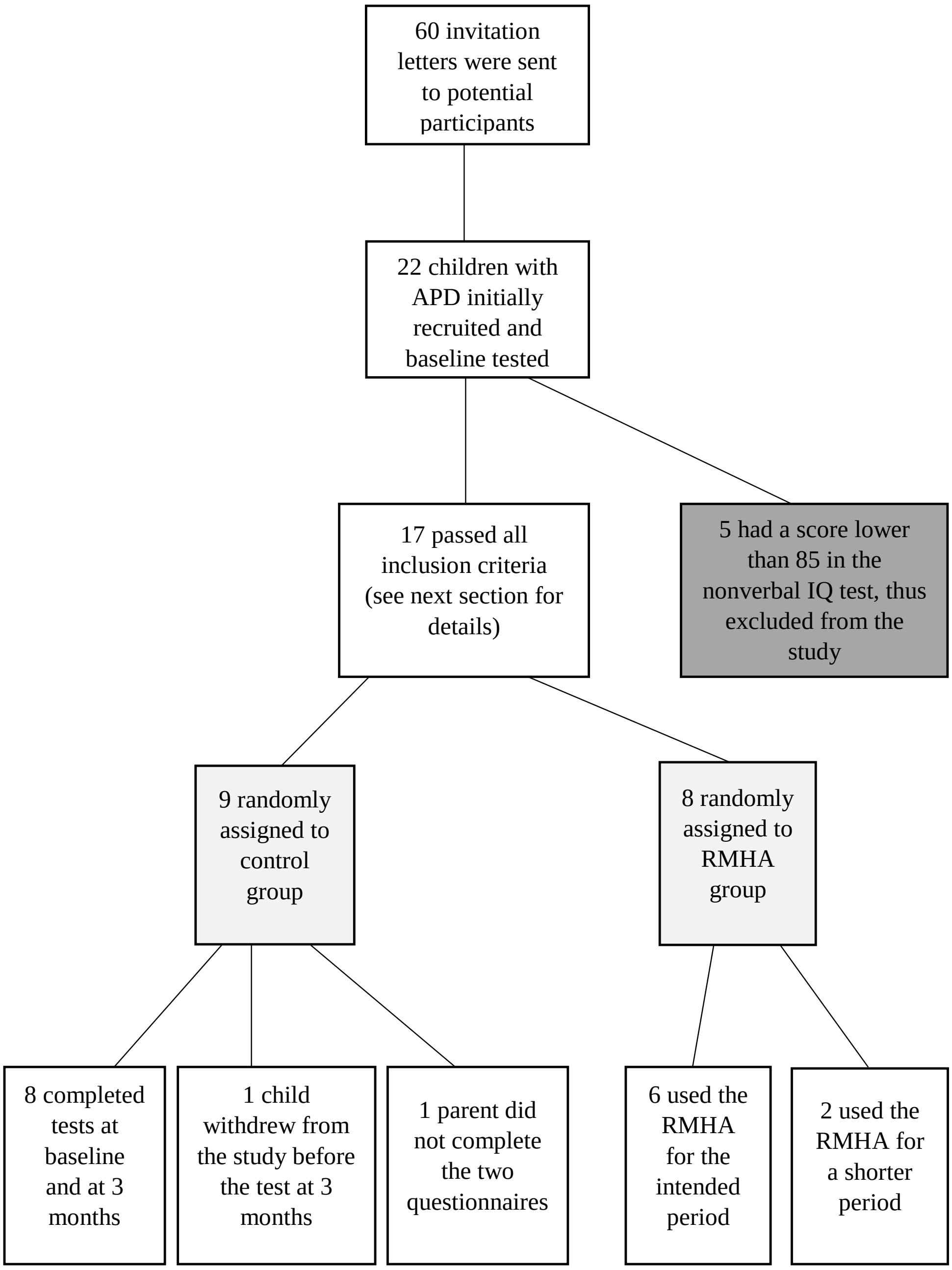

Figure 2.1 – Attrition process
Flow chart showing the attrition process for this study. Dark grey box refers to the participants excluded from the study, while light grey boxes show the final two comparison groups of the study. APD: Auditory Processing Disorder, IQ: Intelligence Quotient, RMHA: Remote Microphone Hearing Aid.

An Intention-to-Treat (ITT) analysis was chosen to prevent overestimation of the intervention's efficacy, to reduce bias caused by deletion of data and to maintain a large-enough sample size (Gupta, 2011). Therefore, the baseline data of the participant who withdrew before the tests at 3 months (see Figure 2.1) were included in the final analysis. The sample size had 10 girls and 7 boys between 7 years 6 months and 11 years 9 months, and mean age 9 years 7 months. All participants were native English speakers, with the majority (n = 14) coming from Greater London, while the rest (n = 3) came from other areas in England.

Inclusion and exclusion criteria

The inclusion criteria of the study were adopted from the diagnostic protocol used in GOSH (GOSH, 2018). These closely follow the ASHA (2005b) position statement. As discussed earlier there are no universally accepted criteria for diagnosis in APD research (see Appendix A). This means that findings from the study cannot be generalised to APD populations, but only to children diagnosed under these specific criteria used in GOSH. Nevertheless, the criteria used here aimed to minimise the effect that age, cognitive ability, hearing thresholds, co-occurring disorders and language might have on the presentation of APD.

In detail, the inclusion criteria used in the study were the following:

1. Diagnosis of APD based on routine clinical tests at GOSH. These tests were administered by qualified audiologists as part of the diagnostic protocol at the hospital. The routine clinical tests need to demonstrate the following:
    a. Reported parental SIN and other listening difficulties, and
    b. Normal peripheral hearing and middle ear functions (i.e. air conduction PTA below 20 dB in all octave frequencies between 250 Hz and 8 KHz [BSA, 2012], middle-ear pressure between -150 to +50 daPa, middle-ear admittance between 0.3 to 1.6 cm^3 and ear-canal volumes between 0.4 to 1.0 cm^3 [BSA, 2013]).
    c. Abnormal performance on the Auditory Figure Ground (AFG) test of the SCAN-3 C (a SIN test; i.e. at the 1st percentile score as per UK norms), and
    d. Abnormal performance (-2 SDs from the normative mean) on at least one AP test (i.e. frequency pattern test, dichotic digits test, duration pattern test, gaps-in-noise or random gap detection pattern tests), or a score of -3 SDs from the normative mean on one AP test, and/ or

    e. Abnormal performance (-2 SDs from the mean) on the Spatial advantage and Total advantage/ High-cue SRT conditions of the LiSN-S test.

2. No neurological or pervasive disorder or developmental delay. Children with a diagnosis of ADHD, epilepsy, ASD, DLD, Down syndrome were excluded from the study[18].
3. Nonverbal cognitive ability score (IQ) of more than 85.
4. Aged between 7-12 years.
5. Native English speakers.

Children who used RMHAs previously were excluded from the study. Inclusion criterion (4) was used as the position statements by the ASHA, AAA and the BSA all highlight that children need to be at least seven years old before they are tested for APD (AAA, 2010; ASHA, 2005b; BSA, 2011b). This is because younger children often show high variability in performance and difficulties comprehending the tasks (AAA, 2010; ASHA, 200b). In addition, the study aimed to avoid a broad age range in the sample. Criteria (2) and (3) were set to minimise the possibility of APD presentation explained by other neurological or cognitive problems and to exclude a link to other pervasive learning disorders or developmental delays. Criterion (5) was used to minimise the effect that possible poor language skills by non-native English speakers may have had on linguistic test scores.

Design

An RCT design was used for this study. Children were randomised into two groups:

- Group A or Control group comprised children with APD who did not receive a RMHA for the study period.
- Group B or RMHA/ intervention group included children with APD who received RMHAs after baseline testing.

---

[18] Great Ormond Street Hospital requires that children undergo such assessments prior to any APD assessment should there be concerns for language or attention deficits. It is recognised here that some children in the sample might qualify for ADHD or DLD diagnosis even without having initial concerns of their attention or language skills. This is a limitation in the study and in the set of criteria used in GOSH and in APD diagnosis in general. Unfortunately, it was not possible to have additional assessments for these disorders during the study.

Stratified randomisation was used (Kang, Ragan, & Park, 2008) in order to balance the characteristics of age and gender between the two groups, as there are indications that these factors might have an influence on AP performance and attentional networks (Coch, Sanders, & Neville, 2005; Yathiraj & Vanaja, 2015). Both groups were randomly constructed based on the two strata of age (9 years 6 months or younger and 9 years 7 months or older) and gender (boys, girls). Permuted blocks within each stratum were used to ensure balance between the two groups, as simple randomisation cannot guarantee that. For this, the allocation ratio was 1:1, block sizes were sizes of 2 and 5 and a computer random number generator was used (with an Excel formula). With this application, the control group comprised 6 girls and 3 boys with a mean age of 9 years and 7 months (SD 18 months, and were between 7 years, 6 months and 11 years, 8 months), while the intervention group had 4 girls and 4 boys with a mean age of 9 years and 6 months (SD 15 months, and were between 7 years, 9 months and 11 years, 1 month). Tests took place at baseline and at 3 months, a period during which group B used the RMHA system. The experimenter was not blind to group allocation which may have introduced bias during testing.

Ethical issues

This study has been reviewed and given a favourable opinion by the Bloomsbury Research Ethics Committee (REC), REC reference: 14/LO/1509. Information sheets for both the parents and their children were posted during the recruiting phase and once communication had been made participants were given a week to decide whether they wished to be part of the study. On their first test visit consent forms were given to parents to sign, while written assent was sought from the children in the presence of their parents. Parents and children were encouraged to ask questions at any point and were only allowed to take part once they understood the purpose and procedures of the study.

Participants had the right to withdraw from the study at any point they wished to. This was outlined on the consent form and explained verbally. There was neither a loss of benefit nor a penalty if participants decided to withdraw from the study. Children and their parents were informed that their information would be anonymised and kept confidential. Data were pseudo-anonymised prior to analysis by using numbered codes. Data were stored in a password protected file on a secure password-protected database and hard copies were locked in a filing unit, in a code-protected room within University College London (UCL), in accordance with the Data Protection Act 1998.

After the completion of tests at 3 months the control group was given the same intervention to use, as RMHA use was part of the management recommendations they received from the audiology consultants at GOSH. This also served as an incentive for the control group to stay committed to the study. A £10 book token was given to each child attending every test session, while parents received a reimbursement of up to £15 for travel expenses.

2.2.2 RMHAs

The device

The RMHA system used in this study was the Mini-mic coupled with the ReSound Up hearing aids (see Figure 1.4 in Chapter 1). The Mini-mics are not considered a standard personal FM system as they broadcast at a lower frequency (2.4 GHz). Despite this, they are still capable of producing a wireless signal to the ear receivers and they offer the same advantages as FM systems, such as reduction of reverberation and consistent signal when teachers move around. Mini-mics have the advantage of being smaller and more portable compared to standard FM systems. They do, however have a smaller range compared to personal FM systems. The Mini-mic can stay connected to the ear instruments within 7m, which was considered adequate for normal-sized classrooms.

Children’s involvement

Children in group B were fitted with the system binaurally. There was no blinding either for the participants nor the researcher during assignment of treatment, as the researcher needed to know to which children he would give the RMHAs, while children could not have been unaware of the intervention as hearing aids are obvious devices worn in the ears. Adding to that, there could not have been a placebo group (i.e. a group that receives a RMHA that is not functioning), as children participants, teachers and parents would have eventually noticed that the device was not functioning. A 3-month period is too long for a non-functioning RMHA to go unnoticed. A placebo group, though, could potentially be used in future studies testing same-day effectiveness of the system, or perhaps even in short-term trials of a few days or weeks.

Children were asked to use the RMHAs daily during school time (5 days per week), only in lecture-based subjects. They were asked to take them off during physical education and recess to protect them from getting damaged or lost. The system had to stay at school, ensuring that it would not be forgotten at home. Two of the children did not use it for the complete duration of 3 months due to delays from the manufacturer in ordering and shipping the devices. Consequently, these two children used the RMHA for a period of 7 weeks. As ITT analysis was used (Gupta, 2011), these children remained in the study and their data were included in the analysis. Electronic communication with parents was initiated by the researcher every month to address any issues related to RMHA use. Communication with the parents of the control group was also initiated monthly to mirror what was being done for the intervention group, to increase engagement and to minimise the risk of participant withdrawal due to long periods without communication.

The RMHA was first presented to children during baseline tests, so that they knew what they were going to be using. Measurements of the size of the external ear and the ear canal opening were taken during this time. These measurements ensure a correct fit and comfort for the children when wearing the hearing aids. The ear piece of the RMHA is small, while a thin transparent tube helps hang the piece on top of the ear. A third shorter tube tucks in inside the outer ear cavity serving as another contact point which helps keep the ear piece more stable. The small size of the hearing aids was something children and parents liked, commenting that the instruments would not draw immediate attention. Additionally, the ear pieces are non-occluding open-ear devices (i.e. the dome that fits the opening of the ear canal has open holes). Using these type of domes instead of occluding closed ones was found to be beneficial as the child can communicate and converse well with classmates and be in touch with environmental sounds while simultaneously receive the clear voice of the teacher (Johnston et al., 2009). Another study on adults with hearing loss also showed that non-occluding open-ear devices were used more by subjects in comparison to occluding closed-ear fittings because of reduced acoustic feedback produced by the former group of devices, as well as elimination of user's own voice (Gnewikow & Moss, 2006).

Children had the option to choose a colour from a wide range of choices offered by the manufacturer. This made the RMHA more personalised, which might have increased children's willingness to use the device. All RMHAs were programmed by the manufacturing company, GN ReSound, as per the company's protocol for devices used by individuals with no hearing loss. This means that no amplification gain was programmed as that is an approach used only in children with peripheral hearing loss (Flexer, 2014). This prevents over-amplification, distortion and noise exposure, which may induce hearing loss (Flexer, 2014).

One-hour school visits were arranged by the researcher to give the device to children and teachers. In attendance were the child, the parent, the teacher(s) and occasionally the Special Educational Needs Co-ordinators (SENCOs). During the session, the following topics were discussed: functioning of the RMHA, connecting the microphone to the ear pieces, changing batteries, charging, troubleshooting, cleaning and storing the system. Participants were encouraged to ask questions about the RMHA and at the end of the discussion the device was fitted and tested. No problems arose during any of the school visits. A detailed guide with the topics covered during the session was handed to teachers and parents and electronic copies of the document were sent upon request. After completion of the trial all participants got to keep the RMHA as this was part of the recommendations by the APD clinic at GOSH. As noted earlier, the control group also received RMHAs after study completion at 3 months. For that reason, school visits for children in the control group were organised by the researcher after completion of the study, following the same procedure as described above.

In two cases, the RMHA was used alongside a classroom sound-field FM system. Personal RMHAs can work normally without interferences when used in combination with a sound-field FM system (Smaldino, Crandell, & Flexer, 2005), hence these children continued taking part in the study using both systems simultaneously. Finally, parents were asked prior to the start of the study not to use any AT intervention during the study period regardless of the group their children would be assigned to.

Teachers' involvement

Teachers were provided with an information sheet and a consent form during the researcher's school visit. They had the opportunity to review the information and sign and post the consent form within a week. As children already agreed to take part in the study, all teachers were also on board with taking part. Teachers' participation included them wearing the microphone during classroom time and completing the SIFTER questionnaire. The researcher liaised with teachers, SENCOs and headmasters before the visits to ensure that schools were aware of their student's agreement for participation and to arrange the school visit.

During school visits, teachers were shown how to wear the microphone by using the special lanyard provided in the manufacturer's box. Using this non-adjustable lanyard meant that it kept the microphone at a constant distance of about 20cm from the mouth of the teacher. The use of the lanyard ensured homogeneity across teachers' use of the RMHA, as speech was received and enhanced in the same way across all cases. In addition, teachers were asked to talk in their normal voice, at a natural volume, without exaggerating pronunciation and to avoid introducing noise close to the microphone. Furthermore, they were shown the correct way the microphone should face to prevent obstructing the receivers. Finally, they were instructed to mute the microphone when they had to privately talk to other students or classroom visitors so that they minimise distractions for the child wearing the ear pieces. The same guideline document given to parents was handed to teachers, too.

2.2.3 Tests and questionnaires

Pure tone audiometry

During the first test visit, children were screened for normal hearing (see details under inclusion criteria in Section 2.2.1) and normal cognitive ability. The PTA is used to assess the hearing threshold of individuals. Participants required to have normal peripheral hearing to be included in the study. The PTA was performed in a double-walled sound booth, using a calibrated GSI 61 clinical audiometer. Children in the study already had PTA tests during their APD assessment, hence it was expected that all children would return results within normal levels. The duration of the test was approximately 15 minutes.

Cognitive ability

As children's cognitive ability could affect their performance on complex auditory behavioural tests (AAA, 2010), normal levels in a cognitive test was required in order for children to be included in the trial. Their cognitive ability was measured through the Wechsler NonVerbal (WNV) Scale of Ability (Wechsler & Naglieri, 2006). The test reduces verbal content so that the effect of possible language difficulties on children's performance on this test is minimised (Wechsler & Naglieri, 2006). Two subtests were used; Matrices, measuring general ability and perceptual reasoning and Spatial span forward and Spatial span backward, which measure general ability and visual working memory. The WNV test was administered on a large desk and children were seated opposite the tester. The raw scores from each test were converted into *T* scores (which have a mean of 50 and a SD of 10). Then a composite full scale score was calculated by adding up the two *T* scores from Matrices and Spatial Span. The composite full scale score has a mean of 100 and a SD of 15 and it reflects the score from a standard IQ test (Wechsler & Naglieri, 2006). Only children scoring higher than 85 on the composite full scale score were included in the study. Five out of twenty-two children failed to pass this threshold, resulting in their exclusion from the study (see attrition chart in Figure 2.1). The duration of the test was approximately 20 minutes.

Speech-in-noise perception

The test used to measure children's SIN perception was the SCAN-3 test for children (R. W. Keith, 2009). This test has several conditions, but the study used two of them. Both measure the ability of the child to recognise speech words in background speech-babble noise. A description of these two conditions is given in Table 2.1.

Table 2.1 – SCAN-3 C conditions
The two SCAN-3 C conditions, AFG +12 dB and AFG 0 dB, and what they assess. AFG: Auditory Figure Ground.

SCAN-3 C	Task description
AFG +12 dB	Assesses the child's ability to recognise words in the presence of competing background speech-babble noise. The target voice is 12 dB greater in intensity than the background noise.
AFG 0 dB	Assesses the child's ability to recognise words in the presence of competing background speech-babble noise. The target voice has the same intensity as the background noise.
	Total duration: 20 minutes

Both conditions assess each ear separately. Raw scores from the left and right ear for each test were combined to create a composite raw score which was then converted into a scaled score, with a mean of 10 and a SD of 3. The test was administered via Sennheiser HD 215 headphones, with children and the tester sitting opposite each other on a large desk. A ProBook HP laptop was used to play the tracks, with the laptop monitor facing away from the child to eliminate distractions. The laptop was connected to a calibrated GSI 61 clinical audiometer set to a constant intensity level of 50 dB, as per the SCAN-3 C manual instructions (R. W. Keith, 2009).

Attention

Different types of auditory and visual attention were measured through the TEACh test, which is a validated attention test (Manly et al., 1999). Types of selective, sustained and divided attention were measured by the different subtests summarised in Table 2.2 below. The TEACh test minimises the effect other skills (e.g. memory, language, comprehension) might have on test results (Manly et al., 1999). There is no statistically significant correlation between the first three tasks (see Table 2.2) and IQ scores from the WISC III, while for the latter two there are statistically significant weak relationships with the WISC, r = .21 for Walk, don’t walk and r = .17 for Code transmission (Manly et al., 1999).

Table 2.2 – TEACh subtests
The six TEACh conditions, what they measure and what children need to do in each of them. DT: Dual Task, TEACh: Test of Everyday Attention for Children.

TEACh subtest	Type of attention measured and task description
Sky search	Measures selective visual attention. Children find visual targets among other visual distractors on an A3-sized paper.
Score!	Measures sustained auditory attention. Children count the number of sounds they hear. As the task is long, they need to self-sustain their attention. The sounds resemble computer game sounds of spaceships firing. No other details on these sounds are provided in the TEACh manual.
Sky search DT	Measures divided auditory-visual attention in a dual task. Children need to find visual targets among distractors, while at the same time count the number of sounds they hear.
Score DT	Measures divided auditory attention. This is another dual task. Children have to count the number of sounds they hear while attempting to find a target word in a recorded news report.
Walk, don't walk	Measures sustained auditory attention and response inhibition. Children put a mark on a special A4-sized paper after each sound they hear while paying attention not to mark when they hear a different sound (being similar in its first part to the target sound).
Code transmission	Measures sustained auditory attention. Children sustain their attention in a monotonous task similar to an n-back task. They need to repeat the number that comes before a specific string of numbers (named the code) each time they detect that string.
	Total duration: 60 minutes

As discussed in the literature review earlier, the RMHA is hypothesised to bring improvements only in the Sus-AA and Div-AVA tasks, (i.e. in the Score!, Sky search DT, Walk don’t walk and Code transmission conditions). The tasks of Sel-VA and Div-AA (Sky search and Score DT) serve as control conditions and are expected to remain unchanged following the intervention period. Test-retest reliability (with age as a control factor) was calculated for these subtests and the test-retest correlation coefficients are given in Table 2.3 below (Manly et al., 1999).

Table 2.3 – Test-retest correlation coefficients for TEACh
Test-retest correlation coefficient for each TEACh subtest (Manly et al., 1999). DT: Dual Task, TEACh: Test of Everyday Attention for Children.

TEACh subtest	Test-retest correlation coefficient
Sky search attention score	.75
Score!	.76
Sky search DT	.81
Score DT	.71
Walk, don't walk	.71
Code transmission	.78

Raw scores for each test were initially collected and then converted into age-scaled scores (mean of 10, SD of 3). The test was administered through a CD player connected to two ALTEC LANSING speakers. The setting was similar to the one used for the SIN test (large desk with the child sitting opposite the tester) but instead of headphones, speakers were used. The speakers were facing the child from the front centre and the sound level was measured from the child's seat at ear level with a calibrated Casella CEL-450C sound level meter at 60 dB Sound Pressure Level (SPL) using a 10-second average of a warble sound.

Working memory

The memory test administered was the AWMA, a validated computer-based working memory test (Alloway, 2007). The tasks chosen measure verbal short-term memory and auditory verbal working memory. While they both refer to the ability to hold verbal information in memory for brief periods, the latter type of memory also deals with manipulation of this information (Alloway, 2007). Table 2.4 outlines the three different subtests of the AWMA.

Table 2.4 – AWMA subtests
The three subtests of the AWMA and what they measure. AWMA: Automated Working Memory Assessment.

AWMA subtest	Type of memory measured and task description
Digit recall	Measures verbal short-term memory. Verbal short-term memory stores information based on spoken language (in this case digits). Children repeat back the digits they heard in the order they were presented.
Listening recall	Measures verbal working memory. This subtest assesses verbal working memory and its ability to store information. Children listen to sentences. They are asked to repeat back words they stored in their memory
Listening recall processing	Measures verbal working memory. This subtest assesses verbal working memory and its ability to store and manipulate verbal information. This subtest is combined with the above Listening recall subtest. Children need to say if the sentence they heard previously is true or false (see Listening recall subtest).
	Total duration: 25 minutes

The test-retest reliability was examined in 128 individuals between the ages of 4 years 10 months and 22 years 5 months and within a period of 4 weeks (Alloway, 2007). Test reliability of the AWMA is good, with correlation coefficients of .89 for Digit recall, .88 for Listening recall and .84 for Listening recall processing. The test was administered on a laptop through its installed software. Raw scores were automatically transformed and presented into standard scores (with a mean of 100 and SD of 15) in a generated final report (Alloway, 2007). The setting was the same as that of the TEACh test (child sitting opposite tester, loudspeakers facing the child from the front centre). The laptop had its monitor facing away from the child to minimise distractions. The sound level was calibrated using the same instrument and procedure as in the TEACh test.

Questionnaires

Parents were given two questionnaires to complete while children were tested. The two parental questionnaires analysed were the CHAPS (Smoski, Brunt, & Tannahill, 1998) and the CCC-2 (Bishop, 2003). The former questionnaire assesses aspects of children's listening and cognition, while the CCC-2 is a 70-item questionnaire that screens for language and communication problems. The CHAPS was selected because it measures the responses of parents on how their children perform in different listening conditions and assesses aspects of their cognition, as well (e.g. auditory attention span and auditory memory). The Noise and Multiple inputs subscales from CHAPS were monitored as they respectively measure listening-in-noise and listening when visual forms of input are present. In addition, the Auditory attention span and the Auditory memory sequencing subscales were used (Smoski et al., 1998), mirroring the behavioural tests of TEACh (i.e. Sus-AA) and AWMA, respectively. Even though the CHAPS converts question scores into means for each subscale (i.e. sum of scores divided by number of questions for each subscale), this raw score is not converted into standardised scores as there are no normative data for the questionnaire. Despite this, the CHAPS was included in the study because it is a questionnaire used for screening APD in clinics and its subscales match the behavioural outcome measures of the study. The CHAPS conditions are described in Table 2.5 below.

The CCC-2 questionnaire assesses different language and communication aspects in children (Bishop, 2003). Even though the study is not using behavioural language and communication measures, the inclusion of the CCC-2 helps monitor different language and communication aspects in children based on parental responses. Inclusion of additional behavioural tests on language and communication would increase testing time for children and was thus decided to use the parental CCC-2 questionnaire instead. Moreover, aspects of language and communication were not the primary focus of this study. The CCC-2 has good reliability, with all of the clustered items having internal consistency *alpha* between .65 and .80. Raw scores from the questionnaire were transformed into scaled scores (mean of 10, SD of 3) using normative tables from the manual (Bishop, 2003). The different conditions used from the CCC-2 questionnaire are outlined in Table 2.5.

Table 2.5 – CHAPS and CCC-2 subscales

The different subscales of the CHAPS and CCC-2 questionnaires and what they assess (Bishop, 2003; Smoski et al., 1998). CCC-2: Children's Communication Checklist – 2, CHAPS: Children's Auditory Performance Scale, DLD: Developmental Language Development, PLI: Pragmatic Language Impairment.

CHAPS subscale	Subscale description
Noise	Assess children's performance on each of these listening conditions.
Multiple inputs	
Auditory memory sequencing	Assess the two cognitive functions of auditory memory sequencing and auditory attention span in children.
Auditory attention span	
	Duration: 15 minutes
CCC-2 subscale	
Speech	Assess aspects of standard language such as structure, vocabulary and discourse. These subscales may be found impaired in children with DLD but are not used to diagnose DLD.
Syntax	
Semantics	
Coherence	
Inappropriate initiation	Assess pragmatic aspects of communication (non-standard language) not readily tested by conventional standard language tests. These subscales may be deficient in children who have PLI, but do not necessarily diagnose PLI.
Stereotyped language	
Use of context	
Nonverbal communication	
Social relations	Assess behaviours that may be found impaired in cases of ASD but do not diagnose ASD. They might warrant further investigation, though.
Interests	
	Duration: 15 minutes

Instead of analysing the individual CCC-2 subscales, composite scores were created. Even though the General Communication Composite (GCC) is used in the manual as a composite score, it was elected to create two separate composite scores of standard and non-standard language aspects. This is because the GCC score is the sum of all eight subscales shown in Table 2.5, while the study aimed to separately examine the effects of RMHAs on standard and non-standard language aspects. More details on this are given later in Results (Section 2.3.6).

Finally, the SIFTER questionnaire was given to teachers to complete. This is a 15-item questionnaire and is used as a screening tool targeting educational risk (Wilson et al., 2011). It has five subscales but the ones this study focused on were the Attention and Communication subscales. The questionnaire was left with teachers to complete within a week but this resulted in a low return rate (6 out of 17), while 3 of the returned questionnaires were received several weeks later and upon request. Due to the low return rate and delayed completion, the SIFTER was excluded from the analysis.

Testing procedure

During testing children's parent or parents were also seated in the room. Parents were asked to keep quiet, put their phones on quiet mode and avoid other distractions. As this was a long test battery (more than 2 hours), regular breaks were given. The three behavioural tests and the conditions within each test were randomised (using simple, computer-generated randomisation with Excel formulas) for each test session and for each participant in order to control for order effects. The tasks of the TEACh test are set in a hierarchical order, building each task based on the previous one, therefore task randomisation for this test was not used.

Verification of RMHAs

Verification of RMHAs gives an objective feedback on the performance of the RMHA (Haastrup & Dworsack-Dodge, 2011). One of the shortcomings of the present study was that verification of the devices was performed at the end instead of at the beginning of the intervention period, as this was not considered during the initial design. In addition, as the verification procedure was not properly followed the first time participants were asked to send their instruments again for check after the study had been completed. There were 3 instruments received from the 8 given out to the intervention group. Verification of the Mini-mic and hearing aids was carried out using the Aurical test chamber by Otometrics (calibrated annually by Otometrics) and the OTOsuite software (ran on an HP Desktop with Windows 7) following the standard procedure for verification of hearing instruments with wireless transmitters (Haastrup & Dworsack-Dodge, 2011; Otometrics, 2015). During the procedure for verification there were two measurements that were obtained, the reference and the transparency measurement (Haastrup & Dworsack-Dodge, 2011). The signal used throughout the verification process was the International Speech Test

Signal (ISTS; Holube, Fredelake, Vlaming, & Kollmeier, 2010) and for each measurement the signal was played for 30 seconds.

The reference measurement measures the hearing aid output with a 65 dB SPL input. To obtain this measurement, the hearing aid was attached to the coupler inside the Aurical test chamber. The chamber has a reference microphone which was placed as close as possible to the front microphone of the hearing aid (without touching it). With the test chamber closed the reference measurement was taken using the ISTS signal played via the OTOsuite software (Haastrup & Dworsack-Dodge, 2011). To obtain the transparency measurement the hearing aid was removed from the chamber and replaced with the microphone of the RMHA system (in this study the Mini-mic). The reference microphone of the chamber was placed as close as possible to the microphone of the Mini-mic (again without touching it). The coupler of the test chamber is outside the chamber unit and connected to the chamber via a cable. The hearing instrument was then attached to that external coupler, the chamber was closed and another measurement at 65 dB SPL was taken using the ISTS through the OTOsuite software. This is the transparency measurement and it was compared to the reference. According to the AAA (2011) guidelines on RMHA technologies, the transparency measurement should be the same as the reference with a +/- 2 dB tolerance.

The procedure described above was followed to verify the Mini-mic and hearing aids used in this study. The reference measurement was 64 dB SPL and the transparency measurement was 63 dB SPL, which was within the tolerance range the AAA (2011) sets (i.e. +/- 2 dB compared to the 64 dB SPL reference measurement) and indicates that the system was functioning as intended. The same measurements were obtained for all three devices tested. Even though the rest of the five devices in the intervention group were not returned for verification, it is expected that they would function similarly as they were programmed and sent with the same order prior to commencement of the study.

2.2.4 Statistical analysis

Data were analysed in SPSS 22 statistics software, using a mixed (or split-plot) ANOVA, with group (i.e. control or RMHA) as the between-subjects factor and time (i.e. baseline and 3 months) the within-subjects factor. Post-hoc correction for multiple comparisons was made using Bonferroni correction for the dependent variable of time and the interaction between group and time. There was no correction for multiple comparisons for the multiple administered tests in this study. While correction for multiple comparisons reduces the chance for Type I error, it increases the chance for Type II error (Armstrong, 2014), and consequently reduces the statistical power of the study (McLaughlin & Sainani, 2014). Adjustment procedures such as the Bonferroni or Benjamin-Hochberg assume that the outcome measures are independent of each other (McDonald, 2014), while the former procedure can severely over-correct when the outcome measures are linked between them (Schulz & Grimes, 2005). In the current study, different outcome measures are not expected to be entirely independent between them. As discussed earlier, there is evidence to show that selective attention influences SIN (Mesgarani & Chang, 2012; O’Sullivan et al., 2015; Zion Golumbic et al., 2013), while the subscales of the CHAPS reflect the behavioural tests of Sus-AA, memory, SIN. For these reasons it was decided not to correct for multiple comparisons. This may have an impact on the results and their interpretation, while the possibility of making a Type I error will be discussed later in the interpretation of the findings. A way to deal with multiplicity concerns is to create composite scores, where appropriate (Schulz & Grimes, 2005). Therefore, to compensate for the lack of multiple comparisons some outcome measures will be merged together to reduce the comparisons. Specifically, the three Sus-AA tests of the TEACh (i.e. Score!, Walk Don’t Walk, Code Transmission) will be averaged to create a composite score. Similarly, the CCC-2 subscales will be averaged and grouped into two composite scores, the standard language (i.e. Speech, Syntax, Semantics, Context) and non-standard language scores (i.e. Inappropriate Initiation, Stereotyped Language, Use of Context, Nonverbal communication). Details on each of these tasks/ subscales are given in the description of each test/ questionnaire (See Section 2.2.3). Moreover, if the intervention group returns statistically significant improved scores over the control group in the behavioural tests and in the questionnaires, then correlation analyses using Pearson's correlation will be applied, as outlined in the hypotheses (see Section 2.1.2).

2.3 Results

A post-hoc power calculation was performed on the actual sample size (*n* = 17) to detect 1 SD effect size at 5% significance, which resulted in a statistical power of 49%. Due to the limited recruitment from GOSH and time constraints the study was carried out despite the reduced statistical power. The implications of this are outlined later in Discussion.

One participant from the control group did not come back for a retest, thus results for that child were collected only at baseline. In addition, a parent from the control group did not complete the two questionnaires at 3 months (see Figure 2.1). As an ITT analysis was used (Gupta, 2011) – see Section 2.2.1 under Sample size for more details – these cases were not removed from the analysis and their missing data were imputed using a multiple imputation model for missing data (Sterne et al., 2009). Five pooled imputations were initially calculated and then averaged in order to draw the final imputed value for the above missing data. Other methods of controlling for missing data were considered, such as removing cases with missing data (Sterne et al., 2009). As an ITT analysis was followed in this study to reduce the chance of overestimation of any effects, removing cases with missing values was not used. Single imputation models were also considered, such as replacing missing data with the mean of the observed values, or replacing missing data with the last measured value (He, 2010; Sterne et al., 2009). However, such methods usually underestimate the standard errors of the estimates, as single imputation models assume that the missing value is known with certainty, when in fact it is not (He, 2010; Sterne et al., 2009). In multiple imputation strategies, though, several sets of possible values are combined to yield an inference (He, 2010) and this was considered to be the most appropriate method for the study design.

2.3.1 Pure-tone audiometry and nonverbal IQ

All children had normal hearing thresholds, below 20 dB Hearing Level (HL) in each frequency between 250 Hz and 8 KHz (BSA, 2012), except for two, who had some frequencies between 20 to 30 dB HL (30 dB HL in the right and left ear at 250 Hz for one participant, and for the other 25 dB HL in the right ear at 250 Hz and 20 dB HL in the left ear at 500 Hz). Both children were reported to have a cold when they came in for the tests. Despite this, these two cases were included in the study after confirming that they previously had normal PTAs and normal peripheral hearing based on their medical notes from GOSH. Their below average performance in the PTA was attributed to their illness, which was expected to cause a temporary drop in hearing levels. An independent samples t-test was conducted to compare the PTA scores between the two groups (right and left ear separately). No differences were found between the scores of the two groups in the right ear, $t(13) = .63$, $p = .534$, $d = .30$ (95% Confidence Interval [CI] [-.65, 1.25]) nor in the left ear, $t(13) = -1.65$, $p = .121$, , $d = -.80$ (95% CI [-1.78, .20]). Furthermore, only children with a score above 85 on the WNV test were included in the study. The 17 children meeting this criterion gave a mean composite full scale score of 99.4 (SD = 11.7) on the WNV test (Wechsler & Naglieri, 2006).

2.3.2 Speech-in-noise test

Outliers were checked by looking for values in the studentised residuals beyond +/- 3 SDs. No outlying values were observed in any of the two SCAN-3 C conditions. Normal distribution was recorded in both conditions using the Shapiro-Wilk test of normality. The assumptions of Homogeneity of variances (using the Levene’s test of equality of error variance) and Equality of covariance matrices were met in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 2.2 shows the results for the two conditions of the SCAN-3 C. No significant changes were observed in neither the AFG +12 dB nor the AFG 0 dB.

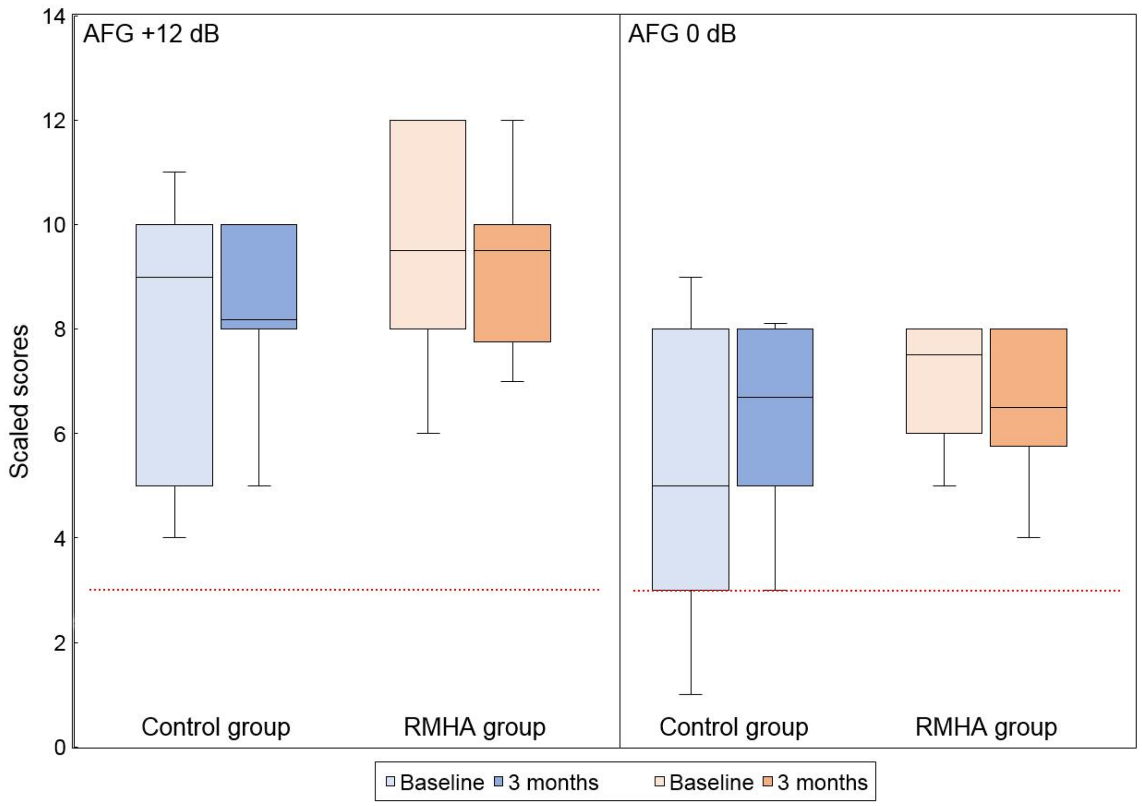

Figure 2.2 – Results of the SCAN-3 C
Boxplots of scaled scores of the SCAN-3 C AFG +12 dB and AFG 0 dB, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 3 scaled scores). AFG: Auditory Figure Ground, RMHA: Remote Microphone Hearing Aid.

Table 2.6 presents the mean values of the control and intervention group at baseline and at 3 months.

Table 2.6 – SCAN-3 C mean scores and SDs
The mean scores and SDs for the AFG +12 dB and AFG 0 dB in the two groups at baseline and at 3 months. Disordered performance: ≤ 3, borderline performance: > 3 and ≤ 7, and normal performance: > 7. AFG: Auditory Figure Ground, SD: Standard Deviation.

SCAN-3 C	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
AFG +12 dB	7.77 (2.58)	8.42 (1.59)	9.62 (2.26)	9.12 (1.72)
AFG 0 dB	5.22 (2.72)	6.30 (1.72)	7.50 (2.13)	7.00 (2.44)

Table 2.7 presents the *p*-values, effect sizes and CI of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. As observed, there was neither statistical significance in the interaction between group and time nor in the main effect of group and main effect of time in any of the conditions.

Table 2.7 – SCAN-3 C significance
The *p*-values, *partial* η^2 (η_p^2) for the interaction between group and time and for the main effect of group and time in the two SCAN-3 C conditions. Confidence intervals of the effect size are also provided. AFG: Auditory Figure Ground, CI: Confidence Interval. $^\dagger p < .10$

Condition	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
AFG +12 dB	.353	.058	.000, .335	$.065^\dagger$	.209	.000, .490	.731	.008	.000, .225
AFG 0 dB	.266	.082	.000, .367	.171	.121	.000, .410	.884	.001	.000, .052

2.3.3 Attention

There were no outliers in any of the conditions of the TEACh test (examined by looking for values in the studentised residuals beyond +/- 3 SDs). All conditions had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene’s test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

In the two experimental conditions no significant changes were observed over 3 months in any of the two groups. Figure 2.3 summarises these results.

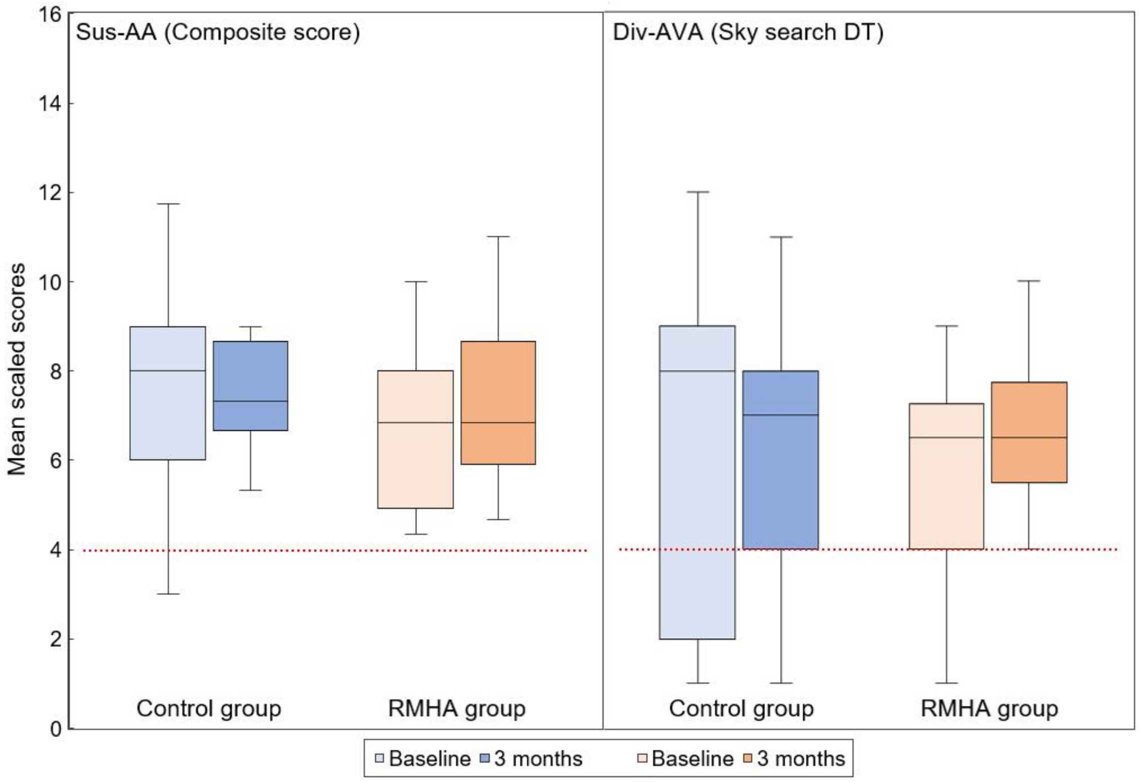

Figure 2.3 – Results of the TEACh experimental conditions
Boxplots of scaled scores of the composite score of the three Sus-AA tasks and the task measuring Div-AVA, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AVA: Divided Auditory-Visual Attention, DT: Dual Task, RMHA: Remote Microphone Hearing Aid, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children.

In the two control conditions measuring Div-AA and Sel-VA, no statistically significant interaction was found between group and time. Figure 2.4 summarises the findings.

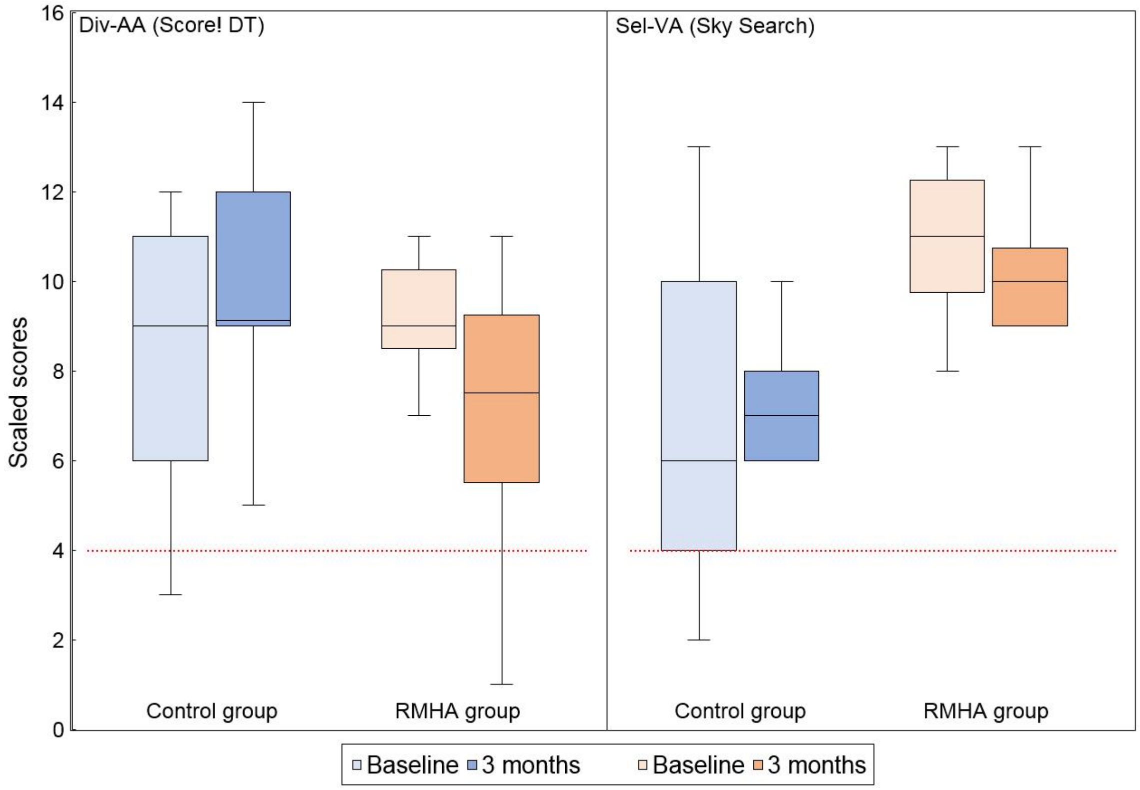

Figure 2.4 – Results of the TEACh control conditions
Boxplots of scaled scores of the tasks measuring Div-AA and Sel-VA, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AA: Divided Auditory Attention, DT: Dual Task, RMHA: Remote Microphone Hearing Aid, Sel-VA: Selective Visual Attention, TEACh: Test of Everyday Attention for Children.

Table 2.8 presents the mean values of the control and intervention group at baseline and at 3 months.

Table 2.8 – TEACh mean scores and SDs
The mean scores and SDs for the experimental and control conditions of the TEACh test in the two groups at baseline and at 3 months. Disordered performance: ≤ 4, borderline performance: > 4 and ≤ 7, and normal performance: > 7. Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone hearing aid, SD: Standard Deviation, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children.

TEACh	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Sus-AA	7.40 (2.58)	7.22 (1.81)	6.79 (2.07)	6.83 (2.82)
Div-AVA	6.33 (4.09)	6.22 (3.30)	5.50 (3.02)	6.75 (2.31)
Sel-VA	7.33 (4.12)	7.30 (1.46)	10.87 (1.88)	10.00 (2.26)
Div-AA	8.44 (3.08)	9.56 (2.87)	9.00 (3.16)	7.00 (3.29)

Table 2.9 presents the *p*-values, effect sizes and CI of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. There were no statistically

significant interactions between group and time in any of the scores in the different TEACh conditions. There was only statistical significance in the main effect of group for the Sel-VA task, $F(1, 15) = 8.45$, $p = .011$, $\eta_p^2 = .360$. This means that if the two scores at baseline and 3 months are collapsed across time, a significant difference between the two groups is observed. In other words, the intervention group had better scores than the control group regardless of the time variable (i.e. baseline and post-intervention scores merged together for each group).

Table 2.9 – TEACh significance
The *p*-values, η_p^2 and CI of the effect sizes for the interaction between group and time and for main effects of group and time in the TEACh conditions. The three Sus-AA tasks were averaged to create a composite score (see Section 2.2.4). CI: Confidence Interval, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children. $^{\dagger}p < .10$, $^{*}p < .05$

Type of attention	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Sus-AA (composite score)	.856	.002	.000, .079	.609	.018	.000, .252	.911	.001	.000, .031
Div-AVA	.526	.027	.000, .273	.901	.001	.000, .038	.596	.019	.000, .255
Sel-VA	.572	.022	.000, .261	.011*	.360	.023, .603	.544	.025	.000, .269
Div-AA	$.085^{\dagger}$	.185	.000, .470	.432	.042	.000, .300	.612	.018	.000, .261

2.3.4 *Working memory*

Outliers were checked by looking for values in the studentised residuals beyond +/- 3 SDs and this did not reveal any outlying values in any of the three AWMA conditions. Normal distribution was observed in all conditions using the Shapiro-Wilk test of normality. Homogeneity of variances (using the Levene’s test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 2.5 presents the findings of the three AWMA conditions.

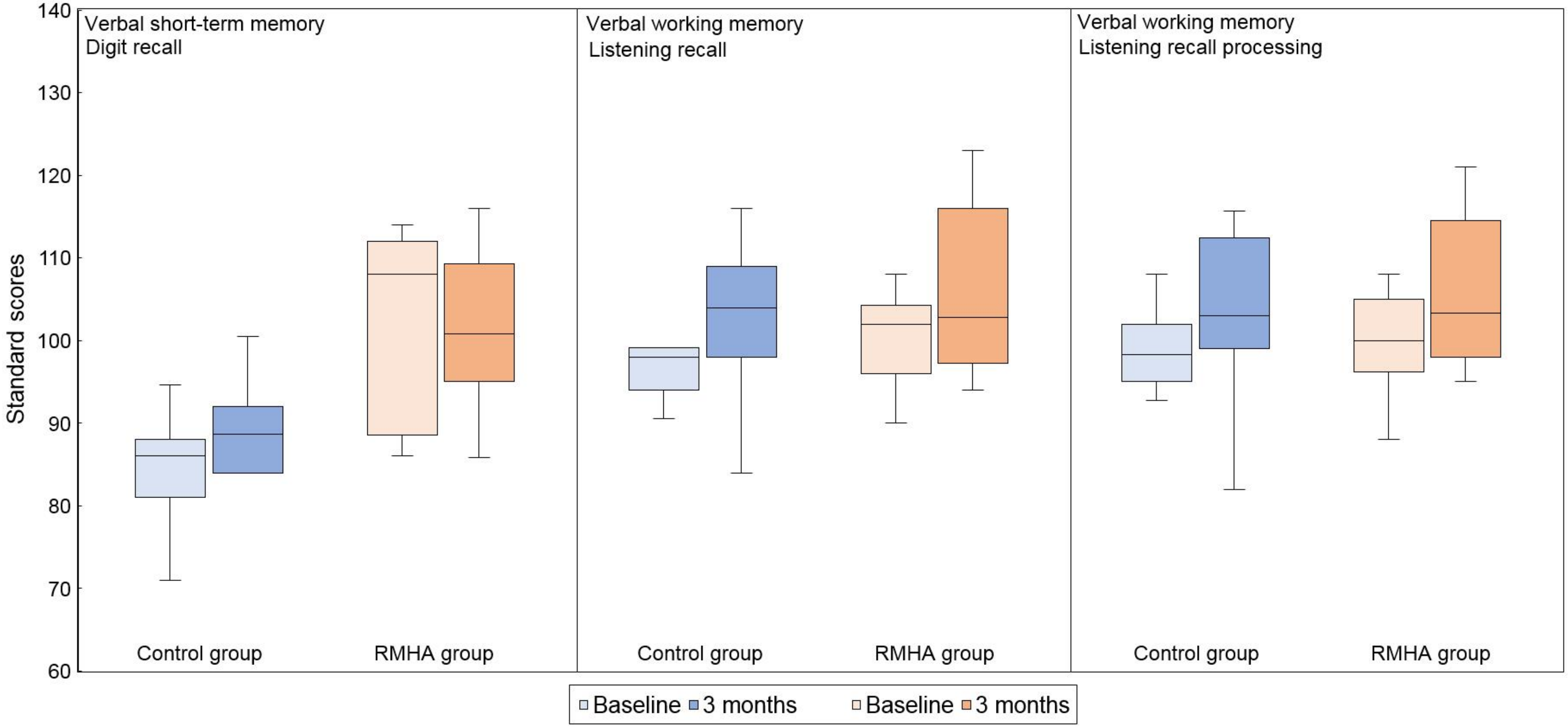

Figure 2.5 – Results of the AWMA

Boxplots of standard scores of the three AWMA conditions, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. Abnormal performance on this test was considered when the standard score was below 60. None of the mean standard scores of the two groups were below that. AWMA: Automated Working Memory Assessment, RMHA: Remote Microphone Hearing Aid.

Table 2.10 presents the mean values of the control and intervention group at baseline and at 3 months.

Table 2.10 – AWMA mean scores and SDs
The mean scores and SDs for the three conditions of the AWMA for the two groups at baseline and at 3 months. Disordered performance: ≤ 60, borderline performance: > 60 and ≤ 85, and normal performance: > 85. AWMA: Automated Working Memory Assessment, RMHA: Remote Microphone Hearing Aid, SD: Standard Deviation.

AWMA	Mean (SD)			
	Control group		RMHA group	
	Baseline	3 months	Baseline	3 months
Digit recall	85.62 (11.89)	85.98 (11.24)	102.08 (12.40)	101.55 (10.50)
Listening recall	97.24 (8.26)	103.13 (10.40)	99.86 (11.93)	105.91 (11.26)
Listening recall processing	98.90 (11.10)	104.69 (14.92)	100.12 (11.19)	105.83 (9.95)

Table 2.11 presents the p-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. There was no statistically significant interaction between group and time in any of the AWMA conditions. For the main effect of group, only the Digit recall condition showed statistically significant difference between groups, $F(1, 15) = 10.92$, $p = .005$, $\eta_p^2 = .421$. This means that if the two scores at baseline and 3 months were collapsed across time, a significant difference between the two groups would be observed. In other words, the intervention group had higher scores than the control group regardless of the time variable (i.e. baseline and post-intervention scores merged together for each group). For the main effect of time only the Listening recall condition returned statistically significant scores, $F(1, 15) = 6.24$, $p = .025$, $\eta_p^2 = .294$. This means that if the two groups were collapsed together, a significant difference over time would be observed. In other words, the scores at 3 months were higher than at baseline regardless of the group variable (i.e. control and intervention group merged together).

Table 2.11 – AWMA significance
The *p*-values, η_p^2 and CIs of the η_p^2 for the interaction between group and time and for the main effect of group and time in the three AWMA conditions. AWMA: Automated Working Memory Assessment, CI: Confidence Interval. *$p < .05$, **$p < .01$

Condition	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Digit recall	.876	.002	.000, .059	.005**	.421	.053, .644	.975	.000	.000, .002
Listening recall	.974	.000	.000, .002	.558	.023	.000, .265	.025*	.294	.001, .556
Listening recall processing	.990	.000	.000, .000	.810	.004	.000, .131	.098	.172	.000, .458

2.3.5 CHAPS questionnaire

No outliers were observed in any of the experimental conditions of the CHAPS (by checking for values in the studentised residuals beyond +/- 3 SDs). Normality of distribution was met in all conditions using the Shapiro-Wilk test of normality. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 2.6 presents the findings of the four CHAPS conditions.

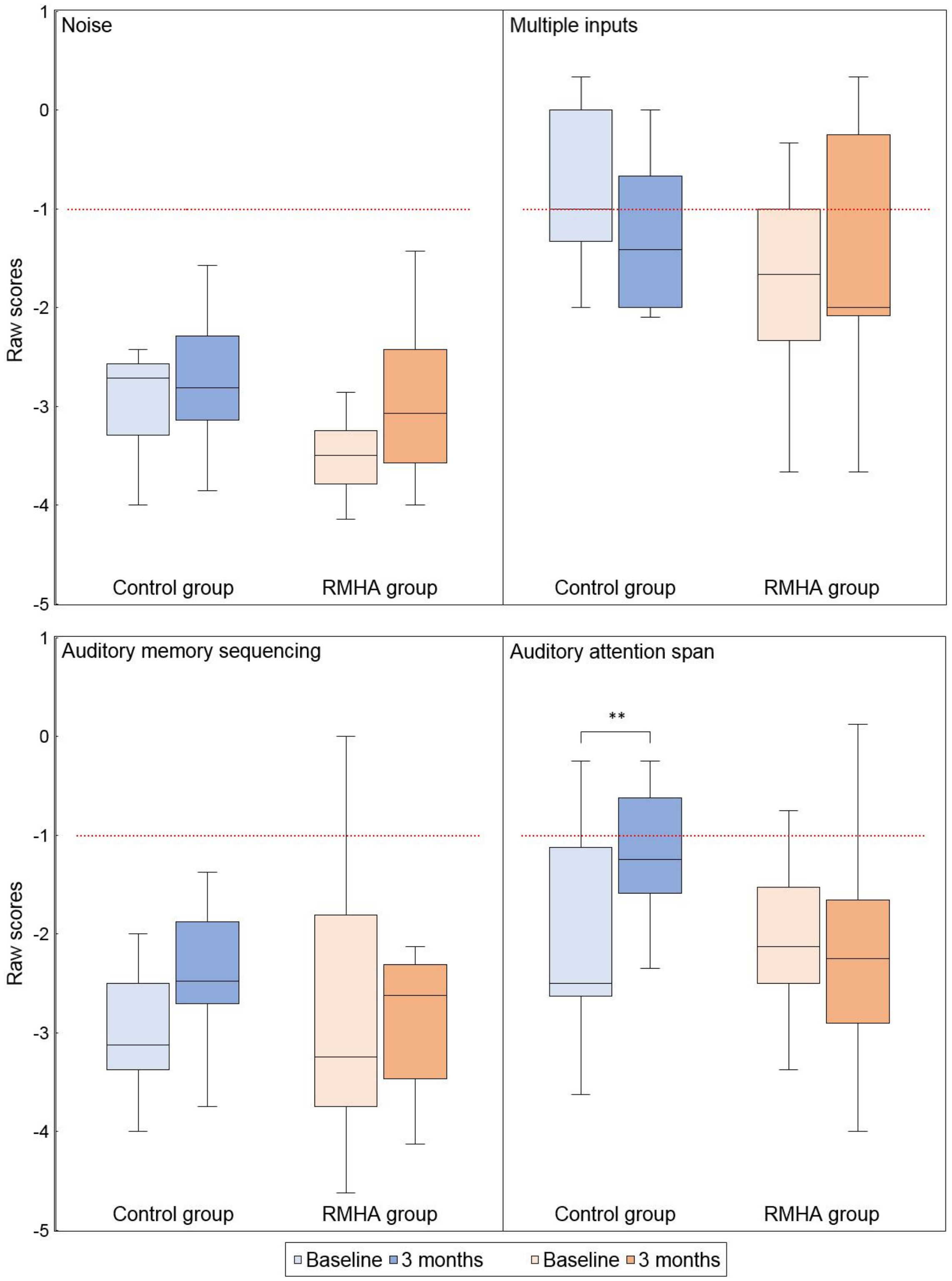

Figure 2.6 – Results of the CHAPS

Boxplots of raw scores of the four subscales of CHAPS, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. Even though this was a non-standardised questionnaire, the CHAPS manual did mention that it considered abnormal performance to be below -1 on the raw scores. Therefore, the red dotted line indicates this cut-off. CHAPS: Children's Auditory Performance Scale, RMHA: Remote Microphone Hearing Aid. $^{**}p < .01$

Table 2.12 presents the mean values of the control and intervention group at baseline and at 3 months.

Table 2.12 – CHAPS mean scores and SDs
The mean scores and SDs for the experimental and control conditions of the CHAPS for the two groups at baseline and at 3 months. Disordered performance: ≤ -1, and normal performance: > -1. CHAPS: Children Auditory Performance Scale, RMHA: Remote Microphone Hearing Aid, SD: Standard Deviation.

CHAPS	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Noise	-2.95 (0.92)	-2.76 (0.70)	-3.51 (0.43)	-2.94 (0.91)
Multiple inputs	-0.81 (0.80)	-1.26 (0.76)	-1.75 (1.05)	-1.50 (1.36)
Auditory memory sequencing	-3.11 (0.88)	-2.35 (0.76)	-2.73 (1.53)	-2.65 (1.31)
Auditory attention span	-2.11 (1.06)	-1.17 (0.68)	-2.03 (0.90)	-2.20 (1.26)

Table 2.13 below presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. There was statistically significant interaction between group and time only in the Auditory attention span condition, $F(1, 15) = 5.10$, $p = .039$, $\eta_p^2 = .254$. For the simple main effects of group, there was no statistically significant difference between the two groups, $F(1, 15) = 4.48$, $p = .051$, $\eta_p^2 = .230$, meaning the scores between the two groups did not differ at baseline (or at 3 months). There was statistical significance in the simple main effect of time in the control group, $F(1, 8) = 13.48$, $p = .017$, $\eta_p^2 = .627$, but not in the control group (see Table 2.13 for details). This means that the scores significantly improved from baseline to 3 months in the control group.

Table 2.13 – CHAPS significance
The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the simple main effects of group and time and main effects of group and time in the four CHAPS conditions. CHAPS: Children's Auditory Performance Scale, CI: Confidence Interval. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

Condition	**Interaction**		**Simple main effect**				**Main effect**			
	Group* Time	η_p^2 **[95% CI of η_p^2]**	**Group**	η_p^2 **[95% CI of η_p^2]**	**Time**	η_p^2 **[95% CI of η_p^2]/** ***MD*** **[95% CI of *MD*]**	**Group**	η_p^2 **[95% CI of η_p^2]**	**Time**	η_p^2 **[95% CI of η_p^2]**
Noise	.376	.053 [.000, .328]	-	-	-	-	.250	.087 [.000, .373]	.088	.182 [.000, .467]
Multiple inputs	.330	.063 [.000, .343]	-	-	-	-	.112	.159 [.000, .447]	.779	.005 [.000, .170]
Auditory memory sequencing	.335	.062 [.000, .341]	-	-	-	-	.933	.000 [.000, .017]	.239	.091 [.000, .378]
Auditory attention span	.039*	.254 [.000, .525]	.869 **(baseline)** .051† (**3 months)**	.002 [.000, .066] .230 [.000, .507]	.006** (**Control**) .006** (**0-3 months**) .705 (**RMHA**)	.627 [.087, .677] -.938 [-1.527, -.349] .022 [.000, .367]	-	-	-	-

2.3.6 CCC-2 questionnaire

The CCC-2 subscales were grouped to create two composite scores. The first composite score was created by averaging the subscales that reflect standard (or structural) language aspects (i.e. Speech, Syntax, Semantics and Coherence subscales). The second composite score resulted from the averaging of the subscales that represent non-standard (or non-structural) language aspects (i.e. Inappropriate Initiation, Stereotyped Language, Use of Context and Non-verbal Communication subscales). These scores were not used to identify children at risk of DLD or Pragmatic Language Impairment (PLI). These composite scores did not intend to identify such disorders and were only used to examine the structural and non-structural language aspects of the children in the sample. The CCC-2 has specific scores and procedures that help identify DLD or PLI in children and these are described in an upcoming study in Chapter 5 (see Section 5.2.3 for details).

No outliers were found in any of the two composite scores (by checking for values in the studentised residuals beyond +/- 3 SDs). Normality of distribution was checked using the Shapiro-Wilk test, with the standard language composite score being normally distributed and the non-standard language composite score not meeting this assumption. Transformations of this composite score did not produce normally distributed results and was therefore decided to run the analysis without transformations. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in both composite scores ($p > .05$ and $p > .001$ respectively).

Figure 2.7 summarises the results of the two composite scores of standard and non-standard language aspects.

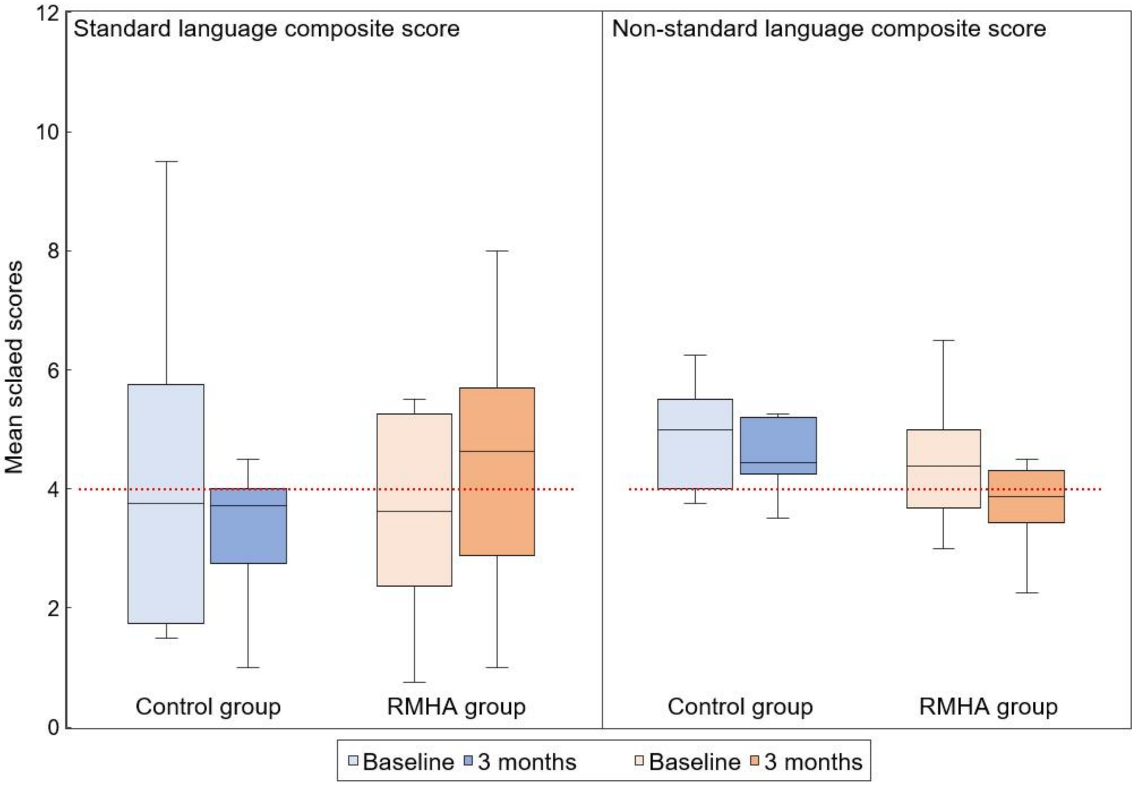

Figure 2.7 – Results of the two composite scores of the CCC-2
Boxplots of scaled scores of the two composite scores of standard language and non-standard language of the CCC-2 questionnaire, for 9 controls and 8 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). CCC-2: Children's Communication Checklist – 2, RMHA: Remote Microphone Hearing Aid.

Table 2.14 presents the mean values of the control and intervention group at baseline and at 3 months.

Table 2.14 – CCC-2 mean scores and SDs
The mean scores and SDs for the two composite scores of the CCC-2 for the two groups at baseline and at 3 months. The GCC and SIDC scores are also presented for additional descriptive information. Disordered performance: ≤ 4, borderline performance: > 4 and ≤ 7, and normal performance: > 7. CCC-2: Children's Communication Checklist – 2, GCC: General Communication Composite, RMHA: Remote Microphone Hearing Aid, SD: Standard Deviation, SIDC: Social Interaction Deviance Composite.

CCC-2	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Standard language composite score	4.11 (2.64)	3.31 (1.12)	3.56 (1.76)	4.50 (2.34)
Non-standard language composite score	4.97 (2.44)	4.93 (1.62)	4.50 (1.16)	4.31 (2.24)
GCC score	36.44 (17.99)	29.67 (10.94)	32.25 (8.95)	35.25 (15.30)
SIDC score	6.67 (6.11)	12.62 (15.59)	2.13 (7.93)	0.13 (11.06)

Results did not reveal statistically significant interactions between group and time in any of the two CCC-2 composite scores. For the main effect of group and main effect of time, again no statistically significant differences were observed. Table 2.15 below summarises this information.

Table 2.15 – CCC-2 composite scores significance
The *p*-values, η_p^2 and CI of η_p^2 of the interaction between group and time and of the main effect of group and main effect of time in the two composite scores of the CCC-2 questionnaire. CCC-2: Children's Communication Checklist - 2, CI: Confidence Interval.

Condition	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Standard language composite score	.083	.187	.000, .472	.725	.009	.000, .228	.878	.002	.000, .057
Non-standard language composite score	.885	.001	.000, .052	.506	.030	.000, .279	.823	.003	.000, .116

As results from the behavioural tests showed no statistically significant improvement in the RMHA group over the control group in the behavioural and questionnaire scores, correlation analyses were not carried out as initially planned.

2.4 Discussion

This study did not reveal statistically significant interactions between group (i.e. control, RMHA) and time (i.e. baseline, 3 months) in any of the outcome measures. However, it identified a number of limitations to the design that may explain this non-significant presentation of results and that can be addressed in the next study. First, the results from the statistical analysis will be discussed followed by a discussion of the implications of the study's limitations.

2.4.1 Speech-in-noise

The two SCAN-3 C conditions did not show improvement over 3 months of RMHA use. This is the first study to look at the effects of RMHAs on SIN performance when tested unaided. Previous research, even though claiming enhancements of the auditory system, it did not actually test children's performance without hearing aids (Johnston et al., 2009). The only study that tested children in unaided conditions was the one by Sharma et al. (2012) which used the HINT as the SIN outcome measure and it did not show significant improvements on that test. Nonetheless, their intervention groups included RMHA plus AT making it difficult to determine the contribution that each intervention had on the HINT scores.

Both groups had higher scores in the AFG +12 dB condition compared to the AFG 0 dB. This is not surprising since the AFG +12 dB presents the target voice at 12 dB higher than the background babble, while the AFG 0 dB has both target and noise at the same level. A ceiling effect could be the reason for the lack of improvement in the AFG +12 dB condition, as both groups had high baseline scores. Nonetheless, the baseline scores in the AFG 0 dB condition were not as high but still remained unchanged over the intervention period. The internal consistency reliability coefficients of the two SCAN-3 C conditions ranged between .50 to .59 (depending on the age group) in the AFG +12 dB and between .63 to .67 in the AFG 0 dB (R. W. Keith, 2009). This moderate test-retest reliability of the SCAN-3 C conditions could possibly explain the non-significant results. Other SIN tests with better test reliability should be used in upcoming studies.

2.4.2 Attention

Experimental conditions

No previous research has looked at the effects of RMHAs on aspects of attention in children with APD. Results did not reveal significant improvements in the experimental conditions (i.e. Sus-AA and Div-AVA) following a 3-month intervention. Sustained auditory attention was measured using a composite score from three different tasks, the Score!, Walk don’t walk and Code transmission. The latter two tasks were found to correlate with the WISC III, a standardised IQ test (Manly et al., 1999). Significant Pearson’s correlations were reported between Walk don’t walk/ Code transmission tasks and the WISC III four-subtest scaled score. While these correlations are weak, they indicate that other cognitive factors may have an influence on these two tasks. Code transmission is similar to an n-back task, which is used in memory tests, and could thus be affected by memory factors, whereas the Walk don’t walk task measures response inhibition in addition to Sus-AA (Manly et al., 1999). On the other hand, the Score! Task (i.e. the other test measuring Sus-AA) did not have a significant correlation with the WISC III score. It could thus be possible that the composite score had influences from other cognitive factors and perhaps it was not measuring pure Sus-AA abilities. From these non-significant findings it can be argued that neither Sus-AA nor Div-AVA (when tested unaided) benefited after children with APD used RMHAs for 3 months.

Control conditions

The control conditions of Sel-VA and Div-AA were hypothesised to remain unchanged post-intervention, since RMHAs were not expected to assist these two types of attention. These hypotheses were confirmed, since no significant changes were recorded in these two tests after 3 months of RMHA use.

2.4.3 Working memory

The three AWMA conditions did not improve after children used RMHAs for 3 months. Test reliability is good for all three AWMA tests with correlation coefficients ranging between .84 and .89 (Alloway, 2007). It is thus unlikely that the lack of improvement was due to compromised test reliability. These non-significant findings are in accordance with findings from two previous studies that used memory measures on children with APD following RMHA interventions (Sharma et al., 2012; Umat et al., 2011). Umat et al.'s (2011) intervention period was 12 weeks and Sharma et al.'s (2012) was 5 weeks. Therefore, a 3-month RMHA intervention does not appear to bring lasting benefits in AWM in children with APD. Looking at Figure 2.5, it is observed that in the Listening recall and Listening recall processing tasks, both groups exhibited a similar pattern of improved scores from baseline to 3 months. Practice effect may explain this pattern of scores in these two tests despite the good test-retest reliability of the AWMA conditions.

Future studies need to further investigate the effects of RMHAs on AWM using differential memory tasks, such as digit span forward or backward. Sharma et al. (2012) administered those two tasks but they used AT alongside a RMHA intervention, making it difficult to determine the effect of RMHAs on working memory. Thus, even though there are converging evidence that AWM may not benefit from RMHA use, there are a few points that need to be further examined in future studies, such as possible use of longer trial periods and differential memory tests.

2.4.4 Questionnaires

In the CHAPS questionnaire, parents of controls rated a significantly better performance for their children in the Auditory attention span subscale at 3 months in comparison to baseline. Despite the fact there was no change in performance in the Auditory attention span condition for the RMHA group, anecdotal reports from parents (5 out of 8 RMHA parents) reported improvement in attentive behaviours in their child at home. There are various studies that report validity problems with the CHAPS (Lam & Sanchez, 2007; Sharma et al., 2009; Wilson et al., 2011), thus it might not be the best tool to measure specific aspects of children's development (i.e. SIN performance, memory, attention). Adding to that, the questionnaire is using raw instead of standardised scores and there is no mention of its test-retest reliability (Smoski et al., 1998). The significant improvement in the Auditory attention span in the control group may thus be attributed to maturation, while the non-significant results in the other two experimental conditions (i.e. Noise and Multiple inputs) could be a result of poor test reliability. Regarding the control condition (i.e. Auditory memory sequencing), it was found to remain unchanged as expected. However, given the issues raised about the CHAPS's test reliability and lack of standardised scores, memory subscales in questionnaires should be further explored in future research.

Furthermore, the composite score of the four standard language conditions of the CCC-2 (i.e. Speech, Syntax, Semantics and Coherence) did not show significant improvement. The questionnaire was completed by parents, thus they judged their children's performance in these subscales at home when RMHAs were not used. It can thus be supported that using RMHAs for 3 months does not bring significant long-term improvements in standard language aspects in children with APD. The composite score of the four non-standard language conditions (i.e. Inappropriate initiation, Stereotyped language, Use of context and Nonverbal communication) also remained unchanged over the study period in both groups, which confirmed the study's hypotheses. This latter finding is in line with Sharma et al.'s (2012) study results. They used a test on children with APD measuring various non-standard linguistic aspects (e.g. Figurative language, Use of context and Drawing inferences from given information) which did not show significant change over a 6-week intervention period. Their study groups used both AT and RMHAs, thus making it difficult to attribute the lack of change in test scores to the use of RMHAs alone. However, it is not expected that the clearer signal induced by the system would have an impact on non-standard language aspects, such as sarcasm, use of context or nonverbal communication. It is noted though, that in the present study the composite score of these four non-standard language subscales did not meet the assumption of normal distribution. Violation of this assumption may increase the chance of making a Type I error (i.e. finding a significant effect when no actual effect exists), but since no significant effects were observed in this composite score it is expected that there was no inflation of Type I error. Finally, this study did not use children's questionnaires and this limitation is noted. The perspective of the user is important, and this is something that will be addressed in the upcoming study.

2.4.5 Limitations

Sample size and intervention period

The limited sample size could be contributing to the lack of statistically significant findings. The power calculation indicated that for significant results to be produced, 15 more children should have been recruited. The 60 invitations sent were the maximum that could have been sent, as all other cases identified from the GOSH patient database did not meet the inclusion criteria of the study (see Section 2.2.1). A post-hoc power calculation was performed on the actual sample size ($n = 17$) to detect 1 SD effect size at 5% significance, which resulted in a statistical power of 49%. This means that the study was well under-powered and the chance of detecting a statistically significant difference (if a difference exists) was down to almost 50%. Another consequence of low power is the imprecision of the effect size (Nakagawa & Cuthill, 2007) which can be inflated, thus showing a greater magnitude than the actual effect size of the population. Additionally, a low powered study also increases the probability of getting a false negative effect. This means that true significant effects in the population have an increased chance of showing as non-significant in low powered studies. Therefore, interpretation of the results cannot be safely made, and a larger sample sized study is required to make sense of the findings.

The lack of significance in the attention tests could be interpreted as RMHAs not having an impact on top-down cognitive functions, such as attention. However, the study period of 3 months might have been too short for RMHAs to yield statistically significant results. Another RCT by Purdy et al. (2009) included children with reading delay who used RMHAs for 6 weeks. It demonstrated improvements in reading questionnaire scores completed by teachers and increased scores in children's self-reports on their classroom listening environment. Nevertheless, there was no significant change in behavioural reading test scores and the authors concluded that a longer intervention period might be required to influence standardised reading tests (Purdy et al., 2009). Similarly, the period of 3 months in this study may have been too short for significant changes in the SCAN-3 C and TEACh experimental conditions to take place. Improvements in higher-order functions, such as attention, might require longer intervention periods to reveal statistically significant results. It should be noted however, that significant changes in questionnaire scores were not recorded in the present study in the corresponding SIN and attention subscales. This could mean that a longer RMHA intervention period for children with APD is required for changes in SIN, Sus-AA and Div-AVA to take place altogether.

Mini-mic

Post-intervention verification of the Mini-mic showed that the devices received were working as intended (see Section 2.2.3, Verification of RMHAs). Even though intended performance was confirmed for the three devices received post-intervention, there were still five devices that were not received, and thus not checked. It is assumed that they would perform the same as the tested devices, but this cannot be confirmed. In addition, the Mini-mic has a range of 7m which was considered adequate for classroom use. However, when the distance between the Mini-mic and ear instruments was close to 7m this would weaken the signal (and its quality) and in some occasions, would temporarily disconnect for a second. This variation in transmittance at longer distances may have compromised the findings. Furthermore, it was not considered to measure the total time the RMHA was used by children at the end of the trial, therefore it is not known for how long each student used the system in the classroom. These limitations of the limited transmission range and lack of monitoring of the total time the RMHAs were used, will be taken into account and addressed in the next study.

Auditory training

Parents in both groups were asked during the first test session not to use other interventions with their children during the study period. While it was expected that participants would adhere to this request, at the post-intervention test session it was made known that 2 of the 9 control children used AT for some time during the 3 months of the trial. There are a number of studies that showed significant improvements (i.e. in SIN, memory and AP task performance) in children with APD after AT interventions (Filippini, Befi-Lopes, & Schochat, 2012; Loo, Rosen, & Bamiou, 2016; Sharma et al., 2012; Tawfik, Mohamed Hassan, & Mesallamy, 2015). The fact that some control children in the present study used AT may have influenced the results. When participants from either group (control or intervention) do not comply to the study protocol, it increases the chance of underestimation of the efficacy of the intervention under investigation (Gupta, 2011). This could explain the significant improvement noted by parents in the CHAPS Auditory attention span condition. As an ITT analysis was used (Gupta, 2011), these children were not excluded from the analysis.

Fatigue

The total testing time for each test session was between 2 to 2,5 hours. Test administration was long, and it was observed that children were easily tired and had reduced motivation after the first half of testing. Even though test randomisation was applied to minimise order effects, this long test administration could have had an impact on the overall test scores. Therefore, in the next study the aim will be to reduce testing time.

Researcher bias

Furthermore, the results at baseline had to be reviewed by the researcher before the retest session at 3 months in order to make comparisons with results obtained from a previous study. This meant that the researcher did not remain blind to the results during administration of the tests at 3 months, which could have introduced researcher bias. Participants and their parents remained blind to the results until all tests were completed. In the next study, the experimenter will remain blind to the results until all test sessions are completed to reduce any observer bias that may have been introduced in this study.

Multiple comparisons

As noted earlier, there was no correction for multiple comparisons and the reasons were explained in Section 2.2.4. There, it was also discussed that a way to compensate for multiplicity is to create composite scores (Schulz & Grimes, 2005). In that same section the outcome measures that were averaged together to create composite scores were detailed (See Section 2.2.4). Even if some outcome measures were averaged together to minimise the analyses, there were still a number of analyses conducted on several outcome measures. Thus, the chance of making a Type I error (detecting a significant effect where no such effect is truly present) was increased. The only significant effect detected in the study was the improved scores in the CHAPS Auditory attention span subscale observed only in the control group from baseline to 3 months. Thus, that significant finding may have been a false positive effect, meaning that there is a chance that no such effect truly existed.

Diagnostic criteria

In this study specific APD criteria were set in order to identify children with APD. These criteria are used in the APD clinic at GOSH (GOSH, 2018) and they closely follow the identification process supported by the ASHA (2005b; see details in Section 2.2.1 under Inclusion and exclusion criteria). Given the controversies in diagnosing APD discussed earlier in the Introduction (Section 1.4.1), following these specific criteria has its limitations. Firstly, referral to the APD clinic at GOSH might depend on the APD awareness of GPs or medical and educational professionals in certain areas in England, therefore children referred for APD might be coming from specific areas of England. In addition, if referring professionals are not well-informed on APD then referrals might be missed. Conversely, if professionals are unaware of other possible causes of the presenting symptoms (discussed earlier in Section 1.3.3) then perhaps other diagnoses might be missed, which may also be the cause of or explain the APD symptoms. The APD clinic at GOSH, though, does exclude cases with diagnoses of ADHD, ASD and DLD (GOSH, 2018). The inclusion of other criteria in the GOSH protocol also attempts to limit the influence of other factors such as age, performance on intelligence tests, hearing thresholds, and language (see details in Section 2.2.1 under Inclusion and exclusion criteria). However, there is heterogeneity in terms of the APD diagnostic criteria used by different clinics worldwide. As discussed earlier in the Introduction (see Section 1.4), differential diagnostic criteria directly impact definition, prevalence, diagnosis and management. The fact that in this present work specific APD criteria have been used, which differ from other position statements (e.g. AAA, European perspective), means that any findings can only be interpreted in the context of those specified diagnostic criteria. This is a problem APD research faces in general (see discussion in Section 1.4.2) and can only be addressed with a universally accepted APD position statement. Therefore, any results from the study should be interpreted with caution having in mind that the findings may apply only to this certain group of children (i.e. children referred for concerns regarding their listening at this specific APD clinic, that uses specific APD criteria, that in turn are different from criteria used by other APD clinics worldwide).

2.5 Conclusion

For the period of 3 months, RMHAs used at school did not have a lasting positive effect on SIN, Sus-AA or Div-AVA. Control conditions (i.e. AWM, Div-AA and Sel-VA) remained unchanged as hypothesised. However, a number of limitations were identified in this study; compromised sample size resulting in low statistical power, short intervention period, test reliability/ practice effect problems with some tests, no monitoring of the total time RMHAs were used, use of AT by some participants, possible fatigue in children due to long testing time and researcher bias. These limitations make interpretation of the findings difficult and will thus be addressed in the next study.

CHAPTER 3 – STUDY 2: EFFECTS OF RMHAs ON ATTENTION AND LISTENING-IN-NOISE ABILITY IN CHILDREN WITH APD

3.1 Introduction

3.1.1 RMHAs and attention

As discussed in the introduction of Chapter 2 (Section 2.1.1), no studies have investigated behavioural attention measures in children with APD after a RMHA intervention. The study in Chapter 2 did not reveal significant improvements in the Sus-AA and Div-AVA behavioural tests following a 3-month RMHA intervention. However, the study faced key limitations, such as small sample size which impacted statistical power, compromised transmission range of the RMHA and non-compliance to the study protocol by some control children. Moreover, the study period of 3 months may have been a short period for lasting effects on cognitive functions, such as attention, to take place. Therefore, further examination of Sus-AA and Div-AVA needs to be carried out in a longer intervention trial.

3.1.2 RMHAs, listening-in-noise and spatial listening

As discussed in the previous study (in Chapter 2), speech perception in noise, an ability found abnormal in children with APD compared to typically developing peers (Cameron et al., 2006; Cameron & Dillon, 2008; Lagacé et al., 2011), has not been adequately studied in RMHA trials. Even though there are studies that claim long-term benefits in the auditory system after RMHA or directional microphone hearing aid use (Johnston et al., 2009; Kuk et al., 2008), these are not backed up by evidence, since testing was performed only under aided conditions. In the study in Chapter 2, the speech in babble noise measure (i.e the SCAN-3 C test) that was used did not show significant improvements over time, which may be explained by the moderate test reliability of the test (R. W. Keith, 2009). New measures of SIN ability with better test reliability are required to further test the hypothesis of SIN improvements after RMHA use.

3.1.3 Research questions and hypotheses

As the study in Chapter 2 had a short intervention period, measures of attention will further be investigated here in a longer trial of 6 months. As such, the effects of RMHA use on the same aspects of attention employed previously (i.e. Sus-AA and Div-AVA) will be examined. In this new study total testing time was reduced as fatigue in children from long test sessions may have previously influenced the overall results. A couple of attention subtests used before (i.e. Walk don't walk and Code transmission) were found to significantly correlate with the WISC cognitive ability test score (Manly et al., 1999) and was thus decided to be removed from this study. At the same time, as the AWM test from Chapter 2 did not show significant changes as hypothesised and with indications of possible practice effects, this test was also removed from the test procedure to further reduce testing time.

In search of new SIN measures, the LiSN-S test was deemed a good test to replace the SCAN-3 C. While the LiSN-S presents only slightly better test-retest reliability (*r* = .80 in one of its conditions) than the SCAN-3 C (Cameron & Dillon, 2007b), it does offer other advantages. Firstly, the three derived measures of the LiSN-S minimise the influence of linguistic factors on the scores (Cameron & Dillon, 2007a) and it is argued that this control of factors might extend to control of cognitive factors, too (Tomlin et al., 2015). Secondly, it could be supported that the LiSN-S is a tool that better reflects real-life performance, as the child is listening for entire sentences instead of single words that the SCAN-3 C requires children to do. Lastly, it allows for the testing of different SIN conditions, such as when target and distractor emanate from the same source (measured by the Low-cue SRT and Talker advantage conditions) and when there is spatial separation between target and distractor (measured by the Spatial advantage and High-cue SRT/ Total advantage conditions). The effect of RMHAs on listening-in-spatialised-noise has not yet been examined in children with APD. It can be argued that as the use of RMHAs brings the speech of the microphone user straight into the child's ears (and hence eliminating the spatial element of listening to the talker), the LiSN-S spatial scores will not improve and might even show worse performance. This argument is based on two studies discussed in detail in Chapter 1 (Section 1.5.2), where it was shown that deficit-specific training significantly improved spatial listening conditions on the LiSN-S but standard SIN training did not (Cameron et al., 2012; Lotfi, Moossavi, et al., 2016).

Finally, the same questionnaires employed previously will be used again (i.e. CHAPS, CCC-2, SIFTER), with the addition of the LIFE-R questionnaire which will be completed by children. The LIFE-R questionnaire assesses the ability of children to listen to their teacher in different listening situations at school. The scores of the questionnaire are expected to improve in the intervention group, as children will be using the RMHA at school. The detailed hypotheses are outlined below. All tests mentioned are described in depth in the Methodology section.

1. The group of children with APD who use RMHAs will show improved scores after 3 and 6 months compared to the APD control group in:
    a. The Sus-AA and Div-AVA TEACh tasks (unaided).
    b. The LiSN-S Low-cue SRT and Talker advantage scores (unaided).
    c. The Noise, Multiple inputs and Auditory attention span subscales of the CHAPS (completed by parents).
    d. The composite score of the four standard language conditions of the CCC-2 questionnaire (completed by parents; Speech, Syntax, Semantics, Coherence).
    e. The Attention and Communication subscales of the SIFTER questionnaire (completed by teachers).
    f. The LIFE-R questions 1 to 9 and the total questionnaire score (completed by children).
2. The group of children with APD who use RMHAs will not show improved scores when compared to controls at 3 and 6 months in:
    a. The Sel-VA and Div-AA tasks of the TEACh test (unaided).
    b. The Spatial advantage and High-cue SRT/ Total advantage conditions of the LiSN-S test (unaided).
    c. The Auditory memory sequencing subscale of the CHAPS.

d. The composite score of the four non-standard language subscales of the CCC-2 (Inappropriate initiation, Stereotyped language, Use of context, Nonverbal communication).

If hypotheses 1a-f are confirmed, then the following additional hypotheses will also be tested:

3. Improved scores in the Sus-AA test will correlate with improved scores in the Auditory attention span subscale of the CHAPS and the Attention subscale of the SIFTER in the RMHA group.
4. Improved scores in the LiSN-S Low-cue SRT and Talker advantage conditions will correlate with improved scores in the Noise subscale of the CHAPS in the RMHA group.

3.2 Methodology

The methodology used in the current chapter closely follows the one from Chapter 2. For brevity, only changes from the previous study or additions will be described here.

3.2.1 Participants

Sample size

A power sample calculation of *n* = 32 (total sample size) was calculated based on an estimated 1 SD effect size, *n* = 16/*f*2, 80% power at 5% significance – where *n* = sample size per group, *f* = effect size. Despite this calculation, the total sample size for this study was 26 children, aged 7-12. This was the maximum number of children that could have been included in the study from the initial 100 invitations sent. Invitations were sent to all cases that met age criteria and were identified with APD from GOSH since the previous study. This means that it was not possible to invite more children at this stage. There were 43 responses to the initial call but only 29 children met the criteria for APD diagnosis (see Section 2.2.1 for details), while the 14 excluded cases had either used RMHAs/ AT or did not want to commit to the trial when they learned about the details of it. Figure 3.1 outlines the attrition process in a flowchart. Diagnostic APD criteria used in the clinic are the same as in the previous study (see Chapter 2, Section 2.2.1). An ITT analysis was chosen to prevent overestimation of the intervention's efficacy, to reduce bias caused by deletion of data and to maintain a large-enough sample size (Gupta, 2011). Three children had a non-verbal IQ score lower than 85 which led to their exclusion, bringing the sample size down to 26 children.

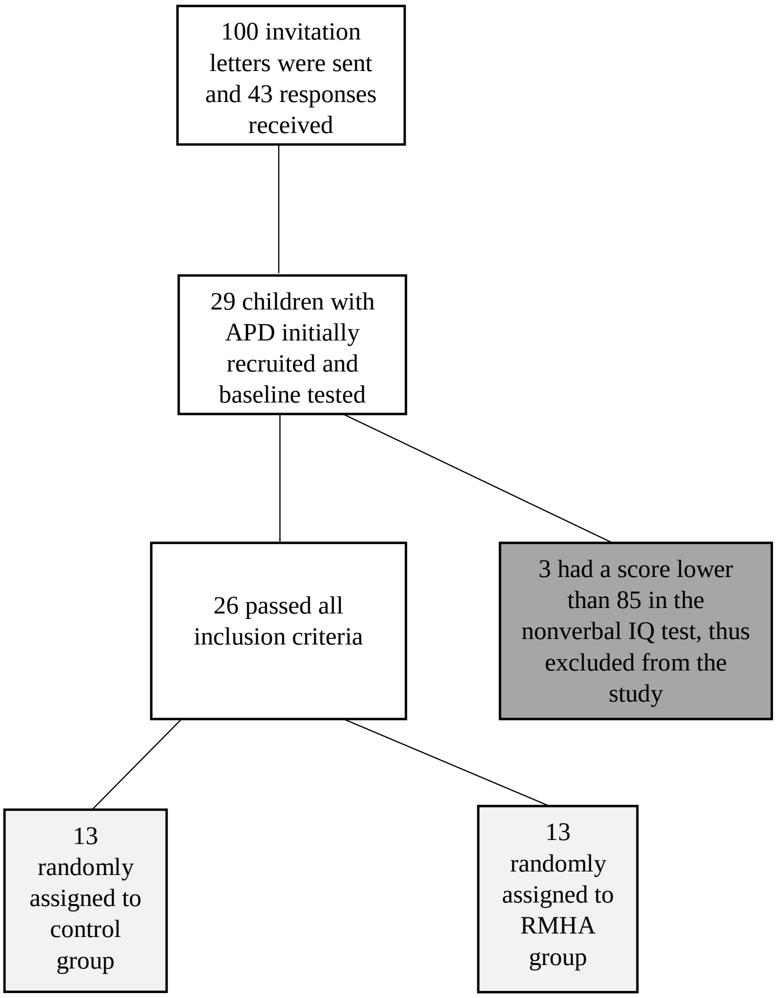

Figure 3.1 – Attrition process
Flow chart showing the attrition process for this study. Dark grey box refers to the participants excluded from the study, while light grey boxes show the final two comparison groups of the study. APD: Auditory Processing Disorder, IQ: Intelligence Quotient, RMHA: Remote Microphone Hearing Aid.

There were 8 girls and 18 boys between 7 years 3 months and 11 years 7 months, with a mean age of 9 years 8 months. All participants were native English speakers (2 were bilingual), with the majority (*n* = 23) coming from Greater London, while the rest (*n* = 3) came from other areas in England. Information on gender and age along with mean non-verbal IQ scores and mean PTA scores (for right and left ears) are presented below in Table 3.1 in individual scores.

Table 3.1 – Individual data
Individual data for the participants in the control and intervention groups and mean and SDs. The presented information include gender, age, performance in the WNV and mean PTA scores for both ears. PTA: Pure Tone Audiometry, SD: Standard Deviation, WNV: Wechsler Non-Verbal.

Control children	**Gender**	**Age (years: months)**	**WNV (standard score)**	**PTA-R (dB)**	**PTA-L (dB)**
c-01	female	8:6	109	2.5	1.7
c-02	male	11:4	102	5.8	2.5
c-03	male	9:8	97	10	8.3
c-04	female	8:8	102	10.8	8.3
c-05	male	10:2	108	5.8	12.5
c-06	male	11:7	102	9.2	9.2
c-07	male	11:5	98	0	3.3
c-08	male	7:5	99	4.2	4.2
c-09	female	8:9	99	5.8	5.8
c-10	male	8:6	123	7.5	9.2
c-11	female	11:1	136	8.3	0.8
c-12	male	9:10	109	10	7.5
c-13	male	8:9	107	3.3	5
Mean (SD)	*4 females 9 males*	*9:8 (16.3 months)*	*107 (11.1)*	*6.4 (3.3)*	*6 (3.5)*

Intervention children	**Gender**	**Age (years: months)**	**WNV (standard score)**	**PTA-R (dB)**	**PTA-L (dB)**
i-01	female	9:4	89	8.3	8.3
i-02	male	10:11	92	5.8	5.8
i-03	female	8:6	97	10.8	9.6
i-04	male	8:1	95	7.5	9.2
i-05	male	10:2	86	6.7	5.8
i-06	male	8:0	109	4.2	4.2
i-07	female	11:3	101	8.3	4.2
i-08	male	7:3	105	5	11.7
i-09	male	11:0	102	7.5	10
i-10	male	9:8	96	10.8	9.2
i-11	female	9:6	112	1.7	-0.8
i-12	male	11:5	114	7.5	7.5
i-13	male	9:10	93	11.7	10
Mean (SD)	*4 females 9 males*	*9:7 (16.1 months)*	*99 (8.7)*	*7.3 (2.8)*	*7.2 (3.4)*

Inclusion and exclusion criteria

Inclusion and exclusion criteria are the same as in Chapter 2 (see Section 2.2.1).

Design

An RCT design was used and children were randomised in two groups:

- Group A or Control group included children with APD who did not receive a RMHA for the study period.
- Group B or RMHA/ intervention group comprised children with APD who received RMHAs after baseline testing.

Stratified randomisation was used (Kang et al., 2008) in order to balance the characteristics of age and gender between the two groups, as there are indications that these factors might have an influence on auditory performance and attention (Coch et al., 2005; Yathiraj & Vanaja, 2015). Both groups were randomly constructed based on the two strata of age (9 years 4 months or younger, and 9 years 5 months or older) and gender (boys, girls). Permuted blocks within each stratum were used to ensure balance between the two groups. For this, the allocation ratio was 1:1, block sizes were sizes of 3, 6, 7 and 13 and a computer random number generator was used (with Excel formula). With this application, each group comprised 4 girls and 9 boys with the control group having a mean age of 9 years 8 months, while the intervention group had a mean age of 9 years 7 months. Children in both groups were tested at baseline, at 3 months and at 6 months.

Participants were only recruited subject to agreement by their parents prior to the start of the study that their children would not use any other auditory type intervention during the study period regardless of the group they would fall into. Parents and children were also informed that with completion of the study they would receive a 12-week computer-based AT programme (compiled by the researcher), as an incentive for participants in both groups to remain in the study for the whole 6 months.

Ethical issues

An amendment was submitted to the Bloomsbury REC outlining changes made in the research protocol for this study. The study has been reviewed and given a favourable opinion by the REC,

REC reference: 14/LO/1509. The rest of the points under this section were the same as in Chapter 2, Section 2.2.1. A £10 book token was given to each child attending every test session and parents received a reimbursement of up to £10 for travel expenses.

3.2.2 RMHAs

The device

The RMHA used in this study was the Micro-mic (see Figure 3.2) coupled with the ReSound Up hearing aids (see Figure 1.4 in Chapter 1).

Figure 3.2 – Micro-mic RMHA
The Micro-mic RMHA by GN ReSound. RMHA: Remote Microphone Hearing Aid.

The RMHAs used in the study were manufactured by GN ReSound and they were the Micro-mic line. The Micro-mic is not considered a standard personal FM system as it broadcasts at a lower frequency (2.4 GHz). Despite this, it is still capable of producing a wireless signal to the ear receivers and it offers the same advantages as FM systems, such as reduction of reverberation and consistent signal throughout its range. The Micro-mic has the advantage of being smaller and more portable compared to a standard FM system. It can stay connected to the ear instruments within 25m, which is a substantial increase in range compared to the Mini-mic that was used in the previous study in Chapter 2 and reached up to 7m. This ensures a more stable signal even if the teacher is moving around in the classroom.

As the verification procedure was not properly followed prior to the start of the study, the researcher had to ask participants to send their instruments again for check after the study had been completed. There were 8 instruments received from the 13 given out to the intervention group. Verification of the Mini-mic and hearing aids was carried out using the same procedures described in detail in Section 2.2.3 (under Verification of RMHAs). The reference measurement was 64 dB SPL and the transparency measurement was 66 dB SPL, which was within the tolerance range the AAA (2011) sets (i.e. +/- 2 dB compared to the 64 dB SPL reference measurement) and indicates that the system was functioning as intended. The same measurements were obtained for all three devices tested. The same measurements were obtained for all 8 devices tested. Even though the rest of the 5 devices in the intervention group were not returned for verification, it is expected that they would function similarly as they were programmed and sent with the same order prior to commencement of the study.

The total and mean per day time (in hours) the RMHAs were used throughout the study by the intervention group are presented below in Table 3.2. Participants were asked to bring the hearing instruments with them in the final test day. The hearing aid was then wirelessly connected to the ReSound Airlink device. The ReSound Aventa software was used to record the information in Table 3.2 through the data log of the hearing aids.

Table 3.2 – RMHAs usage time
The total time in hours the RMHAs were used during intervention per case and the average hours used per day. Total means of these two measurements are given at the bottom row. RMHA: Remote Microphone Hearing Aid.

Child	Hours used (total)	Hours used (per day)
i-01	580	4.8
i-02	186	1.6
i-03	231	1.9
i-04	742	6.2
i-05	63	0.5
i-06	630	5.3
i-07	60	0.5
i-08	344	2.9
i-09	277	2.3
i-10	79	0.7
i-11	61	0.5
i-12	264	2.2
i-13	75	0.6
Total mean	276	2.3

The ‘per day’ hours used were calculated by considering that the 6-month intervention period had approximately 120 school days, thus dividing the total hours used by 120. Lack of reporting of the actual time that RMHAs were used by previous APD studies makes it difficult to calculate an average time that would be considered sufficient to bring about benefits. Some APD studies have mentioned that children were asked to use the RMHA 4-5 hours daily but the actual time the system was used was not recorded (Umat et al., 2011), while other APD studies only stated that they asked participants to use it daily at school without giving specific hours of daily use (Johnston et al., 2009; Kuk et al., 2008; Sharma et al., 2012). Evidence from other populations (i.e. children with learning difficulties) showed that a 2-hour daily use of RMHAs for 6 months improved their attentive behaviours as judged by speech-language pathologists through observations (Blake et al., 1991; see Section 1.8.2, under RMHAs and other disorders, for more details on the study). It was thus considered that at least 2 hours' use per day would be sufficient for potential improvements to take place.

Children's involvement

Children in group B were fitted with the system binaurally. As in the previous study in Chapter 2, there was no blinding neither for the participants nor for the researcher for reasons explained in Section 2.2.2. Also, another tester for the three time-points would have helped reduce bias, however it was not possible to recruit that person due to budgetary constraints. Children were asked to use the RMHAs daily during school time 5 days per week, only in lecture-based subjects. They were asked to take them off during physical education and breaks to protect them from getting damaged or lost. The system had to stay at school, ensuring that it would not be forgotten at home. Electronic communication with parents was initiated by the researcher monthly, to address any issues related to RMHA use. Communication with parents of the control group was also initiated monthly in order to mirror what was being done for the intervention group, to increase engagement and to minimise the risk of participant withdrawal due to long periods without communication. At the same time, during this exchange the researcher was making sure that children in neither group were using other interventions (i.e. AT), something which was not done in the study in Chapter 2. More details on the ear instruments, size measurements, colour choices, programming, type of device (i.e. non-occluding) and the procedure followed for the school visits to give the RMHA are outlined in Chapter 2, Section 2.2.2. After completion of the trial, all participants got to keep the RMHA, as this was part of the recommendations by the APD clinic at GOSH. The control group received RMHAs after study completion at 6 months, as well. For that reason, school visits for children in the control group were also organised at 6 months.

Teachers’ involvement

This section follows the same procedure as in Chapter 2 (see Section 2.2.2).

3.2.3 Tests and questionnaires

Pure tone audiometry

The same guidelines were followed as in Chapter 2, Section 2.2.3. The duration of the test was approximately 15 minutes.

Cognitive ability

The same WNV cognitive test was administered as in the previous study (see Chapter 2, Section 2.2.3 for details). Only children scoring higher than 85 in the composite full scale score were included in the study. Three out of twenty-nine children failed to pass this threshold, resulting in their exclusion from the study (see attrition chart in Figure 3.1). The duration of the test was approximately 20 minutes.

Speech-in-noise perception

The SCAN-3 C test was replaced in this study by the LiSN-S test. As explained previously, it has some advantages over the SCAN-3 C test. Its derived measures can control for language and possibly cognitive factors (Cameron & Dillon, 2007a; Tomlin et al., 2015), while the test offers the opportunity to test both SIN and spatial listening skills. There were four test conditions where each presented up to 30 different sentences. The relative position of distractors in relation to target speech would change (0^o or 90^o) in combination with the pitch (distractor voices being same or different than the target) in each test condition. Test condition 1 had different voices at +/- 90^o, test condition 2 had same voices at +/- 90^o, test condition 3 had different voices at +/- 0^o, and test condition 4 had same voices at +/- 0^o. Based on these four test conditions, there were five different scores calculated. The way these scores were calculated and what each one assesses are outlined in Table 3.3 below.

Table 3.3 – LiSN-S conditions
The five LiSN-S conditions and what they assess. LiSN-S: Listening in Spatialised Noise – Sentences, SRT: Speech Reception Threshold.

LiSN-S	Calculation method	Task description
Low-cue SRT	The score of test condition 4.	The competing speech comes from the same direction (front) and has the same voice as the target speech. Assesses SIN ability when pitch and spatial cues are eliminated.
High-cue SRT	The score of test condition 1.	The competing speech is spatially separated and has different voices from the target speech. Assesses the ability to use both spatial and pitch cues to focus on the target speech.
Talker advantage	The difference in scores between test conditions 3 and 4.	Assesses the SIN ability of the child, as it examines how well the child can distinguish voices from different talkers.
Spatial advantage	The difference in scores between test conditions 2 and 4.	Assesses the advantage the child gets when distracting voices are moved from the front (same source as target) to the side.
Total advantage	The difference in scores between test conditions 1 and 4.	Assesses the same abilities as the High-cue SRT condition (i.e. ability to use spatial and pitch cues to focus on the target speech).
		Total duration: 20 minutes

Because of the way the three advantage measures were calculated (see Table 3.3 above), they are considered to minimise the influence of language, learning and communication variables as stated earlier (Cameron & Dillon, 2007a; Tomlin et al., 2015). Therefore, the Total advantage score may measure the same abilities as the High-cue SRT score, but the former is considered to minimise linguistic and learning factors.

The LiSN-S test ran on a ProBook HP laptop via the Phonak soundcard with the child wearing Sennheiser HD125 headphones. This kept competing speech at a constant level of 55 dB SPL and the frontal target sentences at an initial 62 dB SPL. The target sentences were then adjusted automatically 2 dB up or down if at least 50% of the sentence words were incorrect or correct, respectively. Using head transfer function, competing speech was either coming from the front or side (i.e. 90º), and the distractor speech alternated between different/ same talkers (than/ as the target speech) depending on the condition (see above for details on this). During the test, the child sat opposite the tester and the laptop screen was facing away from the child to minimise distractions. Scores were calculated automatically in the LiSN-S software at the end of all conditions and presented in the form of standardised dB scores and z scores.

The test language of the LiSN-S may have a negative influence on children's performance as the test is presented in North-American English. Previous research has shown that using the North-American SIN SCAN-3 C test on UK children had a negative influence on children's performance compared to North-American norms, and this was due to accent effects (Dawes & Bishop, 2007; Marriage, King, Briggs, & Lutman, 2001). Thus, similar accent effects may cause UK children to underperform in the LiSN-S test (North-American version) as well. Even though there are no published UK normative data, there is a work in preparation that has looked into this matter. The North-American version of the LiSN-S was administered in 48 UK children and then their raw scores were compared against the raw scores of the 72 North-American norms (Murphy & Bamiou, 2018). The comparison did not find significant differences in any of the LiSN-S conditions, except for the Talker advantage condition. In the Talker advantage condition the UK children had a mean z score of -.035 compared to the 0 mean z score of the North-American norms and was thus concluded by the researchers that adding .035 to the obtained z score value would adjust scores and neutralise accent effects (Murphy & Bamiou, 2018). Following this suggestion, .035 in each z score obtained was added in the present study and these adjusted values were used to analyse the Talker advantage condition.

Attention

The test used for measuring different types of attention was the TEACh test, as described in detail under Section 2.2.3 in Chapter 2. Types of selective, sustained and divided attention were measured by the different TEACh subtests summarised in Table 3.4 below. The Walk don't walk and Code transmission conditions, previously used in Chapter 2, were not used in the current study. This was done to reduce testing time and because these two conditions were found to correlate with cognitive ability measures (Manly et al., 1999).

Table 3.4 – TEACh subtests
The four TEACh conditions, what they measure and what children need to do in each of them. DT: Dual Task, TEACh: Test of Everyday Attention for Children.

TEACh subtest	Type of attention measured and task description
Sky search	Measures selective visual attention. Children find visual targets among other visual distractors in an A3-sized paper.
Score!	Measures sustained auditory attention. Children count the number of sounds they hear. As the task is long, they need to self-sustain their attention. The sounds resemble computer game sounds of spaceships firing. No other details on these sounds are provided in the TEACh manual.
Sky search DT	Measures divided auditory-visual attention in a dual task. Children need to find visual targets among distractors in an A3-sized paper, while at the same time count the number of sounds they hear.
Score DT	Measures divided auditory attention. This is another dual task. Children count the number of sounds they hear while they attempt to find a target word in a recorded news report.
	Total duration: 45 minutes

As discussed in the aims and hypotheses of the study, the RMHA is expected to produce improvements only in the Sus-AA and Div-AVA tasks (i.e. in the Score! and Sky search DT conditions, respectively). The tasks of Sel-VA and Div-AA (i.e. Sky search and Score DT) serve as control conditions and are hypothesised to remain unchanged following the intervention period. Test-retest reliability, score calculation and test administration are outlined in the previous chapter, in Section 2.2.3.

Questionnaires

Parents were given the CHAPS and CCC-2 questionnaires to complete, while teachers were given the SIFTER questionnaire (see Section 2.2.3 in Chapter 2 for details). The SIFTER questionnaire was left with teachers to complete within a week, but this resulted in a low return rate (9 out of 26), while 5 of the returned questionnaires were received several weeks later and upon request. Due to the low return rate and delayed completion, the SIFTER was excluded from the analysis. It was not possible for teachers to complete the questionnaire during the researcher's school visit, as they only had a limited time to attend the RMHA session and completion of the questionnaire during the RMHA session would have resulted in rushed answers.

In this study, children also completed a questionnaire, the LIFE-R (K. L. Anderson et al., 2011). The LIFE-R questionnaire assesses how well the student listens in the classroom under different noisy or acoustically-challenging situations (e.g. traffic noise, classroom noise, teacher moving around etc.). Children completed this questionnaire during the test session. If any parts of the questionnaire were not understood by the child, they were discussed and further clarified. The questions were specific to each group of the study, meaning that children in the control group were asked how well they hear their teacher in these listening situations (without any mention of hearing aids) whereas children in the intervention group were asked how well they hear their teacher in these listening situations through the hearing aids. It was hypothesised that questions 1 to 8 would show improved scores over time, since these essentially assess children's perception of their listening when the system is used. Question 9 asked children how well they hear the person talking in school assembly. It was decided that during assembly the RMHA was not going to be used, since it would have been more difficult to use it with different speakers having to wear the mic. As discussed in the introduction (see Section 1.8.1 in Chapter 1), SIN ability was expected to improve even unaided, hence it was hypothesised that question 9 would also increase at 3 and 6 months.

The LIFE-R is a non-standardised questionnaire, thus the scores collected and analysed were raw scores. This has some drawbacks, with the main one being that in case of potential improved scores following an intervention, it is not possible to exclude maturation as the reason of the improvement. However, the questions were revised and edited specifically for primary school students (K. L. Anderson et al., 2011) and this makes it easy for a child to comprehend the questions and answer more accurately (Purdy et al., 2009). Purdy et al. (2009) also checked the LIFE-R's test-retest reliability on a sample of 18 typically developing children. They administered it twice, with a 6-week gap in between and without any intervention given to the typically developing controls. Spearman's *r* was between .63 and .74, revealing moderate test reliability (Purdy et al., 2009).

Testing procedure

During testing the child's parent or parents were also seated in the room, but away from the child. Parents were asked to keep quiet, have their phones on silent mode and avoid other distractions.

Testing time was approximately 1 hour 15 minutes and regular breaks were given. The two behavioural tests were randomised (using simple, computer-generated randomisation with Excel formula) for each test session and for each participant to control for order effects. The tasks of both the TEACh and the LiSN-S test are set in a hierarchical order, building each task based on the previous one, therefore task randomisation was not used. In the previous study in Chapter 2, the experimenter was not blind to the baseline data during the retest at 3 months and this may have introduced observer bias. In this current study the researcher was still the one who collected the raw data but did not convert them into scaled scores (which was the format in which data were analysed and giving indications on performance) until the study was completed at 6 months. It was not possible to appoint another researcher for the data collection or data analysis process due to budgetary constraints.

3.2.4 Statistical analysis

Data were analysed in SPSS 22 statistics software, using a mixed design ANOVA, with group (control or RMHA) as the between-subjects factor and time (baseline, at 3 and 6 months) the within-subjects factor. Correction for multiple comparisons was made using Bonferroni correction for the dependent variable (i.e. time) and the interaction between group and time. There was no correction for multiple comparisons for the multiple administered tests in this study and a detailed discussion along with the reasons behind this are given in the previous Chapter in Section 2.2.4. If the intervention group returns statistically significant improved scores over the control group in the behavioural test and in the questionnaires, then correlation analysis using Pearson's correlation will be applied, as outlined in the hypotheses (see Section 3.1.3).

3.3 Results

A post-hoc power calculation was performed on the actual sample size (n = 26) to detect 1 SD effect size at 5% significance, which resulted in a statistical power of 70%. Due to the limited recruitment from GOSH and time constraints the study was carried out despite the reduced statistical power. The implications of this are outlined later in Discussion.

3.3.1 Pure-tone audiometry and nonverbal IQ

All children had normal hearing thresholds, below 20 dB HL in each frequency between 250 Hz and 8 KHz (BSA, 2012). In addition, only children with a score above 85 on the WNV test were included in the study. The 26 children meeting this criterion gave a mean composite full scale score of 102.5 on the WNV test (Wechsler & Naglieri, 2006). After the two groups were constructed the two mean scores of the groups were compared. The controls had a mean of 107 score on the WNV test, while the intervention group had a mean of 99 (see Table 3.1). An independent samples t-test was conducted to compare the IQ scores between the two groups. No differences were found between the scores of the two groups, $t(24) = 1.95$, $p = .063$, $d = .76$ (95% CI [-.04, 1.55]).

3.3.2 Attention

There were no outliers in any of the four conditions of the TEACh test (examined by looking for values in the studentised residuals beyond +/- 3 SDs). All conditions had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) were statistically non-significant ($p > .05$) in the two control conditions (Sel-VA and Div-AA), but statistically significant ($p < .05$) in the two experimental conditions (Sus-AA and Div-AVA). For the Div-AVA condition, natural log transformation was performed, whereas for the Sus-AA none of the transformations that were applied would make the data of this test normally distributed. Also, other assumptions were not met when transformations of the Sus-AA scores were performed. For these reasons, it was decided that the Sus-AA analysis would be carried out despite the significant results in the Levene's test of equality of error variance. The implications of this are outlined later in the Discussion. The last assumption of Equality of covariance matrices was statistically non-significant ($p > .001$) in all four TEACh conditions.

Figure 3.3 presents the results of the two experimental conditions Sus-AA and Div-AVA.

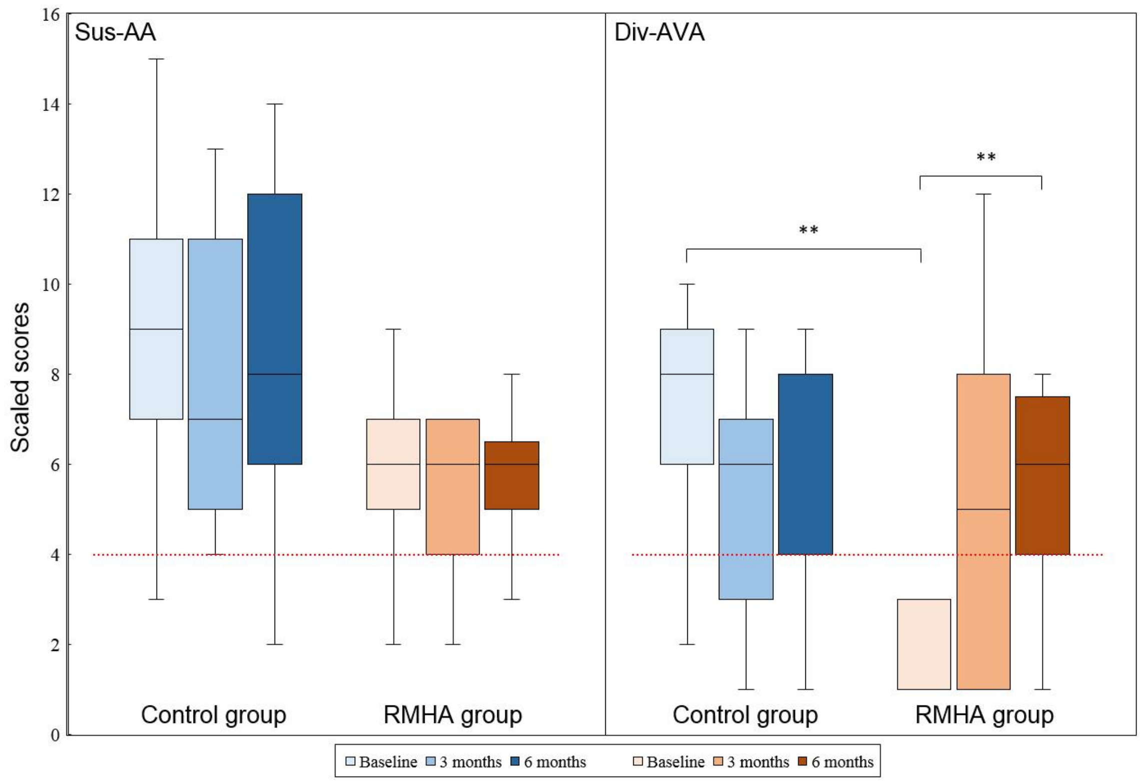

Figure 3.3 – Results of the TEACh experimental conditions
Boxplots of scaled scores of the two experimental TEACh tasks measuring Sus-AA and Div-AVA, for 13 controls and 13 RMHA children, at baseline, 3 and 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children. $^{**}p < .01$

Figure 3.4 below outlines the results of the two control conditions Sel-VA and Div-AA.

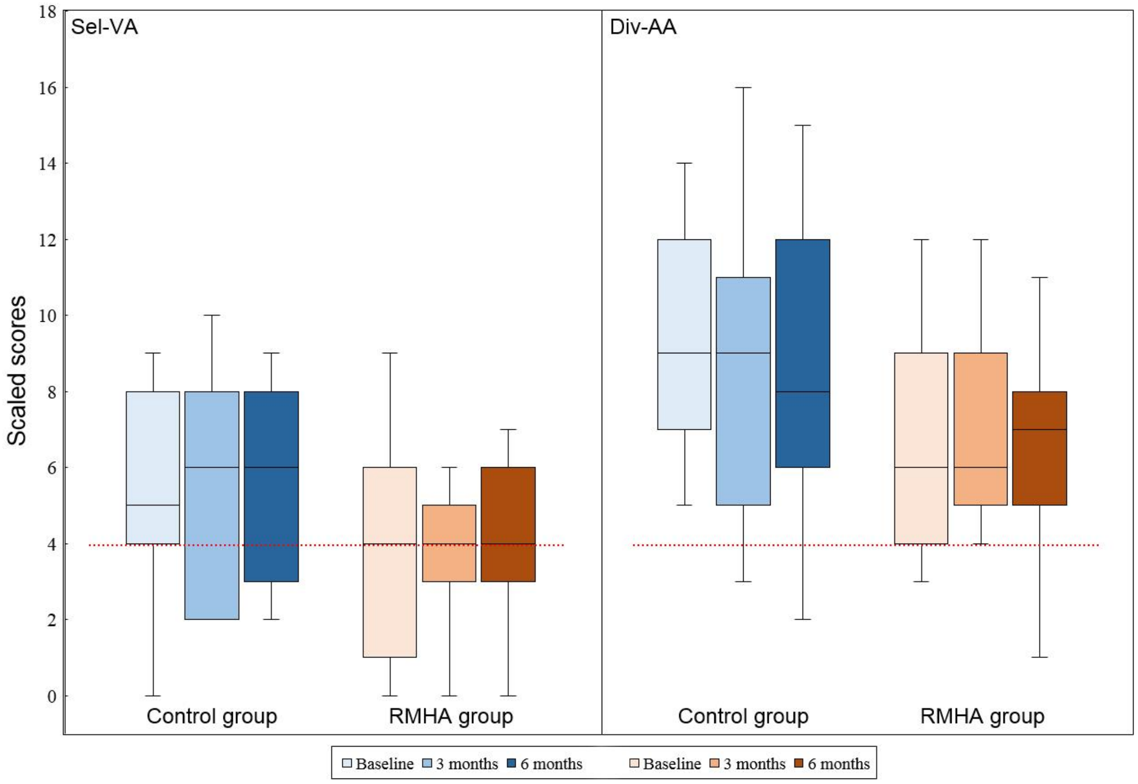

Figure 3.4 – Results of the TEACh control conditions
Boxplots of scaled scores of the two control TEACh tasks measuring Sel-VA and Div-AA, for 13 controls and 13 RMHA children, at baseline, 3 and 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AA: Divided Auditory Attention, RMHA: Remote Microphone Hearing Aid, Sel-VA: Selective Visual Attention, TEACh: Test of Everyday Attention for Children.

Table 3.5 presents the mean values of the control and intervention group at baseline, at 3 and 6 months.

Table 3.5 – TEACh mean scores and SDs
The mean scores and SDs for the conditions of the TEACh test for the two groups at all three time-points. Disordered performance ≤ 4, borderline performance > 4 and ≤ 7, and normal performance > 7. Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone hearing aid, SD: Standard Deviation, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children.

TEACh	**Mean (SD)**					
	Control group			**RMHA group**		
	Baseline	**3 months**	**6 months**	**Baseline**	**3 months**	**6 months**
Sus-AA	9.76 (4.00)	7.76 (3.67)	8.23 (4.22)	6.30 (2.62)	6.61 (3.84)	5.76 (1.87)
Div-AVA	6.61 (3.17)	5.23 (2.83)	5.92 (3.25)	2.38 (2.39)	5.15 (3.67)	5.76 (2.16)
Sel-VA	9.69 (2.01)	10.30 (3.61)	10.92 (1.93)	6.00 (2.85)	8.30 (2.71)	9.61 (2.32)
Div-AA	9.15 (3.13)	8.61 (4.21)	8.69 (4.17)	6.53 (3.12)	6.76 (2.55)	6.07 (3.17)

Table 3.6 below presents the p-values, effect sizes and CIs of the effect sizes of the interaction between group and time (and the significant simple main effects of group and time, if the interaction is significant) and of the main effect of group and main effect of time. There was a significant interaction in the Div-AVA test, $F(2, 48) = 4.47$, $p = .017$, $\eta_p^2 = .157$. This interaction was followed up by looking at the simple main effects of group and time. There was a statistically significant difference in scores between the two groups at baseline, $F(1, 24) = 12.63$, $p = .002$, $\eta_p^2 = .345$, meaning that the control group had significantly better scores at baseline than the intervention group. There was also a statistically significant effect of time in Div-AVA scores of the RMHA group, $F(2, 24) = 7.14$, $p = .004$, $\eta_p^2 = .373$, but not in the scores of the control group (see Table 3.6 for details). To determine the time-point of this difference, pairwise comparisons were run. This showed that in the RMHA group, Div-AVA scores statistically improved from baseline to 6 months, $M = -.481$, Standard Error (SE) = .092, $p = .001$.

Table 3.6 – TEACh significance

The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the simple main effect of group and time and main effects of group and time in the four TEACh conditions. Highlighted in grey is the condition that did not meet the assumption of homogeneity of variances as transformation would not work. CI: Confidence Interval, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Type of attention	**Interaction**		**Simple main effect**				**Main effect**			
	Group* Time	η_p^2 **[95% CI of η_p^2]**	**Group**	η_p^2 **[95% CI of η_p^2]**	**Time**	η_p^2 **[95% CI of η_p^2]/ *MD* [95% CI of *MD*]**	**Group**	η_p^2 **[95% CI of η_p^2]**	**Time**	η_p^2 **[95% CI of η_p^2]/ *MD* [95% CI of *MD*]**
Sus-AA	.242	.057 [.000, .192]	-	-	-	-	.045*	.156 [.000, .397]	.274	.053 [.000, .184]
Div-AVA	.017*	.157 [.004, .319]	.002** (**baseline**)	.345 [.063, .557]	.779 (**Control**)	.021 [.000, .155]	-	-	-	-
			.693 (**3 months**)	.007 [.000, .169]	.004** (**RMHA**)	.373 [.056, .561]				
			.665 (**6 months**)	.008 [.000, .175]	.140 (**0-3months**)	-.333 [-.750, .084]				
					.001** (**0-6months**)	-.481 [-.738, -.224]				
					.945 (**3-6months**)	-.148 [-.541, .245]				
Sel-VA	.074†	.103 [.000, .256]	-	-	-	-	.011*	.243 [.014, .475]	.000***	.312 [.094, .470]
									.086† (**0-3 months**)	-1.462 [-3.077, .154]
									.000*** (**0-6 months**)	-2.423 [-3.648, -1.198]
									.128 (**3-6 months**)	-.962 [-2.117, .194]
Div-AA	.781	.010 [.000, .087]	-	-	-	-	.049*	.151 [.000, .392]	.758	.011 [.000, .092]

In the other three conditions (i.e. Sus-AA, Sel-VA and Div-AA), there was no statistically significant interaction between group and time. There were only statistically significant differences in the main effect of group in all three tasks; Sus-AA task: $F(1, 24) = 4.45$, $p = .045$, $\eta_p^2 = .156$, Sel-VA task: $F(1, 24) = 7.69$, $p = .011$, $\eta_p^2 = .243$, and Div-AA task: $F(1, 24) = 4.28$, $p = .049$, $\eta_p^2 = .151$. This means that if the three scores at baseline, 3 and 6 months are collapsed across time, a significant difference between the two groups is observed. In other words, the control group has better scores than the intervention group, regardless of the time variable (i.e. scores at baseline, at 3 months and at 6 months merged together for each group). Finally, there was a statistically significant difference in the main effect of time in the Sel-VA test, $F(2, 48) = 10.86$, $p < .001$, $\eta_p^2 = .312$. This means that if the two groups are collapsed together, a significant difference over time is observed. In other words, the scores at 6 months are better than at baseline, regardless of the group variable (i.e. control and intervention group merged together) as evidenced by the pairwise comparisons, $M = -2.423$, $SE = .476$, $p < .001$.

Post-hoc analysis controlling for baseline in Div-AVA

As observed earlier, there were significant differences between the intervention and control group in the baseline Div-AVA scores, with the control group scoring significantly better than the intervention group. In intervention studies with clinical control groups there should be no baseline differences as this has a direct impact on the post-intervention scores. The case here was that the intervention group scored worse than the control group and then improved its scores post-intervention. In this instance, it is not possible to know whether the improved performance was due to the intervention or perhaps it was an effect of regression to the mean, placebo effect, or Hawthorne effect (these possibilities are described in more detail later in Discussion). For this reason, control and intervention group should start without prior differences at baseline. Therefore, a post-hoc analysis controlling for baseline was conducted. Specifically, a one-way Analysis of Covariance (ANCOVA) was performed comparing the difference from baseline to 6 months in the Div-AVA scores between the two groups while controlling for the baseline Div-AVA scores.

All assumptions for the one-way ANCOVA were met. In detail, the assumptions of absence of outliers (checked through looking for values outside +/- 3 SDs from the mean on the standardised residuals), normality (checked using the Shapiro-Wilk test on the standardised residuals, $p > .05$), homogeneity of variances (assessed through the Levene's test, $p > .05$), homoscedasticity (checked through scatterplots between the standardised residuals and predicted values), linearity between the dependent variable and the covariate for both levels of the independent variable (assessed through scatterplots between the two variables), and the assumption of homogeneity of regression slopes (assessed through the interaction term, $p > .05$) were all met.

Figure 3.5 below summarises the findings from the one-way ANCOVA analysis.

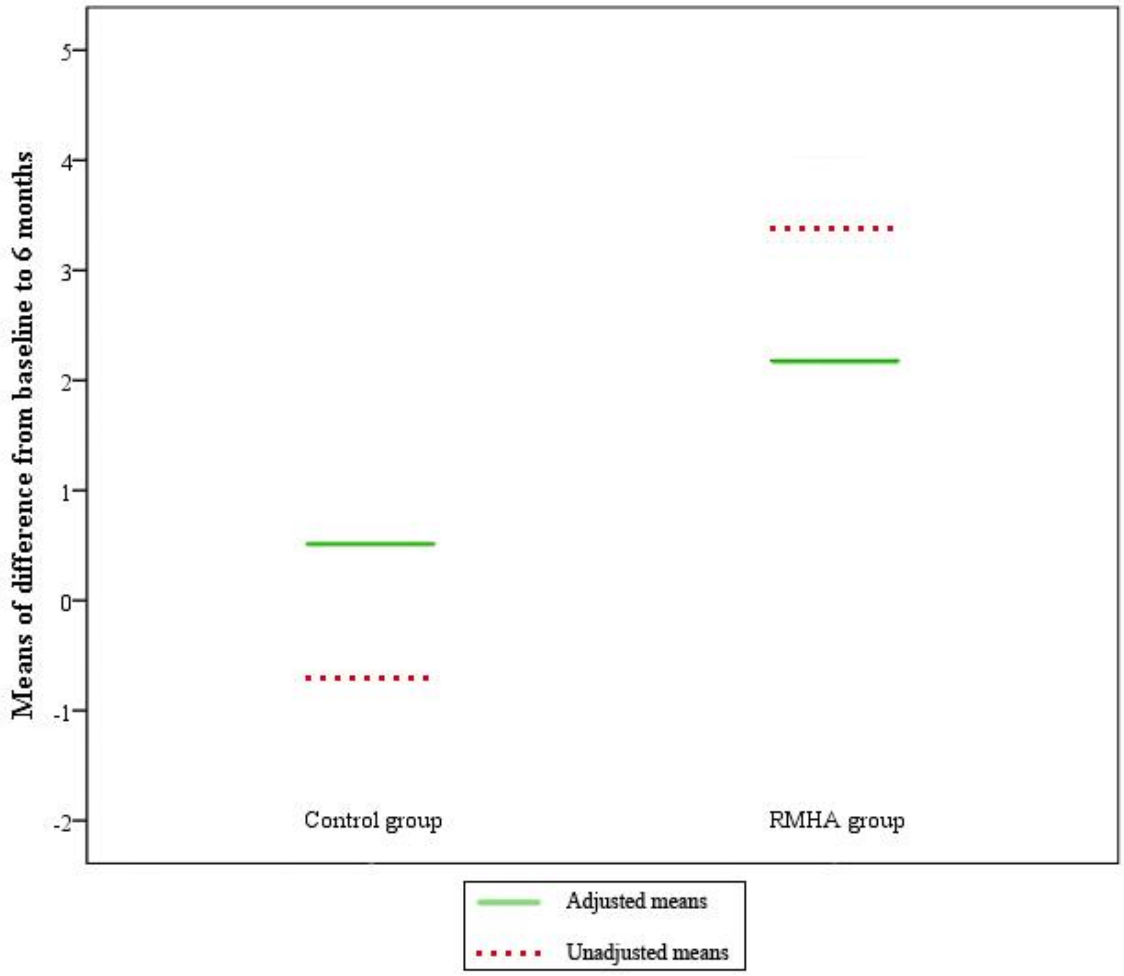

Figure 3.5 – Results from one-way ANCOVA analysis
Adjusted means of the difference from baseline to 6 months in the Div-AVA test for 13 controls and 13 RMHA children are presented in green lines. Please note that since the statistical package only produced the adjusted means and not the adjusted individual values for each participant it was only possible to present the means in the figure above without any additional information (e.g. quartile values, upper and lower observations) in the form of boxplot as was previously done. The unadjusted means (that is, without control for baseline scores) are also included using red dotted lines for visual comparisons with the adjusted means. ANCOVA: Analysis of Covariance, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid.

Table 3.7 below summarises the descriptive statistics from the one-way ANCOVA analysis for the control and RMHA group.

Table 3.7 – Descriptive statistics from the ANCOVA analysis
Adjusted means, SEs, and 95% CIs of the means. ANCOVA: Analysis of Covariance, CI: Confidence Interval, SE: Standard Error.

Test	**Adjusted mean (SE) [95% CI of *M*]**	
	Control group	**RMHA group**
Div-AVA	.513 (.80) [-1.15, 2.17]	2.17 (.80) [.51, 3.84]

The one-way ANCOVA analysis did not reveal a significant difference between the two groups in the adjusted Div-AVA scores (difference from baseline to 6 months), $F(1, 23) = 1.737$, $p = .200$, $\eta_p^2 = .070$, (95% CI [.000, .303]). Thus, after controlling for the baseline scores the difference in scores from baseline to 6 months was no longer statistically significant between the two groups. The implications of this are outlined later in Discussion.

3.3.3 Listening-in-noise

There were no outliers in any of the five LiSN-S conditions (examined by looking for values in the studentised residuals beyond +/- 3 SDs), except for the High-cue SRT (outlier at -3.20 SDs) and Total advantage (outlier at -3.31 SDs). Outliers were removed from the analysis of the High-cue SRT and Total advantage conditions, thus resulting in 13 controls against 12 RMHA children for the analyses of those two conditions. All conditions had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 3.6 below summarises the results of the two experimental conditions (i.e. Low-cue SRT and Talker advantage).

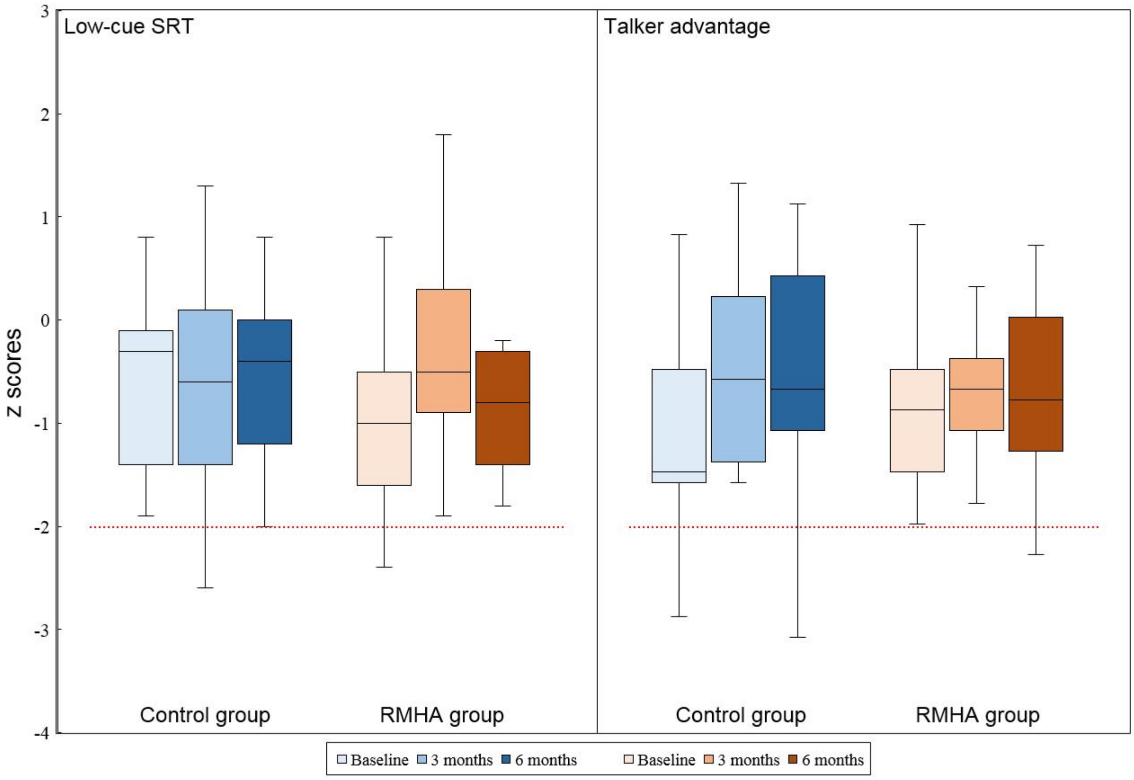

Figure 3.6 – Results of the LiSN-S experimental conditions
Boxplots of z scores of the two experimental LiSN-S tasks measuring SIN, for 13 controls and 13 RMHA children, at baseline, 3 and 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below -2 z scores). LiSN-S: Listening in Spatialised Noise – Sentences, RMHA: Remote Microphone Hearing Aid, SIN: Speech-in-Noise, SRT: Speech Reception Threshold.

Figure 3.7 below summarises the results of the three control conditions Spatial advantage, High-cue SRT and Total advantage.

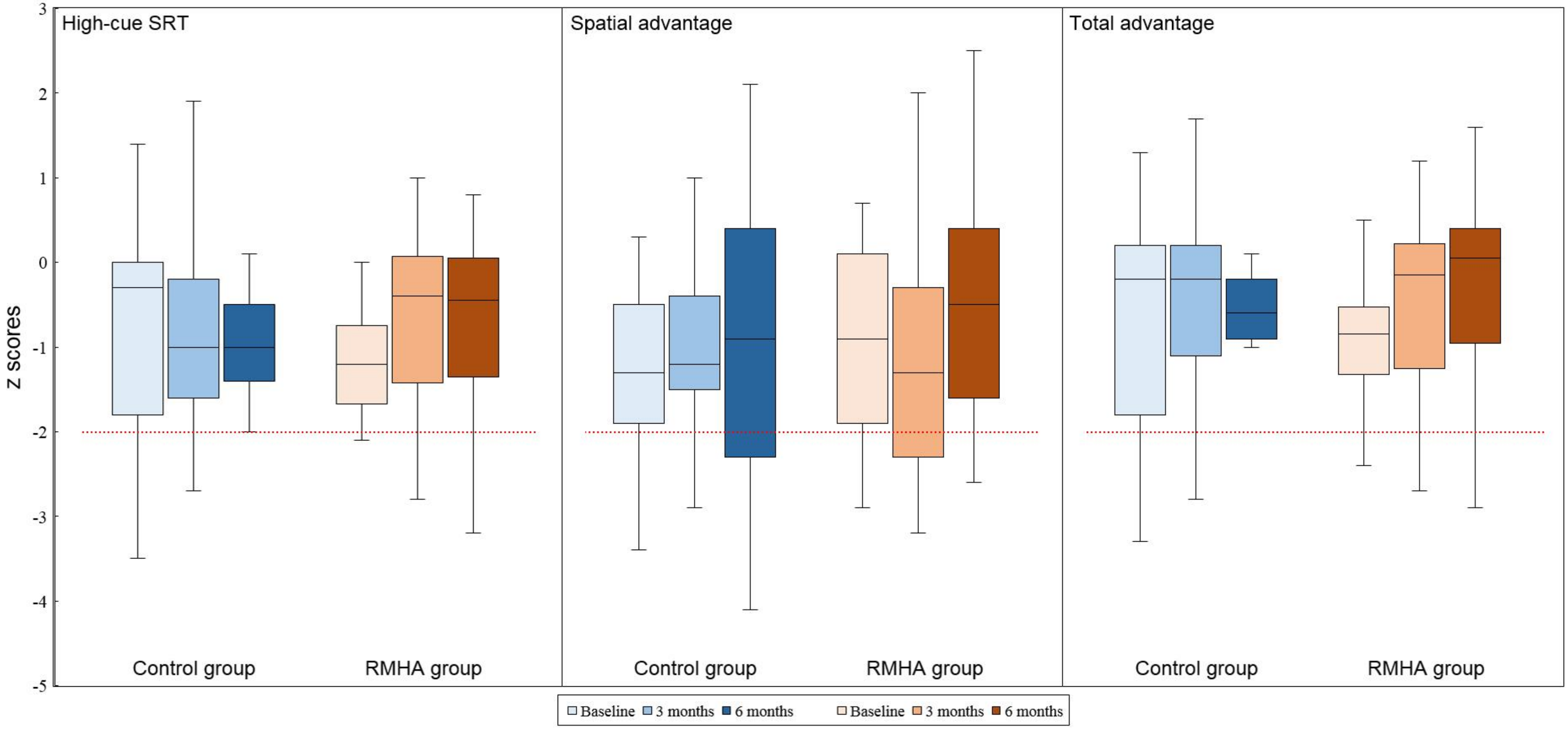

Figure 3.7 – Results of the LiSN-S control conditions
Boxplots of z scores of the three LiSN-S control conditions, for 13 controls and 12 RMHA children (and 13 controls and 13 RMHA children in Spatial advantage condition), at baseline, 3 months and at 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below -2 z scores). LiSN-S: Listening in Spatialised Noise - Sentences, RMHA: Remote Microphone Hearing Aid, SRT: Speech Reception Threshold.

Table 3.8 below presents the mean values of the control and intervention group at baseline, at 3 and 6 months.

Table 3.8 – LiSN-S mean scores and SDs
The mean scores and SDs for the LiSN-S conditions. Disordered performance: ≤ -2, borderline performance: > -2 and ≤ -1, and normal performance: > -1. LiSN-S: Listening in Spatialised Noise – Sentences, RMHA: Remote Microphone hearing aid, SD: Standard Deviation, SRT: Speech Reception Threshold.

LiSN-S	**Mean (SD)**					
	Control group			**RMHA group**		
	Baseline	**3 months**	**6 months**	**Baseline**	**3 months**	**6 months**
Low-cue SRT	-.56 (.90)	-.46 (1.43)	-1.38 (3.33)	-.99 (1.12)	-.63 (1.61)	-1.05 (1.51)
High-cue SRT	-.83 (1.45)	-.70 (1.34)	-1.23 (1.82)	-1.27 (1.02)	-.60 (1.08)	-.68 (1.66)
Talker advantage	-1.01 (1.14)	-.32 (1.00)	-.44 (1.11)	-.84 (.80)	-.59 (.82)	-.68 (.93)
Spatial advantage	-1.50 (1.54)	-1.31 (1.41)	-.98 (1.76)	-.97 (1.07)	-.99 (1.57)	-.53 (1.35)
Total advantage	-.56 (1.34)	-.52 (1.28)	-.63 (1.07)	-.94 (1.12)	-.47 (1.06)	-.30 (1.28)

Table 3.9 below presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. None of the LiSN-S conditions had a significant interaction between group and time. There was also lack of significance in the main effect of group and main effect of time.

Table 3.9 – LiSN-S significance
The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the main effects of group and time in the five LiSN-S conditions. CI: Confidence Interval, LiSN-S: Listening in Spatialised Noise – Sentences, SRT: Speech Reception Threshold. $^{\dagger}p < .10$

LiSN-S	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Low-cue SRT	.562	.019	.000, .115	.875	.001	.000, .039	.252	.056	.000, .189
High-cue SRT	.270	.055	.000, .191	.879	.001	.000, .039	.377	.042	.000, .168
Talker advantage	.512	.027	.000, .137	.704	.006	.000, .166	.084†	.098	.000, .250
Spatial advantage	.937	.003	.000, .052	.363	.035	.000, .243	.219	.061	.000, .198
Total advantage	.417	.037	.000, .160	.995	.000	.000, .000	.494	.030	.000, .145

3.3.4 CHAPS questionnaire

There were no outliers in any of the four CHAPS subscales (i.e. Noise, Multiple inputs, Auditory memory sequencing, Auditory attention span; examined by looking for values in the studentised residuals beyond +/- 3 SDs). All subscales had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) were statistically non-significant ($p > .05$) in all subscales except for the Noise subscale ($p < .05$). For Noise, none of the transformations that were applied would make the Levene's test non-significant, while other assumptions were also not met when transformations were applied. For this reason, it was decided that the analysis for the Noise subscale would be carried out without transformation of scores despite the significant results in the Levene's test of equality of error variance. The implications of this are described later in Discussion. The last assumption of Equality of covariance matrices was statistically non-significant in all four CHAPS subscales ($p > .001$).

Figure 3.8 summarises these findings of the four CHAPS subscales.

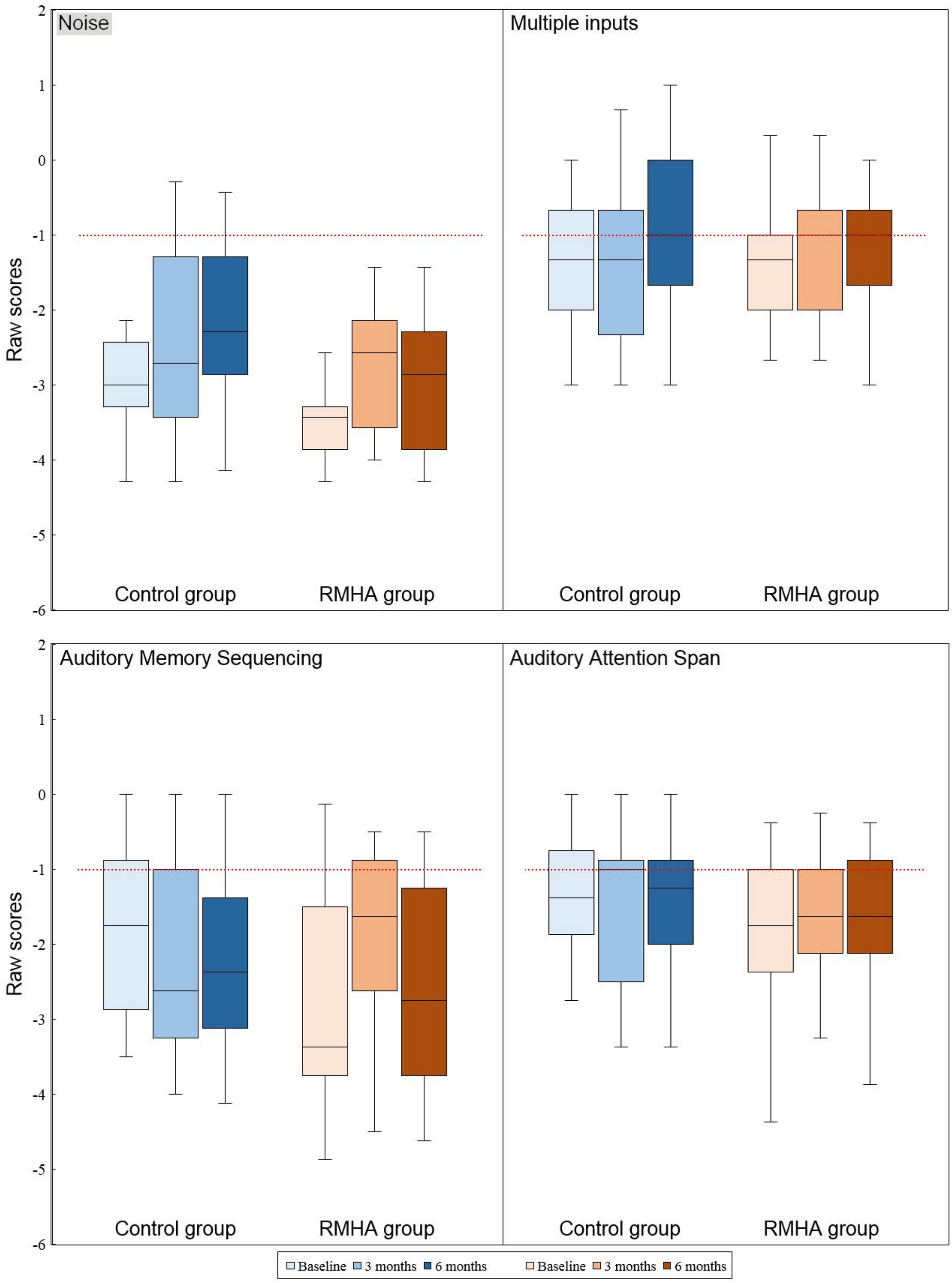

Figure 3.8 – Results of the CHAPS

Boxplots of raw scores of the four CHAPS subscales, for 13 controls and 13 RMHA children, at baseline, at 3 months and at 6 months. The Noise condition highlighted in grey did not meet the assumption of homogeneity of variances (see discussion previously for details). The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. Even though this was a non-standardised questionnaire, the CHAPS manual did mention that it considered abnormal performance to be below -1 on the raw scores. Therefore, the red dotted line indicates this cut-off. CHAPS: Children's Auditory Performance Scale, RMHA: Remote Microphone Hearing Aid.

Table 3.10 presents the mean values of the control and intervention group at baseline, at 3 and 6 months.

Table 3.10 – CHAPS mean scores and SDs

The mean scores and SDs for the conditions of the CHAPS subscales for the two groups at all three time-points. Disordered performance: ≤ -1, and normal performance: > -1. CHAPS: Children's Auditory Performance Scale, RMHA: Remote Microphone hearing aid, SD: Standard Deviation.

CHAPS	**Mean (SD)**					
	Control group			**RMHA group**		
	Baseline	**3 months**	**6 months**	**Baseline**	**3 months**	**6 months**
Noise	-2.85 (1.01)	-2.36 (1.31)	-2.19 (1.07)	-3.29 (.89)	-2.82 (.83)	-2.95 (.97)
Multiple inputs	-1.31 (.99)	-1.33 (1.09)	-1.05 (1.14)	-1.62 (1.21)	-1.21 (1.00)	-1.33 (1.41)
Auditory memory sequencing	-1.77 (1.15)	-2.24 (1.30)	-2.14 (1.35)	-2.63 (1.53)	-2.06 (1.33)	-2.51 (1.41)
Auditory attention span	-1.34 (.86)	-1.62 (1.06)	-1.46 (1.04)	-1.84 (1.16)	-1.57 (.93)	-1.77 (1.17)

Table 3.11 below presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. There were no statistically significant interactions between group and time in any of the four CHAPS conditions. The main effect of group also did not show significant results, whereas the main effect of time was non-significant in all conditions except for the Noise subscale, $F(2, 48) = 5.52$, $p = .007$, $\eta_p^2 = .187$. This means that if the two groups were collapsed together, a significant difference over time was observed. In other words, the scores at 6 months were better than at baseline, regardless of the group variable (i.e. control and intervention group merged together) as evidenced by the pairwise comparisons, $M = -3.50$, $SE = 1.171$, $p = .019$.

Table 3.11 – CHAPS significance

The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the main effects of group and time in the four CHAPS conditions. Highlighted in grey is the condition that did not meet the assumption of homogeneity of variances and transformations would not work. CHAPS: Children's Auditory Performance Scale, CI: Confidence Interval. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

CHAPS	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2/ ***MD***	**95% CI of** η_p^2/ **95% CI of *MD***
Noise	.579	.023	.000, .125	.129	.093	.000, .328	.007**	.187	.017, .351
							.055† (**0-3 months**)	-3.308	-6.674, .058
							.019* (**0-6 months**)	-3.500	-6.515, -.485
							1.000 (**3-6 months**)	-.192	-2.922, 2.537
Multiple inputs	.539	.025	.000, .132	.683	.007	.000, .171	.453	.032	.000, .147
Auditory memory sequencing	.066†	.113	.000, .268	.469	.022	.000, .218	.652	.016	.000, .107
Auditory attention span	.233	.060	.000, .195	.494	.020	.000, .212	.953	.001	.000, .015

3.3.5 CCC-2 questionnaire

The CCC-2 subscales were grouped to create two composite scores. The first composite score was created by averaging the subscales that reflect standard (or structural) language aspects (i.e. Speech, Syntax, Semantics and Coherence subscales). The second composite score resulted from the averaging of the subscales that represent non-standard (or non-structural) language aspects (i.e. Inappropriate Initiation, Stereotyped Language, Use of Context and Non-verbal Communication subscales). These scores were not used to identify children at risk of DLD or PLI as was previously clarified in Chapter 2 (see Section 2.3.6).

Before calculating the composite scores, some of the data from the CCC-2 questionnaire had to be removed or imputed. The CCC-2 data from participant i-12 were removed, as data in all 3 time-points were invalid[19]. Data from case i-12 were not imputed, as all their data were missing and imputing for that would introduce bias. Imputation only took place for participant i-07 at baseline. Their data at that time point were not missing but were also invalid (Bishop, 2003), as explained in the footnote (see footnote 19 below). Hence the analysis included 13 controls and 12 children from the intervention group. There were no outliers in any of the two composite scores (examined by looking for values in the studentised residuals beyond +/- 3 SDs). Both composite scores had normally distributed data checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 3.9 presents the findings of the standard and non-standard language composite scores of the CCC-2 questionnaire.

---

[19] In CCC-2 the positive and negative sum need to be consistent. The positive sum is calculated in the following way: '6 minus the sum of items 51 to 70', and the negative sum in this way: 'the sum of items 1 to 26 plus the sum of items 27 to 50'. When the two scores are inconsistent –according to the Table 3B in page 69 of the CCC-2 manual (Bishop, 2003)– then the answers from the questionnaire are considered invalid.

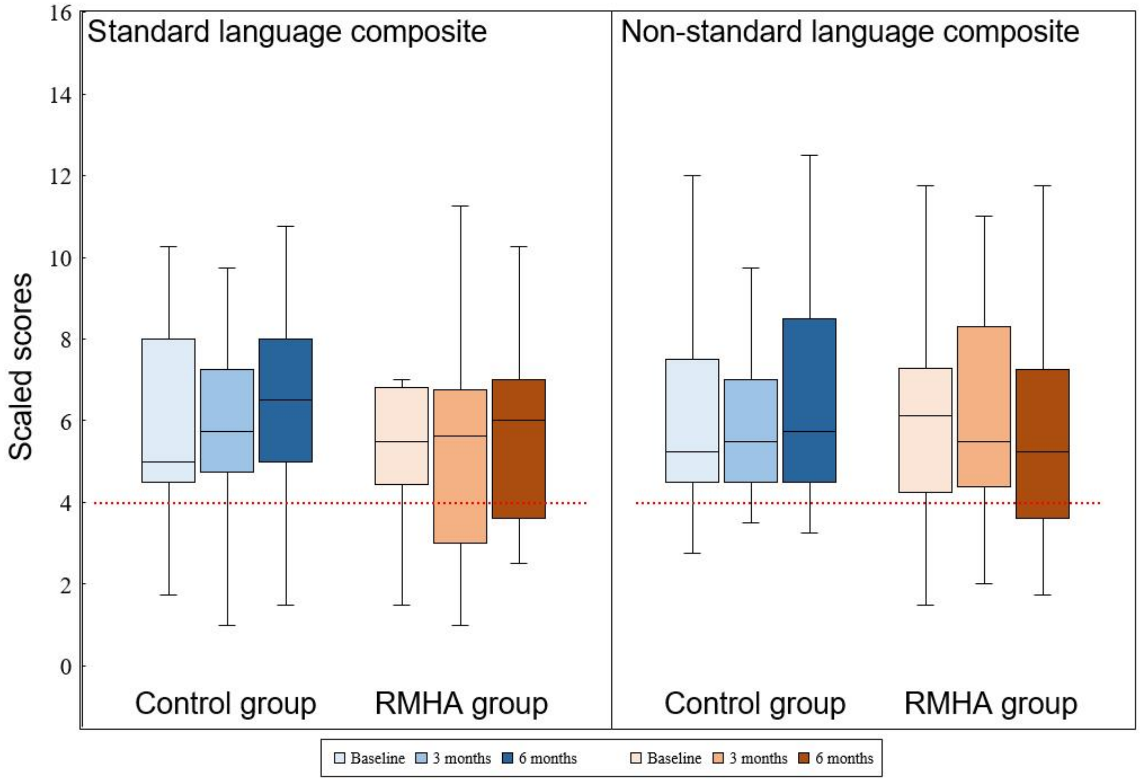

Figure 3.9 – Results of the CCC-2 composite scores
Boxplots of scaled scores of the standard and non-standard language composite scores of the CCC-2 questionnaire, for 13 controls and 12 RMHA children, at baseline, 3 months and at 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). CCC-2: Children's Communication Checklist – 2, RMHA: Remote Microphone Hearing Aid.

Table 3.12 presents the mean values of the control and intervention group at baseline, at 3 and 6 months.

Table 3.12 – CCC-2 mean scores and SDs
The mean scores and SDs for the two composite CCC-2 scores for the two groups at all three time-points. The GCC and SIDC scores are also presented for additional descriptive information. Disordered performance: ≤ 4, borderline performance: > 4 and ≤ 7, and normal performance: > 7. CCC-2: Children's Communication Checklist – 2, GCC: General Communication Composite, RMHA: Remote Microphone hearing aid, SD: Standard Deviation, SIDC: Social Interaction Deviance Composite.

CCC-2	**Mean (SD)**					
	Control group			**RMHA group**		
	Baseline	**3 months**	**6 months**	**Baseline**	**3 months**	**6 months**
Standard language	5.82 (2.65)	6.03 (2.71)	6.50 (2.84)	5.44 (2.56)	5.39 (2.88)	5.58 (2.35)
Non-standard language	6.25 (2.82)	6.38 (2.87)	6.61 (3.06)	6.19 (2.66)	6.08 (2.60)	5.70 (2.83)
GCC score	48.31 (19.70)	49.69 (19.89)	52.46 (21.27)	46.39 (16.55)	45.92 (19.08)	45.17 (16.33)
SIDC score	1.62 (9.43)	2.38 (9.53)	1.54 (9.96)	5.45 (12.63)	3.42 (9.07)	0.50 (12.26)

As seen in Table 3.13 below, there was no significant interaction between group and time in any of the CCC-2 composite scores.

Table 3.13 – CCC-2 composite scores significance
The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the main effect of group and main effect of time of the CCC-2 composite scores. CCC-2: Children's Communication Checklist - 2, CI: Confidence Interval.

CCC-2	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Standard language	.790	.010	.000, .089	.513	.019	.000, .215	.546	.026	.000, .136
Non-standard language	.218	.064	.000, .205	.704	.006	.000, .172	.952	.002	.000, .042

3.3.6 LIFE-R questionnaire

There were no outliers in any of the 9 questions or the total score of the LIFE-R questionnaire (examined by looking for values in the studentised residuals beyond +/- 3 SDs), except for Quiet question (outlier at -4.29 SDs), Teacher moving around question (outlier at -3.28 SDs) and for the Total score (outlier at -3.25 SDs). The outlier of the Total score was removed. For Quiet none of the assumptions discussed below were met and for that reason no analysis was carried out on these scores altogether. For Teacher moving around, it was decided that the analysis would run without removing the outlier, as removal caused other assumptions not to be met. All questions and the Total score had normally distributed data (which was checked using the Shapiro-Wilk test of normality of distribution), except for Quiet (as noted above). Homogeneity of variances (using the Levene's test of equality of error variance) were statistically non-significant ($p > .05$) in all questions except for Traffic noise and Quiet ($p < .05$). For Traffic noise, none of the transformations that were applied would make this test non-significant, while other assumptions were also not met if transformations were applied. For this reason, it was decided that the analysis for Question 1 (Traffic noise) was going to be carried out despite the significant results in the Levene's test of equality of error variance. The implications of this are noted later in Discussion. The last assumption of Equality of covariance matrices was statistically non-significant, $p > .001$, except for Quiet, as noted earlier.

Figures 3.10, 3.11 and 3.12 below graphically present the above findings. The 9 LIFE-R questions and total score are divided in 3 different figures for easier reading.

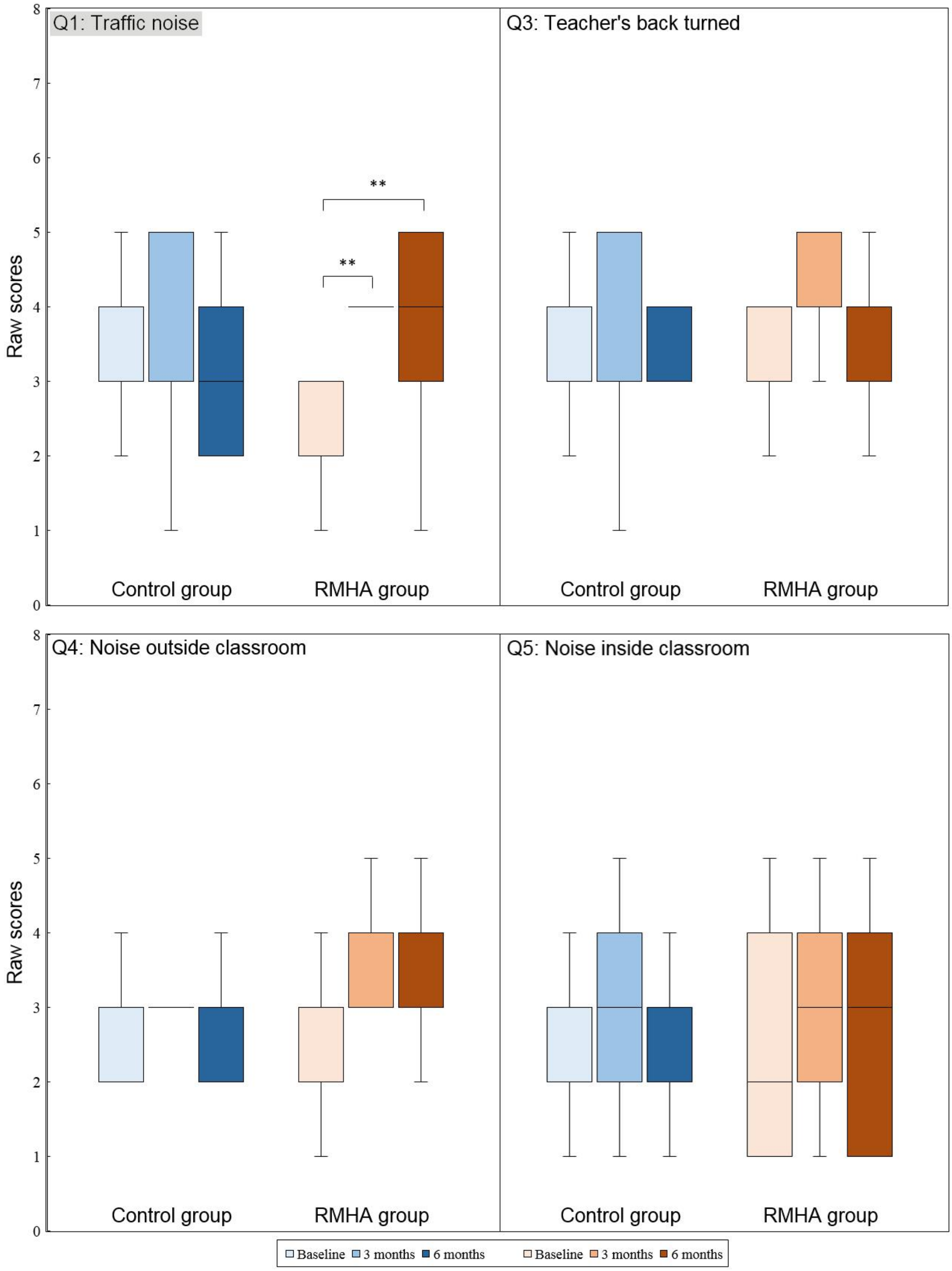

Figure 3.10 – Results of the LIFE-R Q1 and Q3-5
Boxplots of raw scores of questions 1,3,4 and 5 of the LIFE-R questionnaire, for 13 controls and 13 RMHA children, at baseline, 3 months and at 6 months. Question 1 (Traffic noise) highlighted in grey did not meet the assumption of homogeneity of variances (see text for details). The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. This was a non-standardised questionnaire, therefore there was no indication of a cut-off for abnormal performance. LIFE-R: Listening Inventory for Education – Revised, Q: Question, RMHA: Remote Microphone Hearing Aid. $**p < .01$

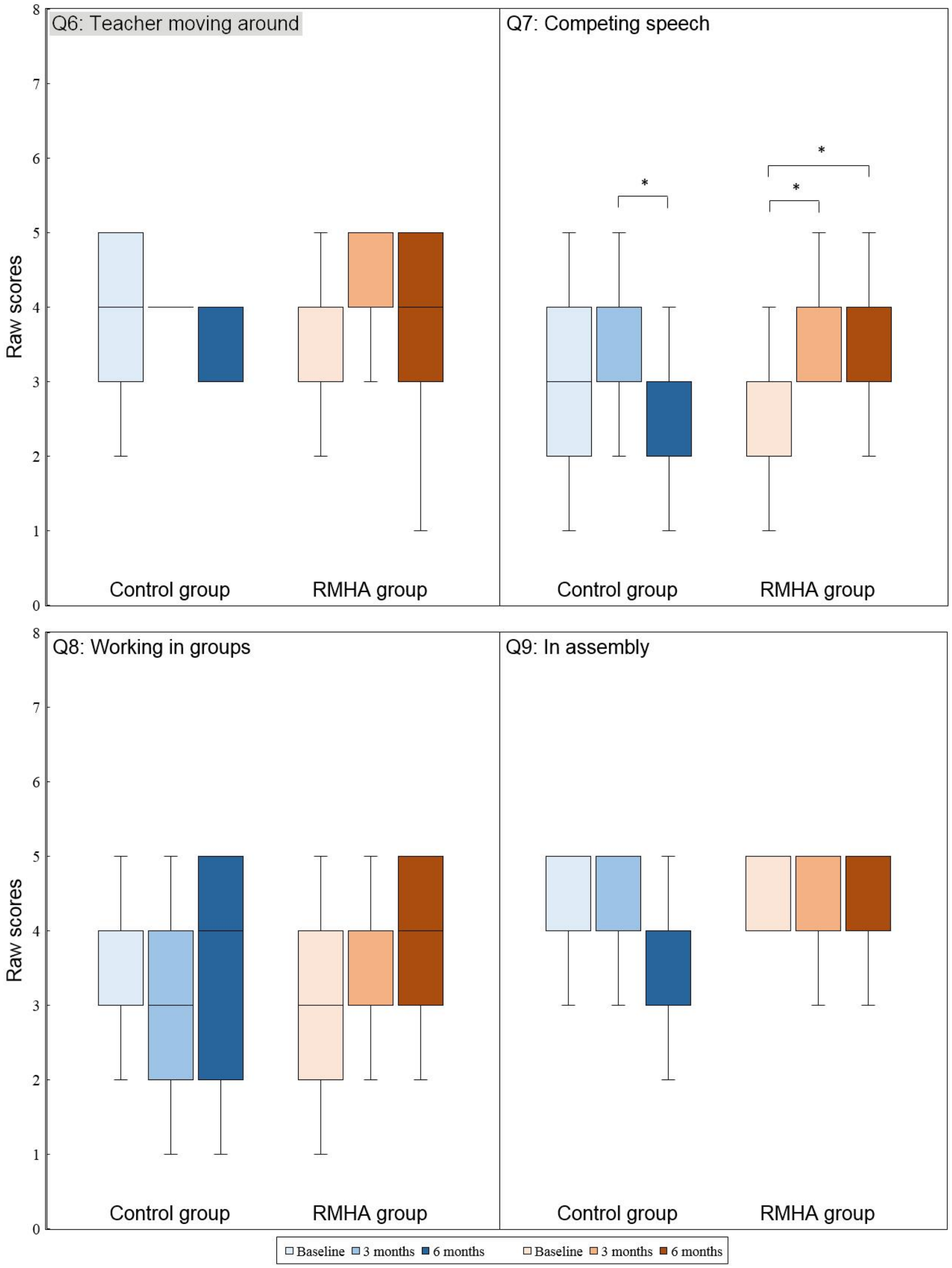

Figure 3.11 – Results of the LIFE-R Q6-9
Boxplots of raw scores of questions 6-9 of the LIFE-R questionnaire, for 13 controls and 13 RMHA children, at baseline, 3 months and at 6 months. Question 6 highlighted in grey did not meet the assumption of absence of outliers (see text for details). The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. This was a non-standardised questionnaire, therefore there was no indication of a cut-off for abnormal performance. LIFE-R: Listening Inventory for Education – Revised, Q: Question, RMHA: Remote Microphone Hearing Aid. *$p < .05$

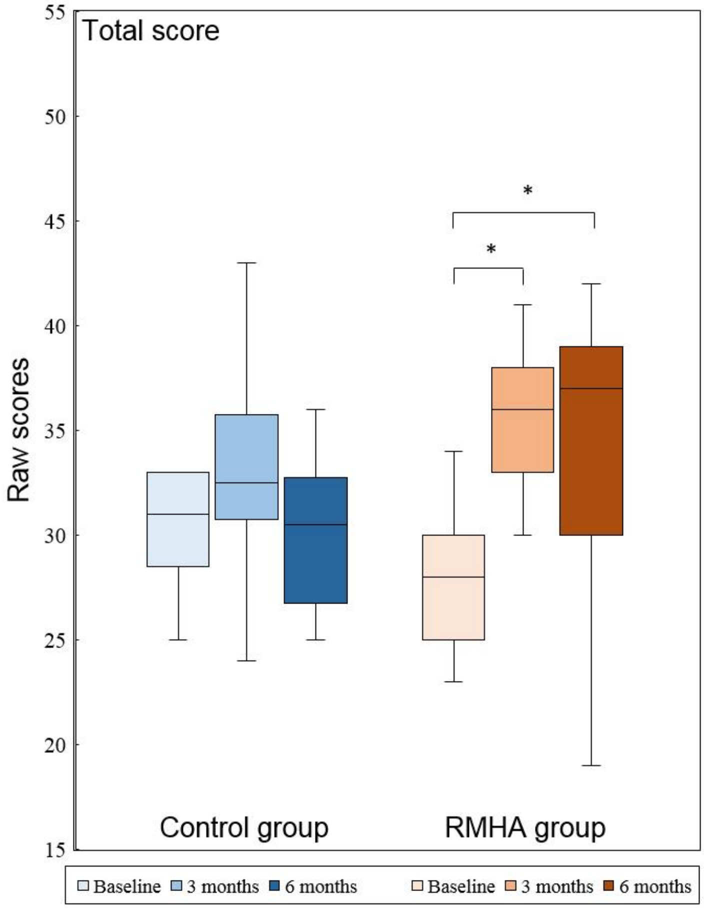

Figure 3.12 – Results of the LIFE-R Total score
Boxplots of raw scores of the LIFE-R questionnaire Total score, for 12 controls and 13 RMHA children, at baseline, 3 months and at 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. This was a non-standardised questionnaire, therefore there was no indication of a cut-off for abnormal performance. LIFE-R: Listening Inventory for Education – Revised, Q: Question, RMHA: Remote Microphone Hearing Aid. **p* < .05

Table 3.14 presents the mean values of the control and intervention group at baseline, at 3 and 6 months.

Table 3.14 – LIFE-R mean scores and SDs
The mean scores and SDs for the LIFE-R questions. This was a non-standardised questionnaire, therefore there was no indication of a cut-off for abnormal performance. LIFE-R: Listening Inventory For Education - Revised, RMHA: Remote Microphone hearing aid, SD: Standard Deviation.

LIFE-R	**Mean (SD)**					
	Control group			**RMHA group**		
	Baseline	**3 months**	**6 months**	**Baseline**	**3 months**	**6 months**
1: Traffic noise	3.15 (1.06)	3.53 (1.26)	3.30 (1.10)	2.61 (1.04)	4.15 (.37)	3.76 1.23)
2: Quiet	4.76 (.59)	4.46 (1.12)	4.69 (.63)	4.53 (.66)	4.92 (.27)	4.69 (.63)
3: Teacher's back turned	3.07 (1.18)	3.61 (1.19)	3.46 (.51)	3.15 (.89)	4.07 (.86)	3.76 (1.01)
4: Noise outside classroom	2.84 (.89)	2.84 (.80)	2.84 (.80)	2.61 (1.12)	3.53 (.66)	3.30 (1.18)
5: Noise inside classroom	2.23 (1.01)	3.07 (1.32)	2.61 (1.12)	2.46 (1.71)	3.15 (1.21)	2.69 (1.37)
6: Teacher moving around	3.84 (.98)	3.92 (.95)	3.46 (.87)	3.38 (1.12)	4.15 (.80)	3.84 (1.14)
7: Competing speech	2.76 (1.23)	3.15 (1.28)	2.53 (1.19)	2.61 (1.19)	3.69 (1.03)	3.46 (1.33)
8: Working in groups	3.46 (1.12)	3.07 (1.11)	3.53 (1.45)	2.92 (1.44)	3.61 (1.04)	3.76 (1.23)
9: In assembly	4.38 (.86)	4.07 (1.18)	3.76 (.92)	4.00 (1.41)	4.30 (1.25)	4.00 (1.29)
Total score	30.91 (4.46)	33.33 (5.59)	30.83 (5.06)	28.30 (7.44)	35.61 (3.45)	33.30 (7.66)

Table 3.15 below summarises the results from the analysis of the eight remaining LIFE-R questions and the total score.

Table 3.15 – LIFE-R significance

The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time, of the simple main effect of group and simple main effect of time, and of the main effect of group and main effect of time of the LIFE-R questions. Highlighted in grey are the conditions that did not meet at least one of the assumptions (see text above for details). CI: Confidence Interval, LIFE-R: Listening Inventory for Education - Revised, RMHA: Remote Microphone Hearing Aid. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

LIFE-R	**Interaction**		**Simple main effect**				**Main effect**			
	Group* Time	η_p^2 **[95% CI of** η_p^2**]**	**Group**	η_p^2 **[95% CI of** η_p^2**]**	**Time**	η_p^2 **[95% CI of** η_p^2**]/ *MD* [95% CI of *MD*]**	**Group**	η_p^2 **[95% CI of** η_p^2**]**	**Time**	η_p^2 **[95% CI of** η_p^2**]/ *MD* [95% CI of *MD*]**
1: Traffic noise	.022*	.147 [.001, .308]	.206 (**baseline**) .106 (**3 months**) .326 (**6 months**)	.066 [.000, .292] .105 [.000, .342] .040 [.000, .253]	.383 (**Controls**) .000*** (**RMHA**) .002** (**0-3 months**) .006** (**0-6 months**) 1.000 (**3-6 months**)	.077 [.000, .270] .481 [.145, .642] -1.538 [-2.462, -.615] -1.154 [-1.977, -.330] .385 [-.684, 1.454]	-	-	-	-
1: Traffic noise - adjusted	.016*	.218 [.006, .454]	.206 (**baseline**) .326 (**6 months**)	.066 [.000, .292] .040 [.000, .253]	.549 (**Controls**) .002** (**RMHA**) .002** (**0-3 months**)	.031 [.000, .308] .558 [.115, .741] -1.154 [-1.799, -.508]	-	-	-	-
2: Quiet	-	-	-	-	-	-	-	-	-	-
3: Teacher's back turned	.713	.014 [.000, .101]	-	-	-	-	.305	.044 [.000, .259]	.010* .034* (**0-3 months**) .054† (**0-6 months**) 1.000 (**3-6 months**)	.175 [.011, .338] -.731 [-1.415, -.046] -.500 [-1.006, .006] .231 [-.374, .836]
4: Noise outside classroom	.102	.091 [.000, .240]	-	-	-	-	.252	.054 [.000, .276]	.102	.091 [.000, .240]

5: Noise inside classroom	.960	.002 [.000, .033]	-	-	-	-	.730	.005 [.000, .160]	.055†	.114 [.000, .269]
6: Teacher moving around	.148	.077	-	-	-	-	.860	.001 [.000, .050]	.128	.082 [.000, .228]
7: Competing speech	.041*	.125 [.000, .283]	.749 (**baseline**) .249 (**3 months**) .075† (**6 months**)	.004 [.000, .146] .055 [.000, .277] .126 [.000, .365]	.040* (**Controls**) .524 (**0-3 months**) 1.000 (**0-6 months**) .015* (**3-6 months**) .013* (**RMHA**) .047* (**0-3 months**) .043* (**0-6 months**) 1.000 (**3-6 months**)	.236 [.000, .446] -.385 [-1.125, .356] .231 [-.411, .872] .615 [.114, 1.117] .306 [.017, .507] -1.077 [-2.142, -.011] -.846 [-1.670, -.023] .231 [-.772, 1.234]	-	-	-	-
8: Working in groups	.218	.061 [.000, .198]	-	-	-	-	.817	.002 [.000, .083]	.332	.045 [.000, .171]
9: In assembly	.464	.028 [.000, .138]	-	-	-	-	.933	.000 [.000, .011]	.464	.028 [.000, .138]
Total score	.042*	.129 [.000, .290]	.304 (**baseline**) .228 (**3 months**) .355 (**6 month**)	.046 [.000, .267] .062 [.000, .292] .037 [.000, .253]	.093† (**Controls**) .002** (**RMHA**) .014* (**0-3 months**) .016* (**0-6 months**) .721 (**3-6 months**)	.217 [.000, .435] .408 [.081, .588] -7.308 [-13.182, -1.434] -5.000 [-9.079, -.921] 2.308 [-2.886, 7.502]	-	-	-	-

Traffic noise had a significant interaction between group and time, *F*(2, 48) = 2.55, $p = .022$, $\eta_p^2 = .147$. This interaction was followed up by examining the simple main effects of group and time. There was no statistically significant difference in scores between the two groups in any of the three time points, but there was a statistically significant simple main effect of time on Traffic noise scores in the RMHA group, *F*(2, 24) = 11.11, $p < .001$, $\eta_p^2 = .481$. To determine the time-point of this difference, pairwise comparisons were run. This showed that for the RMHA group, Traffic noise scores were significantly better from baseline to 3 months, *M* = -1.54, *SE* = .332, $p = .002$ and significantly better from baseline to 6 months, *M* = -1.15, *SE* = .296, $p = .006$. As the assumption of homogeneity of variances for the analysis of the Traffic noise scores was violated, and because of the lack of non-parametric mixed ANOVA equivalents available in SPSS, an additional analysis was done post-hoc. The 3-month time-point was the one that violated the assumption and was thus decided to run a post-hoc analysis by removing this time-point and comparing the scores between the two groups at baseline and at 6 months. An additional row has been added right under the Traffic noise scores in Table 3.15 called Traffic noise – adjusted, which presents the findings. In this way, direct comparisons can be made with the initial analysis that had violated the assumption of homogeneity of variances. In detail, there was a significant interaction between group and time, *F*(1,24) = 6.67, $p = .016$, $\eta_p^2 = .218$. This interaction was followed up by examining the simple main effects of group and time. There was no statistically significant difference in scores between the two groups in any of the two time points (as expected since the same univariate comparisons were run), but there was a statistically significant simple main effect of time in the scores of the RMHA group only, *F*(1, 12) = 15.16, $p = .002$, $\eta_p^2 = .558$. The scores of the RMHA group were significantly better from baseline to 6 months, *M* = -1.54, *SE* = .296, $p = .002$. Therefore, the additional analysis produced similar results to the initial analysis and retained the significant improvement from baseline to 6 months even when removing the time-point that violated the assumption of homogeneity of variances.

Competing speech also had a significant interaction between group and time, *F*(2, 48) = 1.94, $p = .041$, $\eta_p^2 = .125$. This interaction was followed up by examining the simple main effects of group and time. There was no statistically significant difference in scores between the two groups in any of the three time points, but there was a statistically significant effect of time in Competing speech scores in the control group, *F*(2, 24) = 1.26, *p* = .04, η_p^2 = .236, and the RMHA group, *F*(2, 24) = 4.18, *p* = .013, η_p^2 = .306. To determine the time-point of these differences, pairwise comparisons were run. This showed that, first, for the control group Competing speech scores significantly worsened from 3 to 6 months, *M* = .62, *SE* = .180, *p* = .015, whereas for the RMHA group, Competing speech scores significantly improved from baseline to 3 months, *M* = -1.08, *SE* = .383, *p* = .047 and significantly improved, as well, from baseline to 6 months, *M* = -.85, *SE* = .296, *p* = .043.

Furthermore, the Total score had a significant interaction between group and time, *F*(2, 46) = 51.79, *p* = .042, η_p^2 = .129. This interaction was followed up by examining the simple main effects of group and time. There was no statistically significant difference in scores between the two groups in any of the three time points, but there was a statistically significant effect of time in the Total scores for the RMHA group, *F*(2, 24) = 181.41, *p* = .002, η_p^2 = .408. To determine the time-point of these differences, pairwise comparisons were run. This showed that in the RMHA group, total scores significantly improved from baseline to 3 months, *M* = -7.31, *SE* = 2.113, *p* = .014 and significantly improved, as well, from baseline to 6 months, *M* = -5.00, *SE* = 1.468, *p* = .016.

Finally, the main effect of time in the Teacher's back turned question was statistically significant, *F*(2, 48) = 3.63, *p* = .01, η_p^2 = .175. This means that if the two groups are collapsed together, a significant difference over time is observed. In other words, the scores at 3 months were better than at baseline, regardless of the group variable (i.e. control and intervention group merged together) as evidenced by the pairwise comparisons, *M* = -.73, *SE* = .266, *p* = .034.

3.3.7 RMHA total time used

The RMHAs were used on average for 2.3 hours per day. However, 4 children used the RMHA for less than 1 hour per day. Despite these 4 cases with low RMHA use rates, it was decided that they would remain in the analysis as the design used an ITT analysis. Implications of this are outlined in the Discussion below. It should also be noted that a separate exploratory analysis was run by excluding these cases without generating different results and was thus decided that the initial larger sample would be analysed and reported.

3.4 Discussion

3.4.1 Attention

Sustained auditory attention

The subtest measuring Sus-AA remained unchanged over the intervention period, meaning that after 6 months of RMHA use in the classroom Sus-AA did not improve. Since the tests were administered without the assistance of RMHAs, the lack of improvement signifies that there is no lasting effect on auditory attention span, as hypothesised in Section 1.8.2. These findings though, should be interpreted with caution, since the Sus-AA scores did not meet the assumption of homogeneity of variances and transformations were unable to produce scores that would meet the rest of the assumptions. Violation of this assumption increases the probability of making a Type I error, but when samples sizes are equal and normally distributed (the case for Sus-AA here) this probability is decreased (Box, 1953; Nordstokke & Zumbo, 2007). If Sus-AA scores would have shown improvement, then this improvement could have had an increased chance of being a false positive. Nonetheless, this was not the case here as Sus-AA scores did not become better over time. The hypothesis for Sus-AA was based on previous findings from the CHAPS questionnaire that showed that auditory attention span improved over a 6-month hearing aid with directional microphone intervention period in children with APD (Kuk et al., 2008). As noted earlier, the CHAPS questionnaire is not standardised, hence the improvement in the auditory attention span CHAPS scores in Kuk et al.'s (2008) study might have been the product of maturation and not because of the intervention. This means that the hypothesis set in the study could have been falsely based on these findings. It could also be that parents in the intervention group might tend to view (and rate) their children’s attention span as improved, simply because they are using an intervention and not because of genuine improvement. This last point ties in with the placebo or

Hawthorne effect that is discussed in a separate Section later in this Discussion (Section 3.4.4 under Design limitations).

Assuming that the hypothesis is valid, the non-significant results could be explained by the mismatch between dynamic real-life sustained attention situations and monotonous lab-based sustained attention tasks (Head & Helton, 2015). The monotonous Sus-AA task used in the study is different than the other sustained attention situations children face in the classroom. Therefore, one can argue that even if RMHAs might have assisted children to improve their Sus-AA, this was not reflected on the task administered in the lab, as that was a monotonous string of tones that did not match what children were used to sustain their attention on via RMHA use in the classroom. Further work should investigate differential Sus-AA tasks with inclusion of dynamic linguistic stimuli in RMHA intervention trials. Another reason for the lack of significant findings could be that improvement in cognitive test scores might require longer intervention periods for differences to be significant. This point was discussed in more detail in the previous study in Chapter 2, Section 2.4.5. In Purdy et al.'s (2009) study, children with reading delays using RMHAs for 6 weeks showed improvements only in the questionnaires but not in the behavioural tests. As in the current study there were no significant improvements neither in the Sus-AA behavioural test nor in the corresponding CHAPS questionnaire subscale, it can be argued that RMHA intervention may not bring lasting changes in Sus-AA over 6-months of RMHA use.

Divided auditory-visual attention

The scores from the Div-AVA task showed significant improvement from baseline to 6 months only in the intervention group. However, there were also significant differences between the intervention and control group at baseline (see Table 3.6 and Figure 3.3). Children in the intervention group started with significantly worse scores at baseline in comparison to children in the control group. After 3 and 6 months these lower scores in the intervention group improved. This improvement is expected to be an effect of regression to the mean, rather than genuine improvement due to the use of the intervention. This observed significant effect may alternatively be a product of a Hawthorne effect or a placebo effect (see details on this in Section 3.4.4 under

Design limitations). For this reason, a post-hoc analysis was run controlling for the Div-AVA baseline scores and then comparing the difference from baseline to 6 months in the Div-AVA scores between the two groups (see end of Section 3.3.2 for details). This analysis did not reveal any statistically significant differences between the two groups. This means that after controlling for the unbalanced baseline scores between the two groups there were no longer any significant improvement in scores from baseline to 6 months in the intervention group. Therefore, it can be concluded that the use of RMHAs did not provide any long-term benefits to the intervention group over the control group and any initial improvements in Div-AVA scores found prior to controlling for baseline can be attributed to simply a regression to the mean.

Control conditions

Selective visual attention and Div-AA remained unchanged, as initially hypothesised. These two conditions were set as control conditions, since the RMHA was not expected to benefit these two types of attention. Given the limitations in sample size and thus statistical power of the study, it is likely that the lack of significant findings was influenced by the low power.

3.4.2 Listening-in-noise

Speech-in-noise

The two LiSN-S SIN conditions (i.e. Low-cue SRT and Talker advantage) did not show significant improvement over time, contrary to the hypotheses. It was expected that the long-term use of RMHAs by children with APD would bring lasting improvements in their SIN ability, as detailed in Section 1.8.1. These non-significant findings may have been influenced by the moderate test-retest reliability of these conditions (.60 and .30, respectively; Cameron & Dillon, 2007b). However, findings thus far suggest that SIN conditions remain unaffected following RMHA interventions. The study from Chapter 2 using the SCAN-3 C test did not find improvements after a 3-month RMHA trial, while Sharma et al.'s (2012) study employing the HINT test did not show significant improvement in the experimental groups using combined RMHA and AT interventions after 6-weeks of use. But as the previous study in Chapter 2 had other limitations and Sharma et al.'s (2012) intervention groups did not have a RMHA-only group, further investigation is needed to clarify whether RMHA could improve scores in SIN tests after long-term use.

It can also be argued that the hypothesis that long-term RMHA use will increase scores in SIN tests unaided should be re-evaluated. This hypothesis was drawn based on evidence from previous studies linking SIN and selective attention (Mesgarani & Chang, 2012; O'Sullivan et al., 2015; Zion Golumbic et al., 2013) and from studies showing that long-term interventions can have a positive influence on SIN ability (Alain et al., 2007; Song et al., 2012). I therefore hypothesised that RMHAs could have similar effects on SIN, as the improved SNR can help separate target from distractors and long-term use could train the auditory system to perform this separation even without the assistance of RMHAs. Drawing from the study's findings though, it could be argued that RMHAs might not train the auditory system the same way that the interventions from other studies do. This may be because RMHAs are considered a passive intervention (Fey et al., 2011). This argument requires further investigation using electrophysiological studies to examine whether RMHAs activate the same pathways that AT interventions do. A study by Song, Skoe, Banai, and Kraus (2011) found a link between poor behavioural SIN test scores and poor encoding of the fundamental frequency of target speech syllables. This and other paradigms could serve as a basis for future research to build on. For instance, monitoring of SIN tests through behavioural tests can be coupled with such EEG paradigms.

Spatial listening

The initial hypothesis was that spatial listening (measured through the LiSN-S Spatial advantage and High-cue SRT/ Total advantage conditions) would not show significant improvements over time. This hypothesis was confirmed as none of these three LiSN-S conditions was significantly different at 3 and 6 months compared to baseline scores. This could be due to the way RMHAs work; that is, they do not spatially separate target from distractors in the child's ears, as the target voice comes straight into the ears independent of the teacher's position in the classroom (thus stripped of spatial cues). Nevertheless, since the mean scores of the spatial listening conditions in both the control and intervention group were not below the cut-off for disordered performance (see Figure 3.7/ Table 3.8), it could be that RMHAs simply could not have provided any additional benefit to the already within-normal-limits scores. In addition, there were only three children in the RMHA group that had disordered performance in at least one of the LiSN-S spatial listening conditions at baseline, therefore further post-hoc analyses (e.g. correlations between poor and non-poor performers) would not provide any meaningful results due to the small sample size. Even though the findings suggest that spatial listening did not improve (or hindered) by RMHA use, the study had a low power and it could be that significant effects (either positive or negative) have gone undetected because of that (see Section 3.4.4, under Sample size for more details on this).

A recent study used RMHAs on 4 stroke patients for 10 weeks (Koohi, Vickers, Warren, Werring, & Bamiou, 2017). They were tested aided and unaided both at baseline and at 10 weeks with the results showing significant improvement in a spatially separated sentences-in-noise test. Improvement was recorded only in the intervention group, while a group of 5 control stroke patients did not record change in their scores. These improvements were observed both with and without the use of the RMHA, indicating that long-term use of the RMHA system might assist spatial listening. These findings contradict the study’s hypothesis for the LiSN-S spatial conditions. This could be because of the different population studied in Koohi et al.'s (2017) trial compared to the present study (i.e. adult stroke patients vs. children with APD), despite the similar symptoms both samples exhibited (i.e. SIN difficulties in the presence of normal hearing). In any case, the authors attributed the improvement in the unaided condition to possible auditory plasticity induced by the use of RMHAs. However, there are a few methodological shortcomings in the study that should be taken into account. Even though Koohi et al.'s (2017) sample was bootstrapped to combat the small number of patients, the sample size was still small. In addition, the authors did not perform randomisation to divide the groups, which introduced selection bias. These results should be further validated by a larger scale intervention trial (Koohi et al., 2017).

LiSN-S accent effect

Previous studies showed that the North-American SIN SCAN-3 C test used on UK children may have caused children to underperform compared to the North-American norms and this was attributed to differences in the accents between the test and the subjects (Dawes & Bishop, 2007; Marriage et al., 2001). As noted earlier in Methodology, the LiSN-S used in the current study was the North-American version of the test using the corresponding North-American normative data (due to lack of a UK LiSN-S version and lack of UK norms). Accent effects may negatively influence children's performance in the test. A study still in preparation compared the performance of 48 UK children, aged 6-10 (tested under the North-American version), to the performance of the 72 North-American norms, also aged 6-10 (Murphy & Bamiou, 2018). There were no significant differences between the two samples in any of the LiSN-S conditions except for the Talker advantage condition where there was a difference of .035 in z scores (i.e. UK children performed worse than North-American norms). The researchers proposed that adding .035 to the z scores obtained from UK children in the Talker advantage condition only would adjust the scores and neutralise any accent effects. The present study adopted this suggestion and adjusted the values for the Talker advantage scores only. It is thus expected that despite using the LiSN-S North-American version on UK children, effects produced from the different accents between test and subjects would have been minimised. There is, however, the need for development of a British version of the LiSN-S (i.e. recorded in British accent) and normative data collected from the UK.

3.4.3 Questionnaires

CHAPS

The experimental conditions of the CHAPS (i.e. Noise, Multiple inputs and Auditory Attention span) remained unchanged, despite the initial hypotheses of improvement over the intervention period. This contrasts a finding from the previous study in Chapter 2 that showed significant improvement in scores in the Auditory attention span in the control children from baseline to 3 months. But both studies in Chapter 2 and 3 were of low power which may have influenced the results. For instance, studies with low statistical power could miss significant effects or they could inflate non-significant effects (see more details in Section 3.4.4 under Design limitations). In addition, the Noise subscale scores did not meet the assumption of homogeneity of variances. Violation of this assumption increases the probability of making a Type I error, but when samples sizes are equal and normally distributed (the case for CHAPS Noise here) this probability is decreased (Box, 1953; Nordstokke & Zumbo, 2007). If results in the Noise subscale would have shown improvement then interpretation of the findings should have been cautiously made, since an observed positive effect would have had an increased chance of being a false positive. Nonetheless, this was not the case here as the Noise subscale did not present any significant improvement over time.

As discussed earlier, the CHAPS presents with validity issues (Lam & Sanchez, 2007; Sharma et al., 2009; Wilson et al., 2011), meaning that it might not be the best tool for measuring these aspects of children’s development (i.e. SIN, attention). Even in studies where the CHAPS was found to distinguish children with APD from controls, other shortcomings were identified that make the questionnaire an unreliable tool to use (Iliadou & Bamiou, 2012). Specifically, in Iliadou and Bamiou's (2012) study, parents of children with APD gave significantly poorer scores in all conditions compared to parents from a typically developing control group. Nevertheless, parents from another group of children suspected of but not diagnosed with APD also had significantly poorer scores compared to controls in the Noise, Multiple inputs and Auditory attention span subscales. At the same time, the APD group did not differ in the Noise and Multiple inputs subscales from the suspected group (Iliadou & Bamiou, 2012). This means that this tool cannot differentiate between clinically diagnosed and non-diagnosed children, even though it is often used as an APD screening tool. Adding to that, the questionnaire is using raw instead of standardised scores and its test-retest reliability is not validated (Smoski et al., 1998).

Despite the shortcomings of the CHAPS it was used in the study as it was a concise questionnaire that matched all behavioural tests used in the design (i.e. SIN ability, memory, attention). There is no other single questionnaire that was able to examine those subscales together and that was considered an advantage of the CHAPS, especially given that parents already were asked to complete a longer questionnaire (the CCC-2). Moreover, while the abovementioned studies found it to be inconsistent in distinguishing between children with APD and children who were suspected of APD but did not meet APD criteria it has not been used in a similar RMHA randomised controlled design previously. Thus, despite it being a non-normalised tool there was still no information on how it performed in long-term controlled intervention trials. Adding to that, as the current study was of low power any results from all outcome measures cannot be interpreted properly, including the results obtained from the CHAPS. Other possible questionnaires with improved validity (discussed later in the final chapter) may be used in future APD intervention trials, however they need to be validated in children populations as they currently only have been tested on adults. To sum up, considering the shortcomings it can be argued that the CHAPS is not a tool that should be used for screening APD in children and until its test reliability is examined and its data normalised, other questionnaires that measure auditory performance and cognitive factors should be used in APD studies. These questionnaires are presented later in Discussion in Chapter 6.

CCC-2

The composite score of the four standard language conditions of the CCC-2 (i.e. Speech, Syntax, Semantics and Coherence) remained unchanged, which was contrary to the hypothesis. A probable interpretation of the findings would be that the low statistical power (as also discussed earlier) may have had an impact on the results. Positive effects may have gone undetected because of the reduced statistical power and thus it is difficult to draw conclusions from these results. It would be interesting if the CCC-2 was completed by teachers, as well as parents, so that direct comparisons could be drawn between their scores. As teachers view children in the classroom while they use the hearing aids, it could be that they may notice improvement in these language aspects, whereas parents who see their children at home might not. The other finding of non-significant change in the composite score of the four non-standard language CCC-2 conditions is in line both with the results from the previous study (see Chapter 2) and with Sharma et al.'s (2012) findings. In that study, they used a behavioural test (the Comprehensive assessment of spoken language) measuring non-standard language aspects (e.g. Sarcasm, Figurative language, Use of context and Drawing inferences from given information). The scores of children with APD in the intervention groups did not show significant change over a 6-week period. The study groups though, were using both AT and RMHAs, making it difficult to attribute the lack of change in test scores to the use of RMHAs alone. Additionally, the behavioural test used in that study was not a standardised test. But as discussed earlier, it is not expected that the clearer signal produced by the system would have an impact on pragmatic language aspects, such as sarcasm, use of context or nonverbal communication. The low statistical power of the study may also explain the lack of significant findings as already discussed in the previous sections.

LIFE-R

The LIFE-R questionnaire, completed by children, recorded significant improvements in the RMHA group in 2 of the 9 experimental questions and in the Total score. In detail, Question 1 which asked children how well they hear their teacher’s words when there is traffic noise, had significant improvements from baseline to 3 months and from baseline to 6 months in the RMHA group only. This subscale did not meet the assumption of homogeneity of variance, hence no further interpretation of this finding can be given as violation of this assumption increases the chance of making a Type I error (which is finding a significant effect where no such effect truly exists). However, a post-hoc analysis removing the 3-month time-point which was violating this assumption showed that the significant improvement from baseline to 6 months only in the RMHA group was still retained. Question 7, asking children how well they hear their teacher through the hearing aids when another person is simultaneously talking, had improved scores in the RMHA group from baseline to 3 and 6 months. In this question, the control group exhibited significantly worse scores from 3 months to 6 months. Moreover, the Total score of the questionnaire (sum of all scores) showed similar significantly better scores in the intervention group only when comparing the baseline scores to 3 months and the baseline scores to 6 months. The significant improvements observed in Question 7 (i.e. competing speech) could be attributed to the improved SNR that the system was providing during classroom time.

There are a number of limitations to consider regarding these findings. First, the lack of improvement in the rest of the questions poses problems, as it was expected that the use of the RMHA would help improve the SNR for all listening situations. For instance, Questions 3 and 6 (i.e. measuring how well children can hear their teachers when the teachers have their backs turned and when they are moving around the room, respectively) did not show improvements. A possible explanation for this could be that visual cues are important when it comes to listening (Atilgan et al., 2018), even when using RMHAs, as in these two listening situations visual cues were obstructed by either the movement of the teachers or the fact that they did not face the classroom. This finding comes in contrast with Johnston et al.'s (2009) study where they used the LIFE version of the questionnaire and found significant improvements in children’s listening when the teacher’s back was turned after using the RMHA for 5 months. This could be because the current study was under-powered, therefore it was possible that truly significant effects have gone undetected. Low statistical power of a study can also inflate any significant effects, thus detect effects where no such effects would truly be present. This reason could offer an explanation as to why some questions were found to have significantly improved and others have not.

Question 8 is more difficult to interpret as systematic use of RMHAs was not ensured in this listening condition. When children were working in groups (i.e. Question 8) they were not always using the system because the teacher was in close distance when addressing the group, and the RMHA worked optimally when there was some distance between the mic and the hearing aids. When children were listening in assembly (i.e. Question 9), they were asked not to use the RMHA, as that would cause more problems than benefits. This is because in assembly there are different speakers that talk in a short amount of time and that would require frequent changing of the mic, which would introduce noise in the hearing aids and cause unnecessary delays. During completion of the questionnaire, children were asked to respond to these two questions (i.e. 8 and 9) as they normally hear without the use of RMHAs. Having that in mind, it can be argued that RMHA use for 6 months did not bring lasting effects in their listening-in-noise ability, specifically in these two situations (when working in groups and when in assembly). This is in line with the lack of significant findings shown earlier in the SIN conditions of the LiSN-S.

Another limitation is that the LIFE-R questionnaire did not have standardised scores, making it difficult to exclude maturation as an explanation of the improved scores. This can be counter-argued by the lack of improved scores in the controls, as maturation would have been expected to be reflected on their scores, too. At the same time, the fact that in Question 7 there was significant improvement in scores in the intervention group while the control group had significantly worse scores further supports the view that scores were not affected by maturation. Moreover, children in the intervention group might have been inclined to rate better scores after using the system because they could have felt that this is what they were supposed to do, since they were specifically asked to rate how well they listen when using the system. Even though children were allowed to complete the questionnaires on their own, so that they did not feel the pressure to respond in a specific way, and to make them feel comfortable, children would still be aware that the researcher was going to review their answers. This would be comparable to a Hawthorne effect, while as a placebo group was not used in the study these improvements may have also been the product of a placebo effect. Placebo and Hawthorne effect are discussed in more detail in the upcoming Section 3.4.4 under Design limitations.

3.4.4 Other limitations and future research

Other limitations, apart from the ones discussed in each section previously, could have influenced the overall findings of the study. These are outlined in this section.

Sample size

A post-hoc power calculation was performed on the actual sample size (n = 26) to detect 1 SD effect size at 5% significance, which resulted in a statistical power of 70%. This means that the study was under-powered and the chance of detecting a statistically significant difference (if a difference exists) was down to 70%. Another consequence of low power is the imprecision of the effect size (Nakagawa & Cuthill, 2007) which can be inflated, thus showing a greater magnitude than the actual effect size of the population. This may have been the case for the significant findings of this study (i.e. LIFE-R). Additionally, a low powered study also increases the probability of getting a false negative effect. This means that true significant effects in the population have an increased chance of showing as non-significant in low powered studies. Therefore, interpretation of the results cannot be safely made, and a larger sample sized study is required to make sense of the findings.

Duration of the study

This argument has also been discussed earlier in Chapter 2 (see Section 2.4.5) and will be mentioned in brief here. While it was considered that the 6-month intervention period would have been sufficient to yield significant improvements in the behavioural outcome measures, this was not the case based on the majority of findings. Significant improvements were only recorded in some subscales of the LIFE-R questionnaire but the rest of the outcome measures did not change. Even though it can be argued that RMHAs might not have an effect on the rest of the measures, the duration of the intervention period is considered as an alternative explanation of the findings. There were studies that considered this as a possible interpretation of their lack of significance in their outcome measures (Purdy et al., 2009; Sharma et al., 2012). In brief, it is argued that significant improvements in behavioural tests take longer to surface compared to improvements in questionnaires (Purdy et al., 2009). Nonetheless, in the current study improvements in questionnaire subscales (i.e. attention and SIN) did not significantly improve either. Therefore, while it would not be unreasonable to argue that a longer trial period (e.g. 12 months in total) might bring improved behavioural scores in attention and SIN, other possibilities should be considered for this lack of improvement. These possibilities are discussed in the next three sections.

Willingness to use RMHA and total time used

While all children and their parents were on board with the trial and understood what it involved, there were some children who seemed uncomfortable or not as willing to use the RMHA system as others. There were 5 out of the 13 children, that were observed during the school visit or during the follow-up tests, who appeared reluctant to use the device. In two of these cases parents mentioned that their child was feeling uncomfortable the day they were going to receive the system at school. This reluctance to use the system may have had a negative impact on these children’s scores, compromising possible benefits of RMHAs. These cases were not removed from the study, as an ITT analysis was followed. Of those 5 children, none of them was younger than 9 years 8 months. At the same time, all children younger than 9 years 8 months in the intervention group (*n* = 6) were all enthusiastic, smiling and were looking forward to using the system during the school visit of the researcher. It appears that age might play a role in how willing children are to take on a hearing aid intervention. Perhaps older children begin to feel more aware about how they might be perceived by their peers, while for younger children this is not a factor. There are some studies that show that in older adult populations stigmatisation is a barrier for hearing aid uptake (Jenstad & Moon, 2011). It would be interesting to further investigate this in samples of children to check whether findings are in line with the observations that older children are more reluctant to use hearing aid interventions than younger children.

This reluctance by some children to use the RMHA is reflected in the daily average time the RMHA was used (see Table 3.2). Three of the four children who used the RMHA for less than a daily average of 1 hour were part of the group that were sceptical to use the system. The other child with the same low daily usage time stated that the teacher was inconsistent in using the system. This is not uncommon as another study looking at the use of hearing aids in children with hearing loss noted inconsistent use of the prescribed amplification in more than a third of their sample, but the researchers did not discuss the reasons behind this (Fitzpatrick, Durieux-Smith, & Whittingham, 2010). Previous RMHA studies that included children with APD did not mention the total time the system was used in their studies, therefore it is unclear whether they accounted for this factor (Johnston et al., 2009; Sharma et al., 2012; Umat et al., 2011). Future RMHA studies should employ larger sample sizes where they will monitor the time the system is used throughout the study with frequent teacher communication to ensure they also fulfil their part, as that could potentially improve scores in the outcomes measures.

Randomisation

Even though randomisation based on age and gender was considered the best route for dividing the sample into groups, other factors could be selected to perform randomisation. This is said as the Div-AVA baseline scores between children in the control and RMHA group significantly differed. The small sample size or other differences in characteristics that were not accounted for could perhaps explain this significant difference in baseline scores in Div-AVA between the two groups. Since there were no other outcome measures with significant baseline differences between groups, it could be that this had something to do with the specific Div-AVA test, or with Div-AVA per se in this specific sample (as in the previous study in Chapter 2 no such differences in the Div-AVA baseline scores were observed). Future studies could potentially explore this further by obtaining baseline scores in an easier Div-AVA task, which would then be used as a controlled factor in the analysis.

Design limitations

A number of features that make RCTs advantageous over non-randomised designs were followed in this study, such as random allocation to intervention, ITT analysis, predefined outcomes (Sibbald & Roland, 1998). However, other features of RCTs were not adopted and this may have influenced the findings. Specifically, one important feature of RCTs, the double blinding of researchers and participants, was not used in this design. The reason for this is that the researcher had to know which children received the RMHA, as he planned school visits to demonstrate to children and teachers how the system worked. In addition, he was the point of contact for parents, teachers and school staff in case any problems with the RMHAs arose during the study. Due to time and budgetary constraints it was not possible to assign another person to have these elevated duties and responsibilities, therefore double blinding was not performed. This means that both the researcher and the children/ parents were aware of the allocation into control or intervention groups. Preconceived views of participants and researchers may systematically bias the results if double blinding is not used (Sibbald & Roland, 1998). It is thus possible that significant improvements observed in this study (i.e. in the LIFE-R questionnaire) may have been influenced by the researcher or the children using the intervention. Furthermore, another important aspect of RCTs is the use of an inert intervention, also known as placebo control (Misra, 2012). The placebo effect is known as a psychosomatic effect that is produced due to study participation and not due to effectiveness of the intervention (Misra, 2012). By including a placebo group in the design this effect is controlled and researchers can be more confident that observed changes are due to the intervention and not due to the placebo response (Misra, 2012; Snowling & Hulme, 2003). Similarly, the Hawthorne effect is when a treated group is showing significant improvements because of the increased attention the treatment group was given over the control (Snowling & Hulme, 2003). In the present study the placebo or Hawthorne effects were not controlled because it was considered that the use of a placebo treatment (i.e. hearing aids that do not provide an improved signal) would not go unnoticed for 6 months. Possible detection of this (i.e. RMHAs that were not providing an improved signal) may have potentially jeopardised the study and possibly result in the termination of it. Retrospectively, an inclusion of a placebo group would have further limited the statistical power of the study, as more participants would have been required and as discussed earlier it was not possible to recruit more children with APD from the clinic at GOSH. Nevertheless, the lack of control of the placebo effect means that the improved results observed in this study might have been the product of this effect rather than the result of the improved SNR brought about by the RMHAs. This limits the study’s interpretation of the findings. Future studies should explore the possibility of including a placebo group in RMHA trials for a short period first (in order to observe how it is received) before engaging participants in long-term trials.

3.5 Conclusion

From the behavioural tests, only the Div-AVA measure demonstrated significant improvements in the RMHA group, but after controlling for the imbalance at baseline scores this effect was no longer found significant. Therefore, none of the behavioural tests of Sus-AA, Div-AVA or SIN showed long-term improvements after a 6-month use of RMHAs. A finding of interest was that although SIN scores did not improve, the spatial listening conditions of the LiSN-S test remained unchanged, as initially hypothesised. This could mean that long-term use of RMHAs (up to 6 months) does not have a negative impact on spatial listening in children with APD. Finally, when using the RMHA in the classroom, children reported significant improvements in a couple of aspects as assessed by the LIFE-R questionnaire (when listening in traffic noise and when there is competing speech) as well as in the Total score. This means that children found the use of RMHAs helpful in some challenging listening situations in the classroom. Since APD diagnostic criteria are different across the globe any interpretation of the findings is limited only to children that met the same APD diagnostic criteria as the criteria used in this study. But given the study's low power and the lack of control of the placebo or Hawthorne effect these significant effects require validation from larger sample size with improved designs before safe conclusions can be drawn.

CHAPTER 4 – ANALYSIS OF COMBINED DATASET FROM CHAPTERS 2 AND 3

4.1 Introduction

4.1.1 Combined data from studies in Chapters 2 and 3

As previously noted, the post-hoc power calculations of the two previous studies in Chapters 2 and 3 revealed that both studies were compromised in terms of statistical power. Specifically, the study in Chapter 2 had an actual statistical power of 49% while the study in Chapter 3 had a power of 70%. In order to overcome this problem and increase statistical power (and thus the validity of the interpretation of the findings), the matching data from both studies were pooled together. This would increase statistical power and help meet the initial power calculation. In the Methodology below more details are given on the resulting sample size, along with an outline of the compromises that had to be made in this study design and in the outcome measures.

4.1.2 Research questions

The research questions remain the same as previously, as well as the literature from which these questions were drawn from. The detailed hypotheses and research questions for this study are outlined below.

1. The group of children with APD who use RMHAs will show improved scores after 3 months compared to the APD control group in:
    a. The Sus-AA and Div-AVA TEACh tasks (unaided).
    b. The Noise, Multiple inputs and Auditory attention span subscales of the CHAPS (completed by parents).
    c. The composite score of the four standard language conditions of the CCC-2 questionnaire (completed by parents; Speech, Syntax, Semantics, Coherence).
2. The group of children with APD who use RMHAs will not show improved scores when compared to controls after 3 months in:
    a. The Sel-VA and Div-AA tasks of the TEACh test (unaided).

b. The Auditory memory sequencing subscale of the CHAPS.

c. The composite score of the four non-standard language subscales of the CCC-2 (Inappropriate initiation, Stereotyped language, Use of context, Nonverbal communication).

4.2 Methodology

As this analysis of the combined datasets from Chapters 2 and 3 uses the tests and procedures from the studies in those chapters, here there will only be a description of the resulting sample size, the overlapping outcome measures that were used from both studies, and some differences between the two combined datasets. The rest of the topics were already covered in Chapters 2 and 3; inclusion criteria, ethical issues (Sections 2.1.1 and 3.1.1), description of the devices used and children's and teachers' involvements (Sections 2.2.2 and 3.2.2), and description of the tests and procedures followed (Sections 2.2.3 and 3.2.3).

4.2.1 Participants

The resulting sample size from pooling together the datasets from the two studies was 43, which met the previous sample size calculation of 32. Children were already randomised into the control and intervention groups in the previous studies. Table 4.1 below summarises the descriptive statistics for the two resulting groups, control and intervention, including information on group sizes, gender, age, IQ and PTA performance.

Table 4.1 – Descriptive statistics for the two groups
Number of participants, age range, mean age, mean IQ, mean PTA for R and L, and SDs for the control and intervention group. IQ: Intelligence Quotient, L: Left, PTA: Pure Tone Audiometry, R: Right, SD: Standard Deviation.

Group	Total number	Gender	Age range in years: months	Mean age (SD)	Mean IQ in standard score (SD)	Mean PTA R in dB (SD)	Mean PTA L in dB (SD)
Control	22	10 girls 12 boys	7:5 to 11:9	9:9 (1:5)	102.59 (11.08)	6.82 (3.82)	3.87 (5.60)
Intervention	21	8 girls 13 boys	7:3 to 11:5	9:7 (1:3)	100.90 (10.42)	6.80 (4.00)	6.84 (4.71)

As children in the first study in Chapter 2 only used the RMHA for 3 months this combined design only examined the hypotheses for this period, despite the fact that children in the second study in Chapter 3 have used the system for 6 months and were also tested at 6 months. In addition, at least two control children in the first study in Chapter 2 used AT during the intervention period (see details in Section 2.4.5 under Auditory training). As an ITT analysis was followed (Gupta, 2011) these cases were kept in the analysis.

4.2.2 RMHA

Even though the two samples in the previous studies did not use the same microphone, performance of the two systems was almost identical (see Sections 2.2.3 under Verification of RMHAs and 3.2.2 under Device for more details on this). The system used in the second study, however, had a greater transmission range compared to the one in the first study (25m over 7m). This may have given an advantage to children in the second trial as they may have received a more consistent signal when teachers were moving around the classroom.

4.2.3 Tests and questionnaires

The tests used in each of the two previous studies were not all the same. In order to pool the two datasets together to increase the sample size, only the common tests and questionnaires in both studies had to be analysed. Therefore, the outcome measures assessed here were the TEACh (Sus-AA, Div-AVA, Div-AA, and Sel-VA), the CHAPS subscales (Noise, Multiple inputs, Auditory memory sequencing, and Auditory attention span), and the two composite CCC-2 scores (standard language composite and non-standard language composite). Descriptions of all these tests and how they were administered are outlined in detail earlier in Sections 2.2.3 and 3.2.3.

4.2.4 Statistical analysis

Data were analysed in SPSS 22 statistics software, using a mixed design ANOVA, with group (control or RMHA) as the between-subjects factor and time (baseline and 3 months) the within-subjects factor. Correction for multiple comparisons was made using Bonferroni correction for the dependent variable (i.e. time) and the interaction between group and time. There was no correction for multiple comparisons for the multiple administered tests in this analysis for reasons explained earlier in Section 2.2.4.

4.3 Results

4.3.1 Attention

There were no outliers in any of the conditions of the TEACh test (examined by looking for values in the studentised residuals beyond +/- 3 SDs). All conditions had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

Figure 4.1 below presents the findings of the two experimental conditions (Sus-AA, Div-AVA) in boxplots. There was a significant effect observed for the RMHA group. This is further described later.

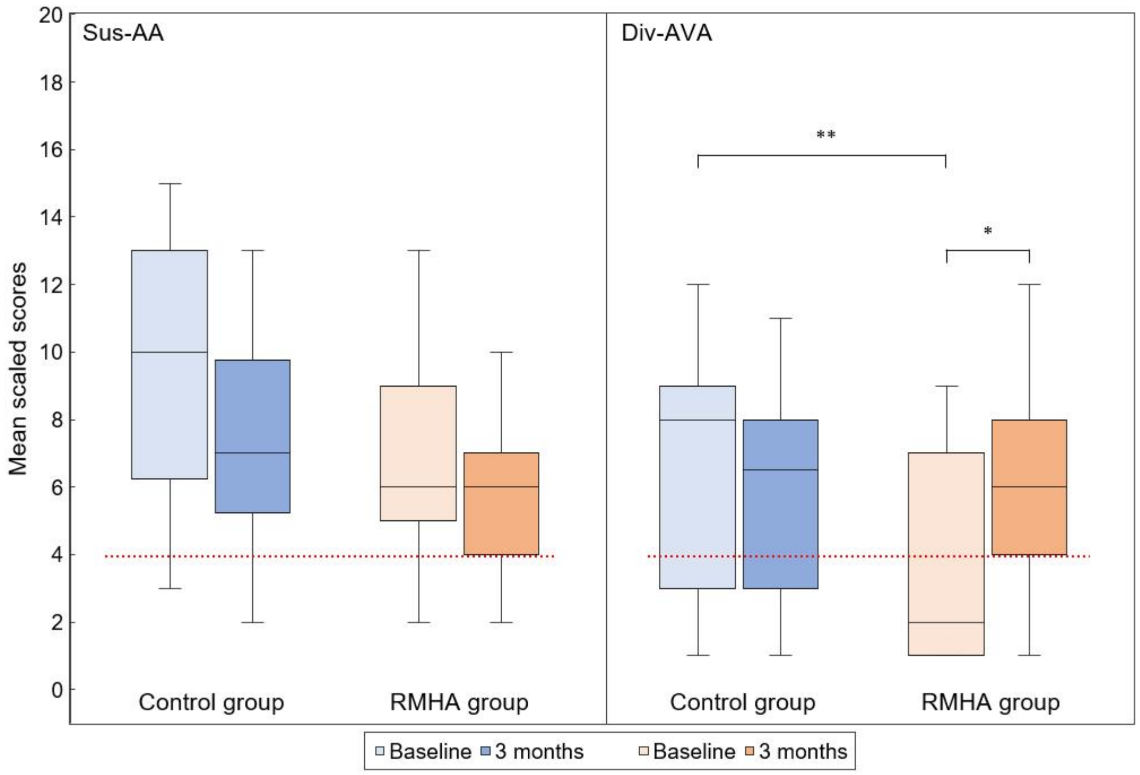

Figure 4.1 – Results of the TEACh experimental conditions
Boxplots of scaled scores of the two experimental TEACh tasks measuring Sus-AA and Div-AVA, for 22 controls and 21 RMHA children, at baseline and 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children. *$p < .05$, **$p < .01$

Figure 4.2 below presents the findings of the two control conditions (Sel-VA, Div-AA). There were no significant findings observed in these conditions.

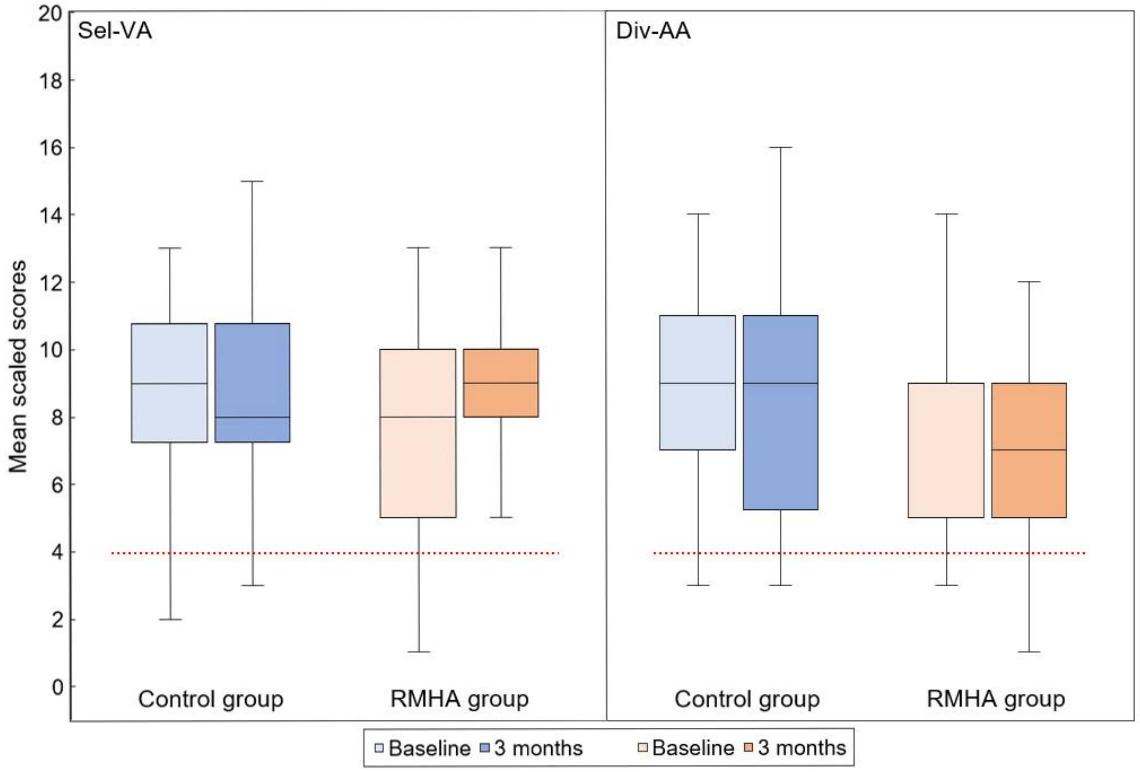

Figure 4.2 – Results of the TEACh control conditions
Boxplots of scaled scores of the two control TEACh tasks measuring Sel-VA and Div-AA, for 22 controls and 21 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). Div-AA: Divided Auditory Attention, RMHA: Remote Microphone Hearing Aid, Sel-VA: Selective Visual Attention, TEACh: Test of Everyday Attention for Children.

Table 4.2 below outlines the mean values of the control and intervention group at baseline and at 3 months.

Table 4.2 – TEACh mean scores and SDs
The mean scores and SDs for the experimental and control conditions of the TEACh test for the two groups at baseline and at 3 months. Disordered performance: ≤ 4, borderline performance: > 4 and ≤ 7, and normal performance: > 7. Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone hearing aid, SD: Standard Deviation, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children.

TEACh	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Sus-AA	9.18 (3.99)	7.47 (3.13)	7.09 (3.12)	6.71 (3.57)
Div-AVA	6.50 (3.48)	5.63 (3.00)	3.57 (3.00)	5.76 (3.25)
Sel-VA	8.72 (3.19)	9.07 (3.25)	7.85 (3.46)	8.95 (2.63)
Div-AA	8.86 (3.05)	9.00 (3.67)	7.47 (3.29)	6.85 (2.78)

Table 4.3 presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time (and the significant simple main effects of group and time, if the interaction is significant) and of the main effect of group and main effect of time. There was a significant interaction in the Div-AVA test, *F*(1, 41) = 5.08, *p* = .029, η_p^2 = .110. This interaction was followed up by looking at the simple main effects of group and time. There was a statistically significant difference in scores between the two groups at baseline, *F*(1, 41) = 8.65, *p* = .005, η_p^2 = .174, meaning that the control group had significantly better scores at baseline compared to the intervention group. There was also a statistically significant effect of time in Div-AVA scores of the RMHA group, *F*(1, 41) = 6.78, *p* = .017, η_p^2 = .253. Therefore, the scores at 3 months were significantly better compared to baseline scores in the RMHA group only, *M* = -2.190, *SE* = .841, *p* = .017.

Table 4.3 – TEACh significance values

The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the simple main effect of group and time and main effects of group and time in the four TEACh conditions. CI: Confidence Interval, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention, TEACh: Test of Everyday Attention for Children. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

Type of attention	**Interaction**		**Simple main effect**				**Main effect**			
	Group* Time	η_p^2 **[95% CI of** η_p^2**]**	**Group**	η_p^2 **[95% CI of** η_p^2**]**	**Time**	η_p^2 **[95% CI of** η_p^2**]/ *MD* [95% CI of *MD*]**	**Group**	η_p^2 **[95% CI of** η_p^2**]**	**Time**	η_p^2 **[95% CI of** η_p^2**]/ *MD* [95% CI of *MD*]**
Sus-AA	.252	.032 [.000, .185]	-	-	-	-	.120	.058 [.000, .228]	$.075^{\dagger}$	.075 [.000, .253]
Div-AVA	.029*	.110 [.000, .297]	.005** (**baseline**)	.174 [.017, .366]	.420 (**Control**)	.031 [.000, .152]	-	-	-	-
			.895 (**3 months**)	.000 [.000, .017]	.017* (**RMHA**)	.253 [.006, .332]				
					.017* (**0-3months**)	-2.190 [-3.945, -.436]				
Sel-VA	.457	.014 [.000, .145]	-	-	-	-	.549	.009 [.000, .131]	.153	.049 [.000, .215]
Div-AA	.470	.013 [.000, .012]	-	-	-	-	.040*	.099 [.000, .283]	.650	.005 [.000, .116]

In the other three conditions (i.e. Sus-AA, Sel-VA and Div-AA), there was no statistically significant interaction between group and time. There was only statistically significant differences in the main effect of group in the Div-AA task, $F(1, 41) = 4.49$, $p = .040$, $\eta_p^2 = .099$. This means that if the scores at baseline and at 3 months are collapsed across time, a significant difference between the two groups is observed for the Div-AA task. In other words, the control group has better scores than the intervention group, regardless of the time variable (i.e. scores at baseline and at 3 months merged together for each group).

Post-hoc analysis controlling for baseline in Div-AVA

As observed earlier, there were significant differences between the intervention and control group in the baseline Div-AVA scores, with the control group scoring significantly better than the intervention group. In intervention studies with clinical control groups there should be no baseline differences as this has a direct impact on the post-intervention scores. The case here was that the intervention group scored worse than the control group at baseline and then improved its scores post-intervention. In this instance, it is not possible to know if the improved performance was due to the intervention or if perhaps it was an effect of regression to the mean, or a placebo/ Hawthorne effect. For this reason, control and intervention group should start without prior differences at baseline. Therefore, a post-hoc analysis controlling for baseline was conducted. Specifically, a one-way ANCOVA was performed comparing the difference from baseline to 3 months in the Div-AVA scores between the two groups while controlling for the baseline Div-AVA scores.

All assumptions for the one-way ANCOVA were met. In detail, the assumptions of absence of outliers (checked through looking for values outside +/- 3 SDs from the mean on the standardised residuals), normality (checked using the Shapiro-Wilk test on the standardised residuals, $p > .05$), homogeneity of variances (assessed through the Levene’s test, $p > .05$), homoscedasticity (checked through scatterplots between the standardised residuals and predicted values), linearity between the dependent variable and the covariate for both levels of the independent variable (assessed through scatterplots between the two variables), and the assumption of homogeneity of regression slopes (assessed through the interaction term, $p > .05$) were all met.

Figure 4.3 below summarises the findings from the one-way ANCOVA analysis.

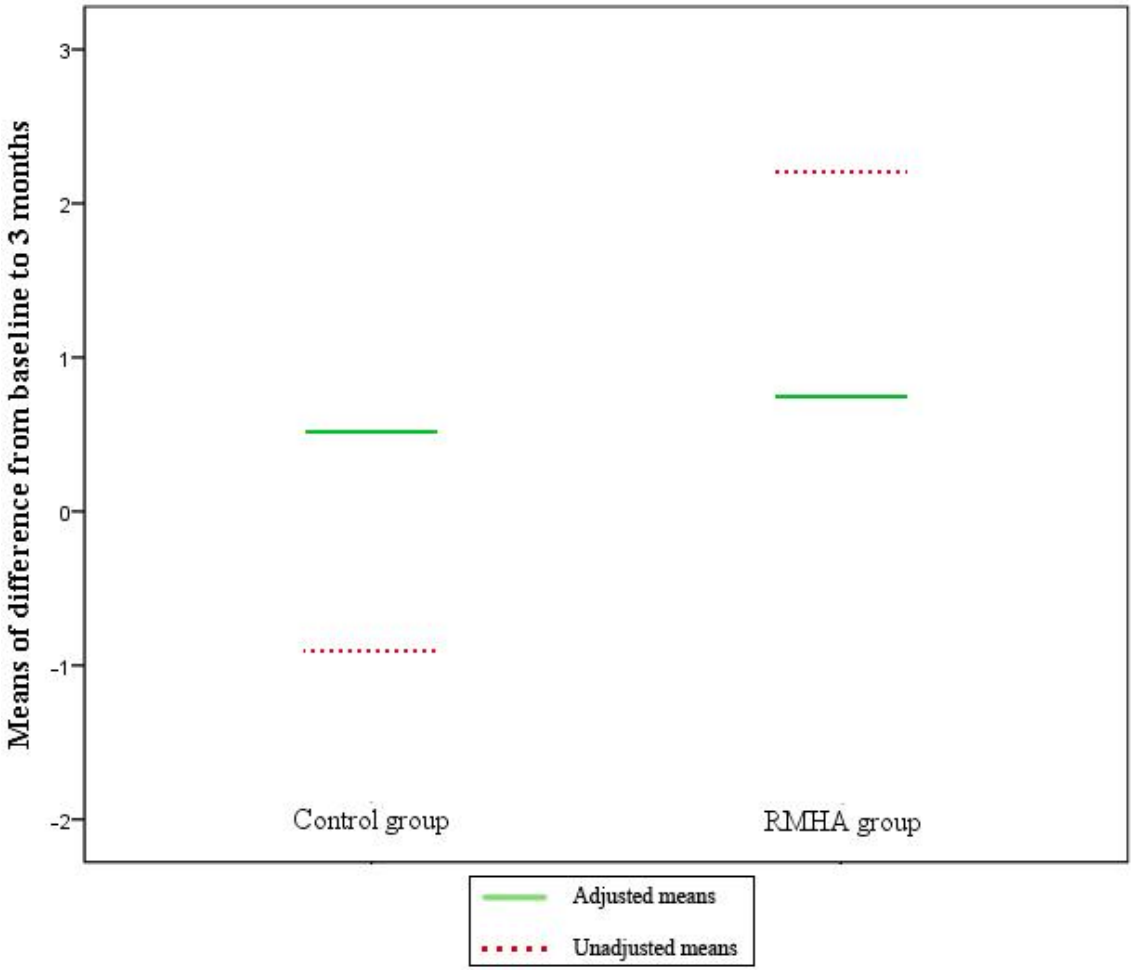

Figure 4.3 – Results from one-way ANCOVA analysis
Adjusted means of the difference from baseline to 3 months in the Div-AVA test for 22 controls and 21 RMHA children are presented in green line. Please note that since the statistical package only produced the adjusted means and not the adjusted individual values for each participant it was only possible to present the means in the figure above without any additional information (e.g. quartile values, upper and lower observations) in the form of boxplot as was previously done. The unadjusted means (that is, without control for baseline scores) are also included using red dotted lines for visual comparisons. ANCOVA: Analysis of Covariance, Div-AVA: Divided Auditory-Visual Attention, RMHA: Remote Microphone Hearing Aid.

Table 4.4 below summarises the descriptive statistics from the one-way ANCOVA analysis for the control and RMHA group.

Table 4.4 - Descriptive statistics from the ANCOVA analysis
Adjusted means, SE, and 95% CIs of the means. ANCOVA: Analysis of Covariance, CI: Confidence Interval, SE: Standard Error.

Test	**Adjusted mean (SE) [95% CI of *M*]**	
	Control group	**RMHA group**
Div-AVA	.517 (.70) [-.915, 1.94]	.743 (.72) [-.72, 2.12]

The one-way ANCOVA analysis did not reveal a significant difference between the two groups in the adjusted Div-AVA scores (difference from baseline to 3 months), $F(1, 40) = .045$, $p = .833$, $\eta_p^2 = .001$, (95% CI [.000, .044]). Thus, after controlling for the baseline scores the difference in scores from baseline to 3 months was no longer statistically significant between the two groups. The implications of this are outlined later in Discussion.

4.3.2 CHAPS

There were no outliers in any of the CHAPS subscales (examined by looking for values in the studentised residuals beyond +/- 3 SDs). All subscales had normally distributed data, checked using the Shapiro-Wilk test of normality of distribution. Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all subscales ($p > .05$ and $p > .001$, respectively).

Figure 4.4 below presents the findings of the CHAPS subscales in boxplots. There were no significant findings observed.

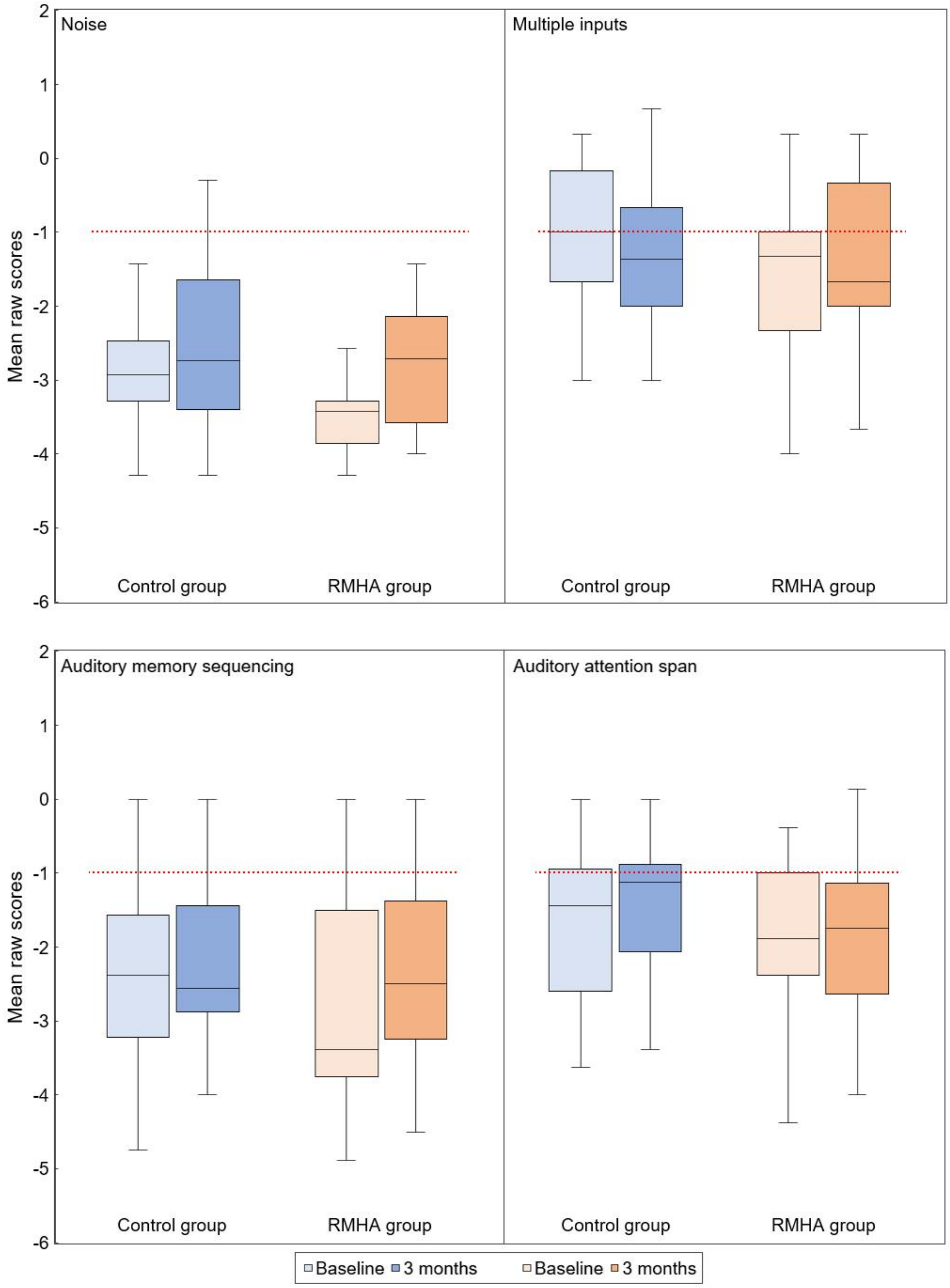

Figure 4.4 – CHAPS subscales results
Boxplots of raw scores of the four CHAPS subscales, for 22 controls and 21 RMHA children, at baseline and at 3 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. Even though this was a non-standardised questionnaire, the CHAPS manual did mention that it considered abnormal performance to be below -1 on the raw scores. Therefore, the red dotted line indicates this cut-off. CHAPS: Children's Auditory Performance Scale, RMHA: Remote Microphone Hearing Aid.

Table 4.5 below outlines the mean values of the control and intervention group at baseline and at 3 months.

Table 4.5 – CHAPS mean scores and SDs
The mean scores and SDs for the experimental and control conditions of the CHAPS for the two groups at baseline and at 3 months. Disordered performance: ≤ -1, and normal performance: > -1. CHAPS: Children Auditory Performance Scale, RMHA: Remote Microphone Hearing Aid, SD: Standard Deviation.

CHAPS	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Noise	-2.88 (.95)	-2.52 (1.10)	-3.37 (.74)	-2.87 (.84)
Multiple inputs	-1.10 (.92)	-1.30 (.94)	-1.66 (1.12)	-1.31 (1.13)
Auditory memory sequencing	-2.31 (1.22)	-2.28 (1.09)	-2.66 (1.49)	-2.28 (1.32)
Auditory attention span	-1.65 (1.00)	-1.43 (.93)	-1.91 (1.04)	-1.80 (1.08)

Table 4.6 below presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. There were no statistically significant interactions between group and time in any of the four CHAPS conditions.

Table 4.6 – CHAPS significance values

The *p*-values, η_p^2 and CIs of the η_p^2 of the interaction between group and time and of the simple main effect of group and time and main effects of group and time in the four CHAPS conditions. CHAPS: Children's Auditory Performance Scale, CI: Confidence Interval. $^{\dagger}p < .10$, $^{**}p < .01$

CHAPS	**Interaction**			**Main effect**					
	Group* Time	**η_p^2**	**95% CI of η_p^2**	**Group**	**η_p^2**	**95% CI of η_p^2**	**Time**	**η_p^2/ *MD***	**95% CI of η_p^2/ 95% CI of *MD***
Noise	.614	.006	.000, .121	.099†	.065	.000, .238	.003**	.195	.025, .386
							.003** (**0-3 months**)	-.434	-.711, -.156
Multiple inputs	.156	.048	.000, .214	.263	.030	.000, .183	.692	.004	.000, .109
Auditory memory sequencing	.392	.018	.000, .156	.610	.006	.000, .122	.315	.025	.000, .171
Auditory attention span	.720	.003	.000, .105	.239	.034	.000, .188	.332	.023	.000, .167

The main effect of group also did not show significant results, whereas the main effect of time was non-significant in all conditions except for the Noise subscale, $F(1, 41) = 9.928$, $p = .003$, $\eta_p^2 = .195$. This means that if the two groups are collapsed together, a significant difference over time is observed. In other words, the scores at 3 months are higher than at baseline, regardless of the group variable (i.e. control and intervention group merged together) as evidenced by the pairwise comparisons, $M = -.434$, $SE = .138$, $p = .003$.

4.3.3 CCC-2

Some of the data from the CCC-2 questionnaire had to be removed or imputed as already noted in the study in Chapter 3 (Section 3.3.5). The CCC-2 data from participant i-12 were removed, as data in all 3 time-points were invalid[20]. Data from case i-12 were not imputed, as all their data were missing and imputing for that would introduce bias. Imputation only took place for participant i-07 at baseline. Their data at that time point were not missing but were also invalid, as explained in the footnote (see footnote 15 below). Hence the analysis included 22 controls and 20 children from the intervention group. There were no outliers in any of the two composite scores (examined by looking for values in the studentised residuals beyond +/- 3 SDs). The standard language composite score had normally distributed data checked using the Shapiro-Wilk test of normality of distribution, while the non-standard language composite did not have normally distributed data in the RMHA group at baseline. All transformations attempted (i.e. square root, natural log, inverse transformation) would not produce normally distributed data at that level and would also make other levels become non-normally distributed. It was thus decided to continue with the analysis and interpretation of the data without transformation. The implications of this are outlined later in Discussion. The last two assumptions of Homogeneity of variances (using the Levene's test of equality of error variance) and Equality of covariance matrices were statistically non-significant in all conditions ($p > .05$ and $p > .001$, respectively).

---

[20] In CCC-2 the positive and negative sum need to be consistent. The positive sum is calculated in the following way: '6 minus the sum of items 51 to 70', and the negative sum in this way: 'the sum of items 1 to 26 plus the sum of items 27 to 50'. When the two scores are inconsistent –according to the Table 3B in page 69 of the CCC-2 manual (Bishop, 2003)– then the answers from the questionnaire are considered invalid.

Figure 4.5 below presents the findings of the CCC-2 composite scores in boxplots. There were no significant findings observed.

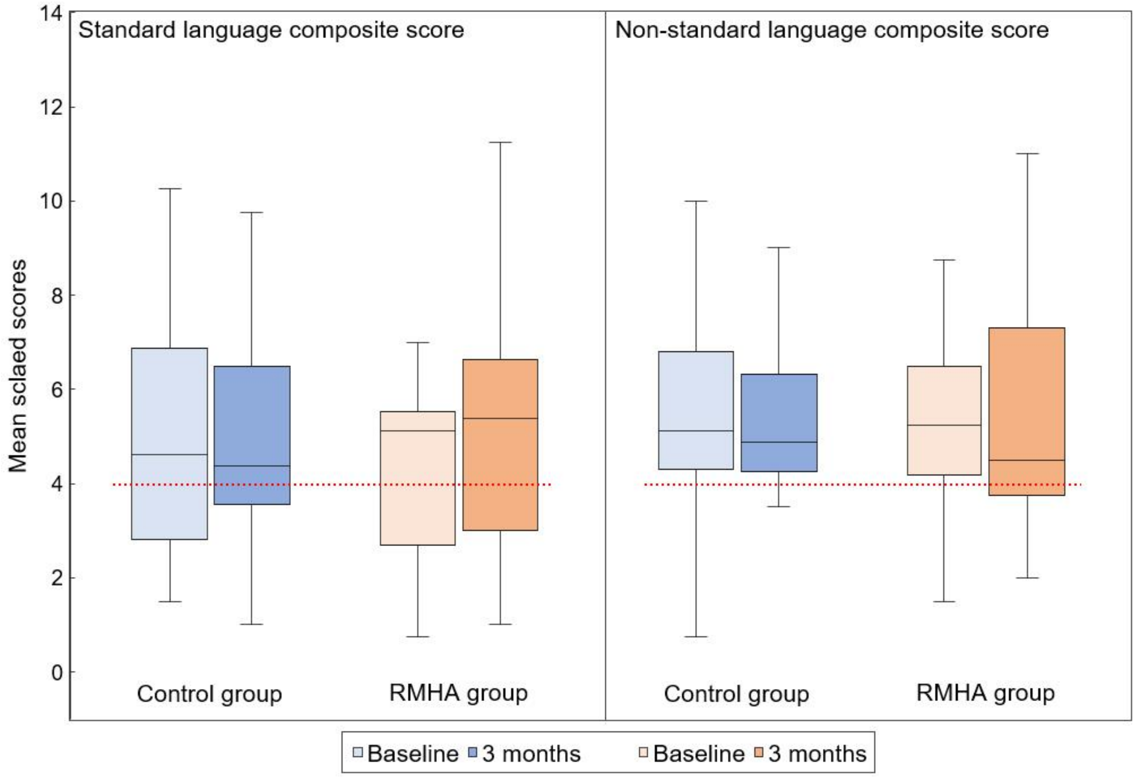

Figure 4.5 – Results of the CCC-2 composite scores
Boxplots of scaled scores of the standard and non-standard language composite scores of the CCC-2 questionnaire, for 22 controls and 20 RMHA children, at baseline, 3 months and at 6 months. The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below 4 scaled scores). CCC-2: Children's Communication Checklist – 2, RMHA: Remote Microphone Hearing Aid.

Table 4.7 below outlines the mean values of the control and intervention group at baseline and at 3 months.

Table 4.7 – CCC-2 mean scores and SDs
The mean scores and SDs for the two composite scores of the CCC-2 for the two groups at baseline and at 3 months. The GCC and SIDC scores are also presented for additional descriptive information. Disordered performance: ≤ 4, borderline performance: > 4 and ≤ 7, and normal performance: > 7. CCC-2: Children's Communication Checklist – 2, GCC: General Communication Composite, RMHA: Remote Microphone Hearing Aid, SD: Standard Deviation, SIDC: Social Interaction Deviance Composite.

CCC-2	**Mean (SD)**			
	Control group		**RMHA group**	
	Baseline	**3 months**	**Baseline**	**3 months**
Standard language composite score	5.12 (2.72)	4.92 (2.56)	4.69 (2.41)	5.03 (2.65)
Non-standard language composite score	5.72 (2.69)	5.79 (2.49)	5.51 (2.30)	5.37 (2.56)
GCC score	43.45 (19.90)	41.50 (19.49)	40.74 (15.63)	41.65 (18.42)
SIDC score	3.68 (8.60)	6.57 (13.36)	4.12 (11.11)	2.10 (10.04)

Table 4.8 below presents the *p*-values, effect sizes and CIs of the effect sizes of the interaction between group and time and of the main effect of group and main effect of time. Results did not reveal statistically significant interactions between group and time in any of the two CCC-2 composite scores. For the main effect of group and main effect of time, again no statistically significant differences were observed.

Table 4.8 – CCC-2 composite scores significance
The *p*-values, η_p^2 and CI of η_p^2 of the interaction between group and time and of the main effect of group and main effect of time in the two composite scores of the CCC-2 questionnaire. CCC-2: Children's Communication Checklist - 2, CI: Confidence Interval.

Condition	**Interaction**			**Main effect**					
	Group* Time	η_p^2	**95% CI of** η_p^2	**Group**	η_p^2	**95% CI of** η_p^2	**Time**	η_p^2	**95% CI of** η_p^2
Standard language composite score	.410	.017	.000, .156	.827	.001	.000, .046	.823	.001	.000, .048
Non-standard language composite score	.686	.004	.000, .113	.675	.004	.000, .114	.875	.001	.000, .025

4.4 Discussion

By pooling the two datasets together the statistical power that was initially calculated has been met. The initial sample size calculation indicated that the sample should have had at least 32 total participants for an 80% power. The combined datasets resulted in a sample size of 43, therefore interpretation of the findings can be made with more confidence now.

4.4.1 Attention

The analysis of the pooled datasets from the studies in Chapters 2 and 3 did not reveal any statistically significant effects in any of the four TEACh subtests (i.e. Sus-AA, Div-AVA, Sel-VA, Div-AA). The experimental conditions of Sus-AA and Div-AVA remained unchanged in the intervention group over a 3-month period of use of the RMHAs. This is contrary to the hypotheses made, but the shortened intervention period of 3 months might explain the non-significant findings. At the end of the first study in Chapter 2 it was argued that any lasting changes in cognitive functions such as attention might require a longer period to take place (see Section 2.4.5 under Sample size and intervention period). Later in the second study in Chapter 3 a longer intervention period of 6 months was adopted. However, as the datasets from the two studies were combined to increase the statistical power in this new analysis only the 3-month time-point could be used, which was a common time-point between the two studies. It would thus be interesting to further examine the effects of RMHAs on Sus-AA and Div-AVA in intervention periods of 6 months or longer.

Furthermore, the Div-AVA scores initially showed significant improvement only in the intervention group from baseline to 3 months. Nevertheless, the children in the intervention group also scored significantly worse at baseline compared to children in the control group. In clinical trials participants in both the intervention and the clinical control groups should have similar baseline scores otherwise interpretation of the results would be biased. In this case the fact that the RMHA group significantly improved after 3 months could simply be because of regression to the mean, a placebo effect or a Hawthorne effect (see Section 3.4.4 under Design limitations for more details on the latter two effects). For this reason, a post-hoc analysis controlling for baseline scores was conducted. The results from this analysis showed that there was no longer a significant improvement when comparing the difference from baseline to 3 months between the two groups. This further supports the argument that the initial significant improvement found for the intervention group was due to a regression to the mean and not because of the use of RMHAs for 3 months.

The two control conditions of Sel-VA and Div-AA also did not record any significant changes. This was in line with the initial hypotheses, as it was not expected that the use of RMHAs would have a positive effect on visual attention or divided listening. This is the first study to examine the lasting effects of RMHA use on behavioural measures of attention. Future studies can use longer intervention periods to test whether RMHAs can have a long-term positive effect on types of attention. Moreover, the current study only focused on unaided testing in these attention behavioural outcome measures. It would thus be interesting in the future to test the effect of RMHAs both on aided and unaided conditions and compare performance on the two conditions. It may be that children only present significant improvements when using the system and that no carry-over effects take place. Overall, the findings showed no improvement in any of the TEACh test conditions. This means that a 3-month use of RMHAs does not provide any improvement in these four measures of attention (Sus-AA, Div-AVA, Sel-VA, Div-AA). At the same time, using the RMHAs for 3 months does not negatively impact these types of attention but longer intervention periods should further examine this notion.

4.4.2 Questionnaires

<u>CHAPS</u>

The experimental conditions of the CHAPS (Noise, Multiple inputs, and Auditory attention span) remained unchanged over the 3-month intervention period. This means that as judged by parents, children using the RMHA did not show any significant improvement in their out-of-school behaviours when listening in noise, and their auditory attention span did not improve either. These findings come in contrast with findings from a previous study that used hearing aids with directional microphone on children with APD for 6 months (Kuk et al., 2008). That study showed that the Auditory attention span subscale of the CHAPS completed by parents significantly improved after 6 months. Despite this, the performance on this subscale after 6 months still remained ‘at risk’. In addition, the study did not use a control group, hence their findings are of low level of evidence. Since parents completed the questionnaire based on their child’s behaviours at home and other non-school environments it means that children were judged while not using the system. Future studies could administer the CHAPS to both parents and teachers. In that way comparisons can be made between children’s behaviours at school (when using the system) and at home (when not using the system).

CCC-2

Similarly, the experimental condition of the CCC-2 questionnaire (i.e. standard language composite score) also remained unchanged over the intervention period. The CCC-2 was completed by parents at home and can thus be supported that a 3-month RMHA use does not have a lasting effect on children’s standard language aspects, such as speech, semantics, syntax and coherence. For the CCC-2 non-standard language composite score, the assumption of normality was not met for the baseline level of the intervention group. Transformations would not make the data normally distributed and thus analysis was run as it was. Interpreting non-normally distributed data from an ANOVA analysis may increase the chance of making a Type I error. This means that there is a higher probability to detect a significant effect where no such effect actually exists. But as this composite score was not found to have any significant effect it is expected that the Type I error was not inflated in this case. This finding of no improvement in the non-standard language CCC-2 composite score is in line with the findings from a previous study. In that study, the researchers used a behavioural test (the Comprehensive assessment of spoken language) measuring non-standard language aspects (e.g. Sarcasm, Figurative language, Use of context and Drawing inferences from given information; Sharma et al., 2012). The scores of children with APD in the intervention groups did not show significant change over a 6-week period. The study groups though, were using both AT and RMHAs, making it difficult to attribute the lack of change in test scores to the use of RMHAs alone. Additionally, the behavioural test used in that study was not a standardised test (Sharma et al., 2012). As discussed earlier, it was not expected that the clearer signal produced by the system would have an impact on non-standard language aspects, such as sarcasm, use of context or nonverbal communication.

4.4.3 Limitations

Intervention period

Pooling the two datasets together met the sample size initial calculations and it improved the statistical power of this analysis over the two previous analyses in the studies in Chapters 2 and 3. Nonetheless, there are other limitations to this analysis that may influence interpretation of the findings. Firstly, as noted earlier the study period might be a short one for any significant long-term effects to take place. Another RCT by Purdy et al. (2009) included children with reading delay who used RMHAs for 6 weeks. It demonstrated improvements in reading questionnaire scores completed by teachers and improved scores in children's self-reports on their classroom listening environment. Nevertheless, there was no significant change in behavioural reading test scores and the authors concluded that a longer intervention period might be required to influence standardised reading tests (Purdy et al., 2009). Similarly, the period of 3 months in this analysis may have been too short for significant changes in the TEACh experimental conditions to take place. Improvements in higher-order functions, such as attention, might require longer intervention periods to reveal statistically significant results.

Differences between the two combined studies

In addition, despite the increased sample size the samples from the two studies were not identical. The intervention groups from the two studies used two different microphones during the study period. Even if it was shown that these two microphones functioned very similarly (see Section 2.2.3 under Verification of RMHAs and Section 3.2.2 under The device for more details on this), there was still a difference in the transmission range between the two devices (7m for the first microphone and 25m for the second). Results in the first study may have been compromised due to the smaller transmission range of the RMHA that children used. Moreover, the set of tests administered in each study were not all the same while the testing time for the first study was longer than the second study. Therefore, results from the first study may have been influenced by the longer testing time which may have made children more tired, having in turn an impact on their test scores as compared to results from the second study. In consequence, pooling the data from the two studies together even though it improved the statistical power may have compromised the overall findings in different ways.

APD diagnostic criteria

As discussed in more detail in Section 1.4, APD does not have universally accepted diagnostic criteria. The criteria used in this study were the same criteria used at GOSH and they closely resemble the ones promoted by the ASHA (2005b). However, there are other APD position statements -such as the AAA (2010) and the ‘European APD consensus’ group (Iliadou et al., 2017)- that argue for the use of different APD diagnostic criteria. The table in Appendix A also shows how different APD research studies used different APD criteria. This is problematic in the field of APD research and practice as knowledge gained from APD studies cannot be generalised in the wider APD population, while prevalence of the disorder and management recommendations also vary according to the criteria used. The results obtained from the current work can therefore only be applied to children who met with the same APD diagnostic criteria used in these studies (see Section 2.2.1 under Inclusion and exclusion criteria for more details on the APD criteria used in the studies).

4.5 Conclusion

The use of RMHAs for 3 months was not enough to produce significant improvements in measures of attention (i.e. Sus-AA, Div-AVA). Any initial positive effects found in Div-AVA scores of the intervention group may be attributed to a regression to the mean rather than to the use of RMHAs. Moreover, use of the RMHAs for 3 months also did not improve parents’ views on their children’s listening and language behaviours. The study period may have been too short for significant results to have taken place, therefore longer trials including large enough sample sizes are required to further examine these effects. Additionally, the inclusion of aided testing alongside unaided testing would be beneficial as comparisons between the two conditions can be made.

CHAPTER 5 – STUDY 3: AUDITORY PROCESSING SKILLS AND TYPES OF ATTENTION: INFORMATION FOR APD MANAGEMENT

5.1 Introduction

The study in this chapter follows closely the work published in the journal ‘Frontiers in Psychology’ (Stavrinos, Iliadou, Edwards, Sirimanna, & Bamiou, 2018).

5.1.1 APD and types of attention

Children suspected of APD are often reported to exhibit poor attentive behaviours, such as poor auditory attention span, inattentiveness and distractibility (AAA, 2010; ASHA, 2005b), but the contribution of attention to AP skills is still not adequately studied. Previous research on children suspected of APD focused on Sus-AA, showing correlations between this type of attention and AP tests (Cameron, Glyde, Dillon, & Whitfield, 2016; Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). Additionally, children with APD demonstrated worse scores in Sus-AA compared to children suspected of APD but not meeting APD criteria (Allen & Allan, 2014).

One study, found that children with dichotic listening deficits (and suspected of but not diagnosed with APD) exhibited worse Div-AA test scores compared to typically developing peers (Martin et al., 2007). Another study looked at the ability of 18 different measures (including a Div-AVA measure) to predict APD (Lam & Sanchez, 2007). The rest of the measures in that study comprised a competing sentences test, an auditory figure ground test, an auditory performance questionnaire and other attention tests. These measures were compared with results from the APD diagnostic tests (that were used to diagnose APD in the study) to examine their sensitivity and specificity as APD screening tools. The Div-AVA measure was not found to be predictive of APD and only the competing sentences (left ear) measure was significantly predicting APD (Lam & Sanchez, 2007). However, some of the children in the study also had an ADHD diagnosis. Even though these latter two studies looked at Div-AA and Div-AVA, the correlation between these divided attention tasks and tests of AP skills has not been investigated yet in a sample of children with confirmed APD diagnosis. Divided auditory attention and the DDT (a standard AP test) both make use of divided listening. At the same time, studies indicate that both Div-AA and dichotic listening use interhemispheric transfer and that deficits in these two functions might be attributed to dysfunction in the corpus callosum (Hutchinson et al., 2008; Musiek & Weihing, 2011; van der Knaap & van der Ham, 2011; Westerhausen & Hugdahl, 2008). Looking into the relationship between tasks measuring these two functions in APD populations, could help characterise the nature of deficits in dichotic listening that these children face. It is less clear how Div-AVA could be related to AP deficits but as described earlier, the auditory and visual modality in bimodal attention tasks use attentional resources from a shared resource pool (Wahn & König, 2017). The compromised auditory modality that children with APD exhibit, could have a negative impact on the function of sharing resources in bimodal attention tasks, which might then manifest in poor performance in Div-AVA tasks.

The relationship between types of visual attention and APD has not been examined consistently thus far. Some trials studied Sus-VA and found it to be poor in children suspected of APD (Gyldenkærne et al., 2014; Sharma et al., 2009), but no further examinations of its relationship to AP skills was made. Another study found that visual alertness (in reaction time) was significantly related to tasks of frequency discrimination and backward masking in children with poor listening skills (Moore et al., 2010). However, the measures used to define these listening skills in Moore et al.'s (2010) study are not routinely used in APD diagnostic batteries, while children in their sample were as young as 6 years old. Children that young show high variability in performance and might not comprehend the psychoacoustic AP tasks (AAA, 2010; ASHA, 200b), which in turn might influence their performance in the tests. Moreover, a study showed that children with APD had worse scores in Sel-VA compared to children suspected of APD but not meeting APD criteria (Allen & Allan, 2014). Further investigation of types of visual attention could help determine the relationship of ADs and AP deficits in children with APD. For instance, possible deficits in visual attention measures could point to broader attentional deficits in both the visual and auditory modality in children diagnosed with APD.

5.1.2 APD diagnostic criteria

A standard APD diagnostic procedure includes audiological tests, electrophysiological tests and AP tests (ASHA, 2005b; Bamiou et al., 2006; Ferguson et al., 2011; Martin & Keith, 2009), but does not routinely include attention measures. A study by Wilson and Arnott (2013), on a sample of 150 Australian children, aged 7-15, compared nine different widely used sets of APD diagnostic criteria and revealed that they had substantial differences in diagnosis rates between them. Despite the evidence of possible interaction between attention and AP skills (Cameron et al., 2016; Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015), none of the nine APD diagnostic batteries in Wilson and Arnott's (2013) study included measures of attention in their criteria. These findings suggest that audiology professionals should consider routinely using attention tests in the APD diagnostic protocol.

5.1.3 Research questions and hypotheses

This study will first investigate the correlations between different types of attention and AP abilities. Secondly, it aims to examine differential diagnostic criteria including attention measures in order to better inform management and treatment strategies for children with AP difficulties. The study first looks into the relationships between different types of attention and AP tests on a sample of children that met specified APD diagnostic criteria, instead of using a sample of children suspected of APD that previous research has focused on up to now. The relationship between auditory and language skills appears to be different between typically developing and disordered groups of children (Grube et al., 2014; Kuppen & Goswami, 2016) and it could be that the relationship between auditory and attention skills also differs depending on the studied population. Extrapolating inferences from APD-suspected to APD-diagnosed children (which literature up to now has done) might lead to inaccurate conclusions regarding the APD population (Iliadou et al., 2016). To address this concern, the study examines correlations only in the subgroup of children who met specified APD diagnostic criteria.

The present study will include measures of different types of attention (e.g. Sus-AA, Div-AA, Div-AVA and Sel-VA) and look at how they relate to AP tests from a standard APD battery (e.g.

DDT, GIN, AFG, LiSN-S and FPT). It is hypothesised that the auditory types of attention will interact with the DDT, as Sus-AA has done in past studies (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015), while divided attention tests are also expected to interact with the DDT as divided attention could be sharing same brain activations with dichotic listening (Hutchinson et al., 2008; Musiek & Weihing, 2011; van der Knaap & van der Ham, 2011; Westerhausen & Hugdahl, 2008). As Sus-AA did not correlate with the rest of the AP tests in other studies (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015) it is not expected that it will either in this study. Since there is no previous evidence that examines the correlation of divided attention with the rest of the AP tests, it is hypothesised that divided attention will not correlate with other AP measures. On Sel-VA, I do not expect that the visual modality will interact with the auditory skills measured by the AP tests. The key hypotheses of the correlation analysis are set below.

1. There will be a significant positive correlation between:
    a. Sus-AA and the DDT.
    b. Div-AA and the DDT.
    c. Div-AVA and the DDT.
2. At the same time, there will be no association between the Sel-VA task and the AP tests.

Next, the study will make comparisons between the yield of the standard APD diagnostic battery and the yield of two other proposed batteries (i.e. an AD battery and a battery testing for APD but excluding cases with ADs). The rationale behind the latter battery was conceived post-hoc and all test batteries are described in detail in the Methodology (Section 5.2.2 under Diagnostic criteria). It is expected that the standard APD battery will have a greater yield than the AD battery, as previous studies reported larger percentages of APD diagnosis over Sus-AA plus Sus-VA deficits; 72% over 41%, respectively in Gyldenkærne et al. (2014) and 49% over 15%, respectively in Sharma et al. (2009). These studies only reported the proportions without statistically comparing them. The hypotheses for this part of the analysis are set below.

3. The diagnostic yield of the standard APD battery:

    a. Will be significantly greater than the diagnostic yield of the AD battery.
    b. Will not be significantly different from the diagnostic yield of the APD without ADs battery.

Finally, a newly proposed battery that identifies children with APD with predominantly inattentive behaviours will be presented and discussed. These behaviours by children with APD are termed here as an ‘inattentive subtype of APD’ (i-APD) for brevity. The term is adapted from the ADHD literature, which refers to a predominantly inattentive subtype of ADHD, as suggested by Chermak et al. (2002). The rationale behind this was conceived post-hoc and it is described in more detail in Methodology below (see Section 5.2.2 under Diagnostic criteria).

5.2 Methodology

5.2.1 Participants

Sample size and inclusion criteria

Eighty-five families of children with reports of SIN difficulties and audiology referrals for APD assessment at GOSH were initially invited with thirty replying. The inclusion criteria of the study were adopted from the diagnostic protocol used in GOSH (GOSH, 2018). These closely follow the ASHA (2005b) position statement. As discussed earlier there are no universally accepted criteria for diagnosis in APD research (see Appendix A). This means that findings from the study cannot be generalised to APD populations, but only to children diagnosed under these specific criteria used in GOSH. Nevertheless, the criteria used here aimed to minimise the effect that age, cognitive ability, hearing thresholds, co-occurring disorders and language might have on the presentation of APD. The following inclusion criteria were set:

1. Referral for APD assessment at GOSH on suspicion of AP deficits.
2. No neurological or pervasive disorder (e.g. epilepsy) or developmental delay.
3. Normal peripheral hearing and middle ear function (i.e. air conduction PTA below 20 dB HL in all octave frequencies between 250 Hz and 8 KHz [BSA, 2012], middle-ear pressure

between -150 to +50 daPa, middle-ear admittance between 0.3 to 1.6 cm^3 and ear-canal volumes between 0.4 to 1.0 cm^3 [BSA, 2013]).

4. Nonverbal cognitive ability score of more than 85 on the WNV scale of ability.
5. Aged between 7-11 years.
6. Native English speakers.
7. No previous use of RMHAs or AT.

Diagnosis of APD was based on routine clinical tests at GOSH. These tests were administered by qualified audiologist as part of the diagnostic protocol at the hospital. The criteria used for identifying children as having APD are reported later in Methodology (see Section 5.2.2 under Diagnostic criteria).

As 3 children scored below 85 on the WNV IQ test, they were excluded from the study, bringing the study sample down to 27 children suspected of APD. In this final sample, there were 9 girls and 18 boys aged 7-11 years with a mean age of 9 years 7 months. All participants had English as first language and three were bilingual, while none had used RMHAs or AT before.

All 27 participants were referred to GOSH for listening difficulties and underwent the APD assessment of the APD clinic in GOSH. Twenty children were given an APD diagnosis while seven did not meet criteria for APD. Thus, there were two subgroups formed, the APD group with 20 children and the non-APD group with 7 children. All 27 participants from both groups were also called in on separate date and tested on some additional tests, described in detail in Section 5.2.2.

Ethical issues

This study was reviewed and given a favourable opinion by the Bloomsbury REC, REC reference: 14/LO/1509. Children and parents were asked to sign assent and consent forms respectively, prior to inclusion in the study. It was explained to participants that they had the right to withdraw from the study at any point. There was neither a loss of benefit nor a penalty if participants decided to withdraw from the study. Children and their parents were informed that their information would be anonymised and kept confidential. Data were pseudo-anonymised prior to analysis by using numbered codes. Data were stored in a password protected file on a secure password-protected database and hard copies were locked in a filing unit in a code-protected room within UCL, in accordance with the data protection act 1998.

5.2.2 Tests and procedure

Tests

Children were tested under the routine procedure for APD assessment at GOSH. First, normal periphery was confirmed through the PTA and tympanogram/ acoustic reflexes tests. Then, children completed AP tests as per the clinic's APD protocol. The AP tests that were administered included the DDT, GIN, FPT and two SCAN-3 C subtests (the AFG +8 dB and AFG 0 dB). They are described in Table 5.1 below.

Table 5.1 – AP tests
Task description of the five AP tests (Guenette, 2006; R. W. Keith, 2009; Musiek & Pinheiro, 1987; Musiek, 2002; Paulovicks, 2008). AFG: Auditory Figure Ground, AP: Auditory Processing, DDT: Dichotic Digits Test, FPT: Frequency Pattern Test, GIN: Gaps-in-Noise.

Test	Task description
DDT	Presents 20 stimuli via headphones, 2 digits per ear, from numbers 1 to 10 except 7. Children are asked to repeat back the numbers from both ears in any order they can remember them.
GIN	Presents 6-second broadband noises in each ear that contain 0-3 silent gaps. Children are asked to detect these gaps and click a button when they do.
FPT	Successive patterns of triplets are played (30 patterns in each ear), with 150 msec low (880 Hz) and high (1122 Hz) tones that have an intertonal interval of 200 msec. Children have to detect and verbally repeat these patterns.
AFG +8 dB	Speech in babble noise test that has the target voice being 8 dB greater in intensity than the background babble noise. Children need to repeat back the words they hear.
AFG 0 dB	Speech in babble noise test that has both the target and the background babble noise at the same intensity. Children need to repeat back the words they hear.

The above AP tests were administered by audiologists at GOSH, as part of the typical test procedure for patients suspected of APD. Hence, this data was collected as part of clinical APD assessments. After that the participants were recruited and three additional tests were administered at UCL Ear Institute on a different day. These additional tests were the WNV scale of ability (see Section 2.2.3), the TEACh (see Section 2.2.3 and Section 3.2.3/ Table 3.4) and the LiSN-S test (see Section 3.2.3/ Table 3.3; Cameron & Dillon, 2011; Manly et al., 1999; Wechsler & Naglieri, 2006). This data was collected for research purposes, and combined with the AP tests collected at GOSH they were used in the analyses in this study. Test order for the tests administered at UCL was randomised for each participant, while the total duration of these three tests was approximately 1 hour 15 minutes. Breaks were given frequently and whenever needed. Finally, the CCC-2 questionnaire was completed by parents on the day (Bishop, 2003). The CCC-2 questionnaire is a tool that identifies presentations that co-occur with APD, such as DLD or PLI (Bishop, 2003; Ferguson et al., 2011). More information on the CCC-2 and its subscales are summarised in Section 2.2.3 and Table 2.5 in Chapter 2. The CCC-2 criteria used in this study to identify children as having DLD, PLI or ASD are described in the next section. This information has been used in the analysis comparing different diagnostic batteries. In the correlation analysis there were two CCC-2 composite scores created in the same way they were compiled in Chapters 2 and 3 (see Section 2.2.3 under Questionnaires for details). These two composite scores, the Standard Language (SL) and Non-Standard Language (NSL) scores, do not identify children as having problems in these aspects (that is, they do not identify children with DLD or PLI) and are only used to group together the subscales that assess these skills (SL and NSL) in children.

Diagnostic criteria

Children in this study were diagnosed with APD following the criteria of the standard APD diagnostic protocol used at GOSH. These criteria identified children as having APD when at least two of the AP test scores were -2 SDs from the mean or when only one test was -3 SDs from the mean (ASHA, 2005b). If all the AP tests were within normal range but disordered score patterns were found on the LiSN-S, then the child was diagnosed with SPD. Spatial processing disorder is a subtype of APD characterised by difficulties segregating targets coming from the front in the presence of simultaneous presentation of distracting sounds from other directions (Cameron & Dillon, 2011). Disordered LiSN-S scores are recorded when the Spatial advantage and High-cue SRT or Total advantage conditions fall 2 SDs below the mean.

There were four batteries used in this study and these are described in Table 5.2 below, along with their criteria for identification. The first three batteries were used in a comparative analysis while the latter, the i-APD battery, was used in a descriptive analysis.

Table 5.2 – Diagnostic batteries
The four batteries used in the study and their criteria for identification and categorisation. AD: Attention Deficit, AP: Auditory Processing, APD: Auditory Processing Disorder, i-APD: inattentive subtype of APD, LiSN-S: Listening in Spatialised Noise – Sentences, SD: Standard Deviation, TEACh: Test of Everyday Attention for Children.

Battery	Criteria for identification/ categorisation
APD battery	At least two AP tests -2 SDs from the mean or one AP test -3 SDs from the mean or the LiSN-S Spatial advantage and High cue/ Total advantage conditions -2 SDs from the mean.
AD battery	At least two TEACh subtest scores -2 SDs from the mean. Children in this category are considered to have attention problems.
APD without ADs battery	Children diagnosed with APD through the standard APD battery (i.e. first battery) but not identified with ADs from the AD battery (i.e. second battery). These children do not present with any attention problems.
i-APD battery	Children diagnosed with APD under the standard APD battery (first battery) and only one TEACh subtest score -2 SDs from the mean. Children in this category are not considered having general attention problems but do have one failed attention test.

Given the co-occurrence of ADs with AP problems, the correlations between attention and AP measures (Gyldenkærne et al., 2014; Sharma et al., 2009) and the evidence that show children with APD perform poorly on attention tests compared to children suspected of APD but not meeting APD criteria (Allen & Allan, 2014), the study aimed to examine the relationship between APD and ADs through comparisons between the two batteries. Therefore, the APD battery was compared with the AD battery to examine whether more children suspected of APD met with APD criteria than AD criteria. After observing a co-occurrence between APD and ADs, a post-hoc comparison between the APD battery and the APD battery excluding ADs was also run. This comparison aimed to examine whether removing cases with ADs from the group meeting APD criteria would reduce the diagnostic yield. Finally, comparing the yield of an AD battery (i.e. at least two failed attention tests) and the yield of the APD without ADs battery (i.e. zero failed attention tests), the cases that fell in between these two categories were examined post-hoc. It was thus considered to propose the i-APD battery that identifies children with APD that have only one failed attention test (i.e. 2 SDs below the mean). This i-APD battery, along with the other two batteries mentioned here, were only used to look into potential useful information in terms of management and did not intend to propose a new APD diagnostic battery. These were simply different ways of looking at the already obtained information from the AP and attention assessments to feed more information to the clinicians.

Using the results from the CCC-2 questionnaire children were identifying as being at risk of DLD, PLI or ASD based on the guidelines provided by the CCC-2 manual (Bishop, 2003). These guidelines make use of two composite scores, the GCC and the Social Interaction Deviance Composite (SIDC) score. The GCC score is the sum of the first eight CCC-2 subscales (described earlier in Section 2.2.3, under Questionnaires). The SIDC score is produced by summing the Inappropriate Initiation, Non-verbal Communication, Social Relations and Interests subscales and subtracting the sum of the first four subscales (i.e. Speech, Syntax, Semantics, Coherence). Children were identified as being at risk of DLD if the SIDC score was between 0 and 9 and the GCC score below 55. Identification of being at risk of PLI was given if the SIDC score was below 0 and the GCC score below 55. Finally, children were identified as being at risk of ASD if the Social Relations and Interests subscales were below the 6th percentile and the GCC score below 55 (Bishop, 2003).

5.2.3 Statistical analyses

Pearson's partial correlation analysis (2-tailed) was used to look into the relationships between types of attention (i.e. Sus-AA, Div-AA, Div-AVA and Sel-VA), AP tests (i.e. DDT, GIN, AFG +8 dB, AFG 0 dB, LiSN-S and FPT) and the CCC-2 scores. Gender and IQ performance were controlled in the analysis. The correlation sample size of the FPT only had 6 children and was thus decided to remove the test from the analysis, as meaningful results cannot be obtained from it. To compare the proportions of identification from the three batteries (i.e. APD, AD, APD without ADs), Cochran's Q test was conducted. This analysis compares percentages of identification between different groups/ conditions.

5.2.4 *Post-hoc comparison analyses between APD and non-APD diagnosed groups*

Post-hoc comparisons between children meeting APD criteria against those who did not were conducted to examine the performance between the two groups on the tests that were not used to assess them for APD. This is because performance in the tests used for APD identification needs to be -2 SDs from the mean to give an APD diagnosis, hence comparisons on these tests between the APD-diagnosed group and the non-APD group (who had scores above -2 SDs from the mean) would be biased. A separate analysis, though, comparing the tests that comprised the APD criteria was also run on the performance of the two groups in the AP tests to examine any possible differences. The TEACh subtests (i.e. Sus-AA, Div-AVA, Div-AA, Sel-VA), the CCC-2 SL and NSL composite scores and the LiSN-S Low-cue SRT and Talker advantage conditions[21] were compared between the two groups using independent samples t-test in the initial analysis, followed by comparisons in performance in the AP tests (i.e. LiSN-S spatial conditions, GIN, DDT, FPT, AFG tests). The APD group comprised 20 children (14 boys, 6 girls, mean age 9 years 7 months [SD = 1 year 3 months], and mean WNV IQ 99.7 [SD = 8.8]), whereas the non-APD group had 7 children (4 boys, 3 girls, mean age 9 years 5 months [SD = 1 year 8 months], and mean WNV IQ 103.4 [SD = 4.4]). A post-hoc power calculation for these t-test comparisons was performed to determine the actual power of the study to detect a 1 SD effect size at 5% significance, revealing a power of 60%, which is smaller than the typical power of 80% (Cohen, 1992). The implications of this are described later in Discussion.

In addition, the percentages of the children in the two groups meeting AD criteria were compared against the percentages of children not meeting AD criteria. Comparisons of the percentages of children in the two groups identified as having DLD, PLI, ASD through the CCC-2 and those not given any identification were also conducted. These comparisons were made using chi-square tests of homogeneity. A post-hoc power calculation for these chi-square analyses was performed to determine the actual power of the study, revealing a power of 41% for the AD comparisons

---

[21] The former assesses the SIN ability of children when spatial and pitch cues are eliminated, while the Talker advantage assesses the ability to distinguish between different talkers. The LiSN-S Low-cue SRT and Talker advantage scores are not used to identify SPD and are therefore used in the current post-hoc analysis.

and 48% for the CCC-2 comparisons, which are both substantially smaller than the typical power of 80% (Cohen, 1992). The implications of this are described later in Discussion.

5.3 Results

The study did not include an a priori power calculation. A post-hoc power calculation was performed to determine the actual power of the study, using $r = .50$, $a = .05$, and $n = 20$. The calculated power was 68%, which is smaller than the typical power of 80% (Cohen, 1992). The implications of this are described later in Discussion.

5.3.1 Audiometry

All participants passed the audiometry tests. They had normal hearing thresholds below 20 dB HL in each frequency between 250 Hz and 8 KHz (BSA, 2012), typical middle-ear pressure between -150 to +50 daPa, normal middle-ear admittance between 0.3 to 1.6 cm^3 and normal ear-canal volumes between 0.4 to 1.0 cm^3 (BSA, 2013).

5.3.2 Correlations between AP and attention tests

Table 5.3 summarises the performance of the 27 children in each of the tests. Children were categorised as Normal when scores were above -1 SD from the mean, Borderline when scores were between -2 and -1 SDs from the mean and Disordered when scores were below -2 SDs from the mean in all tests except the AFG tests. For the AFG tests a score of -2.33 SDs from the mean was required for the performance to be considered disordered as per the guidelines of the test manual (R. W. Keith, 2009).

Table 5.3 – Summary of test results
Name of each test/ subtest, number of participants completing each test, mean z scores and percentages of classification. AFG: Auditory Figure Ground, B: Borderline, D: Disordered, DDT: Dichotic Digits Test, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, FPT: Frequency Pattern Test, GIN: Gaps-in-Noise, LiSN-S: Listening in Spatialized Noise - Sentences, N: Normal, Sel-VA: Selective Visual Attention, SRT: Speech Reception Threshold, Sus-AA: Sustained Auditory Attention, TEACh: Test of Every Attention for Children.

Test Number of children	**Subtest**	**Mean z score**	**N (%)**	**B (%)**	**D (%)**
LiSN-S 27 children	Low-cue SRT	-0.86	56%	33%	11%
	High-cue SRT	-1.15	44%	37%	19%
	Talker advantage	-0.94	48%	41%	11%
	Spatial advantage	-1.24	52%	26%	22%
	Total advantage	-0.77	59%	26%	15%
TEACh 27 children	Sus-AA	-0.86	63%	7%	30%
	Div-AA	-0.75	59%	19%	22%
	Sel-VA	-0.71	63%	30%	7%
	Div-AVA	-1.82	33%	15%	52%
DDT 21 children	Double (Average)	-1.05	52%	14%	33%
GIN 18 children	Average	-0.8	72%	11%	17%
AFG 0 dB 17 children		-1.8	29%	18%	53%
AFG +8 dB 15 children		-1.31	53%	7%	40%
FPT 11 children	Triple (Average)	-0.65	45%	18%	36%

The sample for the Pearson's partial correlation analysis comprised only children who met criteria for APD diagnosis and these were 20 children of the initial 27. Of the seven children who did not meet APD criteria, two did not fail any of the AP tests, four had only one failed AP test (two children failed the AFG 0 dB, one the AFG + 8 dB, and one the DDT) without any borderline performance in the other tests, while one child failed one AP test (i.e. the FPT) and had borderline performance in the AFG 0 dB. Gender and IQ were controlled in this analysis. Age was not controlled as all test scores were converted into z scores. Right and left ear scores of the DDT and GIN were averaged as they did not significantly differ between them, $t(20) = -.06$, $p = .951$, $d = -.01$ (95% CI [-.62, .58]) and $t(17) = -.25$, $p = .815$, $d = -.07$ (95% CI [-.68, .52]), respectively. All assumptions of linearity, absence of outliers (assessed by the Mahalabonis distance, $p > .001$) and normality of distribution (assessed by the Shapiro-Wilk test, $p > .05$) were met. Results from the correlations between tests are presented in Table 5.4. Not all children completed all tests due to either lack of task comprehension or lack of time. There was no correction for multiple comparisons. The study and the analysis were drawn from previous studies reviewed earlier (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015) that have not used a method to correct for their multiple comparisons. Since the current study aimed to have comparable results to those studies corrections for multiple comparisons were not used here either. This had implications in the interpretation of results which are described later in the Discussion.

Table 5.4 – Results of Pearson's partial correlation analysis

Pearson's partial *r* values between attention and AP tests and number of children completing each test. The 95% CI of the correlation coefficient is also presented in brackets for all significant or marginally significant correlations. Gender and IQ performance were controlled in this analysis. Adv.: Advantage, AFG: Auditory Figure Ground, AP: Auditory Processing, CCC-2: Children's Communication Checklist – 2, CI: Confidence Interval, DDT: Dichotic Digits Test, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, GIN: Gaps-in-Noise, IQ: Intelligence Quotient, LiSN-S: Listening in Spatialized Noise – Sentences, NSL: Non-Standard Language, Sel-VA: Selective Visual Attention, SL: Standard Language, Sus-AA: Sustained Auditory Attention. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Sample size	**20**									**19**		**15**	**12**	**11**	**11**
	Sus-AA	**Div-AA**	**Div-AVA**	**Sel-VA**	**LiSN-S Low-cue**	**LiSN-S High-cue**	**LiSN-S Talker Adv.**	**LiSN-S Spatial Adv.**	**LiSN-S Total Adv.**	**CCC-2 SL**	**CCC-2 NSL**	**DDT**	**GIN**	**AFG 0 dB**	**AFG +8 dB**
Sus-AA	1	.76*** [.48, .90]	.51* [.09, .78]	.01	.23	-.01	-.09	.26	-.09	.53* [.10, .79]	.16	.53† [.02, .82]	.53	-.43	-.39
Div-AA		1	.53* [.11, .79]	.27	.25	.08	-.11	-.03	.00	.66** [.29, .86]	.24	.68* [.26, .88]	.27	-.62	-.42
Div-AVA			1	.37	-.15	-.02	.15	-.21	.05	.21	.01	.76** [.41, .92]	.63† [.08, .88]	.00	.14
Sel-VA				1	.07	.16	-.07	-.28	.15	.24	-.04	.45	-.16	-.17	.11
LiSN-S Low-cue					1	.22	.02	.35	-.15	.30	-.05	.09	-.53	.31	-.45
LiSN-S High-cue						1	.11	.12	.93*** [.83, .97]	.24	-.15	-.48	.12	-.31	-.15
LiSN-S Talker Adv.							1	.08	.10	-.05	-.22	.39	-.30	.18	-.04
LiSN-S Spatial Adv.								1	-.03	.09	-.01	-.31	-.12	-.11	-.60
LiSN-S Total Adv.									1	.13	-.14	-.49	.34	-.41	.10
CCC-2 SL										1	.59* [.19, .82]	.51† [-.00, .80]	.00	-.62	-.34
CCC-2 NSL											1	.19	.01	-.22	-.03
DDT												1	.04	-.04	.34
GIN													1	.04	.06
AFG 0 dB														1	-
AFG +8 dB															1

The Div-AA and Div-AVA subtests showed strong associations with the DDT scores, *r*(13) = .68, *p* = .010, and *r*(13) = .76, *p* = .003, respectively, which confirmed the hypotheses. The scatterplots for these two correlations are presented in Figure 5.1 below. The Sel-VA also confirmed the initial hypothesis as it did not present a significant correlation with any of the AP tests. The Sus-AA scores did not demonstrate a significant association with the DDT scores, *r*(13) = .53, *p* = .063, 95% CI [.02, .82], which did not agree with the hypothesis. However, the correlation was marginally non-significantly and the confidence interval wide. The scatterplot for this relationship is presented in Figure 5.2. The same figure also includes the correlations between Sel-VA and one of the AP tests (i.e. the DDT), which as stated earlier did not correlate with any of the AP tests and thus did not have a significant association with the DDT, *r*(13) = .45, *p* = .126. The GIN test also recorded a marginally non-significant correlation with the Div-AA test, *r*(10) = .63, *p* = .053, 95% CI [.08, .88]. Additionally, some of the TEACh subtests intercorrelated between them. Specifically, the TEACh Sus-AA subtest correlated with the Div-AA and Div-AVA subtests, *r*(18) = .76, *p* < .001 and *r*(18) = .51, *p* = .031, respectively, while Div-AA correlated with Div-AVA, *r*(18) = .53, *p* = .024.

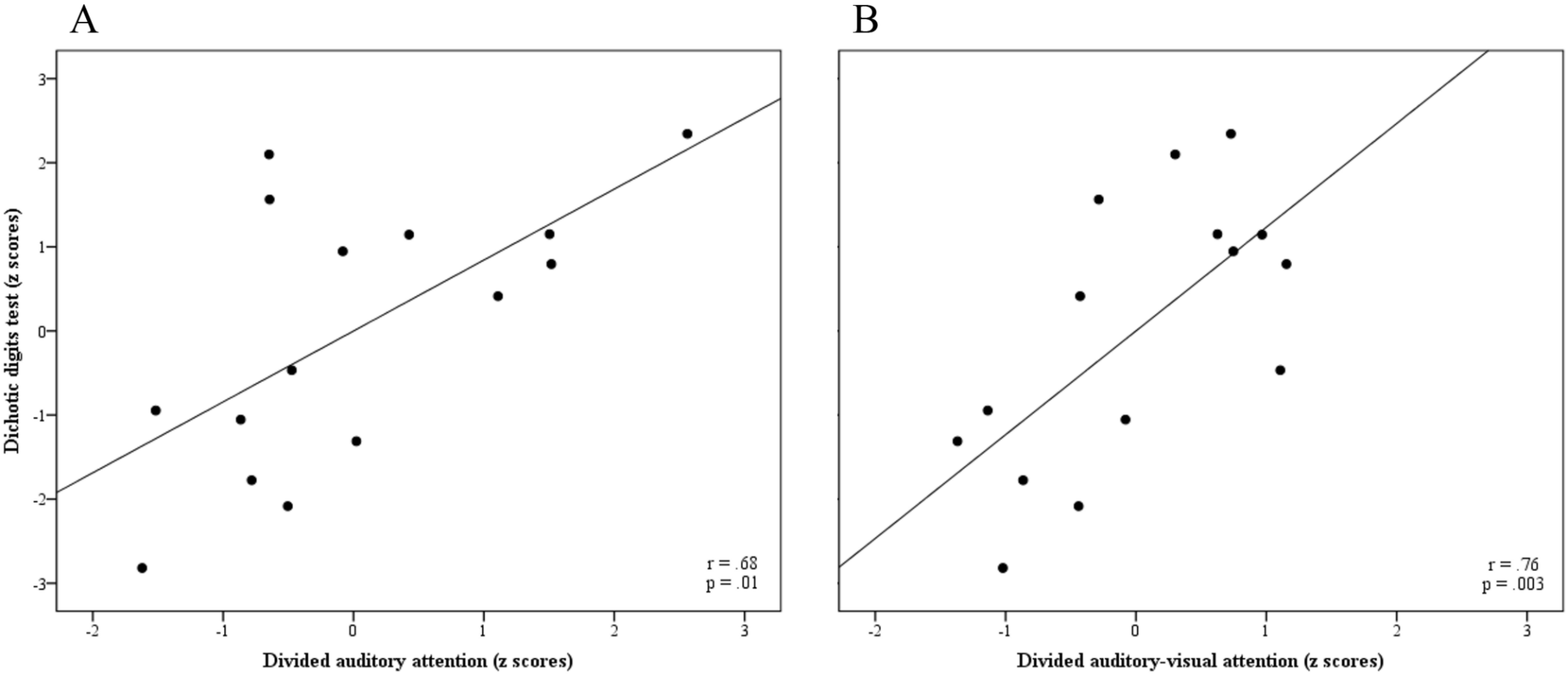

Figure 5.1 – Partial regression plots between the two divided attention tests and the DDT

Partial regression plots (A) between Div-AA and DDT z scores and (B) between Div-AVA and DDT z scores for 15 participants. Pearson’s partial correlation coefficients and significance values are also included in each graph. DDT: Dichotic Digits Test, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention.

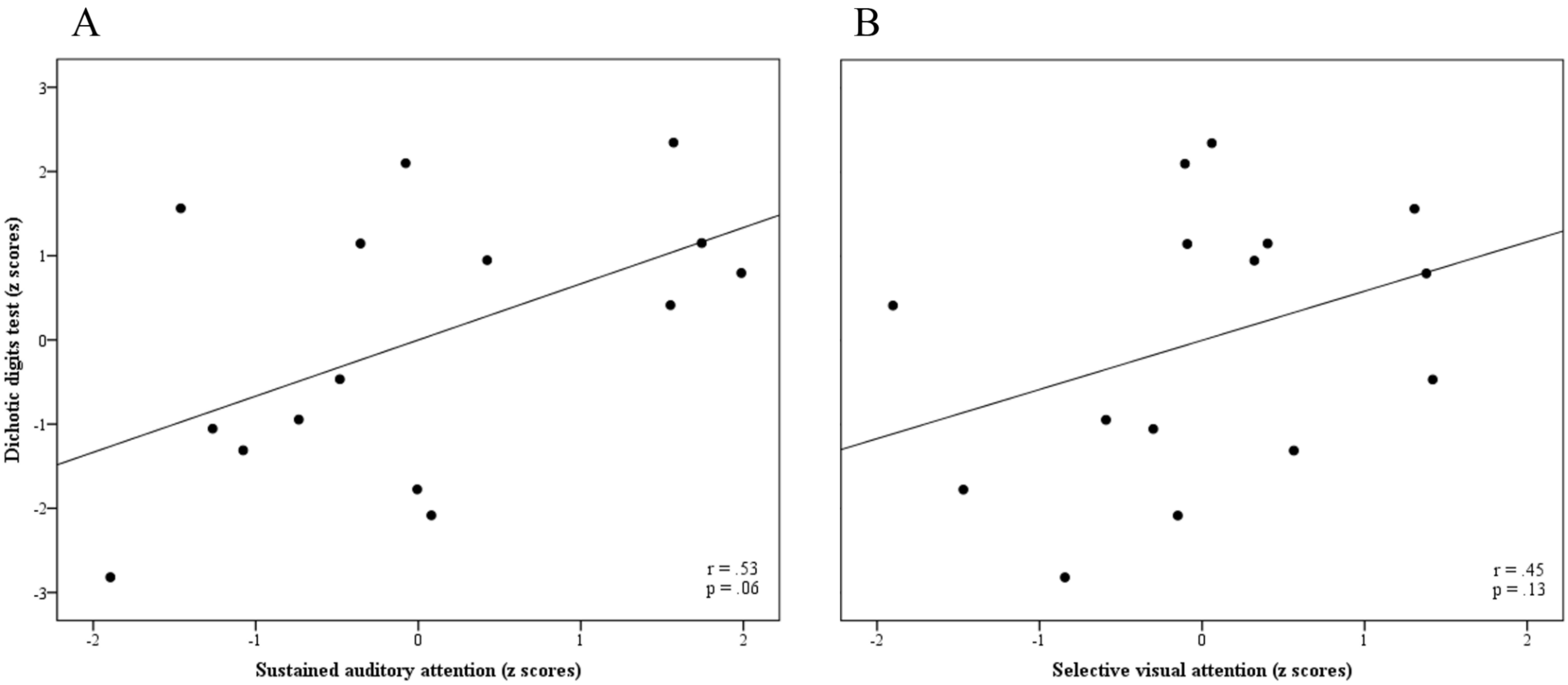

Figure 5.2 – Partial regression plots between Sus-AA/ Sel-VA and the DDT
Partial regression plots (A) between Sus-AA and DDT z scores and (B) between Sel-VA and DDT z scores for 15 participants. Pearson's partial correlation coefficients and significance values are also included in each graph. DDT: Dichotic Digits Test, Sel-VA: Selective Visual Attention, Sus-AA: Sustained Auditory Attention.

Furthermore, the CCC-2 SL scores correlated with both the Sus-AA scores and the Div-AA scores, *r*(17) = .53, *p* = .027 and *r*(17) = .66, *p* = .004, respectively. These two relationships are presented in scatterplots in Figure 5.3 below. The CCC-2 SL had a marginally non-significant association with the DDT scores, *r*(12) = .51, *p* = .093, 95% CI [-.00, .80]. In addition, there were intercorrelations between two LiSN-S conditions and between the two CCC-2 composite scores. In detail, the LiSN-S High-cue SRT correlated with the LiSN-S Total advantage, *r*(18) = .93, *p* < .001 and the CCC-2 SL composite score had a positive relationship with the CCC-2 NSL score, *r*(17) = .59, *p* = .013.

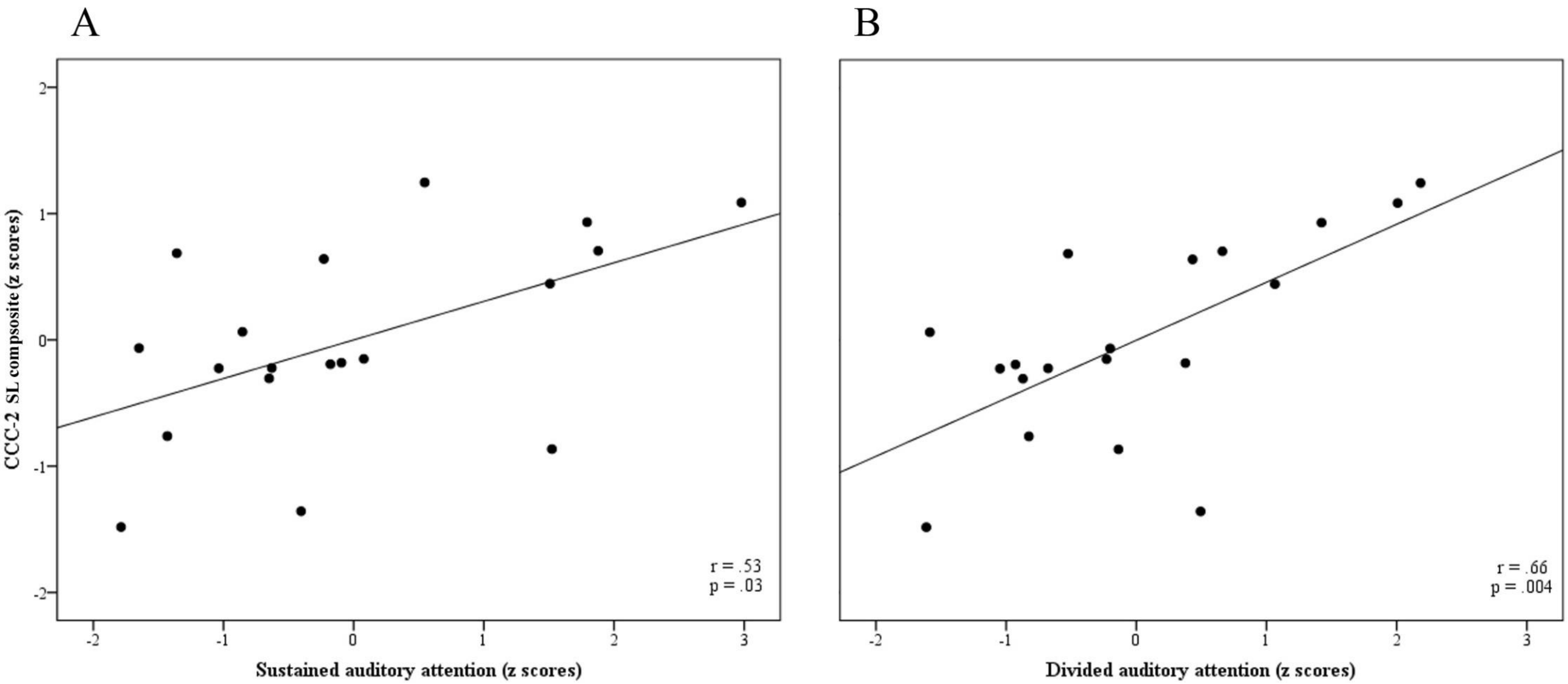

Figure 5.3 – Partial regression plots between Sus-AA/ Div-AA and the CCC-2 SL composite score

Partial regression plots (A) between Sus-AA and CCC-2 SL composite z scores and (B) between Div-AA and CCC-2 SL composite z scores for 19 participants. Pearson's partial correlation coefficients and significance values are also included in each graph. CCC-2: Children's Communication Checklist – 2, Div-AA: Divided Auditory Attention, SL: Standard Language, Sus-AA: Sustained Auditory Attention.

To further explore the marginally non-significant effects (see Table 5.4) on a larger sample and test whether these effects would become significant, post-hoc correlation analyses were run on the entire sample size (*n* = 27) of both children suspected of and diagnosed with APD. Therefore, post-hoc correlation analyses were run only for the three marginally non-significant correlations noted in Table 5.4 (i.e. between Sus-AA and DDT, between the Div-AVA and GIN, and between the CCC-2 SL and the DDT). Results showed that only the Sus-AA and DDT correlation was significant in these analyses, *r*(19) = .47, *p* = .041, 95% CI [.05, .75], while the correlations between Div-AVA and GIN, and between CCC-2 SL and DDT remained marginally non-significant, *r*(14) = .46, *p* = .097, 95% CI [-.04, .77] and *r*(17) = .47, *p* = .057, 95% CI [.02, .76], respectively. This post-hoc analysis though diverges from the study's research question which was to examine these relationships on a sample of children who only met APD diagnostic criteria. This is further discussed later in the Discussion.

5.3.3 Comparison and description of batteries

Table 5.5 presents the disordered performance of each child in the attention and AP tests. It also includes the four batteries used in the second part of the analysis and the cases which they identified, as well as the outcome of the CCC-2 questionnaire. A Cochran's Q test analysis was performed between the percentages of identification of the APD battery, the AD battery and the APD without ADs battery. The two assumptions of sample size adequacy were met (the sample size was larger than 4 [$n \geq 4$] and the sample size multiplied by the number of related groups [k = 3] was greater than 24 [$nk \geq 24$]). The yield of identification was significantly different between the three batteries, $\chi^2(2) = 11.20$, *p* = .004. Post-hoc pairwise comparisons were conducted using Dunn's procedure with Bonferroni correction to identify where the difference lay. The APD battery proportion of diagnoses was significantly higher than the one from the AD battery (74% over 30%, respectively), *p* = .003. The other comparison between the diagnostic yield of the APD battery and that of the APD without ADs battery was non-significant (74% over 44%, respectively), *p* = .085.

Table 5.5 – Individual performance and diagnostic yield of the batteries

Children’s performance in each attention and AP test based on SDs, and identification as disordered according to the four diagnostic batteries in the study; APD, AD, APD without ADs and i-APD. The outcome of the CCC-2 questionnaire is also included. Out of total numbers and percentages are included at the bottom. Black indicates performance dropping -3 SDs from the mean and dark grey -2 SDs from the mean (and -2.33 SDs for the AFG tests). For easier reading, all cases identified with a disorder under each battery and the CCC-2 were highlighted in light grey. Dash indicates incompletion of a test. AD: Attention Deficit, AFG: Auditory Figure Ground, AP: Auditory Processing, APD: Auditory Processing Disorder, ASD: Autism Spectrum Disorder, CCC-2: Children’s Communication Checklist – 2, D: Disordered, DDT: Dichotic Digits Test, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, DLD: Developmental Language Disorder, FPT: Frequency Pattern Test, GIN: Gaps-in-Noise, i-APD: inattentive subtype of APD, L: Left, LiSN-S: Listening in Spatialised Noise – Sentences, PLI: Pragmatic Language Impairment, R: Right, SD: Standard Deviation, Sel-VA: Selective Visual Attention, SPD: Spatial Processing Disorder, SRT: Speech Reception Threshold, Sus-AA: Sustained Auditory Attention, w/o: without.

Child	**Sus-AA**	**Div-AA**	**Div-AVA**	**Sel-VA**	**LiSN-S**			**AFG +8 dB**	**AFG 0 dB**	**DDT**		**GIN**	**FPT**		**APD battery**	**AD battery**	**APD w/o ADs battery**	**i-APD**	**CCC-2**
					High-cue SRT	**Spatial Adv.**	**Total Adv.**			**R**	**L**		**R**	**L**					
C01												-	D	D					PLI
C02					D			-	D			-	D	D	D		D		ASD
C03	D		D					-	D	D	D	D	-	-	D	D			DLD
C04								-	D						D		D		
C05			D					-	D										
C06						D		-	D			D			D		D		
C07		D			D		D	-					D	D	D/SPD		D	D	ASD
C08	D	D	D	D					-		D		-	-	D	D			DLD
C09								D	-	-	-				D		D		
C10			D			D		-	-	D	D	-	D	D	D		D	D	DLD
C11	D		D						-			D			D	D			PLI
C12			D	D		D	D	D	-		D		-	-	D/SPD		D	D	DLD
C13					D	D			-				-	-	SPD		D		PLI
C14	D		D								D	-	-	-	D	D			DLD
C15			D									-	-	-					DLD
C16		D	D					D	-		D	-	-	-	D	D			-
C17			D		D		D	-	D						SPD		D	D	
C18								D	D	-	-	-	-	-	D		D		PLI
C19											D		-	-					
C20	D	D	D		D		D	-	-	D		-	-	-	D/SPD	D			
C21								D	-										DLD
C22	D	D	D						D		D		-	-	D	D			PLI
C23									D				-	-					-
C24					D		D	-		-	-		-	-	SPD		D		DLD
C25	D					D				D			-	-	D		D	D	DLD
C26	D	D	D		D			D	-	D			-	-	D	D			DLD
C27			D							-	-	-	-	-					PLI
Out of total															20/27	8/27	12/27	5/27	18/25
Percentage															74%	30%	44%	18%	72%

There were 3 out of 27 children (11%) who did not meet criteria for typical APD but met criteria for SPD. Additionally, of the 17 children who met criteria for typical APD 3 also met criteria for SPD (18%). Lastly, Table 5.5 also includes the percentage of identification of the i-APD battery. Children identified as having i-APD were 18% of the initial sample (5 out of 27).

5.3.4 Post-hoc comparison analyses

Post-hoc independent samples t-test comparisons were run between the APD ($n = 20$) and non-APD group ($n = 7$) on the TEACh, LiSN-S SIN conditions, and CCC-2 composite scores, and on a separate analysis on the LiSN-S spatial conditions, the GIN, DDT, FPT, and the two AFG tests. Assumptions were met for all outcome measures except for the Div-AVA test[22]. This means that no outliers were observed (checking for values beyond 3 SDs above and below the mean), normality of distribution was met (using the Shapiro-Wilk test, $p > .05$), and the assumption of homogeneity of variances was also met (using the Levene's test of equality of variances, $p > .05$).

Figure 5.4 below summarises the findings from the comparisons on the TEACh, LiSN-S SIN conditions, and CCC-2 scores between the two groups in boxplots.

---

[22] Data were not normally distributed, but transformations would not yield normally distributed data, while transformations also produced outliers. Thus, it was decided to analyse the Div-AVA scores without transformations and highlight this violation of normality of distribution.

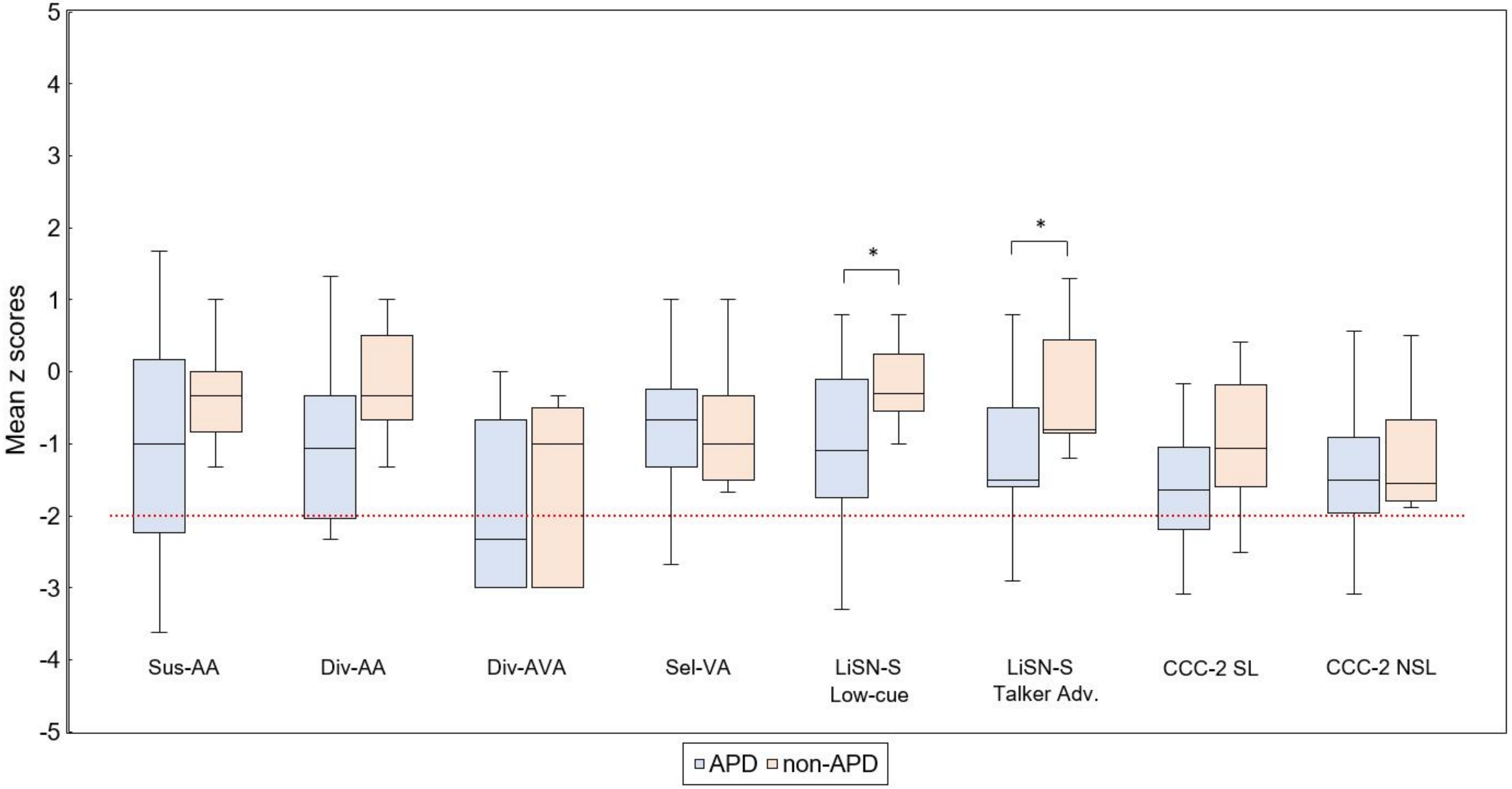

Figure 5.4 – Results for the APD vs. non-APD comparisons
Boxplots of mean z scores on the eight outcome measures for 20 children with APD and 7 non-APD children (and 19 vs. 6 for the two CCC-2 comparisons). The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below -2 z scores). Adv.: Advantage, APD: Auditory Processing Disorder, CCC-2: Children’s Communication Checklist – 2, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, LiSN-S: Listening in Spatialised Noise – Sentences, NSL: Non-Standard Language, SL: Standard Language. *$p < .05$

Figure 5.5 below summarises the findings from the comparisons on the LiSN-S spatial conditions, the GIN, DDT, FPT, and AFG tests between the two groups.

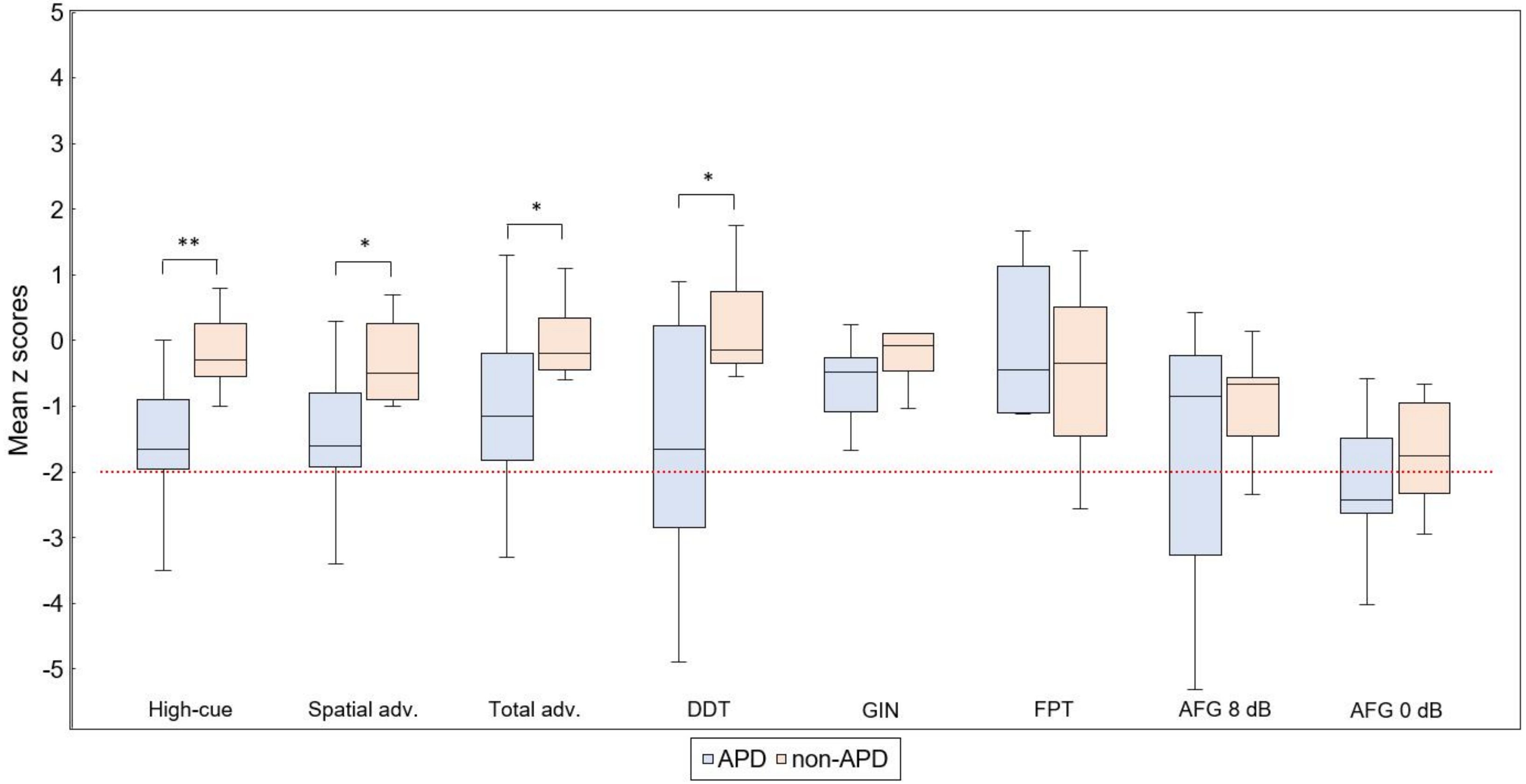

Figure 5.5 – Results for the APD vs. non-APD comparisons on the AP tests
Boxplots of mean z scores on the AP outcome measures for 20 children with APD and 7 non-APD children on the LiSN-S conditions (15 vs. 6 on DDT, 12 vs. 4 on GIN, 6 vs. 3 on FPT, 11 vs. 5 on AFG 8 dB, and 11 vs. 6 on AFG 0 dB). The line inside the boxplots represents the median, while the whiskers indicate the upper and lower observations. The red dotted line indicates the cut-off for abnormal performance (i.e. below -2 z scores, for AFG conditions abnormal performance is below -2.33 z scores). Adv.: Advantage, AFG: Auditory Figure Ground, AP: Auditory Processing, APD: Auditory Processing Disorder, DDT: Dichotic Digits Test, FPT: Frequency Pattern Test, GIN: Gaps-in-Noise, LiSN-S: Listening in Spatialised Noise – Sentences. *$p < .05$, **$p < .01$.

Table 5.6 below presents the descriptive statistics along with the significance values of the eight primary outcome measures and the eight AP tests for the APD and non-APD group.

Table 5.6 – Descriptive statistics and significance values
Mean values, SDs, *p* values, effect sizes and 95% CI of the effect size for the TEACh subtests, the LiSN-S conditions, the CCC-2 composite scores, and the AP tests (i.e.. LiSN-S spatial conditions, DDT, GIN, FPT, and the two AFG tests. AFG: Auditory Figure Ground, AP: Auditory Processing, APD: Auditory Processing Disorder, CCC-2: Children's Communication Checklist – 2, CI: Confidence Interval, DDT: Dichotic Digits Test, Div-AA: Divided Auditory Attention, Div-AVA: Divided Auditory-Visual Attention, FPT: Frequency Pattern Test, GIN: Gaps-in-Noise, LiSN-S: Listening in Spatialised Noise – Sentences, NSL: Non-Standard Language, SD: Standard Deviation, SL: Standard Language, SRT: Speech Reception Threshold. $^{\dagger}p < .10$, $*p < .05$, $*p < .01$

Test Condition	**Mean z scores (SD)**		***p* value**	***MD* [95% CI of *MD*]**
	APD	**Non-APD**		
TEACh				
Sus-AA	-1.02 (1.60)	-.33 (.83)	.288	-.69 [-2.01, .62]
Div-AA	-.97 (1.15)	-.14 (.85)	.096†	-.83 [-1.82, .15]
Div-AVA[23]	-1.90 (1.10)	-1.61 (1.31)	.585	-.28 [-1.33, .76]
Sel-VA	-.70 (1.03)	-.76 (.97)	.891	.06 [-.86, .98]
LiSN-S				
Low-cue SRT	-1.11 (1.12)	-.15 (.65)	.046*	-.95 [-1.88, -.01]
Talker advantage	-1.20 (.83)	-.21 (.97)	.016*	-.98 [-1.77, -.19]
CCC-2				
SL composite	-1.68 (.82)	-.98 (1.09)	.106	-.70 [-1.56, .16]
NSL composite	-1.41 (.98)	-1.12 (.95)	.522	-.29 [-1.25, .65]
AP tests				
High-cue SRT	-1.52 (1.23)	-.10 (.77)	.009**	-1.42 [-2.45, -.39]
Spatial advantage	-1.58 (1.32)	-.30 (.69)	.023*	-1.28 [-2.36, -.19]
Total advantage	-1.05 (1.25)	.01 (.62)	.042*	-1.06 [-2.08, -.03]
DDT	-1.57 (1.74)	.25 (.91)	.026*	-1.82 [-3.41, -.24]
GIN	-.79 (.95)	-.26 (.54)	.318	-.52 [-1.62, .56]
FPT	-.71 (2.58)	-.51 (1.97)	.909	-.20 [-4.26, 3.85]
AFG 8 dB	-1.69 (1.89)	-.97 (.94)	.331	-.71 [-2.24, .81]
AFG 0 dB	-2.17 (1.07)	-1.72 (.90)	.394	-.45 [-1.55, .64]

For the tests not used to diagnose APD, significant differences were found in the two LiSN-S non-spatial conditions Low-cue and Talker advantage, with the APD group scoring significantly lower than the non-APD group, *t*(25) = -2.09, *p* = .046, *M* = -.95 (95% CI, -1.88 to -.01), and *t*(25) = -2.07, *p* = .016, *M* = -.98 (95% CI, -1.77 to -.19). For the AP tests used in diagnosing APD the three LiSN-S spatial conditions and the DDT had significantly lower scores in the APD group than the non-APD group, *t*(25) = -2.83, *p* = .009, *M* = -1.42 (95% CI, -2.45 to -.39) in the High-

---

[23] Not normally distributed data but transformations would not make the data normally distributed, while they also produced outliers.

cue SRT condition, *t*(25) = -2.42, *p* = .023, *M* = -1.28 (95% CI, -2.36 to -.19) in the Spatial advantage condition, *t*(25) = -2.13, *p* = .042, *M* = -1.06 (95% CI, -2.08 to -.03) in the Total advantage condition, and *t*(25) = -2.41, *p* = .026, *M* = -1.82 (95% CI, -3.41 to -.24) in the DDT test.

To compare the percentages of AD identification in the APD and non-APD groups, the chi-square test of homogeneity was initially run. Table 5.7 below summarises the findings.

Table 5.7 – Percentages of AD identification
Percentages and out of total numbers of AD identification for the APD and non-APD group. AD: Attention Deficit, APD: Auditory Processing Disorder.

Group	ADs	No ADs
	Out of total (%)	Out of total (%)
APD	9/20 (45%)	11/20 (55%)
Non-APD	0/7 (0%)	7/7 (100%)

As the sample size was not sufficient (i.e. expected frequency below 5 in two cells), Fisher's exact test was run and interpreted instead. The difference between the two binomial proportions (.45 - .00 = .45) was statistically significant, $\chi^2(1) = 6.49$, *p* = .022, 95% CI [.23, .66]. This means that the APD group had significantly more AD identifications than the non-APD group.

To compare the percentages of identification from the CCC-2 (i.e. DLD, PLI, ASD) in the APD and non-APD groups, the chi-square test of homogeneity was initially run. Table 5.8 below summarises the findings.

Table 5.8 – Percentages of CCC-2 identifications
Percentages and out of total numbers of DLD, PLI, and ASD identifications through the CCC-2 for the APD and non-APD group. Two cases (one from each group) had missing data, hence they were not included in the analysis. APD: Auditory Processing Disorder, ASD: Autism Spectrum Disorder, CCC-2: Children's Communication Checklist – 2, DLD: Developmental Language Disorder, PLI: Pragmatic Language Impairment.

Group	DLD	PLI	ASD	No identification
	Out of total (%)	Out of total (%)	Out of total (%)	Out of total (%)
APD	8/19 (42%)	4/19 (21%)	2/19 (11%)	5/19 (26%)
Non-APD	2/6 (33%)	2/6 (33%)	0/6 (0%)	2/6 (33%)

Since the sample size was not adequate (i.e. expected frequency below 5 in two cells), Fisher's exact test was run and interpreted instead. The difference between the multinomial proportions of each CCC-2 category (.09 for DLD, -.12 for PLI, and .11 for ASD) was not statistically significant, $\chi^2(3) = 1.08$, $p = 1.000$, (95% CI for DLD [-.34, .52], 95% CI for PLI [-.53, .29], 95% CI for ASD [-.02, .24]). Therefore, the APD group did not have significantly more (or fewer) identifications in any of the CCC-2 categories (i.e. DLD, PLI, ASD) compared to the non-APD group.

5.4 Discussion

A post-hoc power calculation was performed on the actual sample size ($n = 20$) to detect 1 SD effect size at 5% significance, which resulted in a statistical power of 68%, while there were instances where the sample size dropped down to 11. This means that the study was under-powered and the chance of detecting a statistically significant difference (if a difference exists) was down to 68% or less. This may have compromised the detection of significant effects, whereas another consequence of low power is the imprecision of the effect size (Nakagawa & Cuthill, 2007) which can be inflated, thus showing a greater magnitude than the actual effect size of the population. This may have been the case for the significant (and marginally non-significant) findings of this analysis which were presented in Table 5.4 and are discussed in the following sections. Therefore, interpretation of the results cannot be safely made, and a larger sample sized study is required to make sense of the findings. Furthermore, as the study did not correct for multiple comparisons there was an increased chance of making a Type I error. This means that an effect could have been detected where no such effect truly exists. This also creates a problem with the interpretation of the findings and any claims of a relationship between variables should be carefully made with this in mind. Therefore, the significant relationships noted in the findings and in the discussion below may have been a result of a false positive effect.

5.4.1 Sustained auditory attention

Literature up to now examined the relationship between Sus-AA and the DDT only in mixed samples of children suspected of and diagnosed with APD and found positive correlations between the two (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). In the current study this relationship was tested on a sample of children only diagnosed with APD and no significant correlation between Sus-AA and the DDT was found. But the *p* value was marginally non-significant (*p* = .063), while the study was underpowered as noted above. Underpowered studies as mentioned earlier may falsely miss a truly significant effect. This could have been the case here. To further examine this relationship between Sus-AA and DDT a post-hoc correlation analysis on the entire sample (including children suspected of APD) was run and was compared to the initial APD-only sample (see Results Section 5.3.2 for the reporting of the findings). The relationship between the two variables in the post-hoc correlation analysis was now significant (*p* = .041). This could mean that the 7 additional children suspected of APD may have helped drive this correlation and that APD-diagnosed children's DDT scores may not have been affected by their sustained attention performance or vice versa. It could also mean, though, that the improved statistical power after the inclusion of 7 more cases (albeit not meeting the study's aims for including only APD-diagnosed children) was sufficient in making a marginally non-significant effect significant.

Furthermore, the study had a small sample size and the correlation coefficient had a wide confidence interval. The combination of these two limiting factors make it impossible to draw safe conclusions on the relationship between Sus-AA and the DDT. The rest of the AP tests (i.e. GIN, LiSN-S, AFG + 8dB and AFG 0 dB) did not correlate with Sus-AA, which agrees with what previous studies found in children suspected of APD (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015). From the 20 children diagnosed with APD, 8 of them (40%) had disordered performance on the Sus-AA task (see Table 5.5). Pairing this finding with the lack of significant association between Sus-AA and AP tests found in this study may indicate that this deficit in sustained attention is inherent in these children and is not influenced by their deficits in AP skills. Heterogeneity of APD could be further supported with this finding as Sus-AA deficits were only present to a subgroup of children diagnosed with APD (40%).

5.4.2 Divided attention

Previous research has not examined the relationship between divided attention and AP skills in APD samples up to now. The present study has found strong correlations between the two measures of divided attention and the DDT. This makes it the first study to examine this relationship and reveal positive correlations between Div-AA/ Div-AVA and the DDT in children with APD. It is not surprising to find correlations between divided attention tasks and the DDT, as both functions demand simultaneous focus on two separate stimuli. The two tests may measure the same psychometric construct of divided auditory attention. After all, both tests make use of dividing focus into two stimuli. However, the tasks employed by the TEACh Div-AA test and the DDT are not identical. The latter involves simultaneous dichotic listening of pairs of digits through headphones. In the TEACh Div-AA task children count the number of specific tones (presented in a random, infrequent pattern) while trying to detect a target word in a news report, a test presented through speakers in the room instead of headphones. The stimuli and the presentation of them are different in these two tests. Therefore, even though both tests might be making use of the same psychometric construct (i.e. divided attention) the use of other functions by each test cannot be excluded since there are differences between the tasks and presentation of the two tests. For example, the DDT because it is presented dichotically through headphones it is used to assess binaural integration skills (Guenette, 2006). On the other hand, as the stimuli in the TEACh Div-AA are presented monaurally through speakers in the room these same binaural integration skills cannot be assessed through administration of the TEACh Div-AA test. Moreover, in the TEACh Div-AA test the fact that one of the tasks involves the counting of random infrequent tones, which is the same task used in the TEACh Sus-AA test, may mean that Sus-AA also contributes to the performance on this test more so than it does for the DDT. This argument can also be supported by the current study’s findings, which showed that Sus-AA had a significant strong relationship with Div-AA but did not significantly correlate with the DDT in children who met APD criteria. Of course, the correlation coefficient between Sus-AA and the DDT was numerically high but marginally non-significant (i.e. p = .063) and given that the study was underpowered this non-significant result may have been the product of the small sample size (as argued earlier in Discussion in Section 5.4.1). Thus, while it does appear that the two tests of Div-AA and DDT may measure the same function of divided attention to some extent, the contribution of other psychometric constructs to each test cannot be excluded.

An alternative interpretation of the findings may be that both functions of dividing attention between tasks (auditory or bimodal tasks) and dichotic listening share or activate the same regions in the brain. Studies have shown that the corpus callosum could be regulating both divided attention (Hutchinson et al., 2008; van der Knaap & van der Ham, 2011) and dichotic listening (Musiek & Weihing, 2011; Westerhausen & Hugdahl, 2008) and can thus be proposed that interhemispheric transfer in the auditory modality occurs in both Div-AA and the DDT. In this case, the DDT would be able to detect poor dichotic listening skills in children, but it would not be possible to exclude deficits in divided attention. It is therefore suggested that the DDT should be administered in combination with divided attention tests during APD assessments. This will help clinicians have a more complete profile of the patient's dichotic listening skills and abilities in divided attention tasks. At the same time, research should further explore the relationship between these two variables in children with APD.

As seen in Table 5.5, there were 6 out of 20 children (30%) with poor Div-AA scores and 11 out of 20 (55%) with poor Div-AVA scores. These large percentages of disordered scores in these two tests could indicate an underlying deficit in divided attention in children diagnosed with APD. As cognitive load appears to increase in divided attention tasks (Mattys & Palmer, 2015), it could be that this subgroup of children with APD that face divided attention deficits have a compromised ability to allocate resources when performing tasks that require attention to divide (Forster, Robertson, Jennings, Asherson, & Lavie, 2014; Lavie, 2005). The fact though that both divided attention tasks intercorrelated but also correlated with the Sus-AA task, could mean that this presentation of deficits in divided attention may be attributed to inherent sustained attention deficits, or that one divided attention task influences the other.

Even though the TEACh manual did not test for correlations between its subtests (Manly et al., 1999), the intercorrelations observed here could be explained by the similar tasks that these tests use. Specifically, the Sus-AA task is the same as one of the auditory tasks in the Div-AA subtest and very similar to the auditory task in the Div-AVA subtest (see Table 3.4 in Chapter 3 for details). The overlap in these attention tasks though, is further substantiated by the literature which shows shared circuit activations between tasks of different types of attention (Sarter et al., 2001). These correlations between Sus-AA, Div-AA and Div-AVA could also be the reason for the larger percentages of disordered performance in these tests (30%, 22% and 52%, respectively; see Table 5.3), as failure in one test might produce poor performance in the other. Finally, the two divided attention measures did not correlate with the rest of the AP tests, indicating that the GIN, LiSN-S and AFG tests are not influenced by divided attention. As noted earlier, these AP tests did not have a significant association with Sus-AA either, suggesting that the GIN, LiSN-S and SCAN-3 C AFG tests are good tools to use in APD assessment, since they appear dissociated from attention influences.

5.4.3 Selective visual attention

Selective visual attention did not correlate with the AP tests in the sample of children diagnosed with APD. This is another novel finding, as research up to now only examined Sus-VA in correlation studies in mixed samples of children suspected of and diagnosed with APD. This finding could be interpreted as an absence of broader ADs in children with APD. However, the small sample size in combination with the wide correlation coefficient confidence interval means that these findings should be interpreted cautiously. In addition, a study has shown that children with APD had worse Sel-VA scores compared to children suspected of but not diagnosed with APD (Allen & Allan, 2014). Interpreting the results from Table 5.5 in the present study may further support the findings from Allen and Allan (2014). The 10% of children diagnosed with APD (2 out of 20) in the current study who failed the visual attention task may appear as a low percentage, especially when compared to the 30%-55% disordered rates of the auditory attention tests. Nonetheless, it is higher than the 2% threshold in the TEACh’s normative data (Manly et al., 1999). The two cases with failed Sel-VA scores had disordered scores in some or all three auditory attention tests. Specifically, one of these two children failed all three auditory attention tests, while the other child had disordered scores in the Div-AVA to go along with his/ her poor Sel-VA poor performance. Even though it is difficult to draw conclusions based on these two cases alone, another study found that in children who performed poorly in AP measures (e.g. frequency discrimination, temporal processing and SIN) those AP measures had a positive relationship to a task of visual alertness (Moore et al., 2010). As the frequency discrimination and temporal processing tests were derived measures, they ensured that variability from procedural and cognitive demands of the tasks were controlled. From this, the authors concluded that attention (along with other cognitive functions such as memory) might be contributing to the AP difficulties. They also suggested that this finding can inform clinical practice which needs to focus on improving attention in children with poor AP skills (Moore et al., 2010). The suggestions from the current study are in line with Moore et al.'s (2010) recommendations, in that the information received from these attention tests should be used to guide management. This is discussed in more detail and in the context of the comparisons between the different batteries in Section 5.4.5 below. Furthermore, another suggestion for future research could be to include paired tests of attention across both modalities. The present study used different types of attention, but they were not paired across the auditory and visual modalities. For example, there was a Sel-VA task but not a Sel-AA task, a Div-AA task but not a divided visual attention one. Therefore, inclusion of different types of attention that cover both modalities could help better characterise the nature of ADs in children with APD and whether these children are at risk of broader ADs.

5.4.4 LiSN-S

The LiSN-S conditions did not significantly correlate either with the attention tests or with the CCC-2 subscales. For the CCC-2 subscales this is explained by the fact that the LiSN-S was developed so that its derived measures (i.e. the three advantage conditions) minimise the influence of language on task performance (Cameron & Dillon, 2007a). Based on the present study's findings, it appears that the LiSN-S conditions might not be influenced by attention either. Another study looked at the relationship between three LiSN-S conditions (the Low-cue SRT, High-cue SRT and Spatial advantage) and sustained attention (auditory and visual) in children suspected of APD and found no correlations between them (Tomlin et al., 2015). The data in the present study further support and extend Tomlin et al.'s (2015) findings, in that other types of attention (e.g. divided attention and Sel-VA) did not associate with LiSN-S conditions either. Even though Sus-AA and the LiSN-S Low-cue SRT subscale did not correlate in either the current or Tomlin et al.'s (2015) study, the National Acoustic Laboratories (NAL; 2010), which developed the LiSN-S test, supports that the Low-cue SRT condition could indicate auditory fatigue in children, as it is the last condition of the test. However, this claim is not supported by research findings. In any case, this relationship between Sus-AA and the Low-cue SRT LiSN-S condition would require further investigation to clarify this argument. As noted earlier, the LiSN-S derived measures were designed to minimise language factors (Cameron & Dillon, 2007a) and it has been supported that this may be extended to the controlling of cognitive factors, too (Tomlin et al., 2015). This would possibly explain the lack of associations that the LiSN-S Talker, Spatial and Total advantage had with the four TEACh subtests (i.e. Sus-AA, Div-AA, Div-AVA and Sel-VA) in the present study.

The LiSN-S Low-cue SRT and Talker advantage scores that measure SIN abilities when spatial cues are eliminated were expected to correlate with the two AFG tests since they also measure SIN. However, such correlations were not observed between these tests. It is likely that this may be due to the inadequate sample size of the study. As discussed earlier, a low powered study can increase the chance of making false negative errors (that is, to miss an effect where an effect actually exists in the population). It is hence possible that in adequately powered studies correlations between the LiSN-S SIN conditions and the AFG tests would be detected. Another interpretation of this could be the fact that the LiSN-S Low-cue SRT condition is said to be sensitive to auditory fatigue (see discussion earlier in the previous paragraph of this section) whereas the AFG tests are not discussed in that context. Moreover, the LiSN-S Talker advantage score is a derived measure that controls language factors (see Section 3.1.3 for details) while the two AFG tests are not derived measures (thus do not control language factors). It could thus be argued that the lack of association between the LiSN-S Talker advantage and the AFG scores may have been influenced by the fact that the former controls for language factors while the latter test does not. In any case, these findings require further examination in studies with higher statistical power.

Examining the intercorrelations between the LiSN-S subscales, it was revealed that the High-cue SRT and Total advantage subscales had a significant strong correlation between them. This is not surprising given that both these subscales are sensitive to pitch and spatial cues (NAL, 2010). Finally, the LiSN-S conditions did not correlate with the rest of the AP tests, which is in line with previous findings (Cameron et al., 2006; Cameron & Dillon, 2008). It is therefore suggested that the LiSN-S is a useful test to use in APD diagnostic batteries, since it appears that its different subscales do not associate with other AP skills, language factors or attention tests. At the same time the LiSN-S identified cases with SPD that did not meet criteria for typical APD. In the current study three children who failed to meet APD criteria did meet criteria for SPD, while three other cases met criteria for both APD and SPD. This further supports the argument for inclusion of the LiSN-S in any APD battery given that the LiSN-S is able to identify a separable subtype of APD.

5.4.5 CCC-2

The CCC-2 SL composite score had positive moderate correlations with Sus-AA and Div-AA. Thus, it could be argued that in this group of children aspects of standard language may be influenced by Sus-AA and Div-AA or vice versa. On the other hand, the CCC-2 NSL composite score did not exhibit significant correlations with any of the attention or AP tests. This suggests that the non-standard language aspects in children diagnosed with APD are not influenced by other attention or AP deficits (or vice versa). Furthermore, the SL and NSL composite scores had a positive correlation between them. This is not surprising as literature has found deficits in pragmatic language in children with DLD (Osman, Shohdi, & Aziz, 2011).

Moreover, there were 14 out of 20 children with APD (70%) that presented with language and communication impairments according to the CCC-2 (see Table 5.5). Other studies did not report these percentages in APD-diagnosed children, but Ferguson et al. (2011) compared the CCC-2 scores between children with APD, DLD and children of typical development. They found that the APD and DLD groups had worse scores than the typically developing group and that the two clinical groups did not differ between them (Ferguson et al., 2011). Ferguson et al.'s (2011) findings along with the large proportion of children with APD with language impairments (i.e. 70%) in the present study’s sample, suggest that children with APD might exhibit similar symptoms to the ones that children with DLD, PLI or ASD exhibit or even that they might present with co-occurring language disorders. The CCC-2 can thus be used in the assessment of APD to identify children that are at risk of language and communication impairments (Ferguson et al., 2011). Clinicians can then use this information to include other professionals in the treatment plan that will help manage these deficits in language and communication. Consequently, it is proposed that the CCC-2 should only be used as a screening tool for language deficits and that it should not be administered as an APD diagnostic measure.

5.4.6 Diagnostic APD criteria and APD management

A newly proposed procedure for managing APD has been examined through the use of different batteries. First, the standard APD battery was compared to the AD battery, showing that the former significantly identified more cases with APD than the latter identified cases with ADs. This shows that more children in the sample would meet criteria for APD diagnosis than criteria for AD diagnosis. This was expected as previous studies, even though not statistically comparing the two proportions, found that in samples of children suspected of APD there were more cases of APD than AD (Gyldenkærne et al., 2014; Sharma et al., 2009). In the present study, to be identified as having ADs, children had to have at least two TEACh subtest scores -2 SDs from the mean. The 8 children meeting this criterion also fulfilled criteria for APD diagnosis, pointing to possible co-occurrence of ADs with APD in these children. It can also be argued that these children's AP deficits might be affecting their attention or vice versa. Therefore, the question is raised, should cases that have poor attention scores not be given a diagnosis of APD and thus be excluded from the APD diagnostic procedure? To address this question, a battery that identified children with APD (based on the standard APD battery) but excluded cases meeting criteria for the AD battery was proposed. The diagnostic yield of this battery was not significantly different from the APD battery, which means that it could be used in APD diagnosis without significantly reducing the percentage of children identified with AP impairments. However, the small sample size used in the comparisons prevents any further interpretation of the results. In any case, the purpose of the study was not to propose these batteries instead of the APD diagnostic battery currently used. To say that, would mean that a subgroup of children who clearly present AP difficulties would be left undiagnosed and hence untreated, since more stringent criteria would be used with a hypothetical use of those batteries. The benefit, though, of using these additional batteries was to better understand the characteristics that make up the difficulties of a child referred for APD assessment. In turn, by using this information more targeted interventions can potentially be given to these children. Therefore, the information obtained from these batteries should be directed at informing management of diagnosed children and not at providing differential diagnoses. The way the information from the batteries is being used to better inform management is further described in the following paragraphs.

The addition of an AD battery next to the standard APD battery helps clinicians have a better picture of the deficits children face and thus better steer management and treatment for these patients. Therefore, the 8 cases that were identified with ADs under the AD battery could receive more diverse but targeted management recommendations to address their multiple needs. As these children might be at risk of more general ADs, it is suggested that they receive recommendations that children with ADHD receive and see a professional who is qualified to diagnose and treat ADHD. An ADHD management plan may include behavioural therapy, teacher and classroom adaptations, special cognitive programmes that target attention and perhaps medication (Wolraich et al., 2011). The AP deficits of these children though, cannot be ignored and hence they should also receive management recommendations that deal with their AP deficits. The APD management plan for these children can focus on their specific AP skills that were found impaired. For instance, if binaural integration was found impaired, targeted AT on dichotic listening can be recommended. Children identified under the APD without ADs battery will follow the management plan that APD clinicians already use, which focuses on environmental modifications to improve SNR in the classroom, RMHA use and bottom-up AT (ASHA, 2005a; BSA, 2011b). Children may also receive top-down AT which usually includes training that targets language and cognition (ASHA, 2005a).

Finally, the i-APD battery identified 5 children with APD (18%) that did not meet criteria for co-occurring ADs but had only one failed attention measure. This battery completes the APD without ADs battery, as it identifies cases that the latter battery might miss. In the sample, the 5 children with i-APD would have been missed by the APD without ADs battery because it requires at least two attention tests to fail, whereas the i-APD can identify a specific attention deficit in only one attention test. These cases with i-APD are said to have an inattentive type of attention but not general attention deficits. The majority of attention tests that were outside normal range in these children with i-APD were mainly tests of divided attention. Four of the five cases had failed scores in either the Div-AA or the Div-AVA task, while Sus-AA deficits were observed in only one child. It could therefore be supported that perhaps the attention deficits in this subgroup of children with APD are limited to divided attention tasks. This may provide useful information to the management plan of these children. Therefore, in addition to the APD management approaches that children with i-APD should receive, they can perhaps also be prescribed with cognitive training that targets the development of divided attention (Kerns, Eso, & Thomson, 1999). However, it can be claimed that because of the intercorrelations observed between the attentions tests (see Results in Section 5.3.2), failure in only one test but not in other intercorrelated tests is unusual and hence poor performance in only one attention test may be due to other factors such as fatigue. This would mean that the information obtained from the i-APD battery may be inaccurate and that they do not represent genuine difficulties in a specific attention test. Therefore, an additional recommendation for children who fail only one attention test could be to ask for a retest at a second time-point on the failed task to ensure that performance is not influenced by fatigue and that they still meet the diagnostic criteria that were set.

5.4.7 Post-hoc comparison analyses

TEACh, LiSN-S and CCC-2 composite scores

The t-test comparisons between the APD and non-APD group on the TEACh, LiSN-S and CCC-2 composite scores showed that only the two LiSN-S SIN conditions (i.e. Low-cue SRT and Talker advantage) were better in the non-APD group than the APD group. All other outcome measures were not significantly different. A previous study with a similar design compared 10 children with APD (mean age 10.3) against 21 non-APD children (mean age 10.1) on the CHAPS Total score, and 8 against 21 on the Fisher’s Auditory Checklist (Dawes, Bishop, Sirimanna, & Bamiou, 2008). Findings from that study did not show any significant differences between the

APD and non-APD groups, with the authors concluding that there were no identifiable characteristics that would help distinguish children in the two groups, other than the AP tests used to categorise them into the groups in the first place. Absolute mean performance in each group in Dawes et al.'s (2008) study was similar and below typical performance for both the APD and non-APD group[24]. The findings by Dawes et al. (2008) along with the results from the current post-hoc analysis indicate that there are no significant differences in the examined outcome measures (i.e. TEACh, CCC-2, CHAPS, Fisher's Auditory Checklist) between children meeting APD criteria and children suspected of APD but not meeting APD criteria. The only exceptions were the LiSN-S Low-cue SRT and Talker advantage conditions, which are tests measuring SIN ability. This may indicate that performance in SIN tasks may be a measure that can distinguish children meeting APD criteria from non-APD children. However, given the small sample size and low power of the study (see relevant Section below for details) any conclusions on this should be carefully made.

AP tests between APD and non-APD group

An additional analysis was run to compare the AP tests often used in APD diagnostic procedures (i.e. DDT, GIN, FPT, LiSN-S spatial conditions, AFG tests) between the APD and non-APD group. The number of children in each group was smaller for most of these tests as opposed to the previous tests discussed above, as not all AP tests were administered in the APD clinic. Children meeting criteria for APD scored worse in all three LiSN-S spatial conditions and in the DDT compared to children suspected of APD but not meeting APD criteria. This may mean that children meeting APD criteria perform worse on these tests, however it can also be argued that children not meeting APD criteria may perform worse (but not poorly enough to warrant a diagnosis) on the rest of the AP tests not found significant (i.e. GIN, FPT, AFG). In any case, these results should be interpreted with caution given that they had even smaller sample sizes than the groups in the analyses of the rest of the outcome measures.

---

[24] In the CHAPS Total score the APD group had a mean score of -2.1, while the non-APD group -2. Scores below -0.05 were considered to be below typical performance (Smoski et al., 1998). In the Fisher's Auditory Checklist, the APD group received a mean of 48% and the non-APD group 47.8%. The cut-off for an abnormal performance was set at 72% or lower (Fisher, 1979).

Percentages of AD identification and CCC-2 classification

Comparing the percentages of AD identification, it was shown that the APD group had a higher percentage of children meeting criteria for AD than the non-APD group. This means that children that met APD criteria were more likely to also meet criteria for AD, whereas children suspected of APD but not meeting APD criteria were less likely to meet criteria for AD. This may indicate that problems in attention may be an area that can distinguish cases with APD from non-APD cases. Nevertheless, the previous comparisons between the two groups on the TEACh subtests showed that none of the four attention tests (i.e. Sus-AA, Div-AVA, Sel-VA, Div-AA) had significantly different scores between the APD and non-APD group. Moreover, as the statistical power for this analysis was low (see following Section for details) there is a chance that these results may have been inflated.

Furthermore, comparisons between the percentages of CCC-2 classifications (i.e. DLD, PLI, ASD) did not differ between the two groups. This may suggest that children that were suspected of APD but did not meet APD criteria could be facing similar language and communication problems that children with APD face, and these problems could even be the cause of the initial concerns by parents and professionals. In addition, as a percentage of children that met APD criteria also met with language and communication impairments identified through the CCC-2 this may mean that these children may have co-occurring language disorders, or that even these language and communication problems may be the cause of their APD symptoms (see discussion in Section 1.3.3 for more details on this).

The percentages of identification for each group are presented in Table 5.8. Direct comparisons of these percentages to percentages of CCC-2 identifications in a sample of children with typical development cannot be made. Despite the fact that during the development of the CCC-2 the questionnaire was given to parents of children with typical development for standardisation, only the mean scores of each subscale were reported in the manual without any mention of percentages of identification in the typical sample (Bishop, 2003). In addition, to the best of our knowledge no published studies were found that used the CCC-2 to report percentages of identification in samples of children with typical development. There are, however, UK prevalence rates for DLD and ASD that are presented in some studies. These are 6.4% for DLD (Scerri et al., 2011) and 1.5% for ASD (Baron-Cohen et al., 2009), whereas for PLI no information on prevalence was found.

Low statistical power

A post-hoc power calculation was performed for the t-test comparisons between APD and non-APD groups, which resulted in a statistical power of 60%. Similarly, a post-hoc power calculation was also conducted for the chi-square comparisons of AD identification between the two groups revealing a statistical power of 41%. This means that the studies were under-powered and the chance of detecting a statistically significant difference (if a difference exists) was 60% for the first post-hoc analyses and below chance (i.e. 41%) for the chi-square analyses. A negative consequence of low power is the increased chance of getting an imprecise effect (Nakagawa & Cuthill, 2007) which may be an inflated significant effect, thus showing a greater magnitude than the actual effect size of the population. This may have been the case for the significant findings of this analysis that showed that the LiSN-S Low-cue SRT and Talker advantage scores were better in the non-APD group compared to the APD group. The inflated effect sizes produced from low-powered studies could also explain the significant differences in AD identification in the APD-diagnosed group compared to the non-APD group. Additionally, a low powered study can increase the probability of getting a false negative effect. This means that true significant effects in the population have a higher chance of showing as non-significant in low powered studies. Thus, non-significant findings recorded in these post-hoc analyses could be a result of the low power. Interpretation of the results cannot be safely made, and a larger sample sized study is required to make sense of these comparisons.

5.4.8 Lack of multiple comparisons

As noted earlier, there was no correction for multiple comparisons in the analyses of this study. The study and the analysis were drawn from previous studies reviewed earlier that have not used a method to correct for their multiple comparisons (Gyldenkærne et al., 2014; Sharma et al., 2009;

Tomlin et al., 2015). Since the current study aimed to have comparable results to those studies corrections for multiple comparisons were not used here either. While correction for multiple comparisons reduces the chance for Type I error, it increases the chance for Type II error (Armstrong, 2014), and consequently reduces the statistical power of the study (McLaughlin & Sainani, 2014). As the statistical power of the study was already low correcting for multiple comparisons would have further reduced the power and this was something the study wished to avoid. Therefore, there are trade-offs with either method used and since the aim of the study was to have comparable results to the studies it drew its design from it was decided not to correct for multiple comparisons acknowledging the shortcomings, as described here in this paragraph. Not correcting for multiple comparisons increases the chance of making a Type I error, which means that a significant effect may be detected where no such effect truly exists. This means that the significant findings in the study (i.e. correlation of divided attention tasks to the DDT) may have been a false positive finding. This should be taken into account when interpreting the study's findings.

5.5 Conclusion

The correlation analyses suggest that the DDT interacts with measures of divided attention in samples of children with APD, but the direction of this interaction cannot be determined from the findings. While Sel-VA was not associated with any of the AP tests and only two cases had deficits in that type of attention, a subgroup of children with APD could still be at risk of general attention deficits (auditory and visual). To further explore the nature of attention deficits in children with APD and their relationship to AP skills, future research should include both auditory and visual attention measures of the same types of attention in larger sample-sized trials. The post-hoc analyses comparing performance between the group of children who met APD criteria and the non-APD group revealed that only SIN measures may be able to distinguish the two groups. Moreover, the APD group had higher percentages of ADs than the non-APD group, but in terms of language and communication problems (i.e. DLD, PLI, ASD) the two groups appear to be similar. This may mean that the language problems that non-APD children face may be the cause of the initial concerns by parents. As all analyses in this study had low power and there was no correction for multiple comparisons, results should be interpreted cautiously and further validated by larger sample-sized studies. The proposed diagnostic procedure that takes into account children’s deficits in attention has implications for clinical practice and APD management. The addition of the AD and i-APD batteries could improve the management of children referred for APD assessment. These diagnostic test batteries better classify children in categories that can address their specific AP and attention needs with targeted interventions. The batteries proposed here do not aim to offer differential diagnoses but only wish to better inform APD management.

CHAPTER 6 – DISCUSSION

Children diagnosed with APD are found to have worse SIN scores than their typically developing peers (Cameron et al., 2006; Cameron & Dillon, 2008; Lagacé et al., 2011). At the same time, they are often reported to have memory problems, poor attention span, and inattentive behaviours (AAA, 2010; ASHA, 2005b). Management of APD needs to be backed up by research evidence of the effectiveness of the proposed interventions and up to now RMHAs, which help improve SNR for children with APD, have not been adequately studied. The first three studies in Chapters 2, 3 and 4 used RCTs to test the effectiveness of RMHAs on working memory, listening-in-noise and attention in children with confirmed APD. The study in Chapter 5 aimed to further investigate the relationship between types of attention and AP skills, which were found to interact between them (Gyldenkærne et al., 2014; Sharma et al., 2009; Tomlin et al., 2015) and to compare performance between children with APD and children suspected of APD but not meeting APD criteria. It also examined a differential approach to assess and manage APD, proposing the inclusion of attention tests in APD diagnostic batteries. The findings of each study were discussed within each respective chapter. In this final chapter the implications of the findings will be discussed, and conclusions and suggestions for future directions will be drawn.

6.1 RMHAs on memory, attention and listening-in-noise ability

The studies in Chapters 2 and 3 were under-powered. Post-hoc power calculations showed that the actual statistical power of the studies was 49% and 70% respectively. This means that any interpretation of the findings from these studies is severely compromised. As discussed in detail in previous sections (Section 2.4.5 under Sample size and intervention period and Section 3.4.4 under Other limitations and future research), low-powered studies have an increased chance of inflating non-significant effects or not detecting significant effects. Therefore, safe conclusions cannot be drawn from these studies. The study in Chapter 4, which combined the data from both these studies, was able to improve the statistical power of the analysis and in turn improve the interpretations of the findings. In the following sections results from all three studies (in Chapters

2, 3 and 4) are discussed, but as noted here any interpretation from the findings of the studies in Chapters 2 and 3 should be made with reservation.

6.1.1 Working memory

Based on findings from the literature (Sharma et al., 2012; Umat et al., 2011), working memory was not expected to improve in the first study in Chapter 2 following a RMHA intervention. Results from that study met the hypothesis, as the 3-month RMHA intervention did not improve children’s score in the AWMA test. A number of limitations were identified that are discussed later, but it is noted here that the test used to measure working memory (i.e. the AWMA) may be prone to practice effects. This is because the scores on two of the three subtests (i.e. Listening recall and Listening recall processing) demonstrated an increasing trend in scores in both the control and the intervention group. This may have compromised the results and thus more studies investigating working memory after RMHA use are required. These new studies should employ better tests that are not sensitive to practice effects. As noted earlier, given the small sample size of the study these non-significant results may have been a product of the low statistical power of the study.

6.1.2 Listening-in-noise

Listening-in-noise was initially tested in the first study in Chapter 2 through the AFG subtests of the SCAN-3 C test. Those results did not reveal significant changes in children with APD after using RMHAs for 3 months, but the findings could have been compromised due to the low power of the study or the moderate test reliability of the AFG subtests (R. W. Keith, 2009). The LiSN-S test used in the next study in Chapter 3 had the advantage of being able to measure both SIN skills (in tasks where the target emanates from the same source as distractors) and spatial listening skills (where the frontal target is spatially separated from distractors coming from the side). This latter listening ability was not expected to change since RMHAs do not create an acoustic environment where spatial listening is trained. The fact that the LiSN-S spatial listening conditions remained unchanged, could mean that using RMHAs for an average of at least 2 hours daily for 6 months does not negatively influence spatial listening in children with APD. At the same time though, RMHAs do not significantly improve scores in conditions measuring SIN, despite claims by previous research that they do (Johnston et al., 2009). Therefore, it was concluded that RMHAs do not improve SIN ability as measured by the AFG subtests of the SCAN-3 C test and the SIN conditions of the LiSN-S test. A positive conclusion is that RMHAs also do not influence spatial listening as measured by the spatial listening conditions of the LiSN-S test, as there were clinical concerns that long-term RMHA use might hinder this ability. Nevertheless, the possibility that longer RMHA trials (e.g. more than 6 months) might negatively impact children's spatial listening skills cannot be excluded. The possibility that the decreased statistical power of the study caused significant findings to appear as non-significant cannot be excluded either. Therefore, future studies with longer trial periods and larger sample sizes should explore this hypothesis.

6.1.3 Attention

In the study in Chapter 3, the 6-month use of RMHAs was expected to improve the test scores in two types of attention. Sustained auditory attention remained unchanged, and this could mean that RMHAs do not have lasting effects on this type of attention, but problems with meeting the statistical assumptions of the Sus-AA scores are noted. It would be interesting for new studies to test Sus-AA in aided conditions, to examine whether sustained attention is improved while RMHA are in use. The study in Chapter 4 which combined the datasets from the two previous studies also did not show any significant difference in Sus-AA performance following a 3-month RMHA intervention. As the initial power calculation was met in this analysis in Chapter 4, interpretation of the study's findings holds more validity.

At the same time, Div-AVA demonstrated significant improvements at 6 months compared to baseline scores in the intervention group only. At the same time, however, the baseline scores of the intervention group were significantly worse than the scores of the control group. After controlling for the baseline scores there was no longer significant improvement from baseline to 6 months in the intervention group. Therefore, RMHA use did not provide any long-term benefit to performance in Div-AVA tasks either. A similar pattern was observed for the combined analysis in Chapter 4 where the increased sample showed the same baseline differences and initial improvement in the Div-AVA scores of the intervention group only. After controlling for the differences at baseline though, there were no longer improvements in the intervention group. Finally, in the studies in Chapters 3 and 4 there were also two other types of attention (i.e. Sel-VA and Div-AA) that were used as control conditions and expected to remain unchanged. This was confirmed by the findings. It was therefore concluded that Sel-VA is not helped by RMHA use as the system enhances the signal in the auditory modality and this is not expected to improve functions in the visual modality. Moreover, RMHAs help enhance the target signal over background noise but do not promote divided listening; thus Div-AA is also not assisted by using the system.

6.1.4 Findings from questionnaires

The subscales of the CHAPS questionnaire assessing SIN ability and Auditory attention span were hypothesised to improve following RMHA use, however they remained unchanged. These subscales were meant to reflect the behavioural tests of SIN and Sus-AA and the fact that they remained unchanged at least suggests consistency across these measures. Nonetheless, the CHAPS faces a number of limitations including use of raw scores instead of standardised scores and moderate test reliability (Smoski et al., 1998). In addition, the literature questions the validity of the CHAPS (Lam & Sanchez, 2007; Sharma et al., 2009; Wilson et al., 2011), while the questionnaire is not capable of distinguishing children with APD from children without APD (Iliadou & Bamiou, 2012). It is therefore concluded that even though the CHAPS is a tool often used in APD screening and in APD research, its use should be reconsidered given its limitations. Questionnaires with increased validity are required for APD screening and for APD research purposes. A study on adults suspected of APD found that the total scores and the majority of the subscales of two questionnaires (i.e. the Speech, Spatial, and Qualities of Hearing Scale and the Modified Amsterdam Inventory for Auditory Disability) had significant positive correlations with speech-in-babble tests, the DDT and the GIN tests, all standard AP tests (Bamiou, Iliadou, Zanchetta, & Spyridakou, 2015). This means that these questionnaires could prove useful tools in APD assessment and in intervention trials and could potentially be used instead of the CHAPS, but first they need to be validated on children populations as they currently have only been used in adult populations.

Contrary to the hypotheses, the experimental composite score of the CCC-2 questionnaire (a composite score of the four standard language subscales of Speech, Syntax, Semantics and Coherence) in the combined study in Chapter 4 remained unchanged in the intervention group. It was expected that the benefits in phonological awareness found in children with dyslexia after using RMHAs (Hornickel et al., 2012) were going to also be observed in this composite CCC-2 score in the present work's analysis with children with APD. The lack of significant findings though, could mean that RMHAs might have different effects on children with dyslexia than on children with APD. Additionally, the control non-standard language composite score of the CCC-2, remained unchanged thus confirming the initial hypothesis. It is therefore concluded that RMHAs do not affect the higher order language functions measured by the CCC-2 and it is argued that their use may only trigger benefits at sensory level. Moreover, it is supported that non-standard language aspects cannot be assisted by simply improving the SNR through RMHA use and perhaps targeted linguistic training can yield non-standard language benefits.

The LIFE-R questionnaire was given to children to complete to assess how well they hear their teachers through the RMHA system in various listening situations in the classroom. Children noted significant improvements at 3 and 6 months when listening through background traffic noise and when listening in the presence of competing speech. For the latter listening condition, controls reported significantly worse scores from 3 to 6 months. The Total score of the questionnaire also demonstrated significant improvement from baseline to 3 and 6 months in the intervention group. While these improvements can be attributed to the use of the RMHA, other conditions (e.g. listening when teacher is moving around or when teacher's back is turned) remained unchanged. As noted earlier in Section 3.4.3, visual cues are important when listening (Atilgan et al., 2018) and that could perhaps be a reason why children did not state improvements in those conditions. Furthermore, children did not report improvements with the RMHA when listening in speech-babble. It was argued that listening in speech-babble noise through the use of

RMHAs might be more challenging than when listening via the RMHA system in the presence of traffic noise, or when there is only one other competing speech. These inconsistencies noted in the findings from the LIFE-R questionnaire may also be due to the low statistical power of the study as discussed earlier.

Overall, the findings in the LIFE-R questionnaire suggest that use of RMHAs could improve children’s listening in some acoustically difficult situations at school, but not all. It iss therefore proposed that other environmental modifications and bottom-up approaches (ASHA, 2005b; BSA, 2011b) should be used alongside RMHAs, in order to maximise the quality of the acoustic input children with APD receive during classroom time. The LIFE-R questionnaire is adjusted to address children and it appears to be a tool that is easily administered and understood by children. But as the questionnaire does not have standardised scores, it is suggested that future studies should collect UK data to produce UK norms for the questionnaire. This would potentially improve the reliability of the LIFE-R and it may be used with more confidence in future APD studies.

The studies from Chapters 2, 3, and 4 are intended to be published in a peer-reviewed journal. The findings have important implications for researchers studying APD and for APD management. Other researchers could draw on the limitations and focus on possible benefits that RMHA may have on aided conditions, instead of looking for long-term improvements. The findings should also be considered by APD clinicians when designing their APD management plan. For instance, RMHAs should not be recommended on the basis of long-term improvement on the aspects of development studied in this thesis (i.e. attention, working memory, SIN), but on immediate improvement of the SNR that positively impacts SIN scores under aided conditions (see results in the study by Johnston et al., 2009).

6.1.5 Limitations and considerations for future directions

The two RCTs in Chapters 2 and 3 identified some limitations that are outlined and discussed here in the context of the current work and future research.

Controlling for factors

As parents might feel inclined to maximise the help their children receive during the study period, researchers need to better monitor participants during interventional trials to ensure that other interventions are not used simultaneously with the target intervention. Studies should clearly state that this process has been followed and that other interventions were not used during the intervention period. Moreover, in RMHA trials researchers should always perform verification or Real Ear Measurement (REM) of the device. This will ensure that the system is functioning as intended. From the four past APD trials using RMHAs or hearing aids, only half of them (Johnston et al., 2009; Kuk et al., 2008) stated that verification of the system was conducted to ensure that the acoustic output of the system was as per the manufacturer's specifications. In the other two studies (Sharma et al., 2012; Umat et al., 2011) there was no mention of this procedure.

Research protocols for RMHA studies should also routinely include the calculation of the total hours the system was used. Low use of the system may compromise any potential benefits that it could provide to the users. The studies by Johnston et al. (2009) and Sharma et al. (2012) only relied on parental reports on the duration the system was used by their children rather than using readings from the hearing aid data log. It is preferable that the data log of the hearing aid is checked, as this can provide an objective measure of the total time used. Parental and/ or teacher reports should still be used so that they provide with an additional or back-up measure. In Umat et al.'s (2011) study, even though the authors mentioned the intended daily time for RMHA use, it was unclear whether they actually looked at the hearing aid log. In the fourth and final APD study, Kuk et al. (2008) did not mention any checks being made to calculate the time the directional microphone hearing aids were used during their trial.

While in the two RCTs children were randomised into groups based on age and gender, the use of IQ as a covariate should perhaps be considered, as well. This is said as the Div-AVA baseline scores in the studies in Chapter 3 and 4 were significantly different between the two groups; controls entered the trial having better scores than the scores of the intervention group. By using the IQ scores to divide children into groups, equally-balanced baseline scores in attention and/ or other cognitive tests between the two groups might be produced.

Furthermore, another factor that might influence results in RMHA studies is the willingness of children to use the hearing aids. As noted in the study in Chapter 3, some children were reluctant to use the system. It is not known how that might have impacted the study findings, as this was not controlled. Excluding children who are reluctant to use the system from trials might introduce selection bias. An alternative way to control this factor is proposed, by administering a questionnaire that rates children's willingness and attitudes towards using RMHAs throughout the intervention period. The scores from the questionnaire could then be included as a covariate in the analysis. Moreover, adherence to the study protocol by teachers should also be monitored in future studies, as it was reported that a teacher was forgetting to systematically use the microphone. Perhaps teachers can complete a daily table stating whether they used the system and for how many periods. This will remind teachers to be consistent and use the RMHA daily. Additionally, researchers can also contact teachers more regularly to remind them to use the system.

Long-term effects

The long-term impact of RMHAs (i.e. following a period where they are not being used after the end of the trial) has not been examined in the current work since this was not included in the research protocol. It would be important for future research to routinely incorporate longitudinal trials, where they will test for the long-term effects of RMHAs. Possible improvements during the intervention period can then be studied longitudinally to test whether the benefit is sustained. In the APD literature, only Umat et al. (2011) retested children 12 months after the end of the trial, but as their results from the statistical analysis were not properly reported (see Section 1.7.2 under Memory for details), it is difficult to draw conclusions on the after-effects of RMHAs on their outcome measure (i.e. working memory). A drawback of longitudinal approaches is that if the time period between the end of the study and re-test is long, it becomes difficult to claim that sustained improvements in the outcome measures were because of the use of the intervention during the study period. Other factors such as maturation, possible use of other interventions, and changes in the participant's environment during that time, might all be helping to sustain the improvements in the outcome measures.

Targeted interventions

Assessment of APD requires a multidisciplinary team (Loo, Bamiou, & Rosen, 2013) and similarly, management of APD needs different professionals to contribute to the construction of a more informed management plan (AAA, 2010). The BSA (2011b) states that management of children with APD is under-informed and that children should be assigned to appropriate management strategies according to their diagnosis. Depending on the assessment of the individual, interventions could focus on improving listening, communication, language and cognitive functions, as well as targeting social and professional development for adults (AAA, 2010). It could therefore be argued that assignment of RMHAs to children with APD in the two studies (in Chapters 2 and 3) should have been done based on children's needs. This deficit-specific assignment of RMHA intervention could have perhaps maximised the benefit that children would receive and in turn may have improved aspects of their development that would have otherwise not improve. Following this deficit-specific assignment of intervention though, would introduce selection bias. Adding to that, Dillon et al. (2012) stress that the use of RMHAs is part of non-specific remediation that should be given to children who are assessed as having difficulties in listening compared to their peers. They support that an APD diagnosis is not always required to prescribe RMHAs, as the aim of its use is to decrease the difficulty in listening in challenging environments for these children. This view is contradicted by others who support that RMHA use should not be routinely recommended to children diagnosed with APD (Rosenberg, 2002) and that management strategies should be based on established research and theories (Pimentel & Inglebret, 2014). Moreover, Weihing, Chermak, and Musiek (2015) highlight the lack of deficit-specific intervention trials and call for the promotion of such trials and the comparison between deficit-specific interventions and general training interventions. It is therefore proposed that future studies should consider assigning participants to different intervention groups depending on their needs and make comparisons between different intervention groups (e.g. different types of AT and RMHAs) and controls. This was partially done by Sharma et al. (2012), but they did not include a RMHA-only group, so the effects of RMHAs compared to the other interventions were not determined.

6.2 Correlation between attention and AP skills, and APD diagnostic process

6.2.1 Attention tests and AP tests

Findings from past studies found an association between Sus-AA and the DDT, but only in samples of children suspected of APD (Gyldenkærne et al., 2014; Sharma et al., 2009). The problem with studying APD-suspected populations instead of APD-diagnosed ones is that inaccurate inferences can be drawn for children with APD from samples of children suspected of APD (Iliadou et al., 2016). Results from the study in Chapter 5 indicated that Sus-AA did not correlate with the DDT in a sample of children with APD diagnosis, contrasting previous findings on APD-suspected children (Gyldenkærne et al., 2014; Sharma et al., 2009). This requires further investigation since in the present study the correlation coefficient confidence interval was wide and the sample size small. The novel finding from this study is that the DDT correlated with two types of divided attention: Div-AA and Div-AVA. There are a few possible interpretations for these correlations, such as likely activation of the corpus callosum during both functions (Hutchinson et al., 2008; Musiek & Weihing, 2011; van der Knaap & van der Ham, 2011; Westerhausen & Hugdahl, 2008), possible measurement of the same psychometric construct, or perhaps one ability influencing the other (divided attention influencing dichotic listening or vice versa). As these interpretations cannot be examined from the study results, these findings can serve as a basis for further exploration of the relationship between types of divided attention and the DDT. Additionally, a percentage of children in the study failed the Sel-VA test (i.e. 10%); a percentage which exceeded the proportion of disordered performance in the test’s norms (Manly et al., 1999). This agrees with another study which revealed worse Sel-VA scores in children diagnosed with APD compared to children suspected of APD but not meeting APD criteria (Allen & Allan, 2014). It is proposed that this may indicate a general attention deficit present in a sub-group of children with APD, but more visual attention measures need to be employed before considering this explanation. As the statistical power of the study was low these findings require further validation from larger sample sized studies. Finally, the LiSN-S conditions did not correlate with the rest of the AP or attention tests and the implications this has on APD diagnosis are discussed next.

6.2.2 APD diagnostic process

Since the GIN, AFG tests and the LiSN-S did not correlate with either the attention tests or between them, it is suggested that they should be considered as good tools to use in APD assessments. On the other hand, as it is still undetermined how the DDT interacts with types of divided attention, it is proposed that (until further investigations are made between the relationship of these two abilities) the DDT is still used as part of the APD assessment, but alongside measures of Div-AA and Div-AVA. In this way, possible divided attention deficits in children failing the DDT (or vice versa) will not be missed and a more complete profile of children's deficits and needs can be constructed.

Furthermore, the study in Chapter 5 examined the usefulness of differential test batteries in the APD test procedure and APD management. With the inclusion of an AD battery patients at risk of having co-occurring ADs along with their AP deficits can be identified. An AD battery can give information on children's specific ADs and can indicate whether they are at risk of broader ADs (i.e. auditory and visual). This can inform APD management and clinicians can propose a management plan targeting children's attentional needs. Moreover, the inclusion of the newly proposed i-APD battery can help identify cases that would have been missed if only the APD and ADs protocols were taken into account. Children with i-APD do not present with general ADs but they do have an attention test with failed scores. It is proposed that these children with i-APD are first retested on that failed test as fatigue may be a reason for their poor performance. If similar performance is recorded during the second test, then it is proposed that children are given targeted interventions based on their failed attention test along with standard APD management strategies.

6.3 Limitations

While the studies in the current work were designed with an attempt to minimise bias, there are a number of key limitations that were not controlled for and that they might have influenced the results and their interpretation. First, the two RCTs did not use a placebo group. As explained earlier (in Section 2.2.2 under Children's involvement and Section 3.4.4 under Design limitations), the use of a placebo group may have caused children in that group, their teachers and their parents to realise that the RMHA was not functioning as intended. This would have potentially led to the termination of the trial and this was something to avoid. Also, the use of a placebo group would have required a larger sample size, and as previously noted there were no more participants that could have been recruited from GOSH. The exclusion of a placebo group, though, means that any significant effect found in these trials (i.e. improved scores in some questions of the LIFE-R questionnaire) cannot be solely attributed to the intervention. This is because the significant effects may have been the product of a placebo effect or a Hawthorne effect. Furthermore, it was discussed earlier (see Section 1.4) that the lack of universally accepted APD diagnostic criteria restricts the interpretation of research findings to only groups of children who meet the same APD criteria. Thus in this case, any significant effects observed in the studies of this thesis cannot be generalised in all populations with APD. The findings are only relevant to children diagnosed with APD based on the same APD criteria used in these studies here.

6.4 Conclusions

Findings from the two clinical trials suggest that the use of RMHAs does not promote long-term improvements in measures such as Sus-AA, Div-AVA, working memory and SIN ability after 6 months of RMHA use. At the same time, spatial listening was unaffected following the RMHA intervention, which is a positive finding and suggests that the system does not impede this listening ability. It is therefore proposed that longer intervention trials are needed (up to 12 months) to further test the effects of RMHAs on these aspects of development in children with APD. It would also be interesting to explore whether Sus-AA and Div-AVA improve under aided conditions (using the RMHA during tests) and compared to performance in unaided conditions.

Findings from the questionnaires indicate that the CHAPS might not be the best tool to use in APD trials and its use in APD assessment should be reconsidered. Other questionnaires (e.g. the Speech, Spatial, and Qualities of Hearing Scale and the Modified Amsterdam Inventory for Auditory Disability) can be considered instead, as they are shown to have better validity and test reliability (Bamiou et al., 2015). These questionnaires, though, need to be first validated in APD children populations, as they were designed for adult populations and have only been used in adult studies. The LIFE-R questionnaire completed by children had some measures that showed improvement over time. When using the RMHA, children reported improved listening in their teacher's words when listening through traffic noise and when faced with situations with competing speech. Therefore, using RMHAs in the classroom benefits children's listening in some acoustically challenging situations and continual use of the system in the classroom should be considered (for at least 6 months, minimum 2 hours daily average use). It should be noted, though, that the low statistical power of the studies means that this significant effect in the LIFE-R questions may have been the cause of an inflated non-significant effect. Additionally, it could be that these significant effects were the product of a placebo or Hawthorne effect as talked about earlier. More research is needed with larger sample sized studies to further validate these findings.

Furthermore, the two RCTs in Chapters 2 and 3 have identified a number of limitations. Drawing from these limitations, it is proposed that future work using RMHAs with children with APD should follow these guidelines:

a. Ensuring that other interventions are not used alongside RMHAs unless part of the research protocol.
b. Verification or REM of the device should be performed and reported so that the functionality of the RMHA is confirmed.
c. Checking the data log of the hearing aid to monitor the total time the system was used during the study period.
d. Administering a questionnaire to children in which they will report their level of satisfaction in using the RMHA. The scores can be used as a control factor in the analysis.

e. Including a placebo group in order to eliminate the influence of a placebo or Hawthorne effect on significant findings. As this has not been done previously in RMHA trials, studies with short-term intervention periods should first examine how a RMHA placebo is received by participants before long-term trials use this. Given the ethical issues surrounding the inclusion of a placebo, an alternative training group could be used instead.

These guidelines aim to minimise bias and control for factors that may be influencing the results. They also ensure uniformity across trials using RMHAs on children with APD.

The main finding from the study in Chapter 5 was that divided attention measures correlated with the DDT. It is worth noting that this is the first study that included divided attention measures in APD trials and the first study that ran correlation analyses in a sample of only APD-diagnosed children. It is suggested that more studies need to cross-validate these findings and further explore the relationship between measures of divided and visual attention and AP skills given the low statistical power of the study. Moreover, comparison of children who met APD criteria against those that were suspected of APD but did not meet APD criteria showed that only two SIN conditions of the LiSN-S were worse in the former group compared to the latter group. Again, further studies with improved statistical power need to examine these comparisons further.

There is still disagreement in the field of APD, especially in terms of how APD should be defined and diagnosed. To that end, in Chapter 5 a protocol proposed, which was aimed at improving the information clinicians receive in order to recommend better targeted management plans. Specifically, the differential test batteries proposed focus on APD management instead of the controversies in diagnosis and definition of APD. Overall, this protocol could help ensure that children with poor scores in AP and attention measures receive targeted management strategies in line with their needs.

REFERENCES

Ahmmed, A. U., Ahmmed, A. A., Bath, J. R., Ferguson, M. A., Plack, C. J., & Moore, D. R. (2014). Assessment of Children With Suspected Auditory Processing Disorder: A Factor Analysis Study. *Ear & Hearing*, *35*(3), 295–305. https://doi.org/10.1097/01.aud.0000441034.02052.0a

Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and Top-Down Influences on Auditory Scene Analysis: Evidence From Event-Related Brain Potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1072–1089. https://doi.org/10.1037/0096-1523.27.5.1072

Alain, C., Snyder, J. S., He, Y., & Reinke, K. S. (2007). Changes in Auditory Cortex Parallel Rapid Perceptual Learning. *Cerebral Cortex*, *17*(5), 1074–1084. https://doi.org/10.1093/cercor/bhl018

Allen, P., & Allan, C. (2014). Auditory Processing Disorders: Relationship to Cognitive Processes and Underlying Auditory Neural Integrity. *International Journal of Pediatric Otorhinolaryngology*, *78*(2), 198–208. https://doi.org/10.1016/j.ijporl.2013.10.048

Alloway, T. P. (2007). *Automated Working Memory Assessment. Manual*. London, United Kingdom: Pearson Assessment.

American Academy of Audiology. (2010). American Academy of Audiology Clinical Practice Guidelines. Diagnosis, Treatment and Management of Children and Adults with Central Auditory Processing Disorder. Retrieved December 28, 2017, from http://www.audiology.org/publications-resources/document-library/central-auditory-processing-disorder

American Academy of Audiology. (2011). Clinical Practice Guidelines: Remote Microphone Hearing Assistance Technologies for Children and Youth from Birth to 21 Years. Retrieved June 23, 2018, from https://audiology-web.s3.amazonaws.com/migrated/HAT_Guidelines_Supplement_A.pdf_53996ef7758497.54419000.pdf

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (5th ed.). Washington, DC; London: American Psychiatric Pub.

American Psychiatric Association. (2017). *DSM-5 Update*. *DSM-5 Update. Supplement to Diagnostic and Statistical Manual of Mental Disorders*. Arlington, VA. Retrieved from https://psychiatryonline.org/pb-assets/dsm/update/DSM5Update_October2017.pdf

American Speech-Language-Hearing Association. (2005a). (Central) Auditory Processing Disorder - The Role of the Audiologist. Retrieved November 27, 2015, from https://www.asha.org/policy/PS2005-00114/

American Speech-Language-Hearing Association. (2005b). (Central) Auditory Processing Disorders. Technical Report. Retrieved December 15, 2015, from http://www.asha.org/policy/TR2005-00043/

Anderson, K. L., & Smaldino, J. J. (1999). Listening Inventories for Education: A Classroom Measurement Tool. *The Hearing Journal*, *52*(10), 74–76.

Anderson, K. L., Smaldino, J. J., & Spangler, C. (2011). LIFE-R The Listening Inventories for Education-Revised. Retrieved September 4, 2017, from http://successforkidswithhearingloss.com/wp-content/uploads/2011/09/LIFE-R-Instruction-Manual.pdf

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). A Dynamic Auditory-Cognitive System Supports Speech-in-Noise Perception in Older Adults. *Hearing Research*, *300*, 18–32. https://doi.org/10.1016/j.heares.2013.03.006

Armstrong, R. A. (2014). When to Use the Bonferroni Correction. *Ophthalmic & Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K. C., & Bizley, J. K. (2018). Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron*, *97*(3), 640–655.e4. https://doi.org/10.1016/j.neuron.2017.12.034

Baddeley, A. (1992). Working Memory. *Science*, *255*, 556–559. https://doi.org/10.4249/scholarpedia.3015

Bamiou, D.-E., Campbell, N., & Sirimanna, T. (2006). Management of Auditory Processing Disorders. *Audiological Medicine*, *4*(1), 46–56. https://doi.org/10.1080/16513860600630498

Bamiou, D.-E., & Iliadou, V. V. (2014). Assessment of Individuals Suspected or Diagnosed With Central Auditory Processing Disorder. A Medical Perspective. In G. D. Chermak & F. E. Musiek (Eds.), *Handbook of Central Auditory Processing Disorder. Volume II. Comprehensive Intervention* (pp. 479–507). San Diego, CA: Plural Publishing, Inc.

Bamiou, D.-E., Iliadou, V. V., Zanchetta, S., & Spyridakou, C. (2015). What Can We Learn About Auditory Processing From Adult Hearing Questionnaires? *Journal of the American Academy of Audiology*, *26*(10), 824–837. https://doi.org/10.3766/jaaa.15009

Bamiou, D.-E., Musiek, F. E., & Luxon, L. M. (2001). Aetiology and Clinical Presentations of Auditory Processing Disorders - A Review. *Archives of Disease in Childhood*, *85*(5), 361–365. Retrieved from http://adc.bmj.com/content/85/5/361.short

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of Autism-Spectrum Conditions: UK School-Based Population Study. *British Journal of Psychiatry*, *194*(6), 500–509. https://doi.org/10.1192/bjp.bp.108.059345

Bellis, T. J. (2011). *Assessment and Management of Central Auditory Processing Disorders in the Educational Setting: From Science to Practice* (2nd ed.). San Diego, CA: Plural Publishing, Inc.

Bellis, T. J., & Bellis, J. D. (2015). Central Auditory Processing Disorders in Children and Adults. In G. G. Celesia & G. Hickok (Eds.), *The Human Auditory System: Fundamental Organization and Clinical Disorders* (1st ed., Vol. 219, pp. 537–556). Oxford, United Kingdom: Elsevier B.V. https://doi.org/10.1097/00020840-199712000-00002

Bellis, T. J., & Ferre, J. M. (1999). Multidimensional Approach to the Differential Diagnosis of Central Auditory Processing Disorders in Children. *Journal of the American Academy of Audiology*, *10*, 319–328. Retrieved from https://www.audiology.org/sites/default/files/journal/JAAA_10_06_03.pdf

Best, V., Ozmeral, E. J., Kopco, N., & Shinn-Cunningham, B. G. (2008). Object Continuity Enhances Selective Auditory Attention. *Proceedings of the National Academy of Sciences, 105*(35), 13174–13178. https://doi.org/10.1073/pnas.0803718105

Bishop, D. V. M. (2003). *The Children's Communication Checklist. CCC-2 Manual* (2nd ed.). London, United Kingdom: Harcourt Assessment.

Bishop, D. V. M. (2004). *Expression, Reception and Recall of Narrative Instrument (ERRNI)*. Oxford, United Kingdom: Pearson Assessment.

Bishop, D. V. M. (2005). *Test for Reception of Grammar—Electronic (TROG-E)*. Oxford, United Kingdom: Pearson Assessment.

Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2012). Auditory Deficit as a Consequence Rather Than Endophenotype of Specific Language Impairment: Electrophysiological Evidence. *PLoS ONE*, *7*(5), 1–11. https://doi.org/10.1371/journal.pone.0035851

Blake, R., Field, B., Foster, C., Platt, F., & Wertz, P. (1991). Effect of FM Auditory Trainers on Attending Behaviors of Learning-Disabled Children. *Language, Speech, and Hearing Services in Schools*, *22*(3), 111–114. https://doi.org/10.1044/0161-1461.2203.111

Box, G. E. P. (1953). Non-Normality and Tests on Variances. *Biometrika*, *40*(3/4), 318. https://doi.org/10.2307/2333350

British Society of Audiology. (2011a). Position Statement. Auditory Processing Disorder (APD). Retrieved November 27, 2015, from http://www.thebsa.org.uk/docs/docsfromold/BSA_APD_PositionPaper_31March11_FINAL.pdf

British Society of Audiology. (2011b). Practice Guidance. An Overview of Current Management of Auditory Processing Disorder (APD). Retrieved December 18, 2015, from http://www.thebsa.org.uk/docs/docsfromold/BSA_APD_Management_1Aug11_FINAL_amended17Oct11.pdf

British Society of Audiology. (2012). Recommended Procedure. Pure-Tone Air-Conduction and Bone-Conduction Threshold Audiometry With and Without Masking. Retrieved February 16, 2017, from http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_RP_PTA_FINAL_24Sept11_MinorAmend06Feb12.pdf

British Society of Audiology. (2013). Recommended Procedure. Tympanometry. Retrieved February 16, 2017, from http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_RP_Tymp_Final_21Aug13_Final.pdf

British Society of Audiology. (2018). *Position Statement and Practice Guidance Auditory Processing Disorder (APD). British Society of Audiology*.

Broadbent, D. E. (1958). *Pereception and Communication*. London, United Kingdom: Plenum Press. https://doi.org/10.1108/eb015727

Brown, L., Sherbenou, R., & Johnson, S. (1997). *Test of Nonverbal Intelligence*. Austin, TX: Pro-Ed.

Brugge, J. F. (2013). Anatomy and Physiology of Auditory Pathways and Cortex. In G. G. Celesia

(Ed.), *Handbook of Clinical Neurophysiology. Volume 10. Disorders of Peripheral and Central Auditory Processing* (pp. 25–59). Amsterdam: Elsevier.

Cacace, A. T., & McFarland, D. J. (2006). Delineating Auditory Processing Disorder (APD) and Attention Deficit Hyperactivity Disorder (ADHD): A Conceptual, Theoretical, and Practical Framework. In T. K. Parthasarathy (Ed.), *An Introduction to Auditory Processing Disorders in Children* (pp. 39–62). New York, NY: Psychology Press. Taylor & Francis Group.

Cameron, S., & Dillon, H. (2007a). Development of the Listening in Spatialized Noise-Sentences Test (LiSN-S). *Ear & Hearing*, *28*(2), 196–211. https://doi.org/10.1080/14992020601164170

Cameron, S., & Dillon, H. (2007b). The Listening in Spatialized Noise-Sentences Test (LISN-S): Test-retest Reliability Study. *International Journal of Audiology*, *46*(3), 145–153. https://doi.org/10.1080/14992020601164170

Cameron, S., & Dillon, H. (2008). The Listening in Spatialized Noise–Sentences Test (LISN-S): Comparison to the Prototype LISN and Results From Children With Either a Suspected (Central) Auditory Processing Disorder or a Confirmed Language Disorder. *Journal of the American Academy of Audiology*, *19*(5), 377–391. https://doi.org/10.3766/jaaa.19.5.2

Cameron, S., & Dillon, H. (2011). Development and Evaluation of the LiSN & Learn Auditory Training Software for Deficit-Specific Remediation of Binaural Processing Deficits in Children: Preliminary Findings. *Journal of the American Academy of Audiology*, *22*(10), 678–696. https://doi.org/10.3766/jaaa.22.10.6

Cameron, S., Dillon, H., Glyde, H., Kanthan, S., & Kania, A. (2014). Prevalence and Remediation of Spatial Processing Disorder (SPD) in Indigenous Children in Regional Australia. *International Journal of Audiology*, *53*(5), 326–335. https://doi.org/10.3109/14992027.2013.871388

Cameron, S., Dillon, H., & Newall, P. (2006). The Listening in Spatialized Noise Test: An Auditory Processing Disorder Study. *Journal of the American Academy of Audiology*, *320*(17), 306–320. https://doi.org/10.3766/jaaa.17.5.2

Cameron, S., Glyde, H., & Dillon, H. (2012). Efficacy of the LiSN & Learn Auditory Training Software: Randomized Blinded Controlled Study. *Audiology Research*, *2*(e15), 86–93. https://doi.org/10.4081/audiores.2012.e15

Cameron, S., Glyde, H., Dillon, H., & Whitfield, J. (2016). Investigating the Interaction Between Dichotic Deficits and Cognitive Abilities Using the Dichotic Digits difference Test (DDdT) Part 2. *Journal of the American Academy of Audiology*, *27*(6), 470–479. https://doi.org/10.3766/jaaa.15085

Chermak, G. D., Bamiou, D.-E., Iliadou, V. V., & Musiek, F. E. (2017). Practical Guidelines to Minimise Language and Cognitive Confounds in the Diagnosis of CAPD : A Brief Tutorial. *International Journal of Audiology*, *0*(0), -000. https://doi.org/10.1080/14992027.2017.1284351

Chermak, G. D., & Bellis, T. J. (2014). Differential Diagnosis of Central Auditory Processing Disorder and Attention-Deficit/ Hyperactivity Disorder. In F. E. Musiek & G. D. Chermak (Eds.), *Handbook of Central Auditory Processing Disorder. Volume I. Auditory Neuroscience and Diagnosis* (2nd ed., pp. 557–590). San Diego, CA: Plural Publishing, Inc.

Chermak, G. D., Bellis, T. J., & Musiek, F. E. (2014). Neurobiology, Cognitive Science, and Intervention. In G. D. Chermak & F. E. Musiek (Eds.), *Handbook of Central Auditory Processing Disorder. Volume II. Comprehensive Intervention* (2nd ed., pp. 3–38). San Diego, CA: Plural Publishing, Inc.

Chermak, G. D., & Musiek, F. E. (1997). *Central Auditory Processing Disorders. New Perspectives*. San Diego, CA: Singular Publishing Group, Inc.

Chermak, G. D., & Musiek, F. E. (2002). Auditory training: Principles and approaches for remediating and managing auditory processing disorders. *Seminars in Hearing*, *23*(4), 297–308. https://doi.org/10.1055/s-2002-35878

Chermak, G. D., & Musiek, F. E. (2007). *Handbook of Central Auditory Processing Disorders, Vol. II. Comprehensive Intervention*. (G. D. Chermak & F. E. Musiek, Eds.). Oxfordshire, United Kingdom: Plural Publishing, Inc.

Chermak, G. D., & Musiek, F. E. (2014). Neurological Substrate of Central Auditory Processing Disorder. In F. E. Musiek & G. D. Chermak (Eds.), *Handbook of Central Auditory Processing Disorder*. *Volume I. Auditory Neuroscience and Diagnosis* (2nd ed., pp. 89–111). San Diego, CA: Plural Publishing, Inc.

Chermak, G. D., Somers, E. K., & Seikel, J. A. (1998). Behavioral Signs of Central Auditory Processing Disorder and Attention Deficit Hyperactivity Disorder. *Journal of the American Academy of Audiology*, *9*(1), 78–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9493945

Chermak, G. D., Tucker, E., & Seikel, J. A. (2002). Behavioral Characteristics of Auditory Processing Disorder and Attention-Deficit Hyperactivity Disorder: Predominantly Inattentive Type. *Journal of the American Academy of Audiology*, *13*, 332–338.

Clopton, B. M., Winfield, J. A., & Flammino, F. J. (1974). Tonotopic Organization: Review and Analysis. *Brain Research*, *76*(1), 1–20. https://doi.org/10.1080/00085006.2011.11092681

Coch, D., Sanders, L. D., & Neville, H. J. (2005). An Event-Related Potential Study of Selective Auditory Attention in Children and Adults. *Journal of Cognitive Neuroscience*, *17*(4), 605–622. https://doi.org/10.1162/0898929053467631

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*(July), 155–159. https://doi.org/10.1038/141613a0

Cowan, J., Rosen, S., & Moore, D. R. (2009). Putting the Auditory Processing Back Into Auditory Processing Disorder in Children. In A. T. Cacace & D. J. McFarland (Eds.), *Controversies in Central Auditory Processing Disorder* (pp. 187–197). San Diego, CA: Plural Publishing, Inc.

Cowan, N. (1998). Visual and Auditory Working Memory Capacity. *Trends in Cognitive Sciences*, *2*(3), 77–78. https://doi.org/10.1016/S1364-6613(98)01144-9

Dawes, P., & Bishop, D. V. M. (2007). The SCAN-C in Testing for Auditory Processing Disorder in a Sample of British Children. *International Journal of Audiology*, *46*(12), 780–786. https://doi.org/10.1080/14992020701545906

Dawes, P., & Bishop, D. V. M. (2009). Auditory Processing Disorder in Relation to Developmental Disorders of Language, Communication and Attention: A Review and

Critique. *International Journal of Language & Communication Disorders*, *44*(4), 440–465. https://doi.org/10.1080/13682820902929073

Dawes, P., & Bishop, D. V. M. (2010). Psychometric Profile of Children With Auditory Processing Disorder and Children With Dyslexia. *Archives of Disease in Childhood*, *95*(1), 432–436. https://doi.org/10.1136/adc.2009.170118

Dawes, P., Bishop, D. V. M., Sirimanna, T., & Bamiou, D.-E. (2008). Profile and Aetiology of Children Diagnosed With Auditory Processing Disorder (APD). *International Journal of Pediatric Otorhinolaryngology*, *72*(4), 483–489. https://doi.org/10.1016/j.ijporl.2007.12.007

Dillon, H., Cameron, S., Glyde, H., Wilson, W. J., & Tomlin, D. (2012). An Opinion on the Assessment of People who may Have an Auditory Processing Disorder. *Journal of the American Academy of Audiology*, *23*(2), 97–105. https://doi.org/10.3766/jaaa.23.2.4

Dodd, B., Holm, A., Orelemans, M., & McCormick, M. (1996). *QUIL: Queensland University Inventory of Literacy*. Brisbane, Australia: The University of Queensland.

Facoetti, A., Corradi, N., Ruffino, M., Gori, S., & Zorzi, M. (2010). Visual Spatial Attention and Speech Segmentation are both Impaired in Preschoolers at Familial Risk for Developmental Dyslexia. *Dyslexia*, *16*, 226–239. https://doi.org/10.1002/dys

Ferguson, M. A., Hall, R. L., & Moore, D. R. (2011). Communication, Listening, Cognitive and Speech Perception Skills in Children With Auditory Processing Disorder (APD) or Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *54*(1), 211–228. https://doi.org/10.1044/1092-4388(2010/09-0167)

Fey, M. E., Richard, G. J., Geffner, D., Kamhi, A. G., Medwetsky, L., Paul, D., … Schooling, T. (2011). Auditory Processing Disorder and Auditory/ Language Interventions: An Evidence-Based Systematic Review. *Language, Speech, and Hearing Services in Schools*, *42*, 246–264. https://doi.org/10.1044/0161-1461(2010/10-0013)APD

Filippini, R., Befi-Lopes, D. M., & Schochat, E. (2012). Efficacy of Auditory Training Using the Auditory Brainstem Response to Complex Sounds: Auditory Processing Disorder and Specific Language Impairment. *Folia Phoniatrica et Logopaedica*, *64*(5), 217–226. https://doi.org/10.1159/000342139

Fisher, L. (1979). *Fisher's Auditory Problems Checklist*. Bemidji, MN: Life Products.

Fitzpatrick, E. M., Durieux-Smith, A., & Whittingham, J. (2010). Clinical Practice for Children With Mild Bilateral and Unilateral Hearing Loss. *Ear & Hearing*, *31*(3), 392–400. https://doi.org/10.1097/AUD.0b013e3181cdb2b9

Flexer, C. A. (2014). Signal Enhancement: Personal FM and Sound-Field Technology. In G. D. Chermak & F. E. Musiek (Eds.), *Handbook of Central Auditory Processing Disorder. Volume II. Comprehensive Intervention* (2nd ed., pp. 355–375). San Diego, CA: Plural Publishing, Inc.

Forster, S., Robertson, D. J., Jennings, A., Asherson, P., & Lavie, N. (2014). Plugging the Attention Deficit: Perceptual Load Counters Increased Distraction in ADHD. *Neuropsychology*, *28*(1), 91–97. https://doi.org/10.1037/neu0000020

Friederichs, E., & Friederichs, P. (2005). Electrophysiologic and Psycho-Acoustic Findings

Following one-Year Application of a Personal ear-Level FM Device in Children With Attention Deficit and Suspected Central Auditory Processing Disorder. *Journal of Educational Audiology*, *12*(1), 31–36.

Gelfand, S. A. (2009). *Essentials of Audiology* (3rd ed.). New York, NY: Thieme Medical Publishers, Inc.

Gnewikow, D., & Moss, M. (2006). Hearing aid Outcomes With Open- and Closed-Canal Fittings. *The Hearing Journal*, *59*(11), 66. https://doi.org/10.1097/01.HJ.0000286220.14835.13

Goswami, U. (2011). A Temporal Sampling Framework for Developmental Dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

Great Ormond Street Hospital. (2018). Auditory Processing Disorder (APD). Clinic at Great Ormond Street Hospital. Referral Criteria & Further Information. Retrieved June 13, 2018, from https://www.gosh.nhs.uk/medical-information/auditory-processing-disorder

Grube, M., Cooper, F. E., Kumar, S., Kelly, T., & Griffiths, T. D. (2014). Exploring the Role of Auditory Analysis in Atypical Compared to Typical Language Development. *Hearing Research*, *308*, 129–140. https://doi.org/10.1016/j.heares.2013.09.015

Guenette, L. A. (2006). How to Administer the Dichotic Digit Test. *The Hearing Journal*, *59*(2), 50.

Gupta, S. (2011). Intention-to-Treat Concept: A Review. *Perspectives in Clinical Research*, *2*(3), 109–112. https://doi.org/10.4103/2229-3485.83221

Gyldenkærne, P., Dillon, H., Sharma, M., & Purdy, S. C. (2014). Attend to This: The Relationship Between Auditory Processing Disorders and Attention Deficits. *Journal of the American Academy of Audiology*, *25*(7), 676–687. https://doi.org/10.3766/jaaa.25.7.6

Haastrup, A., & Dworsack-Dodge, M. (2011). Electroacoustic Evaluation of the Resound Unite ™ Mini Microphone With Otometrics Aurical Hit. Retrieved June 21, 2018, from https://partners.natus.com/asset/resource/file/otometrics/asset/M200862-GB-12 10-Rev A-P8 %282%29.pdf

Hackett, T. A. (2009). Organization of the Central Auditory Pathways in Nonhuman Primates and Humans. In A. T. Cacace & D. J. McFarland (Eds.), *Controversies in Central Auditory Processing Disorder* (pp. 15–45). San Diego, CA: Plural Publishing, Inc.

Hackett, T. A. (2015). Anatomic Organization of the Auditory Cortex. In G. G. Celesia & G. Hickok (Eds.), *The Human Auditory System: Fundamental Organization and Clinical Disorders* (1st ed., Vol. 129, pp. 27–53). Oxford, United Kingdom: Elsevier B.V. https://doi.org/10.1016/B978-0-444-62630-1.00002-0

Hahn, B., Wolkenberg, F. A., Ross, T. J., Myers, C. S., Heishman, S. J., Stein, D. J., … Stein, E. A. (2008). Divided Versus Selective Attention: Evidence for Common Processing Mechanisms. *Brain Research*, *1215*, 137–146. https://doi.org/10.1016/j.brainres.2008.03.058

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The Benefit of Binaural Hearing in a Cocktail Party: Effect of Location and Type of Interferer. *Journal of the Acoustical Society of America*, *115*(2), 833–843. https://doi.org/10.1121/1.1639908

Hayes, S. H., Ding, D., Salvi, R. J., & Allman, B. L. (2013). Anatomy and Physiology of the External, Middle, and Inner ear. In G. G. Celesia (Ed.), *Handbook of Clinical Neurophysiology. Volume 10. Disorders of Peripheral and Central Auditory Processing*. Amsterdam: Elsevier.

He, Y. (2010). Missing Data Analysis Using Multiple Imputation: Getting to the Heart of the Matter. *Circulation: Cardiovascular Quality and Outcomes*, *3*(1), 1–16. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial

Head, J., & Helton, W. S. (2015). Passive Perceptual Learning Versus Active Searching in a Novel Stimuli Vigilance Task. *Experimental Brain Research*, *233*(5), 1481–1489. https://doi.org/10.1007/s00221-015-4222-z

Heine, C., & O'Halloran, R. (2015). Central Auditory Processing Disorder: A Systematic Search and Evaluation of Clinical Practice Guidelines. *Journal of Evaluation in Clinical Practice*, *21*(6), 988–994. https://doi.org/10.1111/jep.12494

Hind, S. E. (2006). Survey of Care Pathway for Auditory Processing Disorder. *Audiological Medicine*, *4*(1), 12–24. https://doi.org/10.1080/16513860500534543

Hind, S. E., Haines-Bazrafshan, R., Benton, C. L., Brassington, W., Towle, B., & Moore, D. R. (2011). Prevalence of Clinical Referrals Having Hearing Thresholds Within Normal Limits. *International Journal of Audiology*, *50*(10), 708–716. https://doi.org/10.3109/14992027.2011.582049

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and Analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, *49*(12), 891–903. https://doi.org/10.3109/14992027.2010.506889

Hornickel, J., Zecker, S. G., Bradlow, A. R., & Kraus, N. (2012). Assistive Listening Devices Drive Neuroplasticity in Children With Dyslexia. *Proceedings of the National Academy of Sciences*, *109*(41), 16731–16736. https://doi.org/10.1073/pnas.1206628109

Hutchinson, A. D., Mathias, J. L., & Banich, M. T. (2008). Corpus Callosum Morphology in Children and Adolescents with Attention Deficit Hyperactivity Disorder: A Meta-analytic Review. *Neuropsychology*, *22*(3), 341–349.

Iliadou, V. V., & Bamiou, D.-E. (2012). Psychometric Evaluation of Children With Auditory Processing Disorder (APD): Comparison With Normal-Hearing and Clinical non-APD Groups. *Journal of Speech, Language, and Hearing Research*, *55*(3), 791–799. https://doi.org/10.1044/1092-4388(2011/11-0035)

Iliadou, V. V., Ptok, M., Grech, H., Pedersen, E. R., Brechmann, A., Deggouj, N., … Bamiou, D.-E. (2017). A European Perspective on Auditory Processing Disorder - Current Knowledge and Future Research Focus. *Frontiers in Neurology*, *8*(November), 1–7. https://doi.org/10.3389/fneur.2017.00622

Iliadou, V. V., Sirimanna, T., & Bamiou, D.-E. (2016). CAPD Is Classified in ICD-10 as H93.25 and Hearing Evaluation—Not Screening—Should Be Implemented in Children With Verified Communication and/or Listening Deficits. *American Journal of Audiology*, *25*(4), 368–370. https://doi.org/10.1044/2016

International Statistical Classification of Diseases and Related Health Problems - Clinical

Modification. (2018). ICD-10-CM Diagnosis Code H93.25. Retrieved June 6, 2018, from https://www.icd10data.com/ICD10CM/Codes/H60-H95/H90-H94/H93-/H93.25

Jenstad, L., & Moon, J. (2011). Systematic Review of Barriers and Facilitators to Hearing Aid Uptake in Older Adults. *Audiology Research*, *1*(e25), 92–96. https://doi.org/10.4081/audiores.2011.e25

Johnston, K. N., John, A. B., Kreisman, N. V, Hall III, J. W., & Crandell, C. C. (2009). Multiple Benefits of Personal FM System use by Children with Auditory Processing Disorder (APD). *International Journal of Audiology*, *48*(6), 371–383. https://doi.org/10.1080/14992020802687516

Jutras, B., Loubert, M., Dupuis, J.-L., Marcoux, C., Dumont, V., & Baril, M. (2007). Applicability of Central Auditory Processing Disorder Models. *American Journal of Audiology*, *16*(December 2007), 100–106. Retrieved from http://aja.pubs.asha.org/article.aspx?articleid=1757481

Kang, M., Ragan, B. G., & Park, J.-H. (2008). Issues in Outcomes Research: An Overview of Randomization Techniques for Clinical Trials. *Journal of Athletic Training*, *43*(2), 215–221. https://doi.org/10.4085/1062-6050-43.2.215

Katz, J. (1992). Classification of Auditory Processing Disorders. In J. Katz, N. A. Stecker, & D. Henderson (Eds.), *Central Auditory Processing: ATransdisciplinary View* (pp. 81–91). St Louis, MO: Mosby Year Book.

Keith, R. W. (2009). *SCAN-3 for Children. Tests for Auditory Processing Disorders. Manual* (2nd ed.). San Antonio, TX: Pearson, Inc.

Keith, W. J., & Purdy, S. C. (2014). Assistive and Therapeutic Effects of Amplification for Auditory Processing Disorder. *Seminars in Hearing*, *35*(1), 27–38. https://doi.org/10.1055/s-0033-1363522

Kerns, K. A., Eso, K., & Thomson, J. (1999). Investigation of a Direct Intervention for Improving Attention in Young Children With ADHD. *Developmental Neuropsychology*, *16*(2), 273–295. https://doi.org/10.1207/S15326942DN1602

Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The Advantage of Knowing Where to Listen. *The Journal of the Acoustical Society of America*, *118*(6), 3804–3815. https://doi.org/10.1121/1.2109187

Kitterick, P. T., Lovett, R. E. S., Goman, A. M., & Summerfield, A. Q. (2011). The AB-York Crescent of Sound: An Apparatus for Assessing Spatial-Listening Skills in Children and Adults. *Cochlear Implants International*, *12*(3), 164–169. https://doi.org/10.1179/146701011X13049348987832

Koohi, N., Vickers, D. A., Warren, J., Werring, D., & Bamiou, D.-E. (2017). Long-Term use Benefits of Personal Frequency-Modulated Systems for Speech in Noise Perception in Patients With Stroke With Auditory Processing Deficits: A non-Randomised Controlled Trial Study. *BMJ Open*, *7*(4), e013003. https://doi.org/10.1136/bmjopen-2016-013003

Korkman, M., Kirk, U., & Kemp, S. (1997). *NEPSY*. Oxford, United Kingdom: Pearson Assessment.

Kuk, F., Jackson, A., Keenan, D., & Lau, C.-C. (2008). Personal Amplification for School-age

Children With Auditory Processing Disorders. *Journal of the American Academy of Audiology*, *19*(6), 465–480. https://doi.org/10.3766/jaaa.19.6.3

Kuppen, S. E. A., & Goswami, U. (2016). Developmental Trajectories for Children with Dyslexia and low IQ Poor Readers. *Developmental Psychology*, *52*(5), 717–734. https://doi.org/10.1037/a0040207

Lagacé, J., Jutras, B., Giguère, C., & Gagné, J.-P. (2011). Speech Perception in Noise: Exploring the Effect of Linguistic Context in Children With and Without Auditory Processing Disorder. *International Journal of Audiology*, *50*(6), 385–395. https://doi.org/10.3109/14992027.2011.553204

Lam, E., & Sanchez, L. (2007). Evaluation of Screening Instruments for Auditory Processing Disorder (APD) in a Sample of Referred Children. *The Australian and New Zealand Journal of Audiology*, *29*(1), 26–39. https://doi.org/10.1375/audi.29.1.26

Lavie, N. (2005). Distracted and Confused?: Selective Attention Under Load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lemos, C. I. C., Tangerino de Souza Jacob, R., Gejão, M. G., Bevilacqua, M. C., Feniman, M. R., & Ferrari, D. V. (2009). Frequency Modulation (FM) System in Auditory Processing Disorder: An Evidence-Based Practice? *Pró-Fono Revista de Atualização Científica*, *21*(3), 243–248.

Liasis, A., Bamiou, D.-E., Campbell, P., Sirimanna, T., Boyd, S., & Towell, A. (2003). Auditory Event-Related Potentials in the Assessment of Auditory Processing Disorders: A Pilot Study. *Neuropediatrics*, *34*(1), 23–29. https://doi.org/10.1055/s-2003-38622

Litovsky, R. Y. (2005). Speech Intelligibility and Spatial Release From Masking in Young Children. *The Journal of the Acoustical Society of America*, *117*(5), 3091–3099. https://doi.org/10.1121/1.1873913

Loo, J. H. Y., Bamiou, D.-E., & Rosen, S. (2013). The Impacts of Language Background and Language-Related Disorders in Auditory Processing Assessment. *Journal of Speech, Language, and Hearing Research*, *56*, 1–12. https://doi.org/10.1044/1092-4388(2012/11-0068)b

Loo, J. H. Y., Rosen, S., & Bamiou, D.-E. (2016). Auditory Training Effects on the Listening Skills of Children With Auditory Processing Disorder. *Ear & Hearing*, *37*(1), 38–47. https://doi.org/10.1097/AUD.0000000000000225

Lotfi, Y., Mehrkian, S., Moossavi, A., Zadeh, S. F., & Sadjedi, H. (2016). Relation Between Working Memory Capacity and Auditory Stream Segregation in Children with Auditory Processing Disorder. *Iranian Journal of Medical Sciences*, *41*(2), 110–117.

Lotfi, Y., Moossavi, A., Abdollahi, F. Z., Bakhshi, E., & Sadjedi, H. (2016). Effects of an Auditory Lateralization Training in Children Suspected to Central Auditory Processing Disorder. *Journal of Audiology & Otology*, *20*(2), 102–108. https://doi.org/10.7874/jao.2016.20.2.102

Madelaine, A., & Wheldall, K. (2002). Further Progress Towards a Standardized Curriculum-Based Measure of Reading: Calibrating a new Passage Reading Test Against the New South Wales Basic Skills Test. *Educational Psychology*, *22*, 461–471.

Maerlender, A. C., Wallis, D. J., & Isquith, P. K. (2004). Psychometric and Behavioral Measures of Central Auditory Function: The Relationship Between Dichotic Listening and Digit Span Tasks. *Child Neuropsychology*, *10*(4), 318–327. https://doi.org/10.1080/09297040490909314

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1999). *TEA-Ch. The Test of Everyday Attention for Children*. London, United Kingdom: Pearson Assessment.

Marriage, J., King, J., Briggs, J., & Lutman, M. E. (2001). The Reliability of the SCAN Test: Results From a Primary School Population in the UK. *British Journal of Audiology*, *35*(3), 199–208. https://doi.org/10.1080/00305364.2001.11745237

Martin, J. S., Jerger, J., & Mehta, J. (2007). Divided-Attention and Directed-Attention Listening Modes in Children With Dichotic Deficits: An Event-Related Potential Study. *Journal of the American Academy of Audiology*, *18*(1), 34–53. https://doi.org/10.3766/jaaa.18.1.4

Martin, J. S., & Keith, R. W. (2009). Central Auditory Processing Disorders. In V. E. Newton (Ed.), *Paediatric Audiological Medicine* (2nd ed., pp. 227–244). Chichester: John Wiley & Sons, Ltd.

Mattys, S. L., & Palmer, S. D. (2015). Divided Attention Disrupts Perceptual Encoding During Speech Recognition. *The Journal of the Acoustical Society of America*, *137*(3), 1464–1472. https://doi.org/10.1121/1.4913507

McDonald, J. H. (2014). *Handbook of Biological Statistics* (3rd editio). Baltimore, MD: Sparky House Publishing.

McLaughlin, M. J., & Sainani, K. L. (2014). Bonferroni, Holm, and Hochberg Corrections: Fun Names, Serious Changes to P Values. *PM and R*, *6*(6), 544–546. https://doi.org/10.1016/j.pmrj.2014.04.006

Mesgarani, N., & Chang, E. F. (2012). Selective Cortical Representation of Attended Speaker in Multi-Talker Speech Perception. *Nature*, *485*(7397), 233–236. https://doi.org/10.1038/nature11020

Mishra, S. K., & Boddupally, S. P. (2017). Auditory Cognitive Training for Pediatric Cochlear Implant Recipients. *Ear & Hearing*, 1–12.

Misra, S. (2012). Randomized Double Blind Placebo Control Studies, the “Gold Standard” in Intervention Based Studies. *Indian Journal of Sexually Transmitted Diseases*, *33*(2), 131–134. https://doi.org/10.4103/0253-7184.102130

Moore, D. R., Ferguson, M. A., Edmondson-Jones, M., Ratib, S., & Riley, A. (2010). Nature of Auditory Processing Disorder in Children. *Pediatrics*, *126*(2), e382–e390. https://doi.org/10.1542/peds.2009-2826

Moore, D. R., Rosen, S., Bamiou, D.-E., Campbell, N. G., & Sirimanna, T. (2013). Evolving Concepts of Developmental Auditory Processing Disorder (APD): A British Society of Audiology APD Special Interest Group “White Paper.” *International Journal of Audiology*, *52*(1), 3–13. https://doi.org/10.3109/14992027.2012.723143

Moossavi, A., Mehrkian, S., Lotfi, Y., Faghihzadeh, S., & Sadjedi, H. (2014). The Relation Between Working Memory Capacity and Auditory Lateralization in Children With Auditory Processing Disorders. *International Journal of Pediatric Otorhinolaryngology*, *78*(11), 1981–1986. https://doi.org/10.1016/j.ijporl.2014.09.003

Murphy, C. F. B., & Bamiou, D.-E. (2018). *Performance of British Children on the LiSN-S*.

Musiek, F. E. (2002). The Frequency Pattern Test: A Guide. *The Hearing Journal*, *55*(6), 58.

Musiek, F. E., & Pinheiro, M. L. (1987). Frequency Patterns in Cochlear, Brainstem, and Cerebral Lesions. *Audiology*. https://doi.org/10.3109/00206098709078409

Musiek, F. E., & Schochat, E. (1998). Auditory Training and Central Auditory Processing Disorders. *Seminars in Hearing*, *9*, 357–366.

Musiek, F. E., & Weihing, J. (2011). Perspectives on Dichotic Listening and the Corpus Callosum. *Brain and Cognition*, *76*(2), 225–232. https://doi.org/https://doi.org/10.1016/j.bandc.2011.03.011

Nakagawa, S., & Cuthill, I. C. (2007). Effect Size, Confidence Interval and Statistical Significance: A Practical Guide for Biologists. *Biological Reviews*, *82*(4), 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

National Acoustic Laboratories. (2010). Interpreting LiSN-S Results and Recommendations for Management. Retrieved November 25, 2017, from https://capd.nal.gov.au/pdf/

National Collaborating Centre for Mental Health. (2009). *Attention Deficit Hyperactivity Disorder: The NICE Guideline on Diagnosis and Management of ADHD in Children, Young People and Adults*. Leicester & London, UK: The British Psychological Society & The Royal College of Psychiatrists. Retrieved from https://www.nice.org.uk/guidance/cg72/evidence/full-guideline-pdf-241963165

Nordstokke, D. W., & Zumbo, B. D. (2007). A Cautionary Tale About Levene's Tests for Equal Variances. *Journal of Educational Research & Policy Studies*, *7*(1), 1–14.

Nourski, K. V, & Brugge, J. F. (2011). Representation of Temporal Sound Features in the Human Auditory Cortex. *Reviews in the Neurosciences*, *22*(2), 187–203.

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., … Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment can be Decoded From Single-Trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706. https://doi.org/10.1093/cercor/bht355

Osman, D. M., Shohdi, S., & Aziz, A. A. (2011). Pragmatic Difficulties in Children With Specific Language Impairment. *International Journal of Pediatric Otorhinolaryngology*, *75*(2), 171–176. https://doi.org/10.1016/j.ijporl.2010.10.028

Otometrics. (2015). AURICAL FreeFit and the Probe Microphone Measurements Module: Reference Manual. Taa: GN Otometrics A/S.

Parasuraman, R. (1998). The Attentive Brain: Issues and Prospects. In R. Parasuraman (Ed.), *The Attentive Brain* (pp. 3–15). Cambridge, MA: The MIT Press.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The Plastic Human Brain Cortex. *Annual Review of Neuroscience*, *28*(1), 377–401. https://doi.org/10.1146/annurev.neuro.27.070203.144216

Paulovicks, J. (2008). The Gaps-in-Noise (GIN) Test and its Diagnostic Significance. *The*

Hearing Journal, *61*(3), 67.

Phillips, D. P. (2014). An Introduction to Central Auditory Neuroscience. In F. E. Musiek & G. D. Chermak (Eds.), *Handbook of Central Auditory Processing Disorder. Volume I. Auditory Neuroscience and Diagnosis* (2nd ed., pp. 113–152). San Diego, CA: Plural Publishing, Inc.

Pickles, J. O. (2015). Auditory Pathways: Anatomy and Physiology. In G. G. Celesia & G. Hickok (Eds.), *Handbook of Clinical Neurology. The Human Auditory System: Fundamental Organization and Clinical Disorders* (1st ed., Vol. 129, pp. 3–25). Oxford, United Kingdom: Elsevier B.V. https://doi.org/10.1016/B978-0-444-62630-1.00001-9

Pimentel, J. T., & Inglebret, E. (2014). Evidence-Based Practice and Treatment Efficacy. In G. D. Chermak & F. E. Musiek (Eds.), *Handbook of Central Auditory Processing Disorder. Volume II. Comprehensive Intervention* (2nd ed., pp. 39–64). San Diego, CA: Plural Publishing, Inc.

Purdy, S. C., Smart, J. L., Baily, M., & Sharma, M. (2009). Do Children With Reading Delay Benefit From the use of Personal FM Systems in the Classroom? *International Journal of Audiology*, *48*(12), 843–852. https://doi.org/10.3109/14992020903140910

Puvvada, K. C., & Simon, J. Z. (2017). Cortical Representations of Speech in a Multitalker Auditory Scene. *Journal of Neuroscience*, *37*(38), 9189–9196. https://doi.org/10.1523/JNEUROSCI.0938-17.2017

Rance, G., & Starr, A. (2016). Auditory Neuropathy Spectrum Disorder. In A. M. Tharpe & R. Seewald (Eds.), *Comprehensive Handbook of Pediatric Audiology* (2nd ed., pp. 227–246). San Diego, CA: Plural Publishing, Inc.

Roach, N. W., & Hogben, J. H. (2004). Attentional Modulation of Visual Processing in Adult Dyslexia. *Psychological Science*, *15*(10), 650–654.

Rocha-Muniz, C. N., Zachi, E. C., Teixeira, R. A. A., Ventura, D. F., Befi-Lopes, D. M., & Schochat, E. (2014). Association Between Language Development and Auditory Processing Disorders. *Brazilian Journal of Otorhinolaryngology*, *80*(3), 231–236. https://doi.org/10.1016/j.bjorl.2014.01.002

Roggia, S. M., & Colares, N. T. (2008). Mismatch Negativity in Patients With (Central) Auditory Processing Disorders. *Brazilian Journal of Otorhinolaryngology*, *74*(5), 705–711.

Rosenberg, G. G. (2002). Classroom Acoustics and Personal FM Technology in Management of Auditory Processing Disorder. *Seminars in Hearing*, *23*(4), 309–317. https://doi.org/10.1055/s-2002-35879

Rosenberg, G. G., Blake-Rahter, P., Heavner, J., Allen, L., Redmond, B. M., Phillips, J., & Stigers, K. (1999). Improving Classroom Acoustics (ICA): A Three-Year FM Sound Field Classroom Amplification Study. *Journal of Educational Audiology*, *7*, 8–28.

Sarter, M., Givens, B., & Bruno, J. P. (2001). The Cognitive Neuroscience of Sustained Attention: Where Top-down Meets Bottom-up. *Brain Research Reviews*, *35*(2), 146–160. https://doi.org/10.1016/S0165-0173(01)00044-3

Scerri, T. S., Morris, A. P., Buckingham, L.-L., Newbury, D. F., Miller, L. L., Monaco, A. P., … Paracchini, S. (2011). DCDC2, KIAA0319 and CMIP are Associated with Reading-Related Traits. *Biological Psychiatry*, *70*(3), 237–245.

https://doi.org/10.1016/j.biopsych.2011.02.005

Schafer, E. C., & Kleineck, M. P. (2009). Improvements in Speech Recognition Using Cochlear Implants and Three Types of FM Systems: A Meta-Analytic Approach. *Journal of Educational Audiology, 15*(1), 4–14.

Schafer, E. C., Mathews, L., Mehta, S., Hill, M., Munoz, A., Bishop, R., & Moloney, M. (2013). Personal FM Systems for Children With Autism Spectrum Disorders (ASD) and/or Attention-Deficit Hyperactivity Disorder (ADHD): An Initial Investigation. *Journal of Communication Disorders*, *46*(1), 30–52. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0021992412001165

Schafer, E. C., Traber, J., Layden, P., Amin, A., Sanders, K., Bryant, D., & Baldus, N. (2014). Use of Wireless Technology for Children With Auditory Processing Disorders, Attention-Deficit Hyperactivity Disorder, and Language Disorders. *Seminars in Hearing*, *35*(3), 193–205. https://doi.org/10.1055/s-0034-1383504

Schulz, K. F., & Grimes, D. A. (2005). Multiplicity in Randomised Trials I: Endponts and Treatments. *The Lancet*, *365*, 1591–1595.

Semel, E., Wiig, E., & Secord, W. (2003). *Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4)*. San Antonio, TX: The Psychological Corporation.

Sharma, M., Purdy, S. C., & Kelly, A. S. (2009). Comorbidity of Auditory Processing, Language, and Reading Disorders. *Journal of Speech, Language, and Hearing Research*, *52*(1), 706–722. https://doi.org/10.1044/1092-4388(2008/07-0226)

Sharma, M., Purdy, S. C., & Kelly, A. S. (2012). A Randomized Control Trial of Interventions in School-Aged Children With Auditory Processing Disorders. *International Journal of Audiology*, *51*(1), 506–518. https://doi.org/10.3109/14992027.2012.670272

Sibbald, B., & Roland, M. (1998). Understanding Controlled Trials: Why are Randomised Controlled Trials Important? *BMJ*, *316*, 201.

Smaldino, J. J., Crandell, C. C., & Flexer, C. A. (2005). Sound Field Amplification: Steps to Placement in the Classroom. In C. C. Crandell, J. J. Smaldino, & C. A. Flexer (Eds.), *Sound Field Amplification. Applications to Speech Perception and Classroom Acoustics* (2nd ed., pp. 142–165). Clifton Park, NY: Thomson Delmar Learning.

Smoski, W. J., Brunt, M. A., & Tannahill, J. C. (1998). *C.H.A.P.S. Instruction Manual*. Tampa, FL: Educational Audiology Association.

Snowling, M. J., & Hulme, C. (2003). A Critique of Claims From Reynolds, Nicolson & Hambly (2003) that DDAT is an Effective Treatment for Children With Reading Difficulties - "Lies, Damned Lies and (Inappropriate) Statistics?" *Dyslexia*, *9*(2), 127–133. https://doi.org/10.1002/dys.245

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of Speech in Noise: Neural Correlates. *Journal of Cognitive Neuroscience*, *23*(9), 2268–2279. https://doi.org/10.1162/jocn.2010.21556

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to Improve Hearing Speech in Noise: Biological Mechanisms. *Cerebral Cortex*, *22*(5), 1180–1190. https://doi.org/10.1093/cercor/bhr196

Stavrinos, G., Iliadou, V. V., Edwards, L., Sirimanna, T., & Bamiou, D.-E. (2018). The Relationship Between Types of Attention and Auditory Processing Skills: Reconsidering Auditory Processing Disorder Diagnosis. *Frontiers in Psychology*, *9*(January), 1–13. https://doi.org/10.3389/fpsyg.2018.00034

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple Imputation for Missing Data in Epidemiological and Clinical Research: Potential and Pitfalls. *BMJ*, *338*, b2393. https://doi.org/10.1136/bmj.b2393

Tallal, P. (2004). Improving Language and Literacy is a Matter of Time. *Nature Reviews. Neuroscience*, *5*(9), 721–728. https://doi.org/10.1038/nrn1499

Tawfik, S., Mohamed Hassan, D., & Mesallamy, R. (2015). Evaluation of Long Term Outcome of Auditory Training Programs in Children With Auditory Processing Disorders. *International Journal of Pediatric Otorhinolaryngology*, 1–7. https://doi.org/10.1016/j.ijporl.2015.11.001

Toe, D. (2009). Managing the Listening Environment: Classroom Acoustics and Assistive Listening Devices. In V. E. Newton (Ed.), *Paediatric Audiological Medicine* (2nd ed., pp. 372–388). Chichester: John Wiley & Sons, Inc.

Tomblin, J. B., Smith, E., & Zhang, X. (1997). Epidemiology of Specific Language Impairment: Prenatal and Perinatal Risk Factors. *Journal of Communication Disorders*, *30*(4), 325–343. https://doi.org/10.1016/S0021-9924(97)00015-4

Tomlin, D., Dillon, H., Sharma, M., & Rance, G. (2015). The Impact of Auditory Processing and Cognitive Abilities in Children. *Ear & Hearing*, *36*(5), 527–542. https://doi.org/10.1097/AUD.0000000000000172

Tremblay, K. L. (2007). Training-Related Changes in the Brain: Evidence From Human Auditory-Evoked Potentials. *Seminars in Hearing*, *28*(2), 120–132. https://doi.org/10.1055/s-2007-973438.

Umat, C., Mukari, S. Z., Ezan, N. F., & Din, N. C. (2011). Changes in Auditory Memory Performance Following the use of Frequency-Modulated System in Children With Suspected Auditory Processing Disorders. *Saudi Medical Journal*, *32*(8), 818–824. Retrieved from http://smj.psmmc.med.sa/index.php/smj/article/view/7248

Updike, C. D. (2006). The Use of FM Systems for Children with Attention Deficit Disorder. *Journal of Educational Audiology*, *13*, 7–14.

van der Knaap, L. J., & van der Ham, I. J. M. (2011). How Does the Corpus Callosum Mediate Interhemispheric Transfer? A review. *Behavioural Brain Research*, *223*(1), 211–221. https://doi.org/10.1016/j.bbr.2011.04.018

Wada, H. (2007). The Ear: Its Structure and Function, Related to Hearing. In M. J. Crocker (Ed.), *Handbook of Noise and Vibration Control* (pp. 277–285). Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470209707.ch20

Wahn, B., & König, P. (2017). Is Attentional Resource Allocation Across Sensory Modalities Task-Dependent? *Advances in Cognitive Psychology*, *13*(1), 83–96. https://doi.org/10.5709/acp-0209-2

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children (WISC-III UK)* (3rd ed.). Oxford, United Kingdom: Pearson Assessment.

Wechsler, D., & Naglieri, J. A. (2006). *Wechler Nonverbal Scale of Ability. Technical and Interpretive Manual*. San Antonio, TX: Pearson, Inc.

Weihing, J., Chermak, G. D., & Musiek, F. E. (2015). Auditory Training for Central Auditory Processing Disorder. *Seminars in Hearing*, *36*(4), 199–215. https://doi.org/10.1055/s-0035-1564458

Westerhausen, R., & Hugdahl, K. (2008). The Corpus Callosum in Dichotic Listening Studies of Hemispheric Asymmetry: A Review of Clinical and Experimental Evidence. *Neuroscience and Biobehavioural Reviews*, *32*(5), 1044–1054. https://doi.org/https://doi.org/10.1016/j.neubiorev.2008.04.005

Wilson, W. J. (2018). Evolving the Concept of APD. *International Journal of Audiology*, *57*(4), 240–248. https://doi.org/10.1080/14992027.2017.1409438

Wilson, W. J., & Arnott, W. (2013). Using Different Criteria to Diagnose (Central) Auditory Processing Disorder: How big a Difference Does it Make? *Journal of Speech Language and Hearing Research*, *56*(1), 63–70. https://doi.org/10.1044/1092-4388(2012/11-0352)b

Wilson, W. J., Jackson, A., Pender, A., Rose, C., Wilson, J., Heine, C., & Khan, A. (2011). The CHAPS, SIFTER, and TAPS-R as Predictors of (C)AP skills and (C)APD. *Journal of Speech, Language, and Hearing Research*, *54*(1), 278–291. https://doi.org/10.1044/1092-4388(2010/09-0273)

Wolraich, M., Brown, L., Brown, R. T., DuPaul, G., Earls, M., Feldman, H. M., … Visser, S. (2011). ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/ Hyperactivity Disorder in Children and Adolescents. *Pediatrics*, *128*(5), 1007–1022. https://doi.org/10.1542/peds.2011-2654

Wong, P. C. M., Uppunda, A. K., Parrish, T. B., & Dhar, S. (2008). Cortical Mechanisms of Speech Perception in Noise. *Journal of Speech, Language, and Hearing Research*, *51*, 1026–1041.

Wright, J. G., Swiontkowski, M. F., & Heckman, J. D. (2003). Introducing Levels of Evidence to The Journal. *The Journal of Bone & Joint Surgery*, *85-A*(1), 1–3.

Wright, T. (2001). Structure and Function of the Ear: its Anatomy and how we Hear. In J. Graham & M. Martin (Eds.), *Ballantyne's Deafness* (6th ed., pp. 26–54). London, United Kingdom: Whurr Publishers.

Yathiraj, A., & Vanaja, C. S. (2015). Age Related Changes in Auditory Processes in Children Aged 6 to 10 Years. *International Journal of Pediatric Otorhinolaryngology*, *79*(8), 1224–1234. https://doi.org/10.1016/j.ijporl.2015.05.018

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., … Schroeder, C. E. (2013). Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a "Cocktail Party." *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

APPENDICES

Appendix A

Studies	**APD Diagnostic criteria used in each study**
Allen & Allan (2014)	Suspected APD initial sample was between ages 7 to 17. Referrals were made from local paediatricians, audiologists, schools, parents and family friends. Two of the sixty-three children from the initial sample (suspected of APD) had hearing thresholds beyond 20 dB HL (at all octave frequencies between 250-8000 Hz), but it is unclear whether these cases were diagnosed with APD. Children had to have normal tympanograms, ABRs and acoustic reflexes. They had to have at least 2 tests at -2 SDs from the mean from these 5 AP tests: Staggered Spondaic Word Test (dichotic listening), Auditory Fusion Test - Revised (temporal gap detection), Pitch Pattern Sequence test (perception of temporal order for 3 tone sequences), Filtered Speech test (perception of monaural speech low pass filtered at 500 Hz), Word in Ipsilateral Competition (an AFG test). All tests were scored based on normative data.
Cameron et al. (2016)	Children in the APD-diagnosed group were between 7 years and 9 years 11 months. Referrals were made from certified primary school teachers, educational psychologists, or paediatricians. Children had to have Australian English as first language, hearing thresholds below 20 dB HL (at all octave frequencies between 250-8000 Hz), normal tympanograms, normal acoustic reflexes. They had to have normal intellectual performance (but no mention of the cut-offs). Children with an ADHD diagnosis were excluded. There was no mention of the APD criteria and cut-offs that were used to identify APD.
Dawes & Bishop (2010)	Children in the APD group had a mean age of 10.4 years. Recruitment was from 4 hospitals in the UK. IQ performance had to be above 80 (standard score). All participants had normal hearing (below 20 dB HL between 250-8000 Hz). APD diagnosis was based on the following: complaint of listening difficulties, normal peripheral hearing, score at -2 SD from the mean on the SCAN-C or SCAN-A and failure of at least one or more AP test (i.e. Pitch Patterns test, Duration Patterns test, Random Gap Detection test; but no mention of cut-offs). All test scores were based on normative data.
Dawes et al. (2008)	Children in the APD group had a mean age of 10.3 years. Recruitment was from GOSH in London. Children had to have normal pure tone audiometry, otoscopy and tympanometry. They had to have -1 SD from the mean on the SCAN test and at least one other failed AP test (i.e. Random Gap Detection, GIN, Pitch Pattern and Duration Pattern test).
Ferguson et al. (2011)	Children had to have normal audiometry (below 20 dB HL at all octave frequencies between 250-8000 Hz), normal middle ear function (more than 150 daPa) and middle ear compliance (more than .2 cc). They had to have English as first language. Children had to have typical symptoms of APD reported by parents (i.e. listening in background noise difficulty, difficulty expressing or clearly using speech, difficulty understanding when listening, difficulty remembering complex instructions, easily distracted and difficulty staying focused).

Filippini et al. (2012)	Children were between 7-13 years old. They had to have normal peripheral hearing (below 20 dB HL at all octave frequencies between 250-8000 Hz), normal otoscopic evaluation, normal middle ear function and normal responses to click-evoked ABR. To be diagnosed with APD they had to fail 2 of the 3 AP tests administered (i.e. SIN test, DDT, Pitch Pattern Sequencing test). There was no mention of cut-offs.
Gyldenkaerne et al. (2014)	Children were between 7-12 years old. They had to have normal peripheral hearing (below 20 dB HL at all octave frequencies between 500-4000 Hz), type A tympanograms, and present contralateral acoustic reflexes (2 KHz). They had to have performance at -2 SDs from the mean on any one of the FPT and GIN tests (both ears).
Iliadou & Bamiou (2012)	Children were between 11 years 4 months and 12 years 7 months. They had to have normal audiometry (below 20 dB HL at all octave frequencies between 500-8000 Hz), normal middle ear pressure (more than -150 mm) and middle ear compliance (more than .3 cc). They had to have at least 2 AP tests with normative scores at -2 SDs from the mean (in at least one ear). At least one of the failed tests had to be a non-speech test. The tests administered were 2 verbal tests (monaural low-redundancy Greek Speech in Babble test, Greek DDT) and 4 non-verbal tests (FPT, Duration Pattern test, Random Gap Detection, and Masking Level Difference). Children did not have ASD or ADHD diagnosis.
Johnston et al. (2009)	Recruited children were between ages 8 years 2 months and 15 years 7 months. They had to have normal peripheral hearing and normal middle-ear function (type A tympanogram). They had to have at least 2 AP tests with scores at -2 SDs from the mean (in at least one ear). The AP tests administered were assessing AFG performance (using the Goldman-Fristoe-Woodcock test of auditory discrimination and synthetic sentence identification), dichotic listening (using the Staggered spondaic words and Dichotic Digits Double Pairs tests), phonemic awareness (using a test of auditory analysis skills and phonemic synthesis test) and auditory sequencing performance (using the duration pattern test and pitch pattern sequence).
Kuk et al. (2008)	Minimum age had to be 7 years old. Children had to have normal hearing sensitivity at 500-4000 Hz, normal tympanogram and acoustic reflexes. At least 2 of the AP test had to be at -2 SDs below the mean. The AP tests administered were: Staggered Spondaic Words, Phonemic Synthesis, Filtered Words, Dichotic Digits, Competing Sentences, Pitch Pattern Sequence, and SIN tests.
Lagace et al. (2011)	Children had to be between 9-12 years old. They had to have normal hearing (15 dB HL and below at all octave frequencies between 500-4000 Hz). They had to have Canadian French as the native language and language of education. They were included only if they did not have conductive or sensorineural hearing loss, a clinical diagnosis of language, attention or learning disorders. They had to have at least 1 AP test with scores below 2 SDs from the mean. The AP tests administered were the Canadian-French version of the Synthetic Identification with Ipsilateral Competing Message test (a SIN test) and the Canadian-French version of the Staggered Spondaic Word test (a SIN and dichotic listening test). Both tests had normative data.

Lam & Sanchez (2007)	Children had to be 7 to 10 years 11 months. They had to have English as their first language. Children with hearing or visual impairments were excluded. Children with diagnoses of language disorder, epilepsy, ASD, head injury were excluded. Cases with ADHD were not excluded (children who took medication had to withhold it 16 hours before the study). Also, cases with learning difficulties or language delays were not excluded. They had to have normal peripheral hearing (below 25 dB HL at frequencies 250, 1000, and 4000 Hz). They had to have scores of -2 SDs from the mean in at least two AP tests (Staggered Spondaic Word, Competing Speech test, DDT, and two auditory memory tests for digits and sentences). The latter two tests were not standardised while the rest had normative data.
Loo et al. (2016)	Children had to be in mainstream schools. They had to be referred based on concerns of listening difficulties. They had to have normal peripheral hearing (below 20 dB HL at all octave frequencies between 250-8000 Hz), normal middle ear function (Type A tympanograms) and normal ipsilateral acoustic reflexes (present at 1000 Hz with a threshold below 100 dB HL). They had to have speech discrimination scores (on the North-western University Auditory Test - 6) of at least 80% correct (both ears at 50 dB HL). They had to have scores of -2 SDs from the mean (in both ears) in at least two (but not all) of the AP tests (i.e. FPT, Duration Pattern test, Temporal resolution test, Random Gap Detection test, Binaural Processing test, Masking Level Difference, Dichotic Speech test, and the DDT). They had to have normal non-verbal IQ (more than 85 - standard score). They were excluded if they had ASD or neurological conditions such as head injuries and brain tumours.
Lotfi, Mehrkian et al. (2016)	Children had to have normal audiometry (below 20 dB HL at all octave frequencies between 250-8000 Hz), normal otoscopy, tympanometry and speech discrimination score. They had to have normal IQ (more than 85 standard score). Children were excluded if they had a history of hearing impairment, ear diseases, and neurological problems. Children had to have poor performance (undetermined cut-offs) in all 3 AP tests administered (i.e. the DDT, Pitch Pattern Sequence test, Monaural Selective Auditory Attention test).
Maerlender et al. (2004)	Children had to be between 7-15 years of age. They had to have normal peripheral hearing (no mention of cut-offs). They had to have scores of 2 SDs below the mean on at least two AP tests (i.e. the DDT, Word recognition test, Duration Pattern Sequencing, Gap detection).
Miller & Wagstaff (2011)	Children had to be monolingual English speakers. They had to have normal hearing, normal cognitive development, no psychiatric disorder, no pervasive developmental disorder, no neurological damage, no uncorrected visual impairment and no motor impairment. Children were between 8 years 5 months and 10 years 1 month. Children had pure-tone thresholds of 25 dB HL or below at all octave frequencies between 250-4000 Hz. They had to score below the age-based cut-offs (provided by the tests), as SDs were not provided, in at least 2 of the AP tests (i.e. FPT, Duration Pattern test, DDT and Staggered Spondaic Word).

Moossavi et al. (2018)	Children were between ages 9-11. They had to have normal hearing (15 dB HL or better at all octave frequencies between 250-8000 Hz), no history of auditory pathology, normal tympanometry, and speech discrimination scores. They had to have normal IQ (above 85 - standard score). They were excluded if they had history of hearing impairment, ear diseases and neurological difficulties. Children had to fail (no cut-offs mentioned) all 3 AP tests (i.e. the DDT, Pitch Pattern Sequencing test and monaural Selective Auditory Attention test).
Rocha-Muniz (2014)	Children were between ages 6-12. Children had to have normal hearing (15 dB HL or better at all octave frequencies between 500-4000 Hz, normal tympanogram, speech recognition scores at or above 88% and no neurological, cognitive or psychiatric disorders. They had to have scores at 2 SDs below the mean on at least two of the AP tests (minimum assessment battery included tests of temporal processing-FPT, monotic SIN and dichotic listening-DDT).
Roggia & Colares (2008)	Children were between 9-14 years of age. They had to have normal hearing (at or below 15 dB HL), speech recognition threshold within tone averages (equal mean values of 500, 1000, 2000 Hz or up to 10 dB HL poorer), speech recognition index at or above 88%, and normal tympanogram (type-A). Children had to have APD diagnosis (based on behavioural assessment) from the Audiology Clinic of the University (i.e. Universidade do Vale do Itajaí in Brazil), without specific tests or cut-off criteria being mentioned.
Schafer et al. (2014)	Children were between the ages of 6 and 11. They had to have normal hearing (below 20 dB HL at all octave frequencies between 250-6000 Hz). Official APD diagnosis was given by licensed professionals. Results of that assessment were not accessible to the investigators. Therefore, there is no mention of specific AP tests or cut-offs.
Sharma et al. (2009)	Children had to have normal hearing at 15 dB HL or better at all octave frequencies between 250-8000 Hz, Type-A tympanograms, ipsilateral 1000 Hz acoustic reflex threshold below 100 dB HL, non-verbal intelligence above 80 (standard score), phoneme scores on both ears at 90% or better (in quiet), normal otoacoustic emission. They had to have a score at 2 SDs below the mean on at least 2 of the AP tests or a score of 3 SDs below the mean on only one of the AP tests (i.e. FPT, DDT, Random Gap Detection test, Masking Level Difference).
Sharma et al. (2012)	Children were between the ages of 7 and 13. Normal non-verbal intelligence (above 80 - standard score). They had to have -2 SDs from the mean on any task of the AP measures (i.e. DDT, FPT, Random Gap Detection, compressed and reverberant HINT sentences, Masking Level Difference).
Tawfik et al. (2015)	Children had to have normal hearing (at 15 dB HL or better at all octave frequencies between 500-4000 Hz), normal middle ear function and excellent speech discrimination (without specified cut-offs). Children were diagnosed as having APD from the Audiology Unit, El-Demerdash Hospital, Ain Shams University, Cairo – Egypt, but without mention of the specific tests and cut-off criteria used.
Tomlin et al. (2015)	Children were between ages 7 and 12. They had to have bilateral normal hearing (at or below 15 dB HL at all octave frequencies between 250-8000 Hz) and normal middle ear function (type-A tympanogram). They had to have normative scores at -2 SDs from the mean on at least 2 AP tests or -3 SDs from the mean on only one AP test (i.e. FPT, DDT, GIN, Masking Level Difference, LiSN-S).

Umat et al. (2011)	Children were between ages 7 and 10. They had to have normal hearing (below 20 dB HL at all octave frequencies between 500-4000 Hz on both ears), bilateral normal middle ear function (type-A tympanogram), and normal IQ (at or better than 80 - standard score). They had to have scores at 2 SDs below the mean in one of the 2 AP tests (i.e. DDT, Pitch Pattern Sequence). Children were excluded if they had poor record on school attendance and if they had ADHD diagnosis. This was to ensure that children would comply with the study's aim (i.e. to use RMHA during school hours and study period).
Wilson et al. (2011)	Children were between ages 6 years 9 months and 14 years 3 months. They had to have Australian English as their first language. They had to be attending a private or public school in Australia. They had to have normal peripheral hearing (at or better than 15 dB HL at all octave frequencies between 250-8000 Hz), normal speech performance-intensity functions, normal tympanogram (higher than .2 ml) and normal tympanic pressure (between -100 and +50 daPa). They had to have completed at least 2 of the AP tests (i.e. Low-pass filtered speech test, Competing sentences test, DDT, FPT). To be labelled at risk of APD, children had to either score -2 SDs from the mean on at least 2 of the AP tests binaurally, or score -2 SDs from the mean on at least 1 of the AP tests binaurally within one or more of the AP domains (i.e. Domain of monaural low-redundancy speech which included the Low-pass filtered speech test, Domain of dichotic speech which included the DDT and Competing sentences, and the Domain of Auditory temporal processing and pattern test which included the FPT). Children were excluded if they had a diagnosis of intellectual, cognitive, attentional, emotional or articulation impairments.