Aachener Beiträge zur
Hörtechnik und Akustik

Karin Loh

Noise Exposure in Pre- and Primary Schools:

Exploring Noise Effects on Young Children

IHTA Institute for Hearing Technology and Acoustics

RWTH AACHEN UNIVERSITY

Noise Exposure in Pre- and Primary Schools: Exploring Noise Effects on Young Children

Von der Fakultät für Elektrotechnik und Informationstechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades einer

DOKTORIN DER INGENIEURWISSENSCHAFTEN

genehmigte Dissertation

vorgelegt von

Karin Loh
M.Sc.

aus Aachen, Deutschland

Berichter:

Universitätsprofessorin Dr.-Ing. Janina Fels
Universitätsprofessorin Kerstin Persson-Waye PhD

Tag der mündlichen Prüfung: 31.01.2025

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

To my beloved father

Abstract

As the awareness of health effects caused by noise has grown in recent years, so has the importance of studying noise in educational buildings. This dissertation focuses specifically on the impact of noise on young children in educational buildings from a children's perspective, a topic that has not been extensively explored. Previous research indicated that noise can significantly impact children's performance and cognition, making this a crucial area of study. However, noise in educational buildings is complex, involving internal background noise, individual distracting sound incidences, external sources, and noise generated by children during activities. Communication in such environments resembles the *cocktail-party effect*. Focusing on children's perspective in these complex acoustic environments, there is limited research on the perceptual and cognitive differences between children and adults regarding noise, and whether adult-based insights apply to children remains uncertain. The dissertation has two primary aims: 1.) addressing the complexity of noise in educational buildings, and 2.) examining noise effects on young children in controlled listening experiments. For the first aim, children's perspectives are integrated by assessing noise scenes through room acoustics and in-situ measurements using head and torso simulators of different anthropometric sizes. Frequency-dependent measures (e.g., psychoacoustic parameters) and their relation to subjective noise ratings by children and adults are explored. Furthermore, the impact of different activities on noise levels is investigated. For the second aim, a child-appropriate listening experiment framework is developed to study young children's cognitive performance, focusing on auditory selective attention in realistic noisy environments. This includes creating and validating a paradigm suited for children, investigating age effects in children aged three to six, and examining noise effects on cognition and physical stress through changes in heart rate variability.

Contents

Glossary

1 Introduction 1

2 Towards Child-Appropriate Noise Assessment Methods 7

2.1 State of the Art . . . 7
2.2 Involved Educational Institutions . . . 13
2.3 Acoustic Measurements in Pre- and Primary schools . . . 14
2.3.1 Room acoustic measurement procedure . . . 14
2.3.2 In-situ noise measurement procedure . . . 17
2.3.3 General measurement results . . . 19
2.3.4 Discussions on child-appropriate acoustic measurement . . . 21
2.3.5 Concluding remarks on acoustic measurements . . . 26
2.4 Influences of Anthropometric Sizes of Adults and Children Head and Torso Simulators (HATSs) . . . 27
2.4.1 Room acoustic parameters . . . 30
2.4.2 Noise parameters . . . 36
2.4.3 Binaural speech transmission index of HATSs . . . 43
2.4.4 Concluding remarks on HATS differences . . . 51
2.5 Children's Noise Perception Survey . . . 52
2.5.1 Questionnaire and interview procedure . . . 52
2.5.2 Differences in age group and room types . . . 54
2.5.3 Correlation of subjective responses and objective measurements . . . 59
2.5.4 Concluding remarks . . . 63
2.6 Activity-based Acoustic Settings (ABAS) . . . 63
2.6.1 Acoustic properties of ABAS . . . 65
2.6.2 Classroom ABAS . . . 66
2.6.3 Playroom ABAS . . . 68
2.6.4 Discussions on ABAS . . . 68
2.6.5 Concluding remarks on ABAS . . . 70
2.7 Insights on Child-Appropriate Noise Assessment . . . 71

3 Children's Auditory Cognition in Noise **73**
3.1 State of the Art . . . 73
3.2 Children's Intentional Switching of Auditory Selective Attention 79
3.2.1 Stimulus material . . . 80
3.2.2 Task . . . 81
3.2.3 Gamification . . . 83
3.2.4 Binaural reproduction . . . 85
3.2.5 Room setup . . . 86
3.2.6 General procedure . . . 87
3.2.7 General experiment design . . . 88
3.3 Child-appropriate Paradigm Validation Study . . . 91
3.3.1 Participants . . . 93
3.3.2 Experiment design in detail . . . 94
3.3.3 Results and discussion . . . 94
3.4 Developmental Differences in Young Children . . . 99
3.4.1 Participants . . . 99
3.4.2 Experiment design in detail . . . 100
3.4.3 Results and discussion . . . 102
3.5 Comparing Physiological and Cognitive Noise Effects . . . 115
3.5.1 Participants . . . 115
3.5.2 Evaluation of heart rate variability . . . 116
3.5.3 Experiment design in detail . . . 117
3.5.4 Results and discussion . . . 118
3.6 Insights on Children's Auditory Cognition in Noise . . . 129

4 Summary and Outlook **133**

A Appendix - Child-Appropriate Noise Assessment **137**
A.1 Complete overview on RA parameters . . . 138
A.2 Complete overview on in-situ parameters . . . 144
A.3 INCH results per age group . . . 155
A.4 ABAS results (prominent-ear method) . . . 158

B Appendix - Children's Auditory Cognition in Noise **161**
B.1 ChildASA noise study (post-hoc analyses) . . . 161

Bibliography **164**

Acknowledgements **177**

Curriculum Vitae **179**

Abbreviations

ABAS	activity-based acoustic setting
A-HATS	adult head and torso simulator
C-HATS	child head and torso simulator
HATS	head and torso simulator
STI	speech transmission index
MTI	modulation transfer index
MTF	modulation transfer function
SPL	sound pressure level
HRTF	head-related transfer function
CR	classroom
PR	playroom
HRV	heart rate variability
RT	response time
ER	error rate

Symbols

General

A-HATS$_{\mathbf{av}}$	binaural parameter for A-HATS using averaging method
A-HATS$_{\mathbf{prom}}$	binaural parameter for A-HATS using prominent-ear method
C-HATS$_{\mathbf{av}}$	binaural parameter for C-HATS using averaging method
C-HATS$_{\mathbf{prom}}$	binaural parameter for C-HATS using prominent-ear method

Room acoustic parameters

$T_{\mathbf{20}}$	reverberation time [s] at the SPL decay of 20 dB
$T_{\mathbf{20,occ}}$	reverberation time [s] at the SPL decay of 20 dB (occupied)
$T_{\mathbf{30}}$	reverberation time [s] at the SPL decay of 30 dB
$C_{\mathbf{50}}$	clarity (for speech conditions) [dB]
$D_{\mathbf{50}}$	definition (for speech conditions) [%]
$U_{\mathbf{50}}$	ratio of useful to detrimental energy [dB]
EDT	early decay time [s]
$T_{\mathbf{S}}$	center time [s]
STI	speech transmission index
$BNL_{\mathbf{A,eq,30s}}$	equivalent A-weighted background noise level [dB]
$IACC_{\mathbf{full}}$	full interaural cross-correlation coefficient
$IACC_{\mathbf{early}}$	early interaural cross-correlation coefficient
$IACC_{\mathbf{late}}$	late interaural cross-correlation coefficient

Noise parameters

$L_{\mathbf{Z}}$	unweighted level [dB]
$L_{\mathbf{Z,10}}$	10th-percentile of unweighted level [dB]
$L_{\mathbf{Z,90}}$	90th-percentile of unweighted level [dB]
$L_{\mathbf{Z,eq}}$	equivalent unweighted level [dB]
$L_{\mathbf{ZF,mean}}$	unweighted level using a fast time weighting [dB]
$L_{\mathbf{Z,eq}}\mathbf{(L.v.H)}$	low-high-frequency ratio of $L_{\mathrm{Z,eq}}$
$L_{\mathbf{A}}$	A-weighted level [dB]

$L_{\mathbf{A,10}}$	10th-percentile of A-weighted level [dB]
$L_{\mathbf{A,90}}$	90th-percentile of A-weighted level [dB]
$L_{\mathbf{A,eq}}$	equivalent A-weighted level [dB]
$L_{\mathbf{AF,mean}}$	A-weighted level using a fast time weighting [dB]
$L_{\mathbf{A,eq}}$**(L.v.H)**	low-high-frequency ratio of $L_{\mathrm{A,eq}}$
$N_{\mathbf{mean}}$	mean loudness [sone]
$N_{\mathbf{5}}$	5th-percentile loudness [sone]
$N_{\mathbf{90}}$	90th-percentile loudness [sone]
$S_{\mathbf{mean}}$	mean sharpness [acum]
$S_{\mathbf{5}}$	5th-percentile sharpness [acum]
$S_{\mathbf{90}}$	90th-percentile sharpness [acum]
$R_{\mathbf{mean}}$	mean roughness [asper]
$R_{\mathbf{5}}$	5th-percentile roughness [asper]
$R_{\mathbf{90}}$	90th-percentile roughness [asper]
$FS_{\mathbf{mean}}$	mean fluctuation strength [vacil]
$FS_{\mathbf{5}}$	5th-percentile fluctuation strength [vacil]
$FS_{\mathbf{90}}$	90th-percentile fluctuation strength [vacil]

Auditory selective attention switching

Sw	switch
Rep	repetition
NoN	no noise
N	noise
C	congruent
SC	semi-congruent
IC	incongruent
LR	left-right position combination
Next	neighbouring position combination
FB	front-back position combination

Heart rate variability

$meanHR$	mean heart rate
$meanNN$	mean NN-intervals
$StdHR$	standard deviation of heart rate
$StdNN$	standard deviation of NN-intervals
$RMSSD$	root mean square of successive differences
$SD1$	Poincaré plot component quantifying short-term HRV
$SD2$	Poincaré plot component quantifying long-term HRV
$SD1/SD2$	ratio of SD1 and SD2

1 Introduction

Noise in educational buildings, such as pre- and primary schools, is a well-known problem for adults and children. Previous research has shown that noise has detrimental effects on children's speech perception (e.g., Klatte et al., 2013; Shield & Dockrell, 2003) and on their cognitive processes (e.g., Jones et al., 2015). Shield and Dockrell (2008) conducted a noise survey in British primary schools that revealed high noise levels in occupied classrooms exceeding 50 dB[A]. However, the acoustic environment, and with this also children's development and performance, is not only impacted by the noise but also by the acoustic design of the rooms. Unfavorable room acoustics can lead to more difficulty for children in following lessons and listening in general. Loh et al. (2022b) summarized in their work the effects of high sound levels and room acoustics:

> [High sound pressure level (SPL)s] can have detrimental effect on adults' well-being at work (e.g., Åhlander et al., 2011) and on children's behavior and development (for reviews, see Klatte et al., 2013; Shield & Dockrell, 2003, 2008). Unfavorable room acoustics, such as long reverberation times relative to the room volume, characterized as being outside the 0.5- 0.8 s optimum range (in occupied rooms) in (Astolfi et al., 2019b), have been shown to lower performance in phoneme identification in adults and children (Neuman & Hochberg, 1983, where reverberation time $T = 0.6\,\mathrm{s}$ were detrimental compared to $T = 0.4\,\mathrm{s}$ or no reverberation), impairment in primary school children's speech perception and listening comprehension (Klatte et al., 2010a), short-term memory (Klatte et al., 2010b), and negative effects on performance, well-being and social climate at school (Klatte et al., 2010c).

Nowadays, children spend a considerable amount of time in educational buildings, starting at three years old and sometimes earlier. Appropriate noise assessment and subsequent control are, therefore, crucial for providing optimal development and learning environments as indicated in the work by Loh et al. (2022b) when reviewing existing literature on classroom acoustics:

> While several studies have reported results of room acoustics measurements (in both occupied and unoccupied rooms), long-term noise measurements are scarce, especially in daycare settings. Further, for characterizing noise and room acoustic measurements in such institutions, two possibilities include using omnidirectional and binaural transducers. The former allows a range of measurements including standardized ones (ANSI Inc., 2002; Astolfi et al., 2019a; Bradley et al., 1999; Building Bulletin 93, 2015; DIN e.V., 2016) while the latter, generally incorporated as microphones near human (who may or may not have freedom of movement) ears, or within ear canals of head and torso simulator (HATS). The latter allow measurements that can represent some of the effects of head, shoulders, and outer ear processing for static listeners (i.e., without head movements). Adult HATS in binaural measurement procedures are relatively common in research settings including in classrooms for children (e.g., Z. Peng et al., 2013; Shinn-Cunningham et al., 2005), and there is at least one example of head and shoulder simulator (Fels et al., 2004; Prodi et al., 2007) that has been qualified to closely represent children aged approximately 3-6 years (hereinafter, referred to as children/child HATS). This paper reports on room acoustics measurements in unoccupied conditions (furnished rooms) and long-term noise measurements during typical hours of occupancy, in several primary schools and daycare centers, using both omnidirectional and binaural transducers. This includes investigating the extent to which relevant acoustic and psychoacoustic parameters vary across the educational settings, and between the transducers, i.e., omnidirectional, adult and child HATS, with the latter two providing first order representations of teachers and students' perception, respectively.

Noise assessment would, therefore, benefit from long-term measurements and consistent measurement methods providing reliable parameters (for a review, see Sala & Rantala, 2016). It further allows the integration of children's perspectives by respecting particular differences between adults and children, such as different anthropometric sizes. These differences affect the physical sound signals (cf. Fels et al., 2004; Fels & Vorländer, 2009) and can lead to a different sound perception between adults and children (cf. Loh et al., 2023b). For this process, it is necessary to incorporate many considerations when choosing suitable room acoustic and noise descriptors. In case of noise and long-term measurements, Loh et al. (2023b) summarized the problem as such:

> Until now, the A-weighted sound pressure level in dB[A] has been the primary descriptor commonly linked to noise effects though the A-weighting was derived from equal loudness curves at 40 phon and does not accurately describe effects at higher levels. Nevertheless, noise consists of a complex

structure including temporal and spectral components, which only the A-weighted sound pressure level might not appropriately address. Psychoacoustic parameters, such as loudness and sharpness, can provide additional insights into spectral components like the high-frequency proportions or the perceived loudness of a sound (Zwicker & Fastl, 2013). Therefore, they might be more appropriate to reflect the subjective impressions of sounds and noises (Genuit & Fiebig, 2014).

To reach higher ecological validity, it seems natural to consider head and torso simulators within acoustic measurements to reflect binaural hearing processes and the natural transmission of sound from the sender to the human ear (Blauert, 1997). With this and considering differences in anthropometric sizes (cf. Fels & Vorländer, 2009), adding children's perspectives on noise perception is possible. Fels et al. (2004) and Fels and Vorländer (2009) indicated in their work that a shift in frequency amplification towards higher frequencies for children exists.

When focusing on children and their daily activities, differences in the sound level introduced by the activities should be considered. As Loh et al. (2023b) summarized in their work, the categorization of activities was mainly defined from an educational perspective. Previous work by Shield et al. (2015), for example, distinguished the activities from an educational perspective and defined four categories: *plenary*, *group work*, *individual work* and *watching/listening*. They also found that noise level strongly depend on the activities, for example, highest noise levels were observed in group work while plenary and individual work were comparable and revealed lower noise levels than group work. In preschools, activities were differentiated between indoor play, outdoor play, singing, and silent activity (cf. Berggren et al., 2008; Noguchi et al., 2019). Picard and Boudreau (1999) chose a more systematic approach by categorizing the activities into structured and unstructured activities and lunch. A further approach was introduced by McAllister et al. (2009) by choosing the time of the day to distinguish the activity. Therefore, no common approach to categorize activities in pre- and primary schools existed, especially when the acoustic point-of-view should be taken into account systematically. Additionally, most studies decided to compare the acoustic properties using parameters based on sound pressure level (e.g., equivalent A-weighted SPL $L_{\mathrm{A,eq}}$ or mean SPLs), which remains questionable to reflect noise perception from a children's perspective sufficiently.

To sum up, and as Loh et al. (2022b) described in their work, it has become of high importance to understand the existing noise and the acoustic environment of educational buildings since a considerable part of children's education and development is happening there. Noise is a complex phenomenon that, on the one hand, is highly subjective from the listener's point of view. On the other hand, it can consist of several acoustic components integrating into an overall perceptional impression rated as annoying. From the perspective of acoustics, noise can comprise general background noise and individual sound incidences that are spatially distributed around the listener and are mostly distracting to the listener. Background noise in educational buildings can be either caused by external sources, such as frequented streets and construction sites, or internal sources, such as ventilation systems and teaching equipment, but also the children themselves can contribute significantly to the overall acoustic situation.

Therefore, communication in these noisy acoustic environments tends to be challenging for children and resembles the well-known *cocktail-party effect* by Cherry (1953). The goal is to understand better the mechanism of the cocktail-party effect within children under noise exposure. It was, therefore, standing to reason to include noise systematically step-by-step into controlled environments for experimental studies and to find paradigms to assess the effects of noise on children's hearing appropriately. Within the cocktail-party situation, auditory selective attention is an essential tool for the listener to orientate and focus on relevant target sounds in complex acoustic environments while blocking distracting sound incidences (for review, see Bronkhorst, 2015). Furthermore, target sound sources often change (especially in educational situations, e.g., after the teacher asks a question, attention must be switched to the answering children). Children must, therefore, successfully develop the ability to intentionally control auditory selective attention switches to cope with the highly complex acoustic (and noisy) situation in educational buildings and the corresponding successful development of communication skills and academic performance (cf. Loh et al., 2022a). Until now, existing paradigms to specifically examine intentional switch of auditory selective attention were extended to close-to-real-life sound representations for listeners in controlled study environments providing spatial sound perception (e.g., Koch et al., 2011; Lawo et al., 2014; Nolden et al., 2019). General background noise was not integrated into the investigations, and children have not been in the study populations yet, especially not young children. Furthermore, it is unclear at which age the intentional switch of auditory selective attention develops.

Noise does not only affect children's cognitive processes (cf., Jones et al., 2015; Klatte et al., 2010c), it can also lead to a bodily reaction, e.g., increased heart rate (Basner et al., 2014), which is also commonly interpreted as noise-induced stress. Therefore, it is reasonable to examine correlating factors of cognitive functions and bodily reactions by conducting combined assessments and analyses. Considering children and focusing on achieving a high level of ecological validity, unobtrusive measurements are desirable to avoid limitations in children's daily activity and environment. For this purpose, measuring electrocardiograms and corresponding stress parameters, e.g., heart rate variability (HRV), could be considered appropriate. HRV is mainly controlled by the autonomic nervous system responsible for activating and relaxing our body (Taelman et al., 2009). Specific HRV parameters, therefore, decrease when being exposed to stress and the body is activated (*Fight-or-Flight Response*, McCarty, 2016) and can be interpreted to derive stress levels.

Aims and Organization of the Dissertation

This dissertation aims to assess noise effects on children's auditory cognition and health by bringing close-to-real-life acoustic scenes into listening experiments on intentional auditory selective attention switching for children between three and ten years old in a controlled and reproducible manner. To achieve this systematically, the dissertation addresses two major objectives. Firstly, the general noise situation in pre- and primary schools is examined, adding children's perspectives:

1. Acoustic measurements were conducted to achieve an inventory of acoustics scenes from educational buildings in Germany, focusing on pre- and primary schools. It is intended to reflect a wide range of acoustic scenes that children between three and ten years are exposed to daily. The acoustic measurements were examined a) from a room acoustic view, focusing on the traditional perspective to access the suitability of rooms and environments for specific purposes, and b) from the perspective that the major noise disturbing the individual is primarily generated inside the room by the people present, considered as the in-situ case (cf. Section 2.3).
2. Focusing on children's perspective, HATSs with different anthropometric sizes were added to the measurement process besides the standard omnidirectional microphone. Measurements were examined to determine the differences in the physical signals introduced by the different anthropometric sizes. For this, room acoustic parameters and noise parameters were investigated across frequency bands and single-value parameters, such as the speech transmission index (STI), were studied in detail in comparison to the parameters of each frequency band to determine the impact of averaging across frequency bands when interpreting acoustic parameters from children's point-of-view (cf. Section 2.4).
3. Subjective noise ratings of adults and children were linked to measured parameters to obtain first insight into the suitability of objective measures to predict noise perception for children (cf. Section 2.5).

4. Concerning the assumption that noise is primarily generated by the people present, a first systematic concept to describe activities in educational buildings with respect to the acoustic properties was developed. For this, adult-children interaction and communication pathways were taken into account. First, categories for pre- and primary schools were defined and examined according to the noise parameters and whether HATS measurements had an impact across the activities (cf. Section 2.6).

Secondly, an appropriate paradigm is developed to examine the effects of noise on children's auditory cognition and health using insights on noise from the first part. For this, the following objectives were defined:

1. Development of a child-appropriate paradigm for children three to ten years old on the intentional switch of auditory selective attention including noisy conditions and close-to-real life sound representations (cf. Section 3.2).
2. Validation of the newly developed paradigm on intentional auditory selective attention switches with children between six and ten (cf. Section 3.3).
3. Investigating age effects on the intentional switching of auditory selective attention with children between three to six (cf. Section 3.4).
4. Examining health effects in children in terms of changes in heart rate variability under noise exposure in a listening experiment on auditory cognition (cf. Section 3.5).

This dissertation's content is organized according to the objectives described in the aforementioned sections.

2

Towards Child-Appropriate Noise Assessment Methods

This chapter will present the acoustic measurements conducted in pre- and primary schools. The main objective of this part of the dissertation was to achieve an inventory of acoustics scenes from German pre- and primary schools and to obtain more insights on the existing noise components from a traditional acoustic assessment perspective using room acoustic parameters and level-based parameters. Additionally, children's point-of-view was added by investigating differences in physical acoustic signals and corresponding signals when measuring using HATS with different anthropometric sizes and parameters, which were developed to predict sound perception more closely to natural hearing and perception, such as psychoacoustic parameters. Furthermore, these parameters were linked to subjective noise rating by adults and children, and the effect of activities on the acoustic properties of the sound environment was investigated.

Parts of this chapter with respect to inventory of the acoustic scenes in German pre- and primary schools was previously published in Loh et al. (2022b), and preliminary results on the activity-based acoustic setting were presented at the conferences *ICBEN 2021* (Loh & Fels, 2021) and *Inter-noise 2023* (Loh et al., 2023b). Respective sections are indicated respectively.

2.1 State of the Art

Parts of the following literature research have been published previously in Loh et al. (2022b), Loh and Fels (2021), and Loh et al. (2023b). For clarity and completeness, the state of the art has been summarized and limited to the context of the dissertation.

Room Acoustic and Noise Measures

The traditional way to assess the suitability of acoustics in rooms for educational purposes was driven from the room acoustic perspective. Previous research and standardization commitees (e.g., ANSI Inc., 2002; Astolfi et al., 2019a; Bradley et al., 1999; DIN e.V., 2016) introduced ranges for room acoustic parameters and regulations that classified and optimized acoustics for good development, education, and learning. However, room acoustic measurements in classrooms revealed a wide range of results, some within and some beyond the optimal values for 'good' acoustics in classrooms as Loh et al. (2022b) summarized in their work:

> In terms of room acoustics, Sala and Rantala (2016) reported reverberation times, STI values (values $\geq$ 0.85 considered adequate for a wide range of hearing and learning conditions for children), and mean background noise level in unoccupied classrooms as 34.5 dB(27-44 dB). Astolfi et al. (2019b) used a consistent measurement setup across classrooms in Italy, which were classified either as rooms with 'good' or 'bad' acoustics according to the occupied rooms' [reverberation time] ($T_{20,\mathrm{occ}}$), as shown in Table 2.1. They also reported clarity index (C_{50} in dB, ratio between the energy arriving in the first 50 ms and the remaining energy) and the ratio of useful to detrimental energy values (U_{50} in dB) to express speech intelligibility, which were highly correlated with $T_{20,\mathrm{occ}}$. For the classrooms with 'good' acoustics, reported values for C_{50} and U_{50} were mostly within the range of optimum values, with $C_{50} \geq$ 3 dB considered good, and $U_{50} \geq$ 1 dB considered optimal; classrooms with 'bad' acoustics had corresponding values outside this optimal range. Persson Waye and Karlberg (2021) reported results from a study in Sweden in unoccupied but furnished rooms, before and after an acoustic intervention. Wang and Brill (2021) reported estimated noise and speech levels from measurements in USA classrooms, along with [reverberation time] and C_{50} values in unoccupied rooms.

Noise measurements further accompanied the room acoustic measurements to obtain insights into the interlinkage between sound/noise development and the room's acoustic design. However, conditions, methods, and parameters chosen to assess noise levels in educational buildings varied strongly across studies as summarized in the work by Loh et al. (2022b):

> As seen in Table 2.1, which refers to measurements in classrooms in mainly primary schools and daycare centers across various countries, the sound pressure levels (SPLs, in decibel) in such educational institutions are considerable. There are also considerable variations between studies due to factors such as: the number of children present (summarized in Sala & Rantala, 2016); the age groups, with daycare centers generally reporting higher levels than primary school classrooms (Picard & Bradley, 2001); activities involved; room

Table 2.1: Summary of noise and room acoustic parameters in previous studies (adapted from Loh et al., 2022b).

Study	**Summary of conditions**	$L_{\mathbf{A,eq}}$ **[dB]**	$T_{\mathbf{20}}$ **[s]**	STI, $C_{\mathbf{50}}$ **[dB]**, $U_{\mathbf{50}}$ **[dB]**
Sala and Rantala, 2016 (others' studies)	Elementary school (approx. 5-12 y/o)	42 - 100	0.7 (0.2 - 1.27)	$STI = 0.68$ (0.44 - 0.81)
	Elementary school (6 y/o)	42 - 100	0.7 (0.2 - 1.27)	$STI = 0.77$ (0.59 - 0.92)
	Preschool (6 y/o)	60 - 85	0.55 (0.41 - 0.85)	$STI = 0.74$ (0.65 - 0.81)
Sala and Rantala, 2016 (own study)	Comprehensiv schools (7-12 y/o)	69 (57 - 89)	0.55 (0.41 - 0.85)	$STI = 0.75$ (0.65 - 0.81)
Astolfi et al., 2019b ('good' acoustics)	Primary school (6-7 y/o)	60 - 75	[0.5 - 0.8]	$2.9 \leq C_{50} \leq 7.6$ $-0.8 \leq U_{50} \leq 4.0$
Astolfi et al., 2019b ('bad' acoustics)	Primary school (6-7 y/o)	62 - 72	$0.5 < \text{RT} < 0.8$	$-2.2 \leq C_{50} \leq 2.7$ $-2.6 \leq U_{50} \leq 0.9$
Persson Waye and Karlberg, 2021 (before)	Preschools (1-5 y/o; dosimeter)	85	0.3 - 0.5	$8 \leq C_{50} \leq 10$
Persson Waye and Karlberg, 2021 (after)	Preschools (1-5 y/o; dosimeter)	83	0.2 - 0.4	$11 \leq C_{50} \leq 13$
Wang and Brill, 2021	K-12 schools (5-16 y/o; SLM)	Speech: 65 (SD: 2.5) Noise: 47 (SD: 3.5)	[0.2 - 1.1]	$-2.0 \leq C_{50} \leq 14.4$
Södersten et al., 2002	Daycare centers (1-6 y/o; binaural)	76 (73 - 78)	-	-
McAllister et al., 2009	Daycare centers (5 y/o; binaural)	83 (82 - 84)	-	-

Note. Reported values are mean and the range (in brackets) unless indicated otherwise.

> acoustics due to excessively low or high reverberation times [...] (summarized in Sala & Rantala, 2016); measurement methods including duration (summarized in Sala & Rantala, 2016; Wang & Brill, 2021), transducer type and locations, e.g., omnidirectional vs. binaural recordings vs. dosimeters (last two rows in Table 2.1), with microphones in front of ears (McAllister et al., 2009; Södersten et al., 2002); and pedagogical aspects. Representing a wide range of such factors, Sala and Rantala (2016) summarized SPLs from studies conducted in Finland, Germany, Sweden, UK, and USA over several years, reporting a range of SPLs from $L_{\mathrm{A,eq}} = 42$ to 100 dB in schools and $L_{\mathrm{A,eq}} =$ 60 to 85 dB in preschools measured for periods ranging from 2 minutes up to 5 working days. Their own investigations included $L_{\mathrm{A,eq}}$ levels as well several percentile levels including $L_{\mathrm{A,90}}$ representing the background noise in occupied classrooms, $L_{\mathrm{A,10}}$ representing the higher levels and $L_{\mathrm{A,50}}$ representing the median level.

Binaural Recordings considering Anthropometric Differences

Especially in terms of methods, noise measurements were rarely conducted using HATS (cf. Table 2.1), while it has become more common and standardized within room acoustic assessment (cf. ISO, 2008). Nevertheless, there were first approaches to integrate a children's perspective in noise assessments by conducting close-to-children measurements as summarized by Loh et al. (2022b):

> [Previous studies] have typically used omnidirectional microphones at fixed locations and/or single-channel noise dosimeters to measure the sound environment in educational institutions. Binaural recordings of children and teachers moving freely within classrooms have been performed in at least two studies (last two rows in Table 2.1) where microphones were placed in front of both ears of teachers and children in preschool classrooms in Sweden; values reported are power averages of left and right ear values. For typical daily activities in classrooms, these values represent a closer representation of hearing levels for both teachers and children. The almost 6 dB difference in the mean $L_{\mathrm{A,eq}}$ values in these two studies using similar measurement methods was partly attributed by the authors to the differences in heights and distances between the teachers and children. The values reported in Södersten et al. (2002) and McAllister et al. (2009) do not include contributions due to self-speech of the participants wearing binaural microphones, which, besides other measurement factors, may partly account for slightly lower values compared to the dosimeter values reported in Persson Waye and Karlberg (2021) which presumably include contributions due to the participant's own

speech. In the latter, significant differences were found between children and personnel amounting to 6 to 8 dB.

[...] While measurements using omnidirectional transducers have several advantages, binaural measurements are a closer representation of hearing conditions. Binaural transducers placed near human ears, as in Södersten et al. (2002) and McAllister et al. (2009), perhaps represent one possibility, with its own set of logistical issues. HATS have limitations in terms of fixed location, and generic head-related transfer function (HRTF); the latter characterizes the frequency-dependent amplifications in the signals when measured at the ear canal entrance (Møller et al., 1995). However, the advantages of HATS include a potentially more robust and repeatable setup compared to putting transducers on humans, with a major limitation being the use of additional equipment that may not be as readily available as individual microphones. Another overhead includes additional binaural analyses due to processing two channels instead of one in general and the potential use of computational expensive binaural models such as those for binaural loudness (Moore & Glasberg, 2007). Yet, to avoid intrusive methods involving humans (especially children), HATS represent a rather convenient middle ground for noise measurements in classrooms, which can be used to augment information provided by standard methods using omnidirectional microphones.

[...] In terms of HATS sizes, children have smaller ears, head, and shoulder sizes than adults, and arguably a HATS representing adult morphology may not represent those of children. Hence, differences in anthropometric sizes between adults and children need to be considered to represent children's perspectives more appropriately. Indeed, different adult HATS can also have different HRTFs, but for the sake of brevity, this is not explored further here, and instead the focus is on comparisons between a selected adult and child HATS. Fels et al. (2004) reported more amplification in the higher frequency bands (starting from 4–5 kHz) for children vs. adults HRTFs. Differences were also observed for different directions in the horizontal and median planes (Fels & Vorländer, 2009). More gain in higher frequencies in children's HRTFs might explain the higher sensitivity to high-frequency sounds reported for children (Persson Waye & Karlberg, 2021). With this in mind, it may be expected that differences between the transducers might also be observable in certain room acoustic and noise parameters, such as [reverberation time] and SPL, when analyzed on a band-by-band basis, especially in the higher frequency bands.

Noise Perception Prediction using Psychoacoustic Parameters

As differences in the physical signal introduced by the different anthropometric sizes of HATS are primarily expected in specific frequency ranges (higher frequency bands), it is standing to reason to investigate differences in the frequency bands. A first approach could be the investigation of level-based parameters, including percentiles, as it has been used in previous research (cf. work listed in Table 2.1) and extend it with respective values for each frequency band of interest. The main question is how single-value approaches by averaging across frequency ranges will mask the differences introduced by the anthropometric differences in the physical signals obtained from adult and child HATS. Regarding single-values predicting noise perception that reflects differences in the spectral, temporal, and spatial aspects of human sound perception, psychoacoustic models and associated parameters should be considered. Though these were developed mainly based on studies focusing on adult perception, they can be considered to be a better representation of human sound perception than level-based parameters as presented in the work by Loh et al. (2022b):

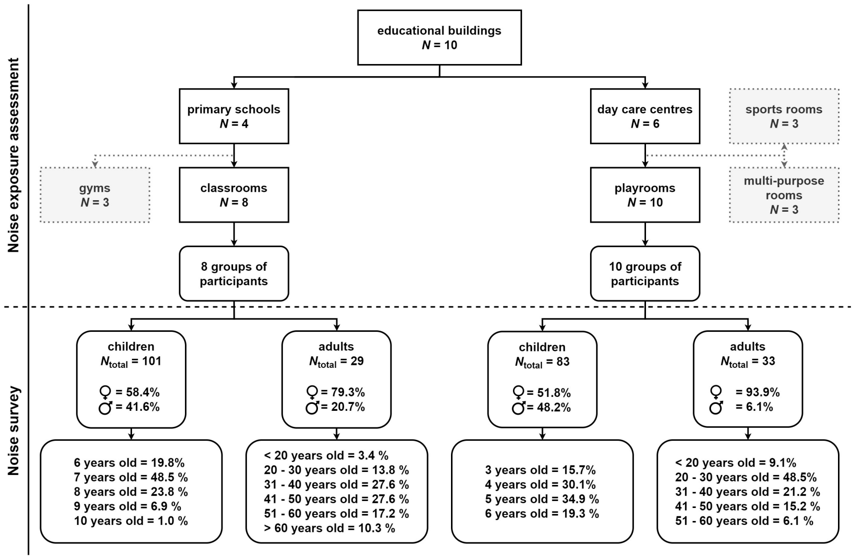
> Psychoacoustic loudness is perhaps the most common (for both stationary and time-varying sounds; ISO 532-1ISO (2017)). However, other psychoacoustic parameters such as sharpness, roughness, fluctuation strength, etc., have been useful in investigations of several subjective attributes of various sound environments for adults. However, the use of psychoacoustic parameters in classroom studies has been very limited, and it is unclear whether there is any benefit in considering psychoacoustic models based on adults' perception to characterize children's perception; psychoacoustic models specifically for children, and adults' models adapted for children are possible too, but not the focus here. Yet, the scope of existing psychoacoustic parameters has the potential to complement and even go beyond investigations that are possible with SPL-based parameters. This includes, but is not limited to, exploring the higher sensitivity of children to high-frequency sounds compared to adults. This is possible by comparing, for instance, SPL of lower vs. higher-frequency octave bands with psychoacoustic sharpness (S) and whether there is a benefit in using one approach over another. Additionally, one may expect higher sharpness values based on measurement with child HATS compared to adult HATS due to higher amplification in higher frequencies for children's HRTFs compared to adults' HRTFs (Fels et al., 2004). Similarly, to explore the effect of fluctuations in the sound environment on human perception, it is possible to compare the performance of SPL-based parameters that quantify the level fluctuations above the ambient SPL (e.g., $L_{\mathrm{A},10} - L_{\mathrm{A},90}$, etc.) and psychoacoustic parameters fluctuation strength (FS) and roughness (R). These

psychoacoustic parameters characterize human perception to slower (FS) and faster (R) amplitude fluctuations and have been shown to be related to annoyance due to air-conditioning, and auditory distraction due to many sounds including speech in office simulations, respectively.

2.2 Involved Educational Institutions

Acoustic measurements and noise surveys were conducted in ten educational institutions in Aachen (Germany), including four primary schools and six daycare centers. This thesis focuses on evaluating classrooms and playrooms, neglecting rooms dedicated to sporting or rollicking activities. In total, $N = 8$ classrooms (CRs) and $N = 10$ playrooms (PRs) were included. An overview of the finally selected rooms and survey participants is given in Figure 2.1.

Figure 2.1: Overview of cooperating educational institutions.



Note. Boxes in grey indicate rooms neglected in this dissertation, though they were acoustically assessed and included in the ethical approval.

During the noise measurements and noise surveys, on average, 22 children (female: 50.0%) and one adult (mostly female) were present in the classrooms, while on average, 15 children (female: 53.5%) and two adults (female: 91.3%) were present in the playroomsren in the primary schools were between six to ten years old, and more than 50% of adults were in the age group between 31 to 50 years. Children in the daycare centers were between three and six years old, and more than 50% of adults were between 21 and 40 years old. The procedure was approved by the Medical Ethics Committee at the RWTH Aachen University, Germany (EK 321/16 and EK 218/18). Furthermore, signed informed consent was collected from all involved educational institutions, teachers, and educators. In the case of children, informed consent was signed by their parents for the participation of their children(for more details, see Loh et al., 2022b).

2.3 Acoustic Measurements in Pre- and Primary schools

A complete acoustic assessment of the measured classrooms and playrooms was published in Loh et al. (2022b). The following section presents the essential parts and insights from this published work explaining the complexity of acoustic measurements with respect to considerations towards child-appropriate measurement setups and procedures. This thesis is tailored to evaluate the differences between measurement and evaluation methods using omni-directional microphones compared to HATSs with different anthropometric sizes. Therefore, and for the sake of clarity, results of the previously published work (Loh et al., 2022b) are not discussed in detail within this thesis.

Descriptives of the individual rooms and information on connected rooms, acoustic treatments, room dimensions, A-weighted ambient background noise levels outside occupied hours, and the average number of people present during the noise measurements within daily educational activities are included in Table 2.2. In daycare centers, playroomss were mainly connected to smaller rooms (e.g., an extra eating or sleeping room). The doors of these rooms were seldom closed, enabling continuous supervision by the educators. Therefore, the room volumes of these smaller rooms were considered part of the overall room volumes while evaluating the room acoustic measurements. These smaller connected rooms are indicated by the additional volumes in Table 2.2.
Furnishings in all rooms corresponded to the purpose of the educational institutions. They remained unchanged not only for the room acoustic assessment but also during the noise measurements within daily activities (for more details, see Loh et al., 2022b).

2.3.1 Room acoustic measurement procedure

The rooms were assessed according to their room acoustic parameters. The work by Loh et al. (2022b) described the measurement procedure as follows:

Table 2.2: Overview of selected classrooms (classroom (CR)s) and playrooms (playroom (PR)s) (Loh et al., 2022b).

Room	**Connected rooms**	**Acoustic treatment**	**Area** A (m^2)	**Height** h (m)	**Add. volumes** A (m^2) × h (m)	**Total volume** (m^3)	**Background noise** (dBA)	**Av. number of people present**			
								Children		**Adults**	
								f	m	f	m
CR01	No	Yes	60.9	3.2		194.9	26.4	7	12	2	-
CR02	No	Yes	71.1	3.0		213.3	25.7	0	17	1	-
CR03	No	No	69.3	3.2		224.6	21.9	13	8	2	-
CR04	No	No	69.3	3.2		224.6	28.0	13	15	1	-
CR05	No	No	66.8	3.6		240.5	32.0	9	13	2	-
CR06	Sloped ceiling	No	56.7	3.4	56.7 × 1.9/2	246.5	25.2	14	9	1	-
CR07	No	Yes	82.7	3.0		248.2	25.6	24	0	1	-
CR08	Sloped ceiling	Yes	61.3	3.0	61.3 × 2.3/2	254.4	25.1	7	13	1	-
PR09	No	No	38.4	3.1		117.0	27.3	7	9	2	-
PR10	Yes	No	32.2	2.8	(2.8 + 16.2 + 8.3) × 2.8	164.8	22.4	10	6	2	1
PR11	Yes	No	54.9	2.8	7.7 × 2.8	172.8	30.0	9	10	3	-
PR12	Yes	No	58.7	3.2	5.3 × 3.2 + 3.1 × 2.0	209.7	25.7	8	6	2	-
PR13	Yes	No	45.5	3.7	12.0 × 3.7	210.5	34.0	4	2	2	-
PR14	Yes	No	49.5	2.7	23.8 × 2.7	194.3	22.2	7	10	3	-
PR15	No	No	44.2	3.0		132.6	29.7	7	7	2	-
PR16	Yes	No	44.1	2.7	(16.2 + 9.4) × 2.7	187.5	23.2	7	5	2	-
PR17	Yes	No	72.4	2.8	19.7 × 2.8	254.2	25.3	10	7	1	1
PR18	Yes	No	49.6	4.0	20.4 × 4.0	280.1	21.9	8	5	2	-

Room acoustic measurements were conducted in unoccupied furnished rooms according to ISO 3382-2 (ISO, 2008) at precision level with two source and six receiver positions. As the sound source, the Institute of Technical Acoustics (ITA)'s 3-way omnidirectional dodecahedron loudspeaker was used. Simultaneous measurements were executed with the ITA adult HATS (Schmitz, 1995) equipped with Schoeps CCM2H microphones, ITA child HATS (Fels et al., 2004) equipped with Sennheiser KE4 microphones, and a 1/2" diffuse field omnidirectional microphone (B&K Type 4134) as a reference. Positions were chosen according to ISO 3382-2 (ISO, 2008) with as little overlap as possible without removing the furnishings inside the rooms. All receivers were positioned to represent standing situations since the chosen positions were quite far from the tables and chairs.
It further represents reasonably the behavior of teachers and educators in the room, who are standing most of the time. The reference microphone was positioned at the height of 1.2 m, the ear axis of the adult HATS was adjusted to 1.5 m, and of the child HATS to 1.0 m height. The measurement signal was an exponential sweep with a duration of 5.944 s, and it was repeated five times per position.

Furthermore, for all six receiver positions, the ambient equivalent A-weighted background noise level over 30 s ($BNL_{\mathrm{A,eq,30s}}$) was measured according to ISO 9568 (ISO, 1993) using the reference microphone (1/2" diffuse field microphone B&K Type 4134).

Room acoustic parameters were obtained according to Loh et al. (2022b):

[The] parameters T_{20}, T_{30}, EDT (early decay time; to potentially represent subjective *reverberance*), C_{50} (clarity index (Bradley et al., 1999)), D_{50} (definition), and T_{S} (center time) were computed according to ISO 3382-2 (ISO, 2008) and ISO 3382-1 (ISO, 2009) for a frequency range of 125 Hz-16 kHz octave bands center frequencies. The A-weighted background noise level over 30 s was evaluated according to ISO 9568 (ISO, 1993) for a frequency range of 31.5 Hz - 16 kHz. The STI was calculated using the indirect method following IEC 60268-16 (IEC, 2012), which computed the STI using the measured impulse response neglecting effects from masking and background noise. Hereby, MATLAB and the ITA toolbox (Berzborn et al., 2017) were used.

Since some of the rooms measured were connected with smaller volumes ([Table 2.2]), the degree of non-linearity in the reverberant energy decay of the measured impulse responses was examined using the method in Annex B of ISO 3382-2 (ISO, 2008). All measurement positions where the degree of curvature of the decay (comparing T_{20} and T_{30}) for the reference microphone exceeded the 10% threshold, signifying substantial deviation from linearity, were removed from further room acoustic analyses (ISO, 2008). [...]

To approximate binaural versions of the standard room acoustic parameters, two approaches were considered. Firstly, the computed parameters from the left and right ear were averaged ($\frac{X_{\text{Left}}+X_{\text{Right}}}{2}$). This method is indicated in the following with $A\text{-}HATS_{av}$/ $C\text{-}HATS_{av}$, representing the values from the adult and child HATSs values, respectively. Secondly, the value from the prominent ear was chosen. In this work, it is assumed to be the higher value out of the left and right ear values. The idea here is that the prominent ear represents the conservative approximation of a binaural model, except for STI in which the higher value or the *better-ear* STI (signifying better signal-to-noise ratio) was used as it has been shown to perform well in relation to a binaural STI model (van Wijngaarden & Drullman, 2008). This method is referred to with $A\text{-}HATS_{prom}$/ $C\text{-}HATS_{prom}$ for the adult and child HATSs, respectively.

2.3.2 In-situ noise measurement procedure

Long-term noise measurements were carried out to assess the daily acoustic situation in the classroom and playrooms. The work by Loh et al. (2022b) described the measurement procedure:

The *in-situ* noise measurements were conducted during the daily activities of children in CRs and PRs. The same equipment as stated for the room acoustic measurements in unoccupied rooms was used to execute the *in-situ* measurements. All three receivers were positioned together in the center (less than 30 cm to each other, cf. Figure 2.2) of the main room of activity so that people were able to move around them. While this potentially introduces acoustic shadowing and interference issues for the transducers, the location of the transducers was due to logistical concerns including ensuring that the measurement equipment did not adversely interfere with the usual behavior of the adults and children (e.g., by attracting too much children's attention). The positioning of measurement equipment was discussed forehand with the teachers, and the study and the equipment were explained to the children [at least one] day before the measurements started in each educational institution.

In-situ measurements were conducted over two days per [classrooms] and [playrooms] during normal daily activities. On the first day, all HATSs were positioned to represent a standing position (ear axis of the adult HATS at 1.5 m, ear axis of the child HATS at 1.0 m, and the omnidirectional microphone at 1.2 m. On the second day, they were positioned to represent a sitting position (ear axis of the adult HATS at 1.2 m, ear axis of the child HATS at 0.8 m, and the omnidirectional microphone at 1.2 m) or to represent a playing height (ear axis of the child HATS at 0.5 m), respectively, according to the dominant scenario of each educational institution. Only periods with children present in the room were considered, where the sound pressure level exceeded 35 dB[Z] (Z-weighted). In other words, a cutoff sound pressure level of 35 dB[Z] was used to distinguish between children's presence and absence in the rooms. This cutoff was based on inspecting several samples within the recordings when children were not present. This resulted in up to 6 hours of recordings on average per room, which were used for further analyses.

Figure 2.2: Example of the centered positioning of the measurement transducers during an *in-situ* measurement in a classroom (Loh et al., 2022b).

In the work by Loh et al. (2022b), three types of parameters were considered in the analyses: 1. parameters based on A-weighted SPL, 2. level-based fluctuation parameters, and 3. a set of psychoacoustic parameters.

For brevity, this work focuses on only two types of noise parameters selected based on the previous work by Loh et al. (2022b):

1. Parameters based on A-weighted SPL (L_{A}) and the unweighted SPL (L_{Z}), including the equivalent SPL ($L_{\mathrm{A,eq}}$ and $L_{\mathrm{Z,eq}}$), SPL using a fast time weighting ($L_{\mathrm{AF,mean}}$ and $L_{\mathrm{ZF,mean}}$), and percentiles ($L_{\mathrm{A,10}}$, $L_{\mathrm{A,90}}$, $L_{\mathrm{Z,10}}$, and $L_{\mathrm{Z,90}}$).

2. Psychoacoustic parameters (cf. Loh et al., 2022b) were computed using the ArtemiS SUITE 14.3 by HEAD Acoustics (Herzogenrath, Germany): *loudness N* for time-varying sounds following ISO 532-1 (ISO, 2017); *sharpness S* according to DIN 45692 (DIN e.V., 2009), *roughness R* and *fluctuation strength FS* according to the Hearing Model by (Sottek, 1993).

To understand the relation of high- and low-frequencies (below and above 1 kHz) in the unweighted equivalent SPL ($L_{\mathrm{Z,eq}}$), especially in comparison to the psychoacoustic parameter sharpness, a low-high-frequency ratio of $L_{\mathrm{Z,eq}}$ was calculated:

$$L_{\mathrm{Z,eq}}(L.v.H) = \frac{mean(L_{\mathrm{Z,eq}}(31.5\,\mathrm{Hz}{-}1\,\mathrm{kHz}))}{mean(L_{\mathrm{Z,eq}}(2\,\mathrm{kHz}{-}16\,\mathrm{kHz}))}$$

The binaural parameters were calculated using the average and prominent ear methods to account for binaural effects, as described in Section 2.3.1. In terms of (A-weighted) levels, "level summation was computed using the left and right ear SPL values instead of the averaging method" (Loh et al., 2022b), to provide an acoustic average; thus, it is further also indicated as A-HATS$_{\mathrm{av}}$/ C-HATS$_{\mathrm{av}}$.

2.3.3 General measurement results

Room acoustics (unoccupied, furnished rooms)

The rooms were evaluated according to the curvature criteria, and positions were excluded if they exceeded the curvature criteria (for complete and detailed results, see Loh et al., 2022b):

> In four [rooms] (CR08, PR10, PR12, and PR16[...]), ≥ 4 positions had to be discarded. In further analyses PR16 was excluded due to especially low number of measurement positions that met the curvature criteria.
> The room acoustics of two classrooms (CR03 and CR04[...]) were noticeably different than the other six (averaged $T_{30} = 0.55$ s and averaged $STI = 0.74$ in these classrooms). Possible explanation could be combined effect of the large room volumes with flat ceilings compared to other larger volumes like CR06, CR07, and the absence of acoustic treatment in these rooms; although CR05 still has comparable room volume and room acoustics to CR06 and CR07 but has no acoustic treatment and has a flat ceiling similar to CR03 and CR04.

Results from the measurements using the adult and child HATS, including both evaluation methods (averaging and prominent-ear), revealed the following observations (Loh et al., 2022b):

> For the mid-frequency octave bands (500 Hz, 1 kHz and 2 kHz center frequencies), values were very similar for all the room acoustics parameters across the measurement methods. Beyond 2 kHz, some deviations can be seen, which can broadly be attributed to the anthropomorphic features (i.e., HRTFs) of the binaural transducers becoming important for smaller wavelengths. In this regard, the values for the adult HATS varied more in comparison to the other measurement methods. Further, for the parameters C_{50}, D_{50} and T_{S}, which are all ratios of early sound energy to late/reverberant energy, the values for the child and adult HATS are similar but deviated from the omnidirectional microphone for the 4 kHz band, and the parameter values for the adult HATS exhibited a distinct deviation in comparison to the corresponding values for the child HATS and the omnidirectional microphone, which have similar values, for the 8 kHz octave band. These deviations starting from the 4 kHz octave band can also be observed in the HRTF magnitude response of child HATS in comparison to adult HATS as presented in Fels et al. (2004). However, the statistical analyses revealed no significant difference between the five measurement methods (Ref vs. A-HATS$_{\mathrm{av}}$ vs. C-HATS$_{\mathrm{av}}$ vs. A-HATS$_{\mathrm{prom}}$ vs. C-HATS$_{\mathrm{prom}}$...).

In-situ acoustics (occupied rooms)

Firstly, the classrooms and playrooms were compared to each other (from Loh et al., 2022b):

> Differences between classrooms in primary schools and playrooms in daycare centers are mainly observable in loudness (average N_{mean} = 11.9 sone (9.3 – 15.3 sone) vs. 10.8 sone (7.5 – 13.9 sone) [...]) and in sharpness (average S_{mean} = 1.5 acum (1.38 – 1.63 acum) vs. 1.46 acum (1.40 – 1.51 acum), and in the percentiles (average N_5= 25.1 vs. 22.9 sone). An increase in loudness is understandably related to increasing $L_{\mathrm{A,eq}}$, with an R^2 of 0.85. Loudness N_{mean} (as y) is predicted from $L_{\mathrm{A,eq}}$ (as x) with the equation $y = -61.90x + 1.12$. Almost no relationship between increasing sharpness and increasing high-frequency content in the *in-situ* sound was found (R^2 of 0.00 for Ref, A-HATS$_{\mathrm{av}}$, and A-HATS$_{\mathrm{prom}}$; R^2 of 0.02 for C-HATS$_{\mathrm{av}}$; R^2 of 0.04 for C-HATS$_{\mathrm{prom}}$).

Secondly, the variation in noise parameters across the measurement methods was examined (from Loh et al., 2022b):

> In terms of the low-high-frequency ratio of $L_{\mathrm{A,eq}}$, the post-hoc analyses revealed significant differences between the Ref and child HATS values while

the Ref and adult HATS values were not significant for both averaging and prominent ear method. Differences between HATSs in $L_{\mathrm{A,eq}}$(L.v.H) for both evaluation methods were significant.
For the psychoacoustic parameters, the post-hoc analyses showed no significant differences for loudness (N_{mean} and N_5) between both HATSs using the averaging method and the omnidirectional microphone, while differences were significant in terms of using the prominent-ear method. However, no differences were found between the adult and child HATS. Considering N_{90}, differences between omnidirectional microphone and adult HATS were significant as well between the two HATSs. For sharpness (S_{mean}), all results from the HATS were significantly different to the omnidirectional microphone and within each other (A-HATS$_{\mathrm{av}}$ vs. C-HATS$_{\mathrm{av}}$ and A-HATS$_{\mathrm{prom}}$ vs. C-HATS$_{\mathrm{prom}}$). However, for S_{90}, no differences between the HATSs with both evaluation methods were observed. For roughness (R_{mean}, R_5 and R_{90}), differences were observed between the omnidirectional microphone and the HATS using averaging method, while differences between the HATS and both evaluation methods were not significant. For fluctuation strength, all measurement methods were significantly different from each other for FS_{90}, while FS_{mean} was only significantly different for omnidirectional microphone vs. C-HATS$_{\mathrm{av}}$; and for FS_5, the only significant difference was between the omnidirectional microphone and the HATSs using the averaging evaluation method.

2.3.4 Discussions on child-appropriate acoustic measurement

The evaluation of the room acoustic measurement revealed the following results that are tailored towards the focus of this thesis (from Loh et al., 2022b):

> Results from this study [...] add to the insights from previous studies across various countries (Astolfi et al., 2019b; Klatte et al., 2013; Persson Waye & Karlberg, 2021; Sala & Rantala, 2016; Wang & Brill, 2021). [...] Overall, playrooms had better room acoustic properties than classrooms, with lower T_{30}, EDT and T_{S}, and higher C_{50}, D_{50} and STI values. Rooms with acoustic treatment understandably yielded better room acoustic values. Playrooms with coupled volumes had better room acoustic properties than single-volume playrooms. [...]
> If the room acoustic classification from the work by Astolfi et al. (2019b) is used, which is based on criteria for both T_{20} (T_{30} values used for the current sample) and C_{50} values, although for occupied rooms, all classrooms except CR03 and CR04 would be classified as having 'good acoustics', while all playrooms except PR10, 11, 14 and 18 would be classified as having 'good

acoustics'. It is likely, however, that some of these values may meet the criteria for 'good acoustics' as per Astolfi et al. (2019b) if the measurements were conducted during occupation, as was shown for university classrooms, especially with mostly reflective surfaces (Choi, 2016). STI (unoccupied) is recommended to be at least ≥ 0.80 and ≥ 0.85 for educational institutions for children without and with hearing, cognition, and/or behavioral issues, respectively (Finnish Standards Association, 2004). For the current sample, except for PR10 and PR11, none of the classrooms or playrooms meet the recommended STI values for even children without hearing and/or learning difficulties. STI, overall, had a strong linear relationship with reverberation time (and other room acoustic parameters), similar to previous studies (Leccese et al., 2018; although not in Sala & Rantala, 2016), which can be used to estimate global STI values based on the simpler way to calculate [reverberation time] values. While this implies that [reverberation time] could perhaps be used as a primary indicator to represent the room acoustics in classrooms, more studies with larger sample sizes are needed to determine the strength of the relationship between [reverberation time] and STI, and their relationship with subjective impressions. U_{50} values, which have been used in several studies of classroom acoustics, were not calculated for the current sample due to the measurement issues. The relevance of the good/bad acoustics in the current sample can further be explored based on subjective impressions of children in these rooms, as was done in Astolfi et al. (2019b), which is proposed for a future study. Nevertheless, based on room acoustics measurements alone, with higher reverberation times in some rooms and lower values for intelligibility than recommended, most classrooms measured in the current study are likely to be not optimal for learning purposes (however, these rooms in occupied conditions might be better suited though it is not examined in this study) and may affect cognitive and behavioral development of children (Klatte et al., 2010b, 2010a, 2010c), and especially for children with hearing loss and/or learning difficulties (Crandell & Smaldino, 2000).

In terms of the evaluation of the level-based parameters, the following insights were derived (from Loh et al., 2022b):

The results [...] add to the previous *in-situ* measurements in primary schools and daycare centers and introduce some level-based sound fluctuation parameters previously used in other fields with multi-talker speech environments like open-plan offices (Yadav et al., 2021). For omnidirectional microphones, the mean $L_{\mathrm{A,eq}}$ values in classrooms and playrooms were almost the same ($\sim 66\,\mathrm{dB}$), with a relatively wider range of values in the latter. These values are within the range of values reported in previous studies, but the range of values in the current sample generally has lower upper limits, i.e., the classrooms and playrooms with higher $L_{\mathrm{A,eq}}$ values had lower $L_{\mathrm{A,eq}}$ values compared to previous studies. This includes $L_{\mathrm{A,eq}}$ values reported from omnidirectional measurements in classrooms of Italy (Astolfi et al., 2019b, 2019a); Finland (Sala & Rantala, 2016) and USA (for the speech levels; Wang & Brill, 2021), where the measurement devices were placed at fixed locations.

Omnidirectional $L_{\mathrm{A,eq}}$ values in the current study [(Loh et al., 2022b)] were around 20 dB lower than values reported in studies wherein children wore dosimeters (Persson Waye & Karlberg, 2021), and in McAllister et al. (2009) where children in daycare centers had a microphone placed near each ear. Compared to McAllister et al. (2009), where children were free to move around, the HATSs in the current study had fixed locations and with microphones at the entrance of the ear canal instead. This, combined with the overall quieter classrooms in the current sample, may partly explain the lower mean binaural $L_{\mathrm{A,eq}}$ values in the current sample of playrooms for the adult and child HATS of around 13 dB, and 14.5 dB, respectively, for the prominent ear values. The mean $L_{\mathrm{A,eq}}$ values calculated using the adult HATS were higher than those for the child HATS, which is opposite to what was reported in McAllister et al. (2009), where they compared values of similar measurement methods using binaural measurements for adults (Södersten et al., 2002) and children. McAllister et al. (2009) had partly attributed their results to children being the primary noise sources, which is also relevant for the current sample. Hence, the counterintuitive finding of higher $L_{\mathrm{A,eq}}$ for the adult compared to child HATS in the current study, which is most likely due to the particular transducer placement, is suggested as a question for future research. At the very least, this comparison highlights the issues in the selection of transducers for child-appropriate *in-situ* studies where the location of the transducers is fixed.

Moreover, $L_{\mathrm{A,eq}}$ values calculated using omnidirectional microphone were significantly different from HATS, and the adult HATS was at least significantly different from child HATS for the prominent ear condition [...]. Based on the octave-band spectra [...], differences between the adult and child HATS (SPLs calculated using level summation for the left and right ear values) and the omnidirectional microphone are largely linked to the 6 dB introduced by the level summation till around 1 kHz, followed by a more complicated trend till 16 kHz. There can be many contributing factors here, including the peak (around 4 kHz) and notches (around 8 and 10 kHz) in the magnitude response of the adult HATS HRTFs. In terms of the prominent ear values for

> the adult and child HATS, no differences to the omnidirectional microphone are observed up to the 2 kHz. Differences around 4 kHz and beyond 8 kHzz are again observable as in the room acoustic parameters, which is in line with the work by Fels et al. (2004) and Fels and Vorländer (2009) explaining these effects with the anthropometric differences between adults and children; however, there are no noticeable differences between the different room conditions (coupled rooms, acoustic treatment and room types). Since SPL above 8 kHz diverged between the adult and child HATS, it can be assumed that the fine structure of the ear played a role in the evaluation and that the $L_{\mathrm{A,eq}}$ could be sensitive to differences introduced by the anthropometric sizes of the ear. Altogether, the spectral variation in the SPL between the different transducers points towards some benefit in using a child HATS over an adult HATS and/or an omnidirectional microphone for $L_{\mathrm{A,eq}}$ values.
>
> For the percentile levels, the $L_{\mathrm{A,10}}$ and $L_{\mathrm{A,90}}$ values in the current study were similar to the ones reported in (Sala & Rantala, 2016), and were within the range of values of previous studies summarized in Sala and Rantala (2016). These percentile levels were, however, not significantly different between the adult and child HATS.

The results regarding the variation in the psychoacoustic parameters revealed the following insights (from Loh et al., 2022b):

> The use of psychoacoustic parameters is not common in classroom acoustics literature. [...] Hence, the discussion here will be limited to a preliminary comparison between the level-based and psychoacoustic parameters.
> Fluctuation Strength (FS) and roughness characterize amplitude modulations up to 20 Hz, and between 15 - 300 Hz, respectively. Roughness has been shown to be related to noise annoyance due to faster sound fluctuations, and FS has been shown to be related to auditory distraction (Schlittmeier et al., 2012). [...]
>
> Mean sharpness values had significant differences between all the transducers. In terms of HATSs, this is consistent with previous findings (Fels et al., 2004) and the expectations of the current study, with higher values of high-frequency content measured with child HATS compared to adult HATS in accordance with anthropometric differences [...]. This observation is supported by the findings regarding the low-high-frequency ratio of $L_{\mathrm{A,eq}}$, though significant differences were only found for the child HATS and between the HATSs. However, these results need validation in terms of children's perception in future studies, where it would be possible to comment on whether the use of HATSs, which introduce measurement and analysis overheads, is sufficiently

> justified over the more traditional use of omnidirectional transducers, which further allow relatively convenient measurement setups.
>
> Overall, the current results show very limited benefit in considering computationally expensive loudness and sharpness over the more traditional and easier to calculate level-based parameters, although there may be some benefit in considering Fluctuation Strength and Roughness, which showed variation with increasing room values for the HATSs.

However, all insights must be considered under the following limitations (from Loh et al., 2022b):

> The measurement methods in this study have several limitations that are generally related to logistical difficulties in conducting measurements with children and/or at educational institutions. For the room acoustic measurements, some of the doors to adjoining rooms were not closed since they are typically left open during daily activities. [...] Room acoustics measurements were not conducted during occupancy, which limits the characterization of rooms to unoccupied conditions only, and further limits comparisons with previous studies.
>
> In terms of the *in-situ* measurements, a major limitation was the fixed locations of the transducers, which was to avoid too many disruptions to the normal activities of children. The close-by positioning of the transducers can lead to shadowing effects that were not examined in detail within this study, which is acknowledged as a limitation. Furthermore, the head directions of the HATSs were static and not changed during the measurement durations, so the effects introduced by head movements could not be analyzed within this study. These issues could be improved in a future study wherein several measurement locations including using a combination of HATS locations and measurements using microphones placed near children's ears, etc., as in McAllister et al. (2009).
>
> Moreover, the number of children in the rooms fluctuated during the day, which could not be monitored. Hence, detailed analyses using the number of children as a factor were not possible and are recommended for future studies. Additionally, the type of noises and associated activities were not specifically analyzed in this study. In further studies, the measurement method could be chosen to allow studying the impact of noise sources (e.g., speech-based, impact sounds, etc.). Finally, since only one example of adult and child HATS were used, results are limited to these HATS. It is likely that using a different HATS will lead to different results to an extent, which can be considered in a separate study.

2.3.5 Concluding remarks on acoustic measurements

Based on the results and discussions, the following conclusions are summarized and tailored to the purpose of this thesis, evaluating the differences between measurement and evaluation methods appropriately adapted for children considering different anthropometric sizes by using HATSs (from Loh et al., 2022b):

> This paper presents pilot results from acoustic assessments in German primary schools and daycare centers with focus on long-term measurements using an adult and a child HATS, and an omnidirectional microphone, besides room acoustic. Main conclusions are as follows:
>
> 1. The room acoustics in both classrooms and playrooms in Germany has a lot of scope for improvements to meet guidelines for 'good acoustics' outlined in recent studies. The current findings point towards the use of typical room acoustics treatment for high-frequency sound absorption to control the reverberation times while ensuring high speech intelligibility.
> 2. Based on omnidirectional measurements, long-term $L_{\mathrm{A,eq}}$ values in classrooms and playrooms are very similar ($\sim 66\,\mathrm{dB}$), which are in general lower and with a smaller range than similar measurements in other countries in Europe and the USA. Similar trends are reported for percentile levels (L_{10}, L_{90}).
> 3. There are some indications that psychoacoustic parameters (especially fluctuation strength and roughness) may be beneficial in complementing SPL-based parameters.
> 4. Overall, while the findings here suggest some benefit in considering child HATS over adult HATS and/or omnidirectional measurements in terms of characterizing *in-situ* noise measurements, especially in the higher frequencies where anthropomorphic details are important, these findings are specific to the current measurement method with fixed locations for the transducers including orientation for the HATSs. More research that considers various transducer locations and orientations is necessary to validate these findings.
> 5. The findings here [...] are considered preliminary and need further studies that characterize children's perception in relation to the wide range of parameters studied here [...].

2.4 Influences of Anthropometric Sizes of Adults and Children Head and Torso Simulators (HATSs)

This section examines the benefits and importance of choosing appropriate measurement methods when assessing noise exposure with respect to children, who presumably have different anthropometric sizes and auditory perceptions than adults. Thus, the pre- and primary school measurements were additionally analyzed and evaluated in detail to address differences specifically between omnidirectional microphones, child and adult head and torso simulators (HATSs).

To account for the directionally independent components within the HATS measurement, all adult and child HATS measured impulse responses and signals were additionally filtered using a diffuse field equalization filter computed from the adult and child HATS, respectively, following Middlebrooks and Green (1990). This process was added after publishing the work by Loh et al. (2022b) aiming to add insights on the impact of the diffuse field equalization, which is one common procedure among others in the post-processing of HATS measurements.

The newly computed descriptives of the room acoustic parameters (T_{20}, T_{30}, EDT, C_{50}, D_{50}, and T_{S}) with respect to the measurement methods are summarized in Table 2.3. The complete parameter tables per room can be found in the Appendix Chapter A.1).

The descriptives of the chosen noise parameters are listed in Table 2.4. For each room and each measurement method, a single value of each noise parameter was computed for the analyses. The full parameter tables with descriptives of each room are added to the Appendix Chapter A.2.

Binaural versions of the standard room acoustic and noise parameters presented in this section were calculated as described in Section 2.3.1 and Section 2.3.2 using the equalized HATS measurements.

Binaural versions of the Z- and A-weighted SPL parameters were calculated using the level summation rule instead of averaging the SPL values of the left and right ear signals. The *prominent ear* version remains unchanged as described in Section 2.3.1.
The binaural versions were evaluated for the Z- and A-weighted SPL. Binaural versions of the A-weighted SPLs were examined conditionally, though binaural effects were presumed to double. The A-weighting of the SPL parameters already accounts for the subjective perception and the influences of the HATS to a certain extent; thus, results from A-weighted SPL must be interpreted cautiously.

Table 2.3: Summary of room acoustic parameter descriptives

RA parameter		**ref**	**A-HATS$_{av}$**	**C-HATS$_{av}$**	**A-HATS$_{prom}$**	**C-HATS$_{prom}$**
	BB	0.54 (0.34 - 0.95)	0.65 (0.34 - 2.34)	0.67 (0.36 - 1.06)	0.66 (0.35 - 2.36)	0.69 (0.38 - 1.14)
T_{20} **[s]**	Mid	0.58 (0.34 - 1.10)	0.58 (0.34 - 1.09)	0.57 (0.33 - 1.09)	0.59 (0.34 - 1.11)	0.59 (0.33 - 1.10)
	L.v.H.	1.28 (0.90 - 1.97)	1.69 (0.90 - 7.71)	1.96 (0.90 - 4.62)	1.70 (0.91 - 7.63)	2.01 (0.91 - 4.89)
	BB	0.56 (0.35 - 0.97)	1.43 (0.46 - 4.32)	2.14 (0.79 - 5.98)	1.48 (0.47 - 4.48)	2.37 (0.81 - 7.37)
T_{30} **[s]**	Mid	0.58 (0.34 - 1.10)	0.80 (0.41 - 2.68)	0.83 (0.34 - 2.92)	0.82 (0.41 - 2.69)	0.91 (0.34 - 2.93)
	L.v.H.	1.29 (0.91 - 1.95)	4.75 (1.24 - 24.34)	4.49 (1.70 - 8.70)	4.78 (1.25 - 24.17)	4.41 (1.75 - 8.83)
	BB	0.51 (0.31 - 0.90)	0.49 (0.29 - 0.90)	0.49 (0.27 - 0.90)	0.52 (0.31 - 0.94)	0.51 (0.30 - 0.93)
EDT **[s]**	Mid	0.56 (0.32 - 1.06)	0.55 (0.31 - 1.05)	0.53 (0.27 - 1.05)	0.57 (0.32 - 1.09)	0.55 (0.30 - 1.08)
	L.v.H.	1.34 (0.91 - 1.94)	1.34 (0.92 - 2.27)	1.43 (0.95 - 2.14)	1.33 (0.92 - 2.22)	1.40 (0.89 - 2.12)
	BB	6.2 (1.6 - 10.4)	7.1 (2.0 - 11.2)	7.2 (2.2 - 11.8)	7.6 (2.5 - 11.8)	7.7 (2.7 - 12.4)
C_{50} **[dB]**	Mid	5.1 (0.3 - 9.5)	5.7 (0.5 - 10.4)	5.7 (-0.0 - 10.8)	6.1 (1.0 - 10.7)	6.1 (0.6 - 11.3)
	L.v.H.	0.6 (-0.1 - 1.0)	0.6 (-0.1 - 1.0)	0.5 (-0.1 - 0.9)	0.6 (0.1 - 0.9)	0.5 (0.0 - 0.9)
	BB	76.9 (58.1 - 89.7)	79.4 (58.6 - 91.8)	79.5 (58.9 - 92.0)	80.9 (61.3 - 92.7)	80.8 (61.7 - 92.9)
D_{50} **[%]**	Mid	74.1 (51.9 - 88.5)	76.3 (52.7 - 91.3)	76.3 (49.9 - 91.8)	78.1 (55.5 - 91.9)	77.8 (53.5 - 92.7)
	L.v.H.	0.9 (0.7 - 1.0)	0.9 (0.7 - 1.0)	0.9 (0.7 - 1.0)	0.9 (0.7 - 1.0)	0.9 (0.7 - 1.0)
	BB	0.03 (0.02 - 0.06)	0.03 (0.02 - 0.06)	0.03 (0.02 - 0.06)	0.03 (0.02 - 0.06)	0.03 (0.02 - 0.06)
T_S **[s]**	Mid	0.04 (0.02 - 0.07)	0.04 (0.02 - 0.07)	0.04 (0.02 - 0.07)	0.04 (0.02 - 0.07)	0.04 (0.02 - 0.08)
	L.v.H.	1.48 (0.96 - 2.15)	1.56 (0.99 - 2.79)	1.75 (1.12 - 2.78)	1.54 (0.95 - 2.70)	1.71 (1.06 - 2.65)

Note. ref = omnidirectional microphone; A-HATS$_{av}$/C-HATS$_{av}$ = adult and child HATS, with the averaging method; A-HATS$_{prom}$/C-HATS$_{prom}$ = adult and child HATS, with the prominent-ear method. BB = broadband (31.5 Hz-16 kHz), Mid = mid-frequencies (500 Hz, 1 kHz, 2 kHz), L.v.H. = low- (31.5 Hz-1 kHz) versus high- (2 kHz-16 kHz) frequency ratio.

Table 2.4: Summary of noise parameter descriptives.

	Ref	A-HATS$_{av}$	C-HATS$_{av}$	A-HATS$_{prom}$	C-HATS$_{prom}$
$L_{Z,eq}$ [dB]	67.3 (61.2 - 69.7)	71.6 (68.5 - 76.5)	70.1 (66.4 - 73.8)	77.9 (64.9 - 91.3)	77.9 (67.0 - 93.2)
$L_{ZF,mean}$ [dB]	59.1 (54.0 - 63.1)	64.9 (59.0 - 71.4)	63.4 (57.0 - 68.6)	60.2 (54.5 - 66.5)	58.6 (52.0 - 63.8)
$L_{Z,10}$ [dB]	68.0 (62.9 - 71.1)	74.8 (71.1 - 80.5)	73.0 (68.4 - 77.1)	70.3 (66.3 - 75.7)	68.3 (63.5 - 72.4)
$L_{Z,90}$ [dB]	50.6 (46.1 - 55.4)	55.1 (47.8 - 62.0)	54.3 (46.9 - 60.6)	50.3 (43.5 - 57.6)	49.3 (41.4 - 55.6)
$L_{Z,eq}$LvH	1.09 (1.05 - 1.13)	1.15 (1.10 - 1.19)	1.07 (1.03 - 1.11)	1.15 (1.10 - 1.21)	1.06 (0.98 - 1.11)
$L_{A,eq}$ [dB(A)]	61.7 (57.9 - 66.2)	69.5 (66.4 - 74.2)	68.4 (64.0 - 72.6)	75.5 (62.5 - 99.4)	73.2 (59.5 - 100.0)
$L_{AF,mean}$ [dB(A)]	52.0 (42.8 - 58.8)	60.6 (51.9 - 67.4)	59.0 (49.2 - 65.4)	56.0 (47.3 - 62.6)	54.3 (44.2 - 60.7)
$L_{A,10}$ [dB(A)]	63.9 (56.9 - 68.6)	72.5 (67.2 - 78.2)	70.8 (63.3 - 75.5)	68.0 (62.8 - 73.4)	66.3 (58.5 - 70.9)
$L_{A,90}$ [dB(A)]	39.9 (28.9 - 49.2)	48.4 (36.9 - 57.1)	47.1 (35.2 - 55.6)	43.6 (32.9 - 52.6)	42.2 (30.0 - 50.7)
$L_{A,eq}$(L.v.H)	0.82 (0.77 - 0.84)	0.89 (0.86 - 0.93)	0.83 (0.79 - 0.87)	0.87 (0.83 - 0.91)	0.80 (0.76 - 0.86)
N_{mean} [[sone]	7.59 (4.48 - 10.73)	8.36 (5.18 - 13.27)	7.99 (4.48 - 11.11)	9.26 (5.80 - 14.40)	8.81 (4.87 - 12.65)
N_5 [sone]	16.96 (12.14 - 21.31)	18.26 (13.81 - 26.24)	17.47 (12.08 - 21.79)	20.23 (15.01 - 28.55)	19.29 (13.12 - 25.22)
N_{90} [sone]	2.65 (0.86 - 4.94)	2.96 (0.98 - 5.28)	2.89 (0.85 - 5.14)	3.31 (1.27 - 5.86)	3.22 (0.94 - 5.59)
S_{mean} [acum]	1.46 (1.29 - 1.62)	1.24 (1.16 - 1.37)	1.44 (1.38 - 1.58)	1.30 (1.23 - 1.42)	1.50 (1.43 - 1.65)
S_5 [acum]	1.99 (1.86 - 2.17)	1.69 (1.57 - 1.82)	1.93 (1.79 - 2.09)	1.78 (1.65 - 1.91)	2.03 (1.89 - 2.20)
S_{90} [acum]	1.13 (0.80 - 1.29)	0.96 (0.84 - 1.10)	1.14 (1.05 - 1.28)	0.99 (0.88 - 1.13)	1.18 (1.08 - 1.32)
R_{mean} [asper]	0.08 (0.06 - 0.10)	0.10 (0.08 - 0.13)	0.08 (0.05 - 0.11)	0.12 (0.10 - 0.14)	0.09 (0.06 - 0.12)
R_5 [asper]	0.16 (0.13 - 0.24)	0.18 (0.16 - 0.20)	0.15 (0.13 - 0.18)	0.21 (0.18 - 0.24)	0.18 (0.15 - 0.21)
R_{90} [asper]	0.03 (0.01 - 0.05)	0.06 (0.04 - 0.08)	0.03 (0.01 - 0.06)	0.06 (0.04 - 0.09)	0.04 (0.01 - 0.07)
FS_{mean} [vacil]	0.07 (0.03 - 0.10)	0.08 (0.05 - 0.10)	0.08 (0.04 - 0.09)	0.09 (0.05 - 0.11)	0.08 (0.05 - 0.11)
FS_5 [vacil]	0.16 (0.10 - 0.20)	0.17 (0.12 - 0.19)	0.16 (0.11 - 0.19)	0.18 (0.14 - 0.22)	0.18 (0.12 - 0.24)
FS_{90} [vacil]	0.03 (0.00 - 0.04)	0.03 (0.01 - 0.05)	0.03 (0.01 - 0.04)	0.04 (0.01 - 0.05)	0.04 (0.01 - 0.05)

Note. Ref = omnidirectional microphone; A-HATS$_{av}$/C-HATS$_{av}$ = adult and child HATS, with the averaging method; A-HATS$_{prom}$/C-HATS$_{prom}$ = adult and child HATS, with the prominent-ear method.

2.4.1 Room acoustic parameters

All room acoustic parameters were statistically examined as single values for each parameter and each room according to the measurement methods. The one-way ANOVA results are summarized in Table 2.5.

Table 2.5: One-way ANOVA results on RA parameters as single values.

RA parameter		Statistical test	Significance
T_{20}	BB	Kruskal-Wallis-Test[a]	$p > .05$
	Mid	Kruskal-Wallis-Test[a]	$p > .05$
	L.v.H.	Kruskal-Wallis-Test[a]	$H(4) = 10.33, p = .035$
T_{30}	BB	Kruskal-Wallis-Test[a]	$H(4) = 36.22, p < .001$
	Mid	Kruskal-Wallis-Test[a]	$p > .05$
	L.v.H.	Kruskal-Wallis-Test[a]	$H(4) = 40.65, p < .001$
EDT	BB	Kruskal-Wallis-Test[a]	$p > .05$
	Mid	Kruskal-Wallis-Test[a]	$p > .05$
	L.v.H.	Kruskal-Wallis-Test[a]	$p > .05$
C_{50}	BB	Kruskal-Wallis-Test[a]	$p > .05$
	Mid	One-Way ANOVA	$p > .05$
	L.v.H.	One-Way ANOVA	$p > .05$
D_{50}	BB	Kruskal-Wallis-Test[a]	$p > .05$
	Mid	Kruskal-Wallis-Test[a]	$p > .05$
	L.v.H.	One-Way ANOVA	$p > .05$
T_S	BB	Kruskal-Wallis-Test[a]	$p > .05$
	Mid	Kruskal-Wallis-Test[a]	$p > .05$
	L.v.H.	One-Way ANOVA	$p > .05$

Note. In case assumptions of normality[a], the respective non-parametric test was chosen. BB = broadband (31.5 Hz-16 kHz), Mid = mid-frequencies (500, 1k, 2 kHz), L.v.H. = low- (31.5 Hz - 1 kHz) versus high- (2 kHz - 16 kHz) frequency ratio.

The results indicate significant effects for the room acoustic parameter T_{30}, calculated for the broadband frequency spectrum and the low- versus high-frequency ratio, as well as for the low-high-frequency ratio of T_{20}.

The post-hoc t-test analyses are summarized in Table 2.6 comparing the five levels to each other. No significant difference was revealed for the low- versus high-frequency ratio of T_{20} within the post-hoc tests between all measurement methods, though the main effect was significant. However, significant differences were yielded in the broadband calculation of T_{30} between the omnidirectional microphone (ref) and both HATSs using the prominent ear method, as well as the reference microphone and the child HATS using the averaging method. Regarding the low-high-frequency ratio of T_{30}, significant differences were found between the omnidirectional microphone and both HATS, independent from the averaging or prominent-ear method.

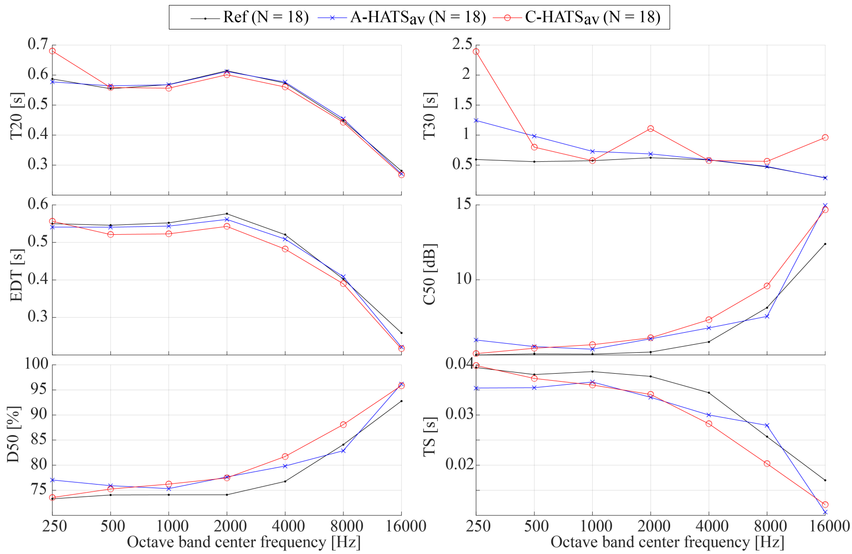
Table 2.6: Post-hoc analyses of RA parameters as single values.

		z	p
	T_{20} L.v.H.		
ref	A-HATS$_{av}$	0.836	>.05
ref	C-HATS$_{av}$	2.482	>.05
A-HATS$_{av}$	C-HATS$_{av}$	-1.646	>.05
ref	A-HATS$_{prom}$	1.142	>.05
ref	C-HATS$_{prom}$	2.686	>.05
A-HATS$_{prom}$	C-HATS$_{prom}$	-1.544	>.05
	T_{30} BB		
ref	A-HATS$_{av}$	2.526	>.05
ref	**C-HATS$_{av}$**	**4.919**	**<.001**
A-HATS$_{av}$	C-HATS$_{av}$	-2.392	>.05
ref	**A-HATS$_{prom}$**	**2.833**	**.046**
ref	**C-HATS$_{prom}$**	**5.289**	**<.001**
A-HATS$_{prom}$	C-HATS$_{prom}$	-2.456	>.05
	T_{30} L.v.H.		
ref	**A-HATS$_{av}$**	**3.796**	**.001**
ref	**C-HATS$_{av}$**	**5.487**	**<.001**
A-HATS$_{av}$	C-HATS$_{av}$	-1.691	>.05
ref	**A-HATS$_{prom}$**	**3.968**	**<.001**
ref	**C-HATS$_{prom}$**	**5.538**	**<.001**
A-HATS$_{prom}$	C-HATS$_{prom}$	-0.797	>.05

Note. Significance values with regard to the p-value have been adjusted by the Bonferroni correction for multiple tests. Significant effects are indicated in bold. ref = omnidirectional microphone; A-HATS$_{av}$/C-HATS$_{av}$ = adult and child HATS, with the averaging method; A-HATS$_{prom}$/C-HATS$_{prom}$ = adult and child HATS, with the prominent-ear method. BB = broadband (31.5 Hz-16 kHz), L.v.H. = low- (31.5 Hz-1 kHz) versus high- (2 kHz-16 kHz) frequency ratio.

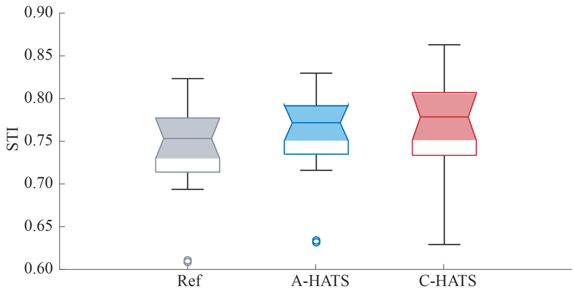
Considering the single values, only the reverberation time with respect to the frequency bands revealed differences between the measurement methods. A detailed examination of the room acoustic parameters over each octave band (e.g., as shown in Figure 2.3 with regard to the averaging method) cannot give explanations, especially in terms of the lower versus the higher frequencies.
It further shows that the reverberation time and the early decay time were quite insensitive to differences between measurement methods across octave bands. The parameters C_{50}, D_{50}, and T_{S}, however, indicated tendencies of larger differences comparing measurement methods across the octave bands, especially towards higher frequency bands, though the single values did not reflect these differences. This leads to the assumption that these parameters analyzed across octave bands might add insights to the standard examination procedure of room acoustic parameters for noise assessment methods appropriate for children.

Figure 2.3: Room acoustic parameters over octave bands for head and torso simulators using the averaging method.



The STI, as a single value calculated following van Wijngaarden and Drullman (2008), across all rooms and concerning the measurement devices, omnidirectional microphone, adult and child HATS, is shown in Figure 2.4. Levene's Test on normality revealed a significant effect for the sample of the reference and the adult head and torso simulator (A-HATS) (both $p < .05$). Thus, the non-parametric Kruskal-Wallis was chosen and revealed an insignificant effect of measurement method $H(2) = 2.75, p > .05$, i.e., no differences between the omni-directional microphone, the A-HATS and the child head and torso simulator (C-HATS) was found.

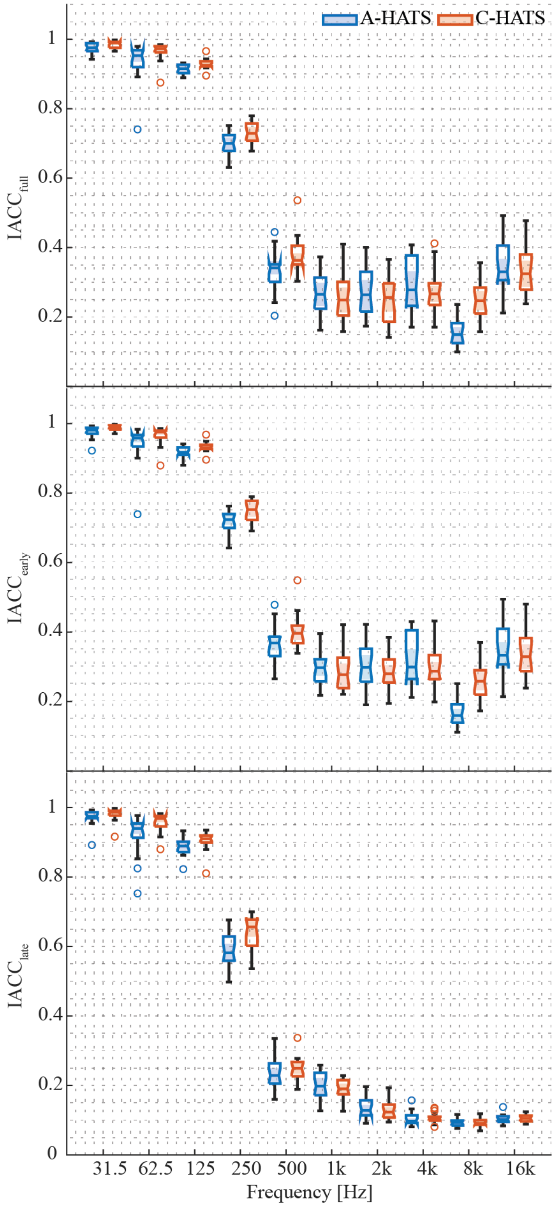
Figure 2.4: STI differences between the measurement methods.



Regarding IACC, single values averaged across all frequency bands for the $IACC_{\text{full}}$, $IACC_{\text{early}}$, and $IACC_{\text{late}}$. were analyzed The $IACC_{\text{early}}$ measured using the A-HATS (M = 0.536, SD = 0.032) compared to the C-HATS (M = 0.550, SD = 0.028) showed an insignificant effect, $t(34) = -1.472$ (two-sided), $p > .05$ as well as for the $IACC_{\text{full}}$, $t(34) = -1.466$ (two-sided), $p > .05$ measured using A-HATS (M = 0.519, SD = 0.035) compared to C-HATS (M = 0.535, SD = 0.031). However, the $IACC_{\text{late}}$ revealed a significant effect, $t(34) = -2.981$ (two-sided), $p = .005$ between A-HATS (M = 0.424, SD = 0.012) and C-HATS (M = 0.435, SD = 0.010).

In general, results indicated little differences between measurement methods when investigating room acoustic parameters using single-values. Even binaural parameters, such as the binaural STI (Van Wijngaarden & Drullman, 2008) and the IACC developed to account for possible differences in binaural hearing yielded differences only for $IACC_{\text{late}}$.
Differences in $IACC$ reflect differences and separation of signals reaching the two ears. Since the A-HATS and the C-HATS had different anthropometric head sizes, it was assumed that at least this parameter would reveal more significant differences. However, IACC (as well as the IACC based on early and late reflec-

Figure 2.5: IACC differences across octave bands between the factor measurement methods.



tions, $IACC_{\text{early}}$ and $IACC_{\text{late}}$, respectively) depends on the sound reflections within a room (Barron, 1971). Since the HATSs were positioned at different positions, the reflections varied significantly from position to position, leading to a high variance in measurement results that might mask the overall difference between the two HATSs. Furthermore, Okano (2002) reported a JND of 5% in IACC based on investigations with adult participants. Though differences were found, it remains unclear whether the differences in IACC between the HATSs are interpretable since they were neglectable in terms of noticeable differences.

A frequency-dependent statistical analysis of the $IACC$ concerning the octave bands with center frequencies in the range of 31.5 Hz to 16 kHz. $IACC_{\text{full}}$, $IACC_{\text{early}}$ and $IACC_{\text{late}}$ were calculated for each octave band and both HATS (cf. Figure 2.5). Independent t-tests were conducted to compare HATS for the IACC values on each octave band. Corresponding results are listed in Table 2.7. Significant differne between both HATS were found for $IACC_{\text{full}}$ and $IACC_{\text{early}}$ in the frequency bands 31.5 Hz, 125 Hz, 250 Hz, and 8 kHz. For $IACC_{\text{late}}$, differences were found in the lower octave bands 62.5 Hz and 250 Hz.

Table 2.7: T-test results for IACC based on octave bands.

	$IACC_{\text{full}}$			$IACC_{\text{early}}$			$IACC_{\text{late}}$		
Hz	*df*	*t*	*p*	*df*	*t*	*p*	*df*	*t*	*p*
31.5	**34**	**-2.274**	**.029**	**34**	**-2.516**	**.017**	34	-1.054	>.05
62.5	34	-2.004	>.05	34	-1.784	>.05	**34**	**-2.458**	**.019**
125	**34**	**-3.274**	**.002**	**34**	**-2.982**	**.005**	34	-1.759	>.05
250	**34**	**-3.133**	**.004**	**34**	**-2.896**	**.007**	**34**	**-2.965**	**.005**
500	**34**	**-2.309**	**.027**	**34**	**-2.196**	**.035**	34	-1.108	>.05
1k	34	0.093	>.05	34	0.142	>.05	34	1.226	>.05
2k	34	0.894	>.05	34	1.068	>.05	34	0.618	>.05
4k	34	0.731	>.05	34	0.854	>.05	34	-.607	>.05
8k	**34**	**-5.603**	**<.001**	**34**	**-5.756**	**<.001**	34	0.004	>.05
16k	34	0.978	>.05	34	0.987	>.05	34	-.528	>.05

Note. Significant effects are indicated in bold. Octave band are denoted with center frequencies. p is calculated *twosided*

Differences in the high-frequency band, 8 kHz, can be explained by the findings reported by Fels et al. (2004) and Fels and Vorländer (2009) where they found significant differences in the HRTF of adults compared to the HRTF of children. However, results indicated differences starting from 500 Hz towards higher frequencies so that the found differences in the lower frequencies, 250 Hz and lower, can not explained by the results by Fels et al. (2004) and Fels and Vorländer (2009).

To summarize findings, single-values of room acoustic parameters cannot represent the full extent of differences between HATS when accounting for differences in anthropometric dimension. Results in this work indicated a strong frequency-dependent difference that was reflected in the IACC to begin with.

To thoroughly examine the frequency-dependent differences, analyses regarding the (binaural) STI in Section 2.4.3 were conducted. The binaural STI specifically takes binaural aspects of hearing into account (e.g., van Wijngaarden & Drullman, 2008) and was evaluated subjectively against the standard STI measured using omni-directional microphones. However, this binaural STI was primarily examined with regard to adults, and children were not in focus. Therefore, this thesis aims to gain more insights on the effect of differences of anthropometric sizes in objective measurement results by comparing the overall singe-value of STI versus binaural STI for both HATS and with different binaural evaluation methods. Furthermore, the STI was examined in detail based on frequency-dependent MTIs to obtain insights between single-values and frequency-dependent analyses concerning anthropometric size differences.

2.4.2 Noise parameters

One-way ANOVAs (and corresponding robust tests or non-parametric alternative tests) were applied for all noise parameters as single values to examine possible differences between the measurement methods (ref vs. A-HATS$_{\mathrm{av}}$ vs. C-HATS$_{\mathrm{av}}$ vs. A-HATS$_{\mathrm{prom}}$ vs. C-HATS$_{\mathrm{prom}}$). The results of the statistical analyses are summarized in Table 2.8. Post-hoc analyses were conducted for all significant main effects of measurement method and significant post-hoc tests are denoted in Table 2.9, Table 2.10, and Table 2.11.

Summing up the significant differences with the level-based parameters, significant main effects of measurement methods were found for all L_{Z} and all L_{A} parameters. Concerning the prominent-ear method, differences between the reference and both HATS were observed for $L_{\mathrm{Z,eq}}$ and $L_{\mathrm{Z,eq}}$(L.v.H). Differences between the two HATS were found for $L_{\mathrm{Z,10}}$, $L_{\mathrm{Z,90}}$, and $L_{\mathrm{Z,eq}}$(L.v.H) for both methods (averaging and prominent-ear method).
Regarding the psychoacoustic parameters, all parameters yielded significant effects of measurement methods except for the loudness parameters N_{mean} and N_{90}.

Regarding the level-based noise parameters, differences were found primarily significant between the omnidirectional and both HATS using either averaging or prominent-ear or both methods. Between the child and adult HATS, significant

Table 2.8: Statistical analysis results on noise parameter as single values.

RA parameter	**Statistical test**	**Significance**
$L_{\mathrm{Z,eq}}$	Welch's F-Test[b]	$F(4, 41.1) = 16.78, p < .001, \eta_p^2 = .492$
$L_{\mathrm{ZF,mean}}$	One-Way ANOVA	$F(4, 89) = 15.30, p < .001, \eta_p^2 = .523$
$L_{\mathrm{Z,10}}$	One-Way ANOVA	$F(4, 89) = 27.76, p < .001, \eta_p^2 = .649$
$L_{\mathrm{Z,90}}$	One-Way ANOVA	$F(4, 89) = 8.71, p < .001, \eta_p^2 = .405$
$L_{\mathrm{Z,eq}}$(L.v.H)LvH	One-Way ANOVA	$F(4, 89) = 56.31, p < .001, \eta_p^2 = .780$
$L_{\mathrm{A,eq}}$	Kruskal-Wallis-Test[a]	$H(4) = 41.76, p < .001$
$L_{\mathrm{AF,mean}}$	One-Way ANOVA	$F(4, 89) = 13.47, p < .001, \eta_p^2 = .495$
$L_{\mathrm{A,10}}$	One-Way ANOVA	$F(4, 89) = 27.67, p < .001, \eta_p^2 = .649$
$L_{\mathrm{A,90}}$	One-Way ANOVA	$F(4, 89) = 7.45, p < .001, \eta_p^2 = .374$
$L_{\mathrm{A,eq}}$(L.v.H)	One-Way ANOVA	$F(4, 89) = 59.88, p < .001, \eta_p^2 = .790$
N_{mean}	One-Way ANOVA	$F(4, 89) = 2.00, p > .05$
N_5	One-Way ANOVA	$F(4, 89) = 3.44, p = .012, \eta_p^2 = .245$
N_{90}	One-Way ANOVA	$F < 1, p > .05$
S_{mean}	One-Way ANOVA	$F(4, 89) = 56.02, p < .001, \eta_p^2 = .780$
S_5	Welch's F-Test[b]	$F(4, 42.2) = 58.53, p < .001, \eta_p^2 = .771$
S_{90}	One-Way ANOVA	$F(4, 89) = 26.66, p < .001, \eta_p^2 = .641$
R_{mean}	One-Way ANOVA	$F(4, 89) = 36.93, p < .001, \eta_p^2 = .706$
R_5	One-Way ANOVA	$F(4, 89) = 26.77, p < .001, \eta_p^2 = .642$
R_{90}	One-Way ANOVA	$F(4, 89) = 30.84, p < .001, \eta_p^2 = .671$
FS_{mean}	One-Way ANOVA	$F(4, 89) = 2.60, p = .042, \eta_p^2 = .208$
FS_5	One-Way ANOVA	$F(4, 89) = 3.54, p = .010, \eta_p^2 = .249$
FS_{90}	Kruskal-Wallis-Test[a]	$H(4) = 10.67, p = .031$

Note. In case assumptions of normality[a] or homogeneity[b] were violated, the respective non-parametric test was chosen. L.v.H. = low- versus high-frequency ratio.

differences were only yielded for the low- versus high-frequency ratio of $L_{\mathrm{Z,eq}}$and $L_{\mathrm{A,eq}}$.

In terms of the psychoacoustic parameters, significant differences were mainly observed between the omnidirectional microphone compared to the adult HATS and between the adult and child HATS. However, no significant differences were found between the omnidirectional microphone and the child HATS. Considering N_5, FS_5, and FS_{90}, differences were yielded between the omnidirectional microphone and the adult HATS using the prominent ear method.

Table 2.9: Significant post-hoc analyses results for the noise parameter concerning the Kruskal-Wallis test.

		z	p
	$L_{\mathbf{A,eq}}$ [dB(A)]		
ref	A-HATS$_{\mathrm{av}}$	4.791	<.001
ref	C-HATS$_{\mathrm{av}}$	3.872	.001
ref	A-HATS$_{\mathrm{prom}}$	5.487	<.001
ref	C-HATS$_{\mathrm{prom}}$	5.435	<.001
	$FS_{\mathbf{90}}$ [vacil]		
ref	A-HATS$_{\mathrm{prom}}$	2.986	.028

Note. Significance values with regard to the p-value have been adjusted by the Bonferroni correction for multiple tests.

From these results, it can be assumed that measuring with HATS can add insights to the noise assessment process. The direction-independent components within the measurement results showed significant differences between the omnidirectional microphone and the HATSs within level-based noise parameters. Thus, when only level-based parameters are applied, information might be missing when interpreting measurement results toward a more plausible and realistic perception of children. With this, it can be questioned whether the commonly used A-weighted sound pressure level, developed based on studies conducted with adult participants, is also suited for noise assessment appropriately for children.

As for the room acoustic parameters, a frequency-dependent investigation of the noise parameters was conducted to reveal more insights into the differences between the omni-directional microphone, the adult HATS, and the child HATS. However, the psychoacoustic parameters already integrated the spectral effects within single values, so the frequency-dependent analysis was only meaningful regarding the level-based parameters of L_{Z} and L_{A} (cf. Figure 2.6).

Table 2.10: Significant post-hoc analyses results for the level-based noise parameter concerning the ANOVA tests.

G_1	G_2	p (G_1 - G_2)	M_{diff}	**[95 %-CI]**
		$L_{\mathbf{Z,eq}}$ **[dB]**		
ref	A-$HATS_{prom}$	<.001	-10.6	[-15.9, -5.3]
ref	C-$HATS_{prom}$	<.001	-10.6	[-16.0, -5.3]
		$L_{\mathbf{ZF,mean}}$ **[dB]**		
ref	A-$HATS_{av}$	<.001	-5.8	[-8.7, -2.9]
ref	C-$HATS_{av}$	<.001	-4.3	[-7.2, -1.4]
		$L_{\mathbf{Z,10}}$ **[dB]**		
ref	A-$HATS_{av}$	<.001	-6.8	[-9.1, -4.5]
ref	C-$HATS_{av}$	<.001	-4.9	[-7.2, -2.6]
		$L_{\mathbf{Z,90}}$ **[dB]**		
ref	A-$HATS_{av}$	.005	-4.5	[-8.1, -0.9]
ref	C-$HATS_{av}$	.042	-3.7	[-7.3, -0.1]
		$L_{\mathbf{Z,eq}}$**(L.v.H)**		
ref	A-$HATS_{av}$	<.001	-0.05	[-0.08, -0.03]
ref	C-$HATS_{av}$	.021	0.03	[0.00, 0.05]
ref	A-$HATS_{prom}$	<.001	-0.06	[-0.08, -0.03]
ref	C-$HATS_{prom}$	<.001	0.04	[0.01, 0.06]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	0.08	[0.06, 0.10]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	0.09	[0.07, 0.12]
		$L_{\mathbf{AF,mean}}$ **[dB(A)]**		
ref	A-$HATS_{av}$	<.001	-8.6	[-12.5, -4.7]
ref	C-$HATS_{av}$	<.001	-7.0	[-10.9, -3.1]
ref	A-$HATS_{prom}$	.040	-4.0	[-7.8, -0.1]
		$L_{\mathbf{A,10}}$ **[dB(A)]**		
ref	A-$HATS_{av}$	<.001	-8.5	[-11.2, -5.9]
ref	C-$HATS_{av}$	<.001	-6.9	[-9.5, -4.2]
ref	A-$HATS_{prom}$	<.001	-4.0	[-6.7, -1.4]
ref	A-$HATS_{av}$	<.001	-8.4	[-13.6, -3.2]
ref	C-$HATS_{av}$	.001	-7.2	[-12.4, -2.0]
		$L_{\mathbf{A,eq}}$**(L.v.H)**		
ref	A-$HATS_{av}$	<.001	-0.08	[-0.10, -0.06]
ref	A-$HATS_{prom}$	<.001	-0.05	[-0.07, -0.03]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	0.07	[0.05, 0.09]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	0.07	[0.05, 0.09]

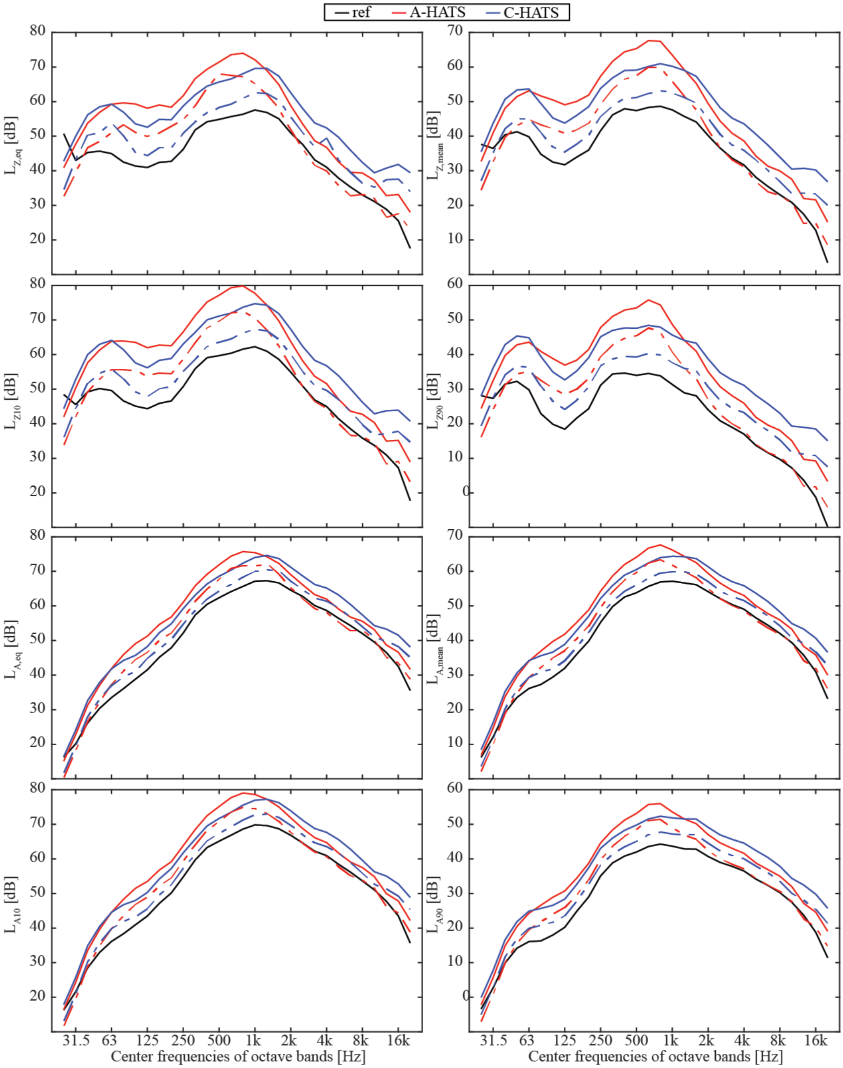
Note. ref = omnidirectional microphone; A-$HATS_{av}$/C-$HATS_{av}$ = adult and child HATS, with the averaging method; A-$HATS_{prom}$/C-$HATS_{prom}$ = adult and child HATS, with the prominent-ear method. L.v.H = low- versus high-frequency ratio.

Table 2.11: Significant post-hoc analyses results for the psychoacoustic noise parameter concerning the ANOVA tests.

G_1	G_2	p (G_1 - G_2)	M_{diff}	**[95 %-CI]**
		N_5 [sone]		
ref	A-$HATS_{prom}$	.018	-3.26	[-6.18, -0.34]
		S_{mean} [acum]		
ref	A-$HATS_{av}$	<.001	0.22	[0.16, 0.28]
ref	A-$HATS_{prom}$	<.001	0.17	[0.10, 0.23]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	-0.20	[-0.26, -0.14]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	-0.21	[-0.27, -0.14]
		S_5 [acum]		
ref	A-$HATS_{av}$	<.001	0.31	[0.22, 0.39]
ref	A-$HATS_{prom}$	<.001	0.21	[0.13, 0.29]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	-0.25	[-0.33, -0.17]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	-0.25	[-0.33, -0.17]
		S_{90} [acum]		
ref	A-$HATS_{av}$	<.001	0.17	[0.09, 0.25]
ref	A-$HATS_{prom}$	<.001	0.13	[0.06, 0.21]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	-0.18	[-0.26, -0.11]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	-0.19	[-0.26, -0.11]
		R_{mean} [asper]		
ref	A-$HATS_{av}$	<.001	-0.03	[-0.04, -0.02]
ref	A-$HATS_{prom}$	<.001	-0.04	[-0.05, -0.03]
ref	C-$HATS_{prom}$	.003	-0.02	[-0.03, 0.00]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	0.03	[0.01, 0.04]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	0.03	[0.01, 0.04]
		R_5 [asper]		
ref	A-$HATS_{av}$	.002	-0.02	[-0.04, -0.01]
ref	A-$HATS_{prom}$	<.001	-0.05	[-0.06, -0.03]
ref	C-$HATS_{prom}$	.001	-0.02	[-0.04, -0.01]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	0.03	[0.01, 0.04]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	0.03	[0.01, 0.04]
		R_{90} [asper]		
ref	A-$HATS_{av}$	<.001	-0.03	[-0.04,-0.02]
ref	A-$HATS_{prom}$	<.001	-0.04	[-0.05, -0.03]
A-$HATS_{av}$	C-$HATS_{av}$	<.001	0.03	[0.01, 0.04]
A-$HATS_{prom}$	C-$HATS_{prom}$	<.001	0.03	[0.01, 0.04]
		FS_5 [vacil]		
ref	A-$HATS_{prom}$	.020	-0.03	[-0.06, 0.00]

Note. ref = omnidirectional microphone; A-$HATS_{av}$/C-$HATS_{av}$ = adult and child HATS, with the averaging method; A-$HATS_{prom}$/C-$HATS_{prom}$ = adult and child HATS, with the prominent-ear method. L.v.H = low- versus high-frequency ratio.

Figure 2.6: SPL parameters over octave bands (average method = solid line, prominent-ear method = dashed line).



The results show deviations of the frequency response curves between the adult HATS and the child HATS as well as the omnidirectional microphone in the high frequencies above 2 kHz for all four level-based parameters. The shape of the frequency response curve of the child HATS and the omnidirectional microphone was similar to each other, especially with a constant offset across all frequency bands of the child HATS with higher levels relative to the omnidirectional microphone. In contrast, the frequency response curve of the adult HATS showed a steeper decrease of levels towards higher frequencies. This finding might explain the differences in the psychoacoustic parameters of sharpness and roughness between the adult HATS and the omnidirectional microphone. In contrast, the differences between the omnidirectional microphone and the child HATS were neglectable.

Adding to the limitations of the general measurement setup as presented in Section 2.3.4, it is important to highlight the fact that the both chosen HATS (ITA adult HATS (Schmitz, 1995) and ITA child HATS (Fels et al., 2004)) within this work did not comprise ear canal simulators. Fels (2008) highlighted in her dissertation the importance of the ear canal contributing to the amplification of the higher frequency range. The ear canal impedances of children up to seven years old differed significantly from those of adults in simulated and measured data. The data of ear canal as well as eardrum impedances indicated a shift of the first resonance from higher to lower frequencies in the range between 2.9 kHz and 9.5 kHz with increasing age. This explains the small differences between HATS found in this thesis. It can be assumed that adding ear canal simulators to the HATS respectively to the anthropometric measures of adults and children would make differences in the higher frequency bands more prominent observable.

Furthermore, adult and child HATS differed significantly in term of the pinna shapes. While the adult HATS provided a detailed outer ear replica, the child HATS only comprised a simplifed model including the pinna and the cavum conchae (cf. Fels & Vorländer, 2009). Spectral amplication introduced by the complex shape of the outer ear were, therefore, not fully represented in the child HATS measurements. This could be an explanation foir the missing expected difference between child HATS and omnidirectional microphone that was observable for adult HATS compared to omnidirectional microphone and child HATS.

Differences between the frequency response curves were more distinct within the L_{Z} than in the L_{A}. With this observation, it can be assumed that possible effects between omnidirectional, adult HATS and child HATS might be removed within the noise assessment due to the A-weighting. It is necessary to examine whether these differences in L_{A} and the psychoacoustic parameters are perceivable and evaluated differently by adults compared to children.

2.4.3 Binaural speech transmission index of HATSs

Following the insights from the previous sections concerning the importance of frequency-dependent differences in room acoustic and noise parameters, the *speech transmission index* (STI) (IEC, 2012) was examined in detail to understand the specific impact of measurements using HATSs. This parameter was chosen since it is a commonly used metric to describe information loss within the speech transmission between speakers and listeners, especially in the context of classrooms(Astolfi et al., 2019a). Preliminary results of the work presented in this section were published as a conference paper at Forum Acusticum 2023 (Loh et al., 2023a).

As described in IEC (2012) and in the work by Loh et al. (2023a):

> [... The STI] is derived from objective assessment methods [and] is calculated by determining and further processing the Modulation Transfer Function (modulation transfer function (MTF)), which contains the effects inside the transmission channel reflected by differences in the intensity envelope of the speech signal (sent signal vs. received signal). Traditionally, the STI within rooms has been assessed using omnidirectional mono microphones without considering the effects of binaural hearing. Recently, several studies have investigated binaural models in combination with the STI measured using head and torso simulators (HATSs). They showed that using HATSs better represents human perception (Schlesinger et al., 2009; van Wijngaarden & Drullman, 2008). A binaural STI (bSTI) comprises advantages of binaural hearing and, therefore, helps to minimize mismatches between the subjective speech intelligibility and the objective STI metric.
>
> It is unknown whether these binaural models, developed based on investigations with adults, also apply to children with considerably different anthropometric sizes. Fels et al. (2004) and Fels and Vorländer (2009) showed that the different anthropometric sizes of children compared to adults result in a different amplification of sound on each octave band, respectively [...]. Following this, a study in Italian classrooms (Prodi et al., 2007) measured STI and interaural cross-correlation (IACC) using an adult HATS, a child HATS, and a mono microphone as receiver devices. They observed some differences in the STI between those three receiver devices. However, differences in IACC were more pronounced.
> This observation was expected since the STI is developed as a single-value metric that does not account for individual differences in each octave band. Though *modulation transfer indices* (MTIs) are calculated for the *modulation*

> *transfer functions* (MTFs) individually for each octave band, the MTIs are summed up finally to the single-value STI and interpreted as such (IEC, 2012).

According to previous works (Bronkhorst, 2015; Lavandier & Best, 2020), there are different methods to account for the binaural hearing aspects within the computation of STI. The MTF coefficients for each octave band are derived for each ear separately, and before calculating the MTIs, specific MTF coefficients are chosen to represent both ears. For this work, three methods were chosen:

1. The mean can be calculated from the MTF coefficients from the left and right ear (*mean method*).
2. The better ear can be chosen, i.e., both ears are directly compared with each other, and the values of the MTF coefficients of the better ear are determined (*prominent-ear method*).
3. The binaural method follows the work by van Wijngaarden and Drullman (2008), which chooses the better MTF coefficients of both ears separately for each octave band (*binaural (band-by-band) method*).

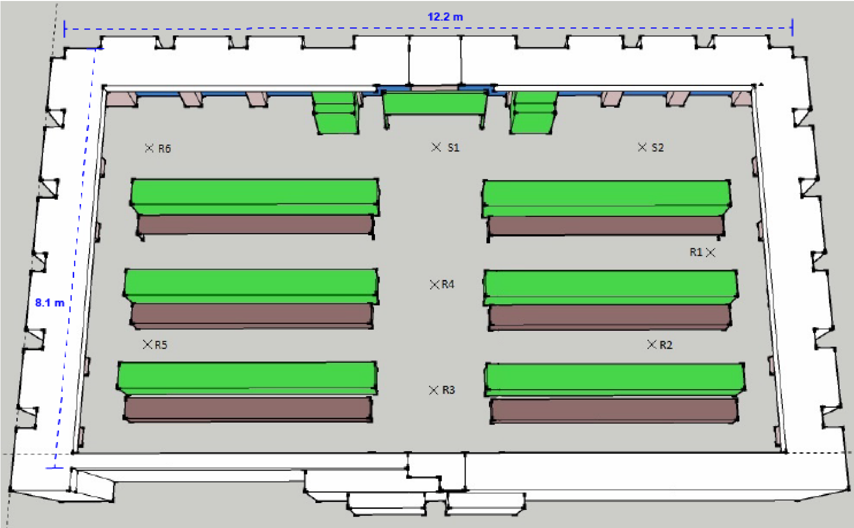
The work in this section explores the effects introduced by measurements using HATSs compared to a mono omni-directional microphone by evaluating the different methods for calculating the binaural STI and the corresponding MTIs. Differences were primarily expected within the MTIs when comparing the different receiver devices that might not be reflected in the final STI value and could be essential to account for children's hearing and speech intelligibility in complex acoustic environments. Additionally, the effect of the positioning of the different receiver devices was investigated to understand the corresponding impact on the MTIs and resulting STIs.

Measurement procedure

The measurements were conducted in the seminar room of the Institute for Hearing Technology and Acoustics (IHTA), RWTH Aachen University in Germany (cf. Loh et al., 2023a). This room was considered exemplary for classrooms and provided optimal conditions for controlled room acoustic measurements. The STI was measured following IEC 60268-16 (IEC, 2012) using the indirect method. Impulse responses were measured with two sender and six receiver positions defined according to DIN 3382-2 (ISO, 2008), resulting in twelve sender-receiver position combinations (cf. Figure 2.7). The measurement signal was a sine sweep covering the frequencies from 20 Hz to 20 kHz.

The measurement procedure is further described in Loh et al. (2023a) and Loh et al. (2023a):

Figure 2.7: Positions of senders (S) and receivers (R) (from Loh et al., 2023a).



As receiver devices, an omnidirectional microphone (B&K 1/2" Type 2669) as the reference mono microphone, the ITA adult HATS (Schmitz, 1995), and the [preschool] child HATS (Fels et al., 2004) were used. The child HATS was designed based on a range of statistically analyzed anthropometric sizes of children. [Compared to the adult HATS,] it has no elaborated ear shape. [... To maximize possible differences between the adult and the child HATS, the preschool child HATS was chosen over the primary school child-sized HATS as it was used in the aforementioned acoustic measurements. To put it into context, the] head size [of the preschool child HATS] is comparable to a [...] child at approx. three to six years old (Fels et al., 2004). Both HATSs' ear axes and the microphone were positioned at 1.15 m over the floor, representing a sitting adult's ear height. All HATSs were facing toward the sender, the omnidirectional dodecahedron loudspeaker (cf. Figure 2.8).
Furthermore, a singer HATS simulating a natural speaker was used to playback modulated pink noise for a direct measurement method for STI [as shown in Figure 2.8]. However, these results are not discussed in this work and are objects to future studies.

MATLAB and the ITA toolbox (Berzborn et al., 2017) were used for the measurement procedure and the following signal post-processing. The adult and child

Figure 2.8: Measurement setup in the seminar room (from Loh et al., 2023a).

Note. Omnidirectional dodecahedron source (front left), singer HATS source (front right), adult HATS receiver (center), and child HATS receiver (rear by the windows).

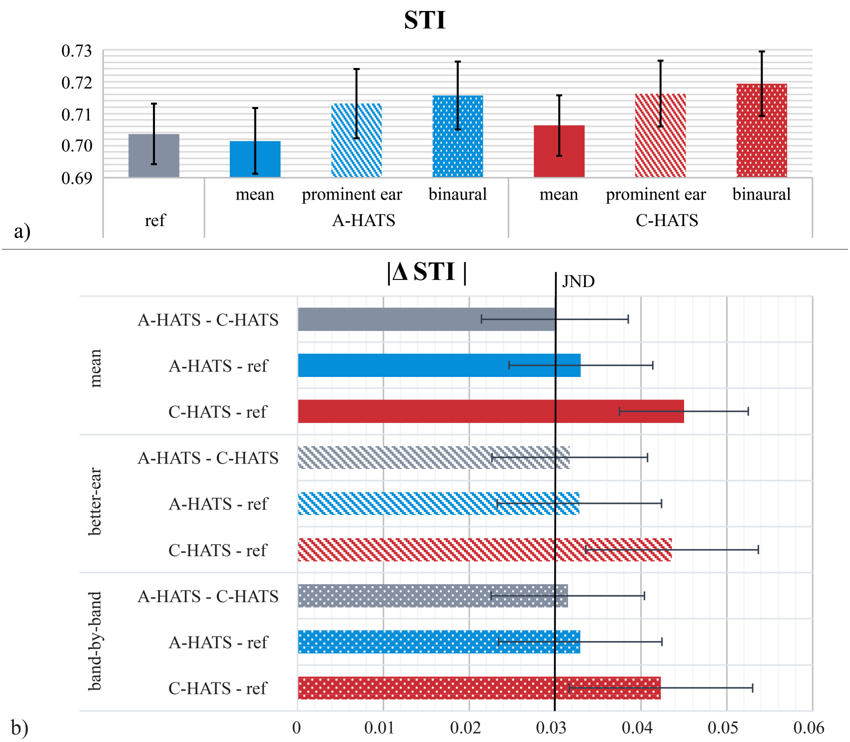
HATS measurements were equalized using their corresponding diffuse field filters to account for the directionally independent components of the measurements as described in Section 2.4.1.

Results and discussion

Starting with the overall STI (cf. Figure 2.9), measurement results were averaged over all twelve sender-receiver position combinations receiving one single value for the room after considering the binaural evaluation methods. STI results for all three measurement devices were generally observed between 0.701 and 0.719. Differences between measurement devices were found between 0.030 and 0.045. The lowest differences were yielded between adult and child HATS, while the highest differences were found between child HATS and the reference microphone. STI results evaluated using the prominent-ear and band-by-band methods were slightly higher for the band-by-band method. However, differences between measurement devices scaled linearly, resulting in the fact that both methods accounting for binaural hearing yielded similar differences between measurement devices.

modulation transfer index (MTI) results of each octave band are summarized in Table 2.12 showing the overall MTI averaged over all twelve sender-receiver-position combinations in the room. Differences of each octave band evaluated

Figure 2.9: Overview of average STI over $N = 12$ sender-receiver-positions and the differences of average STI between the receiver devices.



Note. |ΔSTI| = absolute difference in STI between receiver devices. A-HATS/C-HATS = adult/child HATS, ref = reference microphone. The labels *binaural* and *band-by-band* are used interchangeably for the same method.

Table 2.12: Overview of modulation transfer indices per octave band.

			MTI						
			125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
	ref	M	0.649	0.662	0.667	0.714	0.694	0.687	0.782
		SEM	0.011	0.013	0.011	0.009	0.012	0.011	0.010
mean	A-HATS	M	0.652	0.658	0.664	0.720	0.696	0.702	0.744
		SEM	0.011	0.010	0.011	0.011	0.011	0.014	0.010
	C-HATS	M	0.647	0.654	0.668	0.708	0.700	0.699	0.781
		SEM	0.011	0.013	0.010	0.009	0.010	0.011	0.011
prominent ear	A-HATS	M	0.655	0.676	0.673	0.730	0.707	0.714	0.758
		SEM	0.012	0.012	0.010	0.012	0.012	0.015	0.012
	C-HATS	M	0.649	0.661	0.677	0.718	0.709	0.710	0.794
		SEM	0.012	0.013	0.012	0.011	0.011	0.013	0.012
binaural (band-by-band)	A-HATS	M	0.663	0.680	0.679	0.734	0.709	0.715	0.758
		SEM	0.011	0.012	0.010	0.011	0.012	0.015	0.012
	C-HATS	M	0.655	0.668	0.682	0.722	0.711	0.712	0.795
		SEM	0.011	0.012	0.011	0.011	0.010	0.013	0.012

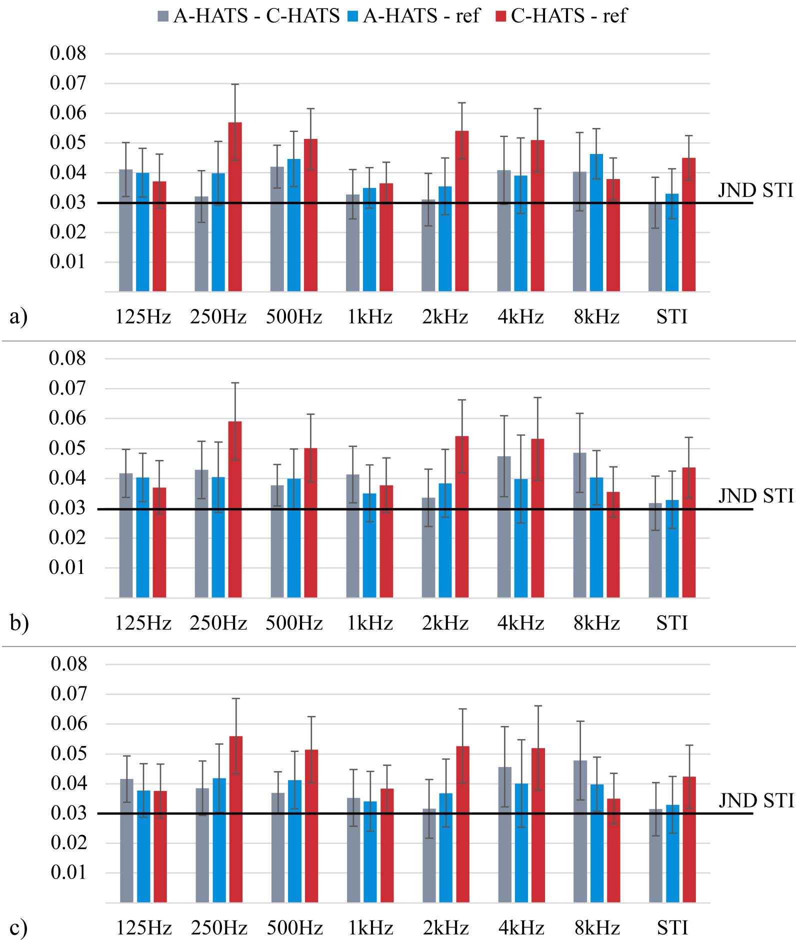
Note. A-HATS = adult HATS, C-HATS = child HATS, ref = reference microphone. Mean (M) and standard error of the mean (SEM) is calculated for $N = 12$ sender-receiver positions.

using all three methods accounting for binaural hearing abilities are shown in Figure 2.10 in direct comparison to the STI differences and the just-noticeable-difference (JND) of 0.03 as defined by Bradley et al. (1999). One-way ANOVA analyses for each evaluation method and each octave band yielded no significant effect of measurement devices (all $p > .05$), except in the mean method for the 8 kHz octave band ($F(2, 35) = 4.333, p = .021, \eta_p^2 = .208$). The post-hoc t-test yielded a significant difference between reference microphone and adult HATS ($p = .043, M_{\text{Diff}} = 0.038$) and a marginal difference between adult and child HATS ($p = .051, M_{\text{Diff}} = -0.037$) with slightly higher MTI for child HATS on the 8 kHz octave band. Though differences between measurement devices using all three binaural evaluation methods were statistically insignificant, it is noticeable that differences in all octave bands and the overall STI were above the JND of 0.03 (Bradley et al., 1999).
Results of this study underline previous findings by Fels et al. (2004) and Fels and Vorländer (2009), stating that the differences in anthropometric head size lead to differences in amplification on the individual octave bands. It was observed that differences in the MTIs of each octave band and all binaural evaluations were found above the JND. Especially the 8 kHz octave band was also found to be statistically different between measurement devices as observed by Fels and Vorländer (2009) stating differences due to anthropometric head sizes were primarily observed in higher frequencies. With this, it can be assumed that the STI as a single value might show differences between omni-directional microphones and HATS. However, it might also neglect individual differences in octave bands, specifically when comparing adult and child HATS. In terms of the binaural evaluation method, it was found that differences between the prominent-ear and band-by-band methods scaled linearly, resulting in similar differences between measurement devices no matter which evaluation method. Furthermore, differences in the MTI values were more specific within the octave bands when comparing the mean and the binaural methods. This observation suggests that choosing a binaural method over averaging left and right ear parameters might add insights into the speech intelligibility interpreted from children's perspectives.

However, results of this sudy must be interpreted with consideration as described in the work by Loh et al. (2023a):

> [...] Results of this study were considered preliminary since the MTIs and (binaural) STIs were derived following the indirect method for STI calculation. This method accounts for existing noises and directivity of the speaker only in a limited way, which might be essential for binaural effects (Lavandier & Best, 2020). In this case, each ear would be positioned slightly differently within

Figure 2.10: Absolute differences in MTI and STI between the receiver devices evaluated over $N = 12$ sender-receiver-positions considering the a) mean, b) prominent ear, and c) binaural (band-by-band) method.



> the directivity of a speaker, which might not be the case if an omnidirectional source like the dodecahedron is used. Furthermore, with regard to the work by [Liang and Yu (2020)], a strong dependency of the bSTI to the angle and positioning of the HATSs within the room and towards the sound source is expected. This effect was reduced in this study by common facing of the HATSs towards the omni-directional source, which explains the slight differences between the left and right ear, resulting in [...] differences between the [binaural STI] and the reference STI.

Additionally, it must be mentioned that the JND by Bradley et al. (1999) was obtained from studies conducted with adult participants. It is questionable whether this JND also holds for children's speech perception since it is unclear whether it is directly comparable to adults' speech perception (Loh et al., 2023a).

Concluding remarks

This study showed the importance of binaural measurement methods when assessing speech intelligibility within classrooms from children's perspectives. Though the STI obtained using HATS indicated how speech intelligibility between children and adults would differ, octave band-specific differences might not be reflected in the STI. The investigation of MTIs revealed that differences between the omnidirectional microphone and adult and child HATSs should not be neglected within the room acoustic assessment, especially regarding differences in the higher frequency bands of the HATSs. Furthermore, binaural evaluation methods can add insights toward child-appropriate speech intelligibility assessments besides averaging left and right ear parameters.

2.4.4 Concluding remarks on HATS differences

The findings of this work indicate the benefit of using HATS and adding binaural evaluations to noise assessment. It adds insight into the spectral differences in the measured sound signals, not only in the noise parameters but also in the room acoustic parameters. To appropriately add the perspective of children, the results of this work showed that using HATS with appropriate anthropometric sizes can reveal additional insights into the differences in room acoustic, level-based parameters, and psychoacoustic parameters that might be introduced by differences in anthropometric sizes, especially when addressing children's perspective in noise assessment in educational buildings.

2.5 Children's Noise Perception Survey

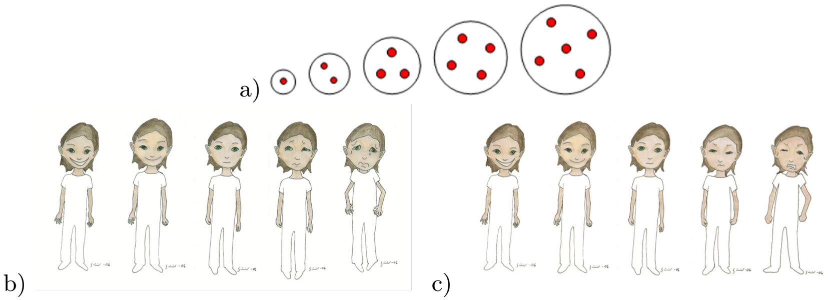
After analyzing the acoustic environments in pre- and primary schools based on objective values derived from physically measured signals with HATSs comprising anthropometric sizes of interest, it was apparent to compare these results with the subjective perspective of people who are daily exposed to these noisy acoustic environments.
This section aimed to assess children's and adults' noise perception using subjective methods, such as a questionnaire, and to evaluate whether the results can be directly linked to the objective measures from the previous sections (cf. Section 2.3 and Section 2.4).
In the first step, the results of the questionnaires of children and adults from preschool and primary schools were compared. From previous work by Persson Waye et al. (2013) is known that children are more sensitive to high-frequency sounds, and the objective of this work was to examine whether differences between adults and children existed, as well as between different educational institutions (preschool and primary schools) and whether there are linked age effects (older children versus younger children) to the responses. Furthermore, problems of the present study design were discussed when linking the results from subjective methods designed for children with objective measurement results.

2.5.1 Questionnaire and interview procedure

For this work, the Inventory of Noise and Children's Health (INCH) by Persson Waye et al. (2013) was chosen to measure adults' and children's sound perceptions. This questionnaire enables the assessment of children's reactions to indoor sounds at a very young age in a daycare center. This study evaluated only data regarding the frequency of perception and emotional responses (sadness and anger) in three different sound categories. Other information collected via the INCH questionnaire was not included in this work.

Participant's reactions to three sound categories, 1) sounds of angry, yelling children, 2) loud and strong sounds, and 3) scraping and screeching sounds, were answered using a five-point Likert scale. Therefore, the answer to a specific sound category is either positive or negative with respect to a neutral center. Visual graphic representations of this bipolar scale were used to allow direct identification and reflection of the younger participants on their perceptions and emotions. The participant's report on the frequency of a specific sound category was represented in circles, including one, *almost never*, to five dots, *very often* (cf. Figure 2.11.a), while the report on emotional reactions was visualized using bodily and facial expressions shown in a drawing of a person representing sadness on the scale *sad/afraid* to *glad/safe* (cf. Figure 2.11.b) and anger on the scale *angry/irritated* to *kind/friendly* (cf. Figure 2.11.c).

Figure 2.11: Excerpt from the INCH questionnaire (Persson Waye et al., 2013).



Note. a) Representation of frequency of occurrence. b) Representation of the emotional reaction *sadness*. c) Representation of the emotional reaction *angriness*.

The original questionnaire in Swedish and English was translated to German under the consultation of experienced educators and teachers with an additional adaptation of the language to be further suitable for children younger than four years old. Children were interviewed according to the questionnaire. The interviewers were instructed to follow the questionnaire's wording as closely as possible. Deviations were only allowed in case the children did not understand the question. Children were asked to point to the graphical representation of their reaction to the question, and the interviewer marked the answer. Children were interviewed individually in a separate room within their daily activities. The interview was executed in parallel to the in-situ acoustic measurements. Hence, it could be reasonably assumed that the children's answers were directly related to the measured sound environment.

The questionnaire was further adapted for adults, containing the same questions. However, coping methods and stress-related questions were asked more appropriately for adults than using figures. However, the questions on the reactions to the sound categories were retained using visual graphic representations to ensure direct comparability between children and adults. The adult questionnaires were handed out to the participants at the beginning of the acoustic measurements in the corresponding educational institutions and filled out by the adult participants.

The procedure was approved with the acoustic measurements within the study protocols EK 321/16 and EK 218/18 by the Medical Ethics Committee at RWTH Aachen University. Signed informed consent was collected from all involved educational institutions, teachers, and educators. In the case of children, informed consent was signed by their parents before the interview of their children (for more details, see Loh et al., 2022b).

Children and adults were recruited from the involved classes and groups and were present most of the time during the in-situ acoustic measurements while participating in normal daily activities. The adult participants were educators and teachers of the corresponding groups of children and were exposed to the same noise as the children. All adults worked in the investigated educational environments for at least six months. The INCH interview was part of the study, and it included acoustic measurements in pre- and primary schools. An overview of all participating institutions and respectively interview participants is shown in Figure 2.1 and in Table 2.13.

Table 2.13: Overview of INCH survey participants.

Age category	**Playrooms**		**Classrooms**	
	$N_{\mathbf{female}}$	$N_{\mathbf{fmale}}$	$N_{\mathbf{female}}$	$N_{\mathbf{fmale}}$
Children				
3 y/o	6	7	-	-
4 y/o	12	13	-	-
5 y/o	17	12	-	-
6 y/o	8	8	10	10
7 y/o	-	-	32	17
8 y/o	-	-	15	9
9 y/o	-	-	2	5
10 y/o	-	-	0	1
Adults				
$<$ 20 y/o	3	-	1	-
20-30 y/o	15	1	3	1
31-40 y/o	7	1	3	5
41-50 y/o	4	-	8	-
51-60 y/o	2	-	5	-

2.5.2 Differences in age group and room types

For the difference analyses, the five-point Likert scale was recoded into a binary score with responses on the negative side, including the neutral center (≤ 3) as 0 and answers on the positive side (> 3) as 1, achieving a binary representation of the answers. Descriptives of the INCH responses according to the binary score are summarized in Table 2.14.

When examining the dataset, it became obvious that it violated several assumptions when conducting a multi-factor ANOVA. However, a nested non-parametric

Table 2.14: Descriptives of the INCH survey (binary score).

A. Perceived frequency of occurrence

		Sound of angry, yelling children		Loud and strong sounds		Scraping and screeching sounds	
	N	% M	% SEM	% M	% SEM	% M	% SEM
Classrooms ($N = 8$)							
Adults	29	20.7	7.7	51.7	9.4	10.3	5.8
Children	101	46.5	5	62.4	4.8	10.9	3.1
Playrooms ($N = 10$)							
Adults	33	50	8.8	81.8	6.8	30.3	8.1
Children	83	54.2	5.5	65.1	5.3	21.7	4.6

Note. Percentage of responses answering 4 and 5, interpreted as "often" and "very often" in terms of frequency of perception.

B. Emotional reaction: sadness

		Sound of angry, yelling children		Loud and strong sounds		Scraping and screeching sounds	
	N	% M	% SEM	% M	% SEM	% M	% SEM
Classrooms ($N = 8$)							
Adults	29	10.3	5.8	6.9	4.8	25	8.2
Children	101	41.6	4.9	52.5	5	31.7	4.7
Playrooms ($N = 10$)							
Adults	33	15.6	6.4	28.1	8	19.4	7
Children	83	47	5.5	43.4	5.5	28.8	5

Note. Percentage of responses answering 4 and 5, interpreted as "much" and "very much" in terms of the emotional reaction sadness.

C. Emotional reaction: angriness

		Sound of angry, yelling children		Loud and strong sounds		Scraping and screeching sounds	
	N	% M	% SEM	% M	% SEM	% M	% SEM
Classrooms ($N = 8$)							
Adults	29	41.4	9.3	41.4	9.3	44.8	9.4
Children	101	52.5	5	50.5	5	34.7	4.8
Playrooms ($N = 10$)							
Adults	33	36.4	8.5	46.9	8.8	27.3	7.9
Children	83	40.7	5.4	46.8	5.5	28.4	5

Note. Percentage of responses answering 4 and 5, interpreted as "much" and "very much" in terms of the emotional reaction angriness.

test procedure using the Chi-square tests was chosen to investigate possible interaction effects.

Therefore, the dataset was split to conduct Chi-square analyses examining differences in the variable *age group* with two levels (adults vs. children) and the variable *room type* with two levels (classrooms vs. playrooms) sequentially. The analyses were conducted for the three response types (1. frequency of perception, 2. sad emotional reaction, and 3. angry emotional reaction) regarding all three noise types (1. sounds of angry, yelling children, 2. loud and strong sounds, and 3. scraping and screeching sounds).

Firstly, the dataset was split according to age groups, such as adults and children. The differences between the two room types, classrooms and playrooms, were then analyzed for all three response types (cf. Table 2.15 A). Significant differences between the two room types, classrooms and playrooms, were found within adults in terms of the perceived frequency of occurrence regarding the noise type *sounds of angry yelling children* and *loud and strong sounds* (both $p < .05$). Significant more adults in playrooms reported a higher frequency of occurrence of the two noise types than in classrooms. Within children, the perceived frequency of occurrence regarding the noise type *scraping and screeching sounds* differed significantly between the two room types ($p < .05$). In playrooms, considerably more children reported a higher frequency of occurrence regarding *scraping and screeching sounds*. Regarding the emotional reaction of sadness, a significant difference was observed within adults ($p < .05$) in terms of *loud and strong sounds*. More adults in playrooms reported feeling sad considering *loud and strong sounds* than in classrooms.

Secondly, the dataset was split according to the room types (classrooms and playrooms). Differences between the age groups, adults and children, were analyzed for all three response types (cf. Table 2.15 B). Significant differences between the age groups, adults and children, were found within classrooms in terms of the perceived frequency of occurrence regarding *sounds of angry yelling children* ($p < .05$). In classrooms, more children than adults reported to perceive sounds of angry yelling children (very) often. In terms of the emotional reaction of sadness, significant differences between adults and children were observed within classrooms regarding two noise types *sounds of angry yelling children* ($p < .01$) and *loud and strong sounds* ($p < .001$). For both noise types, significantly more children than adults in classrooms reported feeling (very) sad when perceiving these noises. However, in playrooms, differences in the emotional reaction of sadness between adults and children were only significant regarding the *sounds of angry yelling children* ($p < .01$). In playrooms, significantly more children than adults reported feeling (very) sad when hearing the sound of angry, yelling children.

Table 2.15: Differences in INCH responses between room types and age groups.

A. Difference between room type (classrooms and playrooms)

	Sounds of angry yelling children				**Loud and strong sounds**				**Scraping and screeching sounds**			
within	χ^2	*df*	*N*	*p*	χ^2	*df*	*N*	*p*	χ^2	*df*	*N*	*p*
Perceived frequency of occurrence												
Adults	**5.668**	**1**	**61**	**.017**	**6.397**	**1**	**62**	**.011**	3.710	1	62	.054
Children	1.076	1	184	.300	0.142	1	184	.707	**3.999**	**1**	**184**	**.046**
Emotional reaction sadness												
Adults	0.372	1	61	.542	**4.638**	**1**	**61**	**.031**	0.273	1	59	.601
Children	0.54	1	184	.463	1.511	1	184	.219	0.182	1	181	.670
Emotional reaction angriness												
Adults	0.164	1	62	.686	0.186	1	61	.666	2.078	1	62	.149
Children	2.483	1	182	.115	0.238	1	180	.626	0.811	1	182	.368

Note. Significance level at $p < .05$. Significant effects are indicated in bold.

B. Difference between age groups (adults and children)

	Sounds of angry yelling children				**Loud and strong sounds**				**Scraping and screeching sounds**			
within	χ^2	*df*	*N*	*p*	χ^2	*df*	*N*	*p*	χ^2	*df*	*N*	*p*
Perceived frequency of occurrence												
Classrooms	**6.232**	**1**	**130**	**.013**	1.065	1	130	.302	0.007	1	130	.933
Playrooms	0.165	1	115	.685	3.147	1	116	.076	0.957	1	116	.328
Emotional reaction sadness												
Classrooms	**9.715**	**1**	**130**	**.002**	**19.176**	**1**	**130**	**<.001**	0.464	1	129	.496
Playrooms	**9.617**	**1**	**115**	**.002**	2.255	1	115	.133	1.022	1	111	.312
Emotional reaction angriness												
Classrooms	1.11	1	130	.292	0.750	1	130	.387	1.001	1	130	.317
Playrooms	0.188	1	114	.665	0.000	1	111	.997	0.015	1	114	.904

Note. Significance level at $p < .05$. Significant effects are indicated in bold.

Regarding the INCH responses and the perceived frequency of different sounds, significant differences in age groups were found for the sounds of angry and yelling children within classrooms. Here, more children ($M = 46.5\%$) than adults ($M = 20.7\%$) reported a more frequent occurrence of the sounds of angry and yelling children. This result indicates slightly higher attention toward these sounds for children than adults regarding the sounds originating directly from children. In case that *yelling children* can be considered high-frequency sounds, this result would support the insights from the work by Persson Waye et al. (2013) states that children are more sensitive to high-frequency sounds. However, the significant difference in age group considering the sad emotional reaction to the sounds of *angry, yelling children* in both room types, classrooms and playrooms, as well as to *strong and loud sounds* in classrooms, indicated more children (more than $M = 41.6\%$) than adults (less than $M = 15.6\%$) to report emotional reaction. This result suggests that children may experience or express a stronger emotional response towards these sounds as compared to adults, in addition to the fact that they perceive these sounds more frequently than adults. However, this result should be treated cautiously since the participant group of adults ($N = 62$) is significantly smaller than that of children ($N = 134$).

Differences between the room types, classrooms and playrooms, were found for the perceived frequency of two sound types *angry, yelling children* (CR: $M = 20\%$ vs. PR: $M = 50\%$) and *loud and strong sounds* (CR: $M = 51.7\%$ vs. PR: $M = 81.8\%$) within adults, where this was significantly more reported in playrooms. Regarding children, a difference between the room types, classrooms and playrooms, was found for the frequency of perception of *scraping and screeching sounds.* This result provided evidence that at least a higher sensitivity to the different sound types existed in playrooms than in classrooms or that preschool-age children are more sensitive to these presumably high-frequency sounds. This is in accordance with the results by Fels et al. (2004). They found that a higher amplification of high-frequency sounds exists due to the head-related transfer functions introduced by children's smaller anthropometric size than adults. The finding also supports the results from an acoustic intervention study previously performed among preschools in the work by Persson Waye et al. (2016) using the same INCH questionnaire among preschool-aged children. A significantly reduced perception of *scraping and screeching sounds* was found in relation to the acoustic intervention, indicating a higher sensitivity to these types of sounds. Furthermore, the reduction of children's perception of scraping and screeching sounds with the intervention was significantly associated with an intensity reduction of the emotional reactions to these sounds, particularly to the angry response. With this, it can be assumed that the sensitivity is coupled with higher awareness and attention to the high-frequency sounds and a higher annoyance in children.

However, there is a major limitation to this investigation, which is due to the small sample sizes regarding the different room types to compare. To link results between existing acoustic environments and subjective perception, only the main rooms of activities where the people spend most of their time, i.e., the classrooms ($N = 8$) and the playrooms ($N = 10$), could be reasonably taken into account. Representative values for each age group and each room had to be calculated, leading to very small and varying sample sizes with, on average, 22 children and one adult in classrooms and 16 children and two adults in playrooms.

2.5.3 Correlation of subjective responses and objective measurements

For the correlation analyses, the five-point Likert scale was examined based on the values evaluated on a metric scale by recoding the responses from 1 to 5 corresponding to the graphical representation of the frequency and emotional reaction as shown in Figure 2.11 from right to left, allowing the interpretation of the value towards increasing annoyance with higher values achieved on the metric scale. Descriptives of the INCH survey based on the metric scale are summarized in Table 2.16.

As objective measurement results, only the noise parameters from the in-situ measurements were considered (level-based metrics and psychoacoustic parameters, as presented in Table 2.4). Results of every subjective response variable and objective parameter were averaged per room to link single values of each room's subjective and objective parameters.

Pearson's correlation coefficients and corresponding significance indicators were calculated to examine possible relations between subjective responses and objective measurement results. Where the assumption could not be met, the non-parametric alternative Spearman's correlation coefficient was evaluated instead. Results for adults were investigated separately from children's responses and were further only correlated with results measured using the reference microphone (omni-directional) and the adult HATS. The respective evaluation was carried out for children but using the child HATS instead. Significant correlation results are summarized in Table 2.17. Any other correlations between noise parameters and INCH responses were insignificant, $p > .05$, and for clarity, not listed individually.

The findings for adults, while preliminary, suggest that there is a connection between the frequency of noise perception and the noise parameter sharpness (mean and percentiles), independent of the measurement device (reference mi-

Table 2.16: Descriptives of the INCH survey (metric scale).

A. Perceived frequency of occurrence

		Sound of angry, yelling children		**Loud and strong sounds**		**Scraping and screeching sounds**	
	N	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*
Classrooms ($N = 8$)							
Adults	29	2.72	0.19	3.45	0.17	2.28	0.18
Children	101	3.45	0.13	3.60	0.13	1.85	0.12
Playrooms ($N = 10$)							
Adults	33	3.63	0.14	4.24	0.13	2.91	0.18
Children	83	3.42	0.19	3.78	0.15	2.25	0.14

Note. Responses were evaluated based on a metric scale.

B. Emotional reaction: sadness

		Sound of angry, yelling children		**Loud and strong sounds**		**Scraping and screeching sounds**	
	N	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*
Classrooms ($N = 8$)							
Adults	29	2.90	0.11	2.79	0.10	3.11	0.16
Children	101	3.53	0.08	3.54	0.08	3.24	0.08
Playrooms ($N = 10$)							
Adults	33	3.09	0.08	3.16	0.13	3.16	0.08
Children	83	3.28	0.16	3.16	0.13	2.83	0.14

Note. PResponses were evaluated based on a metric scale.

C. Emotional reaction: angriness

		Sound of angry, yelling children		**Loud and strong sounds**		**Scraping and screeching sounds**	
	N	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*
Classrooms ($N = 8$)							
Adults	29	3.41	0.15	3.34	0.15	3.34	0.16
Children	101	3.65	0.08	3.59	0.08	3.22	0.09
Playrooms ($N = 10$)							
Adults	33	3.33	0.11	3.47	0.14	3.30	0.11
Children	83	3.33	0.15	3.20	0.14	3.00	0.13

Table 2.17: Significant correlations, denoted in r $(p - value)$, between INCH responses and noise parameters.

Adults				
ref				
	$S_{\mathbf{5}}$			
Frequency - N2	-0.553 (.017)			
Frequency - N3	-0.485 (.041)			
A-HATS$_{\mathbf{av}}$				
	$S_{\mathbf{mean}}$	$S_{\mathbf{5}}$	$S_{\mathbf{90}}$	
Frequency - N1	-0.570 (.013)	-0.602 (.008)		
Frequency - N2	-0.639 (.004)	-0.727 (.001)	-0.469 (.050)	
A-HATS$_{\mathbf{prom}}$				
	$S_{\mathbf{mean}}$	$S_{\mathbf{5}}$		
Frequency - N1	-0.480 (.044)	-0.485 (.042)		
Frequency - N2	-0.578 (.012)	-0.606 (.008)		
	$L_{\mathbf{Z,eq}}$			
Sadness - N3	-0.500 (.035)			
	$L_{\mathbf{Z,eq}}$			
Angriness - N1	-0.582[a] (.011)			
Children				
ref				
	$N_{\mathbf{90}}$	$S_{\mathbf{5}}$	$R_{\mathbf{mean}}$	$R_{\mathbf{90}}$
Frequency - N1	0.522 (.026)			
Sadness - N2		0.480 (.044)		-0.484 (.042)
Angriness - N2			-0.499 (.035)	-0.491 (.038)
C-HATS$_{\mathbf{av}}$				
	$S_{\mathbf{5}}$	$S_{\mathbf{90}}$	$R_{\mathbf{mean}}$	
Sadness - N2	0.529 (.024)			
Sadness - N3		0.541 (.020)		
Angriness - N2			-0.495 (.037)	
C-HATS$_{\mathbf{prom}}$				
	$S_{\mathbf{90}}$	$R_{\mathbf{mean}}$	$FS_{\mathbf{mean}}$	$FS_{\mathbf{5}}$
Sadness - N3	0.513 (.029)			
Angriness - N2		-0.513 (.029)	-0.474 (.047)	-0.533 (.023)
	$L_{\mathbf{Z,eq}}$**(L.v.H)**	$L_{\mathbf{A,eq}}$		
Frequency - N2		-0.515 (.029)		
Sadness - N2		-0.509 (.031)		
Angriness - N1	0.469 (.050)			

Note. [a] Spearman correlation coefficient, when assumption ist violated. N1 = Sound of angry, yelling children; N2 = Loud and strong sounds; N3 = Scraping and screeching sounds.

crophone or adult HATS) and the binaural evaluation method Furthermore, the $L_{\mathrm{Z,eq}}$ showed significant correlation with the emotional reaction of sadness (in terms of scraping and screeching sounds) and angriness (in terms of sounds of angry, yelling children).

These findings must be treated with consideration. The sample sizes of adults participating in the questionnaire per room did not exceed five people per room, which can not be representative for adults in general. However, the educational situation did not allow for a larger sample size since there would always be fewer supervisors than children per group. In Germany, the standard supervisor-to-child ratio is always 1:7.5, so there will always be a small sample size of adult participants exposed to the same sound environment as the children.

In terms of the children, results suggest a correlation between the psychoacoustic parameters and mainly the emotional reaction of sadness and angriness towards the loud and strong sounds as well as the scraping and screeching sounds. While the parameters loudness and sharpness indicated a positive correlation (i.e., the higher the psychoacoustic parameter was measured, the higher the annoyance), a negative correlation was found for roughness and fluctuation strengths. These results led to the assumption that the loud and high-frequency components result in higher annoyance, which aligns with previous findings by Persson Waye and Karlberg (2021). This finding suggests that psychoacoustic parameters were better suited to reflect noise perception, especially regarding child-appropriate evaluation methods. However, there is no specific trend toward an omni-directional measurement device or a child HATS.

Level-based parameters only yielded significant results in terms of $L_{\mathrm{A,eq}}$ and the low-versus-high-frequency ratio of $L_{\mathrm{Z,eq}}$ when measured using the child HATS and evaluated using the prominent-ear method. This observation underlined the findings of the previous sections that choosing appropriate measurement methods is important when addressing children's perspectives on noise and using common parameters, such as level-based parameters.

This investigation is, however, limited by the small sample size of $N = 18$ rooms in educational buildings, all chosen within one city in Germany. Therefore, $N = 18$ rooms cannot be considered a representative sample size, especially under the assumption of neglecting possible age effects since classrooms and playrooms were analyzed together. Furthermore, only one single value of the room acoustic and noise parameters can be applied for this type of correlation analysis. As findings from previous sections suggest, frequency-dependent analysis was not integrated into this study and is subject to future studies.

2.5.4 Concluding remarks

The INCH questionnaire revealed a higher sensitivity in terms of awareness and possible annoyance to high-frequency sounds in children than in adults. The high-frequency sounds in this work are reflected in the *yelling children* and the *scraping and screeching sounds*. Interestingly, this finding was different among the room types, classrooms and playrooms, indicating either a different sensitivity of the children to the sound environments existing in the room or different sound environments dominating the two room types. Furthermore, results indicated a significant correlation between psychoacoustic parameters and the INCH response, pointing towards the fact that psychoacoustic parameters are better suited to represent noise perception of children than level-based parameters.

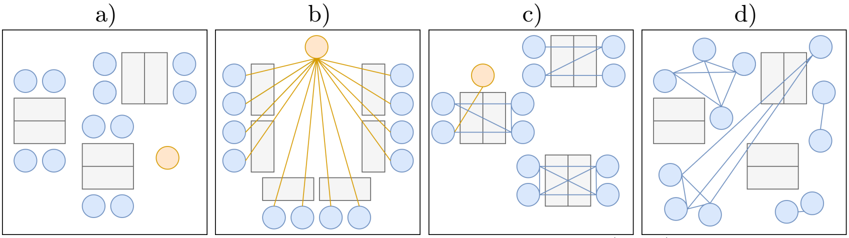
2.6 Activity-based Acoustic Settings (ABAS)

Concerning long-term acoustic measurements, it was standing to reason that the noise and corresponding noise levels would depend on the predominant activity inside a room. Previous work (e.g., Shield et al., 2015) investigated this fact in relation to classroom activities and examined the corresponding noise levels. The activities during lessons were classified into four categories: 1. plenary, 2. group work, 3. individual work, and 4. watching/listening. This classification was based on standard class activities typically applied in an educational context. Analyses on noise levels revealed that group work yielded the highest noise levels, while little difference was found between plenary and individual work.

The goal was to link activity components to specific acoustic properties by creating a more systematic description of activities that includes the acoustic perspectives and goes beyond the educational context. Therefore, the work presented in this section focused on the people's interaction and communication within the activities. Based on this systematic description of activities, the acoustic properties in terms of two level-based and two psychoacoustic noise parameters were examined to gain insights on comparing in-situ measurements using HATS with different anthropometric head sizes.

Description of activity patterns, further denoted as activity-based acoustic setting (ABAS), in classrooms were generated with respect to the specifications of the teacher-pupil interaction and communication pathways, whereas communication pathways were investigated without specific directions in a first step to obtain general insights. With this, four types of activities commonly in primary schools (cf. Figure 2.12) were identified by Loh and Fels (2021):

Figure 2.12: Activity-based acoustic settings in primary schools (6-10 y/o).



Note. Communication pathways of the speaking person (lines) between teacher (orange) and children (blue) in primary schools. a) silent work, b) frontal teaching, c) group work, and d) breaks. Figure adapted from Loh and Fels (2021).

- Frontal teaching - The teacher guides and controls activity and communication. Communication exists mainly between the teacher and an individual child.
- Group work – The teacher supervises but does not control activity and communication. Children are working in groups. The teacher remains in the rooms and participates in discussions within individual groups or moves from group to group. Communication is limited within groups.
- Silent work – Children work individually, and the teacher actively supervises the situation. Communication is unwanted and, thus, almost non-existent. The teacher controls the amount of accepted communication.
- Breaks – The teacher rarely interfered in the children's activities. Communication is not structured and is generally not limited in frequency, interference, or space.

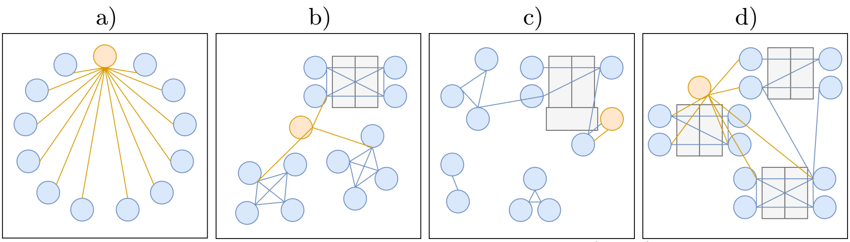
Younger age groups have been rarely examined in previous literature. Similar to the classroom activity-based acoustic patterns, the preschool focus was on the educator-children interaction and corresponding communication pathways. The concept by Loh et al. (2023b) included four main activities for preschools (cf. Figure 2.13):

- Free playtime – Children are free to choose their activity. Communication is limited within groups of interacting children. There is little interaction between groups since the subject of attention differs in each group. The educator does not explicitly control or interfere with the situation.
- Guided playtime – Children are split into small groups with specific instructions on how to interact. The educator controls and supervises the

situation by speaking to each group individually. Communication between children is limited to the groups. Communication between groups is almost non-existent.

- Group circle – All people are sitting in a circle. The educator guides and controls the situation, hands over the right to speak, and no more than one speaks simultaneously.
- Mealtime – Communication is not controlled. The educator takes turns speaking to every group individually and giving instructions to the whole group. Children are chatting uncontrolled within and between groups. Communication is not limited.

Figure 2.13: Activity-based acoustic settings in preschools (3-6 y/o).



Note. Communication pathways of the speaking person (lines) between educators (orange) and children (blue) in preschools. a) group circle, b) guided playtime, c) free playtime, and d) mealtime. Figure adapted from Loh et al. (2023b).

2.6.1 Acoustic properties of ABAS

For the acoustic examination, eight classrooms from four primary schools and four playrooms from three preschools were chosen for a detailed analysis based on the four exemplary activities existing in the range of daily activities in primary schools and preschools, respectively. The measurements were taken from the acoustic inventory described in Section 2.3, where a third-party person was present during the whole measurement period. Continuous monitoring was ensured, and detailed activity protocols were recorded. Specific details on the measurement procedure, ethical approval, data post-processing, and evaluation methods were reported in the work by Loh et al. (2022b). For this study, only periods with children in the room were evaluated. All data were post-processed using MATLAB and the ITA toolbox (Berzborn et al., 2017) to extract the respective time frames of each activity according to the activity protocol.

Unweighted sound pressure level ($L_{Z,eq}$), A-weighted sound pressure level ($L_{A,eq}$), loudness (N_{mean}), and sharpness (S_{mean}) were examined in this context. All parameters were computed using ArtemiS SUITE 14.3 by HEAD Acoustics (Herzogenrath, German) as described in Section 2.3.2. All time frames according to the predefined activity were averaged to achieve a single value.

Binaural effects were included to account for a more natural hearing and sound perception, as described in Section 2.3.2 (In-situ noise measurements) by considering the independent from direction components and calculating binaural parameters for the adult and child HATSs (A-HATS$_{av}$, C-HATS$_{av}$, A-HATS$_{prom}$, and C-HATS$_{prom}$). For brevity and clarity of this thesis, only the results of the average method were presented in detail, and the results of the prominent-ear method were added to the appendix (cf. Appendix A.4) since both results did not differ significantly. Significant differences were observed regarding the parameter sharpness obtained from the child HATS, where the results tended to be higher in the prominent-ear method than in the average method.

Preliminary results were partly presented at theInternational Commission on Biological Effects of Noise (ICBEN) 2021 (Loh & Fels, 2021) and published in a conference paper at Inter-noise 2023 (Loh et al., 2023b).

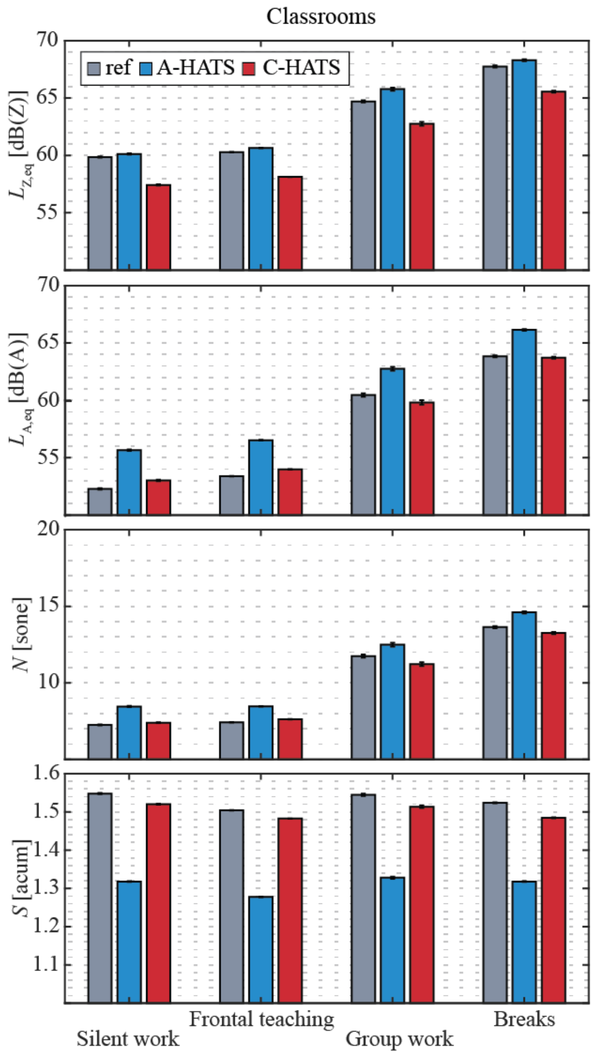
2.6.2 Classroom ABAS

Concerning the average method, the lowest sound pressure levels (both A- and Z-weighted) were found for frontal teaching and silent work, while breaks yielded the highest SPLs (cf. Figure 2.14). Similar results were found for the psychoacoustic noise parameter loudness.

It is noticeable that the A-HATS$_{av}$ yielded the highest SPL, especially in A-weighted SPLs, and loudness values compared to the other receiver devices. In sharpness, though, values of the A-HATS$_{av}$ were lower than those of the C-HATS$_{av}$ and the reference microphone. Values obtained from the reference microphone and the C-HATS$_{av}$ were comparable for A-weighted SPLs and loudness. In contrast, in Z-weighted SPLs and sharpness, the C-HATS$_{av}$ revealed lower values than the reference microphone.

In general, it can be stated that the results yielded different values for each noise parameter concerning the different ABAS. However, the pattern comparing the three receiver devices remained similar across all four ABAS in classrooms with the specific observation that differences between reference microphone and C-HATS$_{av}$ were neglectable.

Figure 2.14: Primary school ABAS: noise parameters (average method).



Note. Comparing results of the four activities, each obtained from the reference microphone (Ref), adult (A-HATS) and child head and torso simulator (C-HATS) evaluated using the average method, respectively.

2.6.3 Playroom ABAS

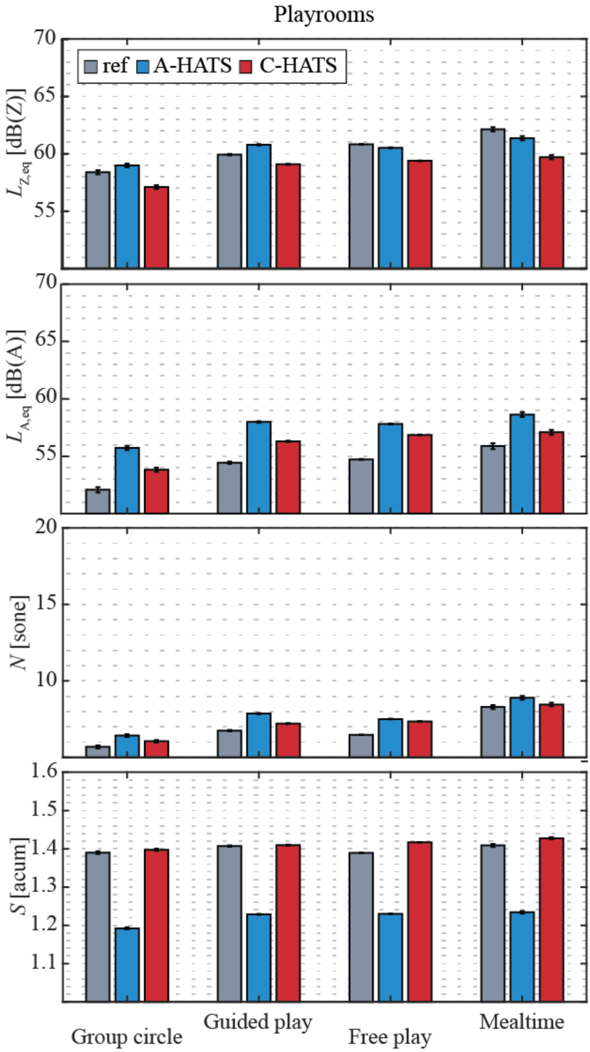
Comparing noise parameters between preschool and primary school activities, it was observable that primary school activities yielded generally higher (A- and Z-weighted) SPL, loudness, and sharpness values independent from the type of activity (cf. Figure 2.15). Values of the activities group work and breaks additionally exceeded 62 dB[Z], 60 dB[A], and 10 sone, while in preschools, these values were not exceeded in any activities. Overall, differences in results between the measurement devices from the primary schools were more pronounced than in preschools. However, variation across activities was smaller in preschools than in classrooms.

Overall, the lowest noise levels were observed for the activity group circle, while the highest were found for mealtime independent from the receiver device. The activities of free play and guided play showed similar results.
As observed for classrooms, the A-HATS$_{av}$ yielded the highest values in (A- and Z-weighted) SPLs and loudness as well as the lowest sharpness values when comparing the three receiver devices. At the same time, results from the reference microphone and the C-HATS$_{av}$ were similar to each other. Also here, the pattern of the relation between the three receiver devices remained comparable across activities with one exception in Z-weighted SPL: towards higher Z-weighted SPLs, the pattern with highest values for the A-HATS$_{av}$ changed to highest values for the reference microphone, while the relative value of C-HATS$_{av}$ remained lowest in all for activities.

2.6.4 Discussions on ABAS

For both educational institutions, differences in all four examined noise parameters were found between the four categories of activity-based acoustic settings defined by the adult-children interaction and communication pathways. As previously assumed, the most controlled and supervised situation with little communication paths (as given in group circles in preschools or silent work in primary schools) yielded the lowest measured noise levels in (A- and Z-weighted) SPLs and loudness. Activity settings with a high degree of freedom in communication, including many interaction and communication pathways, revealed the highest noise levels in (A- and Z-weighted) SPLs and loudness. As loudness is a psychoacoustic parameter that was developed to give insights into the pleasantness and annoyance of an acoustic situation, this might explain a more annoying perception of the acoustic situation with regard to higher loudness values. These results were in line with the observation from previous work (cf. Berggren et al., 2008; Picard & Boudreau, 1999; Shield et al., 2015) and added perspectives from a psychoacoustic point of

Figure 2.15: Preschool ABAS: noise parameters (average method).



Note. Comparing results of the four activities, each obtained from the reference microphone (Ref), adult (A-HATS) and child head and torso simulator (C-HATS) evaluated using the average method.

view. However, the psychoacoustic parameter sharpness did not add information on the differences between the four activity settings in both pre- and primary schools. With this observation, it can be assumed that high-frequency sounds contributed less to the overall activity-based acoustic situations. However, these results are preliminary and must be validated with a subjective evaluation of the individual acoustic settings compared to each other.

With respect to differences between the three receiver devices, preliminary results by Loh et al. (2023b) were validated in this work:

> [Generally] differences between the three measurement devices were found in the results[, leading] to the assumption that parameters obtained using the head and torso simulators can provide additional insights into the natural human sound perception, especially evaluated using the prominent ear method commonly suggested by the research community (Bronkhorst, 2015; van Wijngaarden & Drullman, 2008). Differences were observed compared to the reference microphone [as well as] between the adult and child head and torso simulators, which can be interpreted that different anthropometric sizes between adults and children can lead to different sound perceptions, reflected in all [four] parameters examined in this work aligning to the previous work by Fels et al. (2004) and Fels and Vorländer (2009). [It] shows the importance of considering binaural evaluation methods and measured signals with regard to the individual anthropometric sizes when assessing sound perception from children's perspectives.
>
> Results from this study are considered preliminary and require further investigations into binaural evaluation methods and long-term evaluation parameters. There is little knowledge of how sound is perceived over a long time with children. It is unclear whether children are affected by long-term exposure to the same extent as adults. Furthermore, current psychoacoustic parameters were developed based on studies with adults. It is unknown whether these observations can be directly translated into children's perceptions.

2.6.5 Concluding remarks on ABAS

Overall, we introduced two concepts of activity-based acoustic settings based on the adults-children interactions and corresponding communication pathways with regard to the predominant activities, one for classroom activities in primary schools, i.e., regarding children from six to ten years old, and one for playroom activities in preschools, i.e., children between three to six years old.
Analyses of the acoustic properties of ABAS confirmed assumptions that controlled interaction and communication situations with fewer communication path-

ways led to lower noise levels, reflected in (A- and Z-weighted) SPL and loudness values (cf. Loh et al., 2023b). Additionally, it was observed that the impact of high-frequency sounds in educational situations had less impact than expected since children's voices were considered to contain more high-frequency components than other noise sources in the educational context. Finally, results from this work indicated the importance of integrating binaural measurement and evaluation methods into sound perception assessment using HATS to reflect children's noise perception more appropriately. Binaural methods can add insights to the commonly applied procedures using mono and omni-directional microphones.

2.7 Insights on Child-Appropriate Noise Assessment

This dissertation compared room acoustic and in-situ long-term measurements using head and torso simulators with different anthropometric sizes to investigate possible effects on children's noise perception. In general, it was found that overall room acoustics in German classrooms and playrooms could be improved to meet guidelines for 'good acoustics' by introducing acoustic treatments for high-frequency sound absorption and controlling reverberation times and speech intelligibility. Regarding long-term measurements, it was observed that psychoacoustic parameters can complement insights derived from SPL-based parameters. Additionally, differences in sound measurements using head and torso simulators with adults' compared to children's anthropometric size indicated some benefit in characterizing in-situ noise measurements, especially concerning higher frequency reflected in some psychoacoustic parameters, e.g., sharpness.

These differences in the higher frequency spectrum were specifically observed when examining the room acoustic parameters and in the long-term in-situ measurements with respect to the frequency bands between 250 Hz and 16 kHz. Within these investigations, it was found that specific room acoustic parameters, such as C_{50}, D_{50}, and T_{S}, were more sensitive to differences due to the anthropometric sizes of head and torso simulators than other parameters, such as the reverberation times T_{20}, T_{30}, and EDT. Additionally, findings from detailed examinations of the STI and IACC revealed that single-value parameters might mask potential differences in the higher frequency spectrum when measuring using head and torso simulators. This is especially essential when interpreting results from children's perspective, where differences due to the different anthropometric sizes compared to adults were primarily observed in the higher frequency range. Though STI is conceptualized as a single-value and binaural version of the STI is capable of indicating differences between adult and child HATSs, it was found beneficial to investigate the frequency bands by examining the modulation transfer indices (MTIs), which can additionally reflect differences across the whole frequency bands that were introduced by the differences between the HATSs. These findings were underlined by investigating the long-term measurements concerning the psychoacoustic and SPL-based parameters, taking the different frequency bands into account. Binaural versions of the psychoacoustic parameters added insights to the interpretation of noise perceptions of children compared to adults. The frequency response curves of the long-term measurements obtained from adult and child HATSs and the omnidirectional reference microphone explicitly revealed the differences in the higher frequencies, which explained the differences in the psychoacoustic parameters. The benefit of noise assessment using psychoacoustic parameters over level-based parameters was further underlined by investigating the correlation of subjective noise ratings using the INCH questionnaire and the noise parameters studied in this dissertation. Results revealed more correlations between psychoacoustic parameters and subjective noise ratings than level-based parameters.

The hypothesis that educational activities in pre- and primary schools can significantly influence noise levels was examined in detail within this dissertation by introducing concepts describing activity-based settings from an acoustic point of view. The activities were investigated, focusing on adult-children interaction and communication pathways, which introduced a systematic way to compare activities besides the common way of classifying activities in the educational context. With this, measurement results confirmed assumptions that activities with controlled and less communication pathways led to lower noise levels regarding the level-based parameters and the psychoacoustic parameter loudness. However, differences in high frequency were found to have less impact when examining noise levels in terms of activities, though differences between adult and child HATSs were still observed and can add insights on interpreting the noise perception of children in comparison to adults.

However, results from this dissertation are considered preliminary and require further investigations on children's perceptual evaluations of room acoustic and noise parameters. All acoustic parameters examined in this dissertation were developed from an adult point of view. It is unclear whether children would reveal similar perception ratings regarding the room acoustic and noise parameters when the same analyses were conducted with children participants, and thus, whether objective noise assessment using current room acoustic, level-based, and psychoacoustic parameters is valid for children. This dissertation mainly revealed differences in the physically measured signals when including head and torso simulators with different anthropometric sizes in the measurement procedures and that these can be reflected in the acoustic parameters under specific conditions.

3 Children's Auditory Cognition in Noise

This chapter presents the development of a child-appropriate paradigm on intentional auditory selective attention switching for children three to ten years old. Three studies were conducted to 1) validate the newly developed paradigm, 2) to investigate age effects for young children and at which point the paradigm can reliably examine intentional switching of auditory selective attention and when this cognitive function could be considered fully developed, and 3) to examine health effects reflected in heart rate variability changes when exposed to noise in an auditory cognition paradigm.

Parts of this chapter have been published in Loh et al. (2022a), have been submitted to *Scientific report by Naure* (Loh et al., 2024) and was presented at *Euronoise 2021 in Madeira* including a contribution to the proceedings (Loh et al., 2021).

3.1 State of the Art

Parts of the following literature research have been published previously in Loh et al. (2021) and Loh et al. (2022a), and was submitted to Loh et al. (2024) for review. For clarity and completeness, the state of the art has been summarized and limited to the context of the dissertation.

Intentional Switch of Auditory Selective Attention

For this dissertation, a cognitive function was chosen that is part of the auditory abilities within the cocktail party effect. The cognitive control of intentional switching of auditory selective attention in spatial acoustic setups and complex acoustic environments represented a good choice and has been extensively examined in previous studies conducted with adult participants (Koch et al., 2011; Lawo et al., 2014; Nolden et al., 2019) as summarized by Loh et al. (2022a):

> The authors measured reaction times and error rates to analyze attention switch costs and information processing interference. They used a paradigm

comprising two auditory stimuli (numbers from one to nine without five) that were simultaneously presented to each ear. One of them was indicated as the target stimulus using a preceding cue. The participant's task was then to solve the categorization task (smaller or larger than five) by selectively attending to the target stimulus. With this paradigm, effects due to attention switch were observed when the target's position changed from left to right ear or vice versa in contrast to repeating the target's position. The ability to select relevant information is reflected in the congruency effect. In one trial, stimuli presented by both target and distractor could be either from the same (congruent) or from different categories (incongruent). The congruency effect is thus defined as the differences in responses between incongruent and congruent trials.

Results from these investigations lead to two main findings: On the one hand, auditory attention switches lead to higher reaction time (RT) and worse performance in error rate (compared to repetitions of the attention focus). The so-called attention switch costs are interpreted as a cost to resolve interferences caused by switches of attention. On the other hand, a consistent congruency effect in error rates indicates delayed filter choices during attention switches, which results in worse performance when target and distractor stimuli are from different categories.

Children's Auditory Selective Attention

Since children cannot be considered small adults, it is necessary to understand the differences between adults and children. Loh et al. (2022a) summarized findings on children's auditory selective attention as follows:

Children's selective attention was broadly examined in previous research (Doyle, 1973; Huang-Pollock et al., 2002; Röer et al., 2018). Early theories of selective attention(e.g., Broadbent, 2013) state that incoming information is filtered by basic features before processing semantic content. However, further studies added insights to the early filter theory (Deutsch & Deutsch, 1963; Treisman, 1969) and led to the challenge to better understand the interplay of selective attention and external influences. Different paradigms were therefore introduced to examine abilities to select relevant information and attention control by blocking involuntary attention shifts in particular.

Children's auditory selective attention was specifically analyzed by Doyle (1973) in a dichotic listening experiment with the task to remember and to repeat words presented by a target in the presence of distracting. She observed that younger children (8 years old) have less ability to focus and to inhibit irrelevant information than older children (11 years old). In conclusion, the cognitive ability to distinguish between relevant and irrelevant information did not seem to be fully developed by the age of eight.

Regarding the inhibitory control of attention switching, studies by Dibbets and Jolles (2006)and A. Peng et al. (2018) showed that there was no lack of inhibitory control for young children. Children were not worse than young adults in attentional flexibility. Röer et al. (2018) added to these observations examining children's inhibitory control in the auditory domain, and the influence of irrelevant stimuli. Their results align with studies where a higher susceptibility to distracting sounds in children compared to adults was observed (for a review, see Klatte et al., 2013). In general, results from these studies revealed higher performance costs for children in comparison to adults, while cognitive flexibility remained similar. Additionally, Huang-Pollock et al. (2002) found these performance differences between children and young adults. However, they concluded that the mechanisms for selective attention are already fully developed at an early age. Presumably, performance differences depend on the perceptual load at that stage of information processing at which attentional selection occurs. Another explanation could be that the executive functions or abilities to resolve conflicts were not fully developed at the age of 7–8 years old (Dibbets & Jolles, 2006; A. Peng et al., 2018).

In sum, previous research revealed similar auditory attention flexibility for children and young adults. It was found that the ability to focus on relevant information and to ignore irrelevant information develops in early childhood and improves with increasing age. Moreover, performance differences between children and adults were mostly introduced by less developed abilities to process information sufficiently.

Intentional Control of Auditory Selective Attention in Spatial Hearing

The experiments on intentional switching of auditory selective attention were incrementally adapted towards spatial and sound reproduction methods as described by Loh et al. (2022a):

To reach higher ecological validity, sound reproduction for listening experiments needs to be more realistic and plausible than in dichotic listening experiments, providing an acoustic scene that meets the listener's expectation (Lindau & Weinzierl, 2012). Dichotic listening experiments are highly artificial and cannot reflect real-life listening appropriately. Natural hearing implies, among others, the spatial distribution of sound sources, the possibility of binaural hearing with two ears, and connected cognitive benefits. For this reason, the study by Oberem et al. (2014) extended the dichotic listening paradigm by Koch et al. (2011) to a binaural-listening paradigm on intentional switching of auditory selective attention. On the one hand, they presented sound spatially via a setup with real loudspeakers distributed around the participant. On the other hand, spatial sound sources were first simulated using head-related transfer functions (HRTFs), incorporating spatial information, in a virtual acoustic environment, and then presented via headphones. These spatial sound reproduction methods lead to more plausible sound perception in contrast to a simple headphone representation in dichotic listening experiments with mono-signals on the left and right ear each (Oberem et al., 2016).

In general, comparable results to the dichotic listening paradigm have been found in this study (Oberem et al., 2014). Both significant switch costs and significant congruency effects in reaction time and error rate have been observed and underline the conclusion by Koch et al. (2011). A major advantage of this binaural listening paradigm was the possibility to examine the influence of spatial separation on attention switch and the ability to select relevant information more differentially. Target's location now can adopt even more realistic positions such as front and back beside left and right or even more specific positions such as front-right and back-left. Results from this study and further studies using this binaural listening paradigm (Oberem et al., 2017, 2018) revealed a significant effect of the target's position as well as target–distractor position combination in space. For example, worse performance was found for target and distractor positions in front and back compared to positioning at the left and right sides.

Child-Appropriate Spatial Sound Reproduction in Listening Experiments

Previous research on intentional switching of auditory selective attention with spatial acoustic setup has mainly focused on young adults (18–35 years) and elderly adults (60–75 years) (Oberem et al., 2014, 2016, 2017, 2018). There is little knowledge about whether these findings can be transferred to children since no study to date has examined this cognitive ability in children in comparison to adults as described by Loh et al. (2022a). They also pointed out which considerations are required when taking sound reproduction methods appropriate for children into account (Loh et al., 2022a):

[...] Real-life comparable listening in experiments includes more than the spatial distribution of sound sources. It also comprises plausible sound reproduction methods that consider natural (binaural) hearing cues provided by the head and body of an individual person. Differences in anthropometric sizes of head and body lead to different variances in the head-related transfer functions, which are caused by shadowing, diffraction, and reflection of the sound. Thus, plausible sound reproduction is crucial for the outcome of listening experiments. For example, participants reacted faster and achieved higher accuracy in auditory localization experiments when plausible sound presentations were used (Møller et al., 1996). In other words, if participants can utilize natural cues provided by binaural hearing, performance in listening experiments will increase.

[...] This fact was also investigated regarding the intentional switch of auditory selective attention. Oberem et al. (2014) investigated not only the effect of spatial sound source distribution but also different types of plausible sound reproduction methods. They compared, among others, real sources positioned around the participants in contrast to headphones presentation. The latter included individualized head-related transfer functions that were measured individually per participant incorporating their individual anthropometric head and body. In addition to previous studies, the results revealed significant benefits if more natural sound reproduction methods were chosen.

[...] In terms of children, smaller anthropometric sizes of head and ear are noticeable in comparison to adults. These size differences lead to significant differences in head-related transfer functions and therefore conclude in different sound perceptions (Fels et al., 2004; Fels & Vorländer, 2009). With this in mind, it is essential to include individualization processes in sound reproduction in the listening experiment implementation. For example, these individualization processes adjust head-related transfer functions (Bomhardt & Fels, 2014) according to children's smaller head and ear sizes. Thus, it is possible to provide a more plausible sound perception for children.

[...] Spatial hearing becomes possible mainly through binaural and monaural cues, such as interaural time and level differences and spectral information provided by the head shape and pinna. These are especially important when localizing and identifying sound sources in a spatial setup (Blauert, 1997). It must be noted that these were not available in traditional dichotic-listening paradigms. Interaural time and level differences (as binaural cues) are mainly important to locate in the horizontal plane. Spectral cues (as monaural cues) are required to accurately resolve sound origins in challenging situations when, for example, sound sources are situated in a front-to-back configuration. In

> terms of children, differences in anthropometric sizes might lead to the assumption that also a smaller range of binaural cues is available for children compared to adults. That means that smaller head sizes result in smaller interaural time and level differences, as well as that smaller pinna sizes and shapes provide less spectral cues for children in comparison to adults. Therefore, it might be expected that the binaural and monaural cues arriving at the ears of children are less pronounced, so that children cannot use them to a full extent as adults can. In previous research, localization precision of children at the age of 4–5 years old has been studied using a real-life setup with loudspeakers in the context of the precedence effect (Litovsky & Godar, 2010; Litovsky et al., 1999). By using a design with surrounding loudspeakers, it can be assumed that the participants were able to use their individual binaural and monaural cues to the full extent as given naturally. Even though the localization accuracy of children was lower than that of adults (minimum audible angle [MMA] = approx. 10.2° vs. approx. 3.6°), the results in these studies indicated comparable localization abilities in a perceptual context (Litovsky & Godar, 2010). If choosing spatial positions above the MMA, it can be assumed that the available binaural and monaural cues for children are sufficient to localize in a perceptual context. In the context of auditory selective attention in a spatial setup, the question arises whether cognitive processes are involved besides the given physical cues [...].

Heart rate variability and physiological stress

To meet higher ecological validity in listening experiments, especially in terms of children, influences caused by the experimental setup should be minimized. The usage of unobtrusive measurement methods for physiological parameters, which do not constrain participant movement during the experiment, is undeniable. There is a wide range of commercial products to measure heart rate available for everyday usage with minimal intrusion into everyday activities, such as wearable devices measuring electro-cardiogram data using wireless electrodes (Rodrigues et al., 2022).

HRV parameters, as they were used in this dissertation, were considered to reflect noise-induced bodily stress reactions due to the *fight-or-flight response* by Mc-Carty (2016). The parameters can be separated into time- and frequency-domain HRV parameters since frequency-based HRV parameters are more prone to error, such as measurement disturbance due to body movements (Lackner et al., 2020), and some of them were not recommended for the interpretation of short-term measurements (Chemla et al., 2005), frequency-based HRV parameters were not investigated in this dissertation. In terms of the time-domain HRV parameters, only those suitable to examine ultra-short-terms were considered: *meanHR* and *meanNN* for about 10 s(cf., Baek et al., 2015; Salahuddin et al., 2007; Shaffer et al., 2016), *StdNN* (Munoz et al., 2015; Wu et al., 2020) and *StdHR* (Cipresso et al., 2019) for about 30 s, and *RMSSD* (Kim et al., 2021; Munoz et al., 2015) and *pNN50* (Kim et al., 2021) for approx. 10 s. The parameters *SD1*, *SD2*, and *SD1/SD2* are derived from the Poincare plots visualizing the distribution of differences between RR intervals. *SD1* was found to be suitable for time frames of 15 s (Schaaff & Adam, 2013) and *SD2* for 90 s (Shaffer et al., 2016). *SD1/SD2* were found reliable for time frames of 30 s (Wu et al., 2020).

3.2 Children's Intentional Switching of Auditory Selective Attention

The paradigm to investigate intentional the switch of auditory selective attention was brought into focus to examine children's cognition in complex acoustic environments in educational buildings, as this cognitive function (as it was examined by Oberem et al. (2014)) was considered essential to listen and perform in complex acoustic environments successfully. The paradigm by Oberem et al. (2014) was adapted and extended with features appropriately for children aged three to ten years. It was further ensured that children could solve the paradigm task and that their motivation was maintained throughout the experiment. This section presents the base paradigm to examine intentional switching of auditory selective attention in children and all extensions, including aspects of complex acoustic scenes and methods to appropriately reflect children's hearing as close to real life as possible. Parts of this section were previously published in Loh et al. (2022a) and are submitted for review in Loh et al. (2024).

This dissertation presents the application of this child-appropriate paradigm in three studies investigating the influence of noise, the spatial distribution of sound sources as essential components of complex acoustic environments, and the age and health effects corresponding to these two aspects. For brevity and precision of this work, the short reference names of each study are indicated in brackets in the following list:

1. Validation study including primary school children and adults (*ChildASA - Validation study*) in Section 3.3.
2. Preschool study examining age and developmental effects (*ChildASA - Preschool study*) in Section 3.4.
3. Study on physiological and cognitive noise effects in children (*ChildASA - Noise study*) in Section 3.5.

3.2.1 Stimulus material

For the *ChildASA - Validation study*, a first version of the speech material was developed as described in the work by Loh et al. (2022a):

> [It] consisted of eight German animal names that could be categorized easily as *flying* or *nonflying.* Each word contained two syllables, and all were phonetically dissimilar words (English translations in brackets): *Biene* (bee), *Ente* (duck), *Taube* (pigeon), and *Eule* (owl) versus *Katze* (cat), *Ratte* (rat), *Schlange* (snake), and *Wanze* (bug). The [first version of the] speech material was spoken by one female adult (24 years old) and by one female child (7 years old), both native German speakers.
>
> [It] was recorded with a Zoom H6 hand-held recorder and a diaphragm condenser microphone Neumann TLM 170 (cardioid directivity pattern) under anechoic conditions at 24-bit resolution and 44.1 kHz sampling rate. A time stretching algorithm in the open-source program Audacity was used to adjust all stimuli to the same length of 600 ms (max. modification of length: 39.8%). This length adjustment procedure ensures that all stimuli started and ended synchronously when presented at the same time. All stimuli were further equalized in loudness following the German standard DIN 45631/A1 (DIN e.V., 2010). The loudness of a signal was then calculated using the Zwicker algorithm, which also considers the time-variant properties of a signal and is more correlated to the human subjective evaluation of a signal's specific loudness. The loudness of all signals was then adjusted iteratively to each other based on the computed loudness parameter. [...] The choice to have a female adult speaker, a female child speaker, and a speech-shaped noise with a long-term frequency spectrum of children was made to represent aspects of daily life in educational institutions in a real-life a manner as much as possible (in Germany, most primary school teachers are female).

The first set of stimulus material unintentionally comprised the stimulus *Wanze* (English: bug), which cannot be categorized unambiguously. Though no significant difference between the stimuli was found in the first study validating the paradigm for primary school children, it was decided to create a new set of stimulus material without the stimulus *Wanze.* Aiming to also address preschool sound environments, children younger than six years old were included as speakers as described in Loh et al. (2024):

> [Therefore, the second version of the] speech material [similar to the first version] consisted of eight German animal names containing two syllables, all considered phonetically dissimilar. The animal names could easily be distinguished as *flying* or *nonflying*: *Biene* (bee), *Ente* (duck), *Taube* (pigeon), and

> *Eule* (owl) versus *Katze* (cat), *Ratte* (rat), *Schlange* (rat), *Robbe* (seal). A female adult speaker and a male child speaker spoke each set of animal names. [...] Detailed information on generating the [stimuli material] can be found in the technical report via https:/doi.org/10.18154/RWTH-2023-00740.

To address effects of room acoustics (e.g., unfavorable reverberation time starting from $T_{20} = 0.8\,\mathrm{s}$ as stated by Astolfi et al. (2019b, 2019a)), a stimuli length longer than 600 ms (as in the first and second version of the speech material) was required. Therefore, The stimuli were extended in the next increment of the paradigm design with the two-syllable words indicating the size of the animal presented as the stimulus. It could be either *Große* (English: big) or *Kleine* (English: small). These words were recorded as part of the second version of speech material as provided in https:/doi.org/10.18154/RWTH-2023-00740.

When noise was applied in the study, a speech-shaped noise with a long-term frequency spectrum of children (age: 5–6 years old) was presented at all four noise positions simultaneously (cf. Figure 3.1, right). The recording details to achieve the long-term frequency spectrum of preschool children can be found in the technical report via https:/doi.org/10.18154/RWTH-2023-00740).

The SPL of all four noise positions and the individual target and distractor were calibrated to the following SPLs in each study as indicated in Table 3.1.

Table 3.1: Noise and stimuli SPL calibration of each study.

Study name	Individual T & D SPL	Noise Total SNR
ChildASA - Validation study	65 dB[SPL]	+6 dB
ChildASA - Preschool study	64 dB[SPL]	0 dB
ChildASA - Noise study	64 dB[SPL]	Condition 1: 0 dB
		Condition 2*: +6 dB

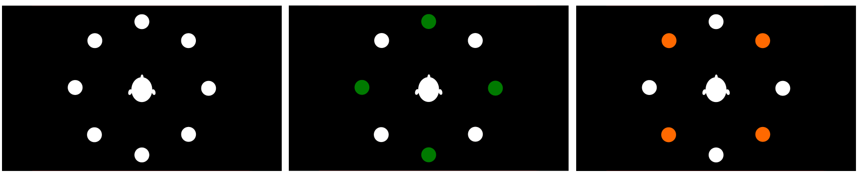
Note. T = target; D = distractor. *examined only in the study with adults, and not further investigated in this dissertation.

3.2.2 Task

Comparable to the speech material, the task evolved incrementally. First, the task from the work by Oberem et al. (2014) was adapted appropriately for children as presented in Loh et al. (2022a) and was used in the *ChildASA - Validation study*:

> [Two stimuli were presented simultaneously,] one of them being he target (T) and the other one being the distractor (D)[. They] were spoken by different speakers, [e.g.,] the target was the female adult voice and the distractor was the female child voice or vice versa. The decision of which voice was applied to the target and the corresponding distractor was made in a randomized manner and changed from trial to trial. Target and distractor were always located at two different and never collocated positions. They could be situated at two of the four possible positions (cf. Figure 3.1, center).
>
> The participants' task was to identify and to categorize the spoken target-stimulus correctly while ignoring the distractor. [...] The target's position was previously indicated using a visual cue (see Figure 3.2.a) displayed on a monitor (22-in. screen, 1.3 m distance). To support children's imagination of space, the scene was shown from an elevated position behind the listener in the cue. Thus, children can directly identify themselves with the character in the center and project the target in the corresponding direction [...]. The two categories, *flying* or *nonflying,* were mapped to two response buttons, which were held in the left and right hand each and pressed by the left and right thumb, respectively. The response mapping was always visible on the bottom of the cue represented as the wing, for flying, and the paw, for nonflying (see Figure 3.2.a).

Figure 3.1: Setup of virtual sound sources and receiver position (from Loh et al., 2022a).



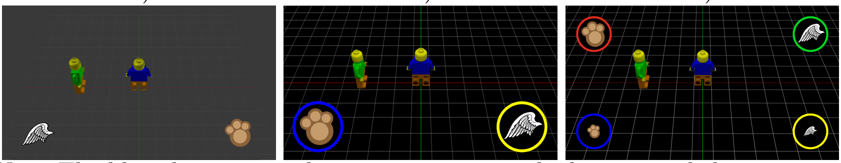
Note. Left: Listener positioned in the center surrounded by eight possible sound source positions. Center: Possible target and distractor source positions are marked in green. Right: Noise positions are marked in orange.

For the *ChildASA - Preschool study*, version two including coloration of the respective buttons was introduced to allowed better identification of the buttons (cf. Figure 3.2.b). In the *ChildASA - Noise study*, the animal names were extended with the words *Große* (English: big) and *Kleine* (English: small), which resulted in four possible response options: big flying, big nonflying, small flying, and small nonflying. The third version of the cue was introduced as shown in Figure 3.2.c.

Hereby, *flying* and *nonflying* were balanced to the left and right, and *big* and *small* were matched to the top and bottom of the cue, respectively.

Figure 3.2: Child-appropriate cue indicating the target's position.

a) b) c)



Note. The blue character in the center represents the listener and the green one shows the direction to focus on. a) 1st version (from Loh et al., 2022a), b) 2nd version (from Loh et al., 2024), c) 3rd version.

The process of an individual trial including the task was described by Loh et al. (2022a) as follows:

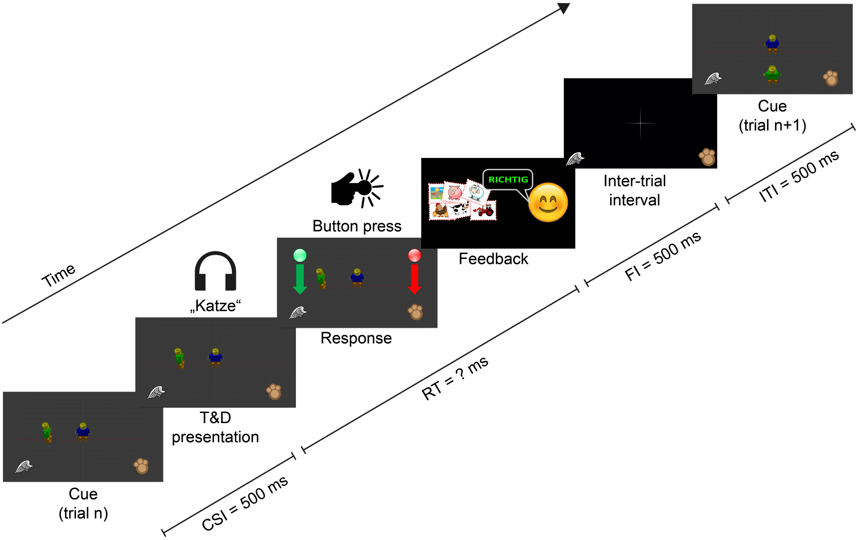
> Each trial started with the onset of the visual cue. After 500 ms, the cue-stimulus-interval, target and distractor were simultaneously presented while the visual cue remained displayed. From this point on, the [reaction time (RT)] was measured until the participant gave a response. Immediately after the response, the feedback was given by either a happy smiley (correct answer) or a sad smiley (wrong answer) in a feedback interval lasting 500 ms. Between feedback and the next cue, the intertrial interval, a fixation cross was shown for 500 ms. Then, the next trial started with the onset of the next cue. [...] The schematic procedure of a trial is shown in Figure 3.3.

3.2.3 Gamification

To ensure reliable cognitive results, it was necessary that children maintained a sufficient level of motivation and stayed focused and attentive throughout the whole experiement. For this purpose, several gamification element were added to the paradigm and the experiment progress, as described in the work by Loh et al. (2022a):

> An extended feedback system was designed to reflect the performance of a participant in every task directly. [... It] integrated a progress visualization in the experiment in order to maintain an appropriate motivation level during the whole experiment [... and the] total performance during one block was implicitly displayed to the participant during the breaks. They could receive

Figure 3.3: Example of a trial structure including all interval durations (from Loh et al., 2022a).



Note. CSI = cue-stimulus-interval; RT = reaction time; FI = feedback-interval; ITI = intertrial-interval.

> three to five out of five stars, depending on their error rates per block. Every participant got at least three stars. If they achieved less than 20% errors, they received four stars. If they reached less than 10% errors, it was indicated with five stars. Furthermore, the overall progress in the experiment was displayed by an increasing number of colorful stickers. They were collected per block throughout the experiment

This feedback system was used in all studies except in the *ChildASA - Preschool study*. Children younger than six years old were expected to loose motivation and attention even faster than children older than six years old, as explained in the work by Loh et al. (2024):

> [The] story and game elements were refined according to the age of three to six. Valkenburg and Piotrowski (2017) inspired to integrate a main character leading through the experiment, explaining each step of the preparational part as well as the task and reflecting the process within the experiment using a narrative storyline.In this [work], a plush rabbit toy with long ears was chosen

> and introduced to the children as *Hasi from the MobiLab* (*Hasi*: German pet name for a rabbit) on the first day in the daycare centres so that children could get familiar with it. *Hasi* was supposed to go on a listening adventure across different stops with the children to see how well they could hear. Each stop comprised another topic, e.g., the zoo or the fire brigade. During the preparational part, *Hasi* was physically going along with the procedure, and when the instruction on the monitor started, *Hasi* was shown on the monitor (cf. Figure 3.4.a).During the experiment, the children collected stickers for every ten completed trials. With increasing numbers of stickers, the process within one block was reflected. Additionally, the six blocks were visually shown in a progress map as *Hasi* travelling from block to block (cf. Figure 3.4.b). [... Direct] feedback was [also adapted and] provided to the participant for every individual response in each trial: *Hasi* was shown with a smiling or sad face. This was combined with an overall performance per block displayed [in the same manner as in the base feedback system] implicitly during the breaks(Loh et al., 2024).

Figure 3.4: Child-friendly experiment components (from Loh et al., 2024).

a) b)

Note. a) *Hasi* the main character, b) Illustration of the storyline.

3.2.4 Binaural reproduction

A major objective of this dissertation was to integrate aspects of children's hearing as close to real life as possible. For this, it was necessary to adjust the sound reproduction methods for the virtual sound environment accordingly, as described in the work by Loh et al. (2022a):

The acoustic virtual environment for the listening experiment was implemented using the Virtual Acoustics integration for MATLAB [(Berzborn et al., 2017)]. To ensure a plausible perception of the virtual sound sources, an individualized set of head-related transfer functions (HRTF) was calculated based on the participant's individual head dimensions (head width, height, and circumference), cf. Figure 3.5 from right to left]) by modifying the interaural time difference cues following Bomhardt and Fels (2014) and the HRTF set of the ITA (Institute of Technical Acoustics, now known as IHTA, Institute for Hearing Technology and Acoustics, RWTH Aachen University) [adult] artificial head (Schmitz, 1995). In other words, the standard HRTF set is morphed according to the individual head dimensions and is more comparable with the individual HRTF set if it would be measured. A static binaural reproduction via open headphones (Sennheiser HD 650) was chosen for this study using a robust headphone equalization following Masiero and Fels (2011). Hence, six headphone transfer functions of every participant were measured using Sennheiser KE3 microphones placed at the entrance of the blocked ear canal using exponential sweeps. After every measurement, the participant was asked to readjust the headphones. All measurements were averaged, and the equalization was finally realized as a minimum-phase filter.The virtual sound sources were set up in a free field condition and at a two-meter distance to the participant. In total, eight virtual sound source position in the horizontal plane were simulated (see Figure 3.1, left). Four out of these eight positions (front at 0°, right at 90°, back at 180°, left at 270°) were intended as possible source positions for the target and distractor (cf. Figure 3.1, center). The remaining virtual sound source positions (at 45°, 135°, 225°, and 315°) simulated a surrounding noise (cf. Figure 3.1, right).

3.2.5 Room setup

All listening experiments presented in this dissertation took place in the mobile hearing laboratory, *MobiLab* (see Figure 3.6). The details were described by Loh et al. (2022a):

The listening experiment took place in a mobile hearing laboratory, *MobiLab*. The MobiLab is a modified trailer, including an acoustically optimized hearing booth. It can be easily set up close to institutions where participants are available (e.g., schools) to allow on-site listening experiments. The hearing booth ($l \times w \times h = 1.86\,\mathrm{m} \times 32.40\,\mathrm{m} \times 31.77\,\mathrm{m}$) ensured a quiet environment during the listening experiment with a sound reduction index $R'_w = 35\,\mathrm{dB}$ (Pausch & Fels, 2019). For the [studies], the Mobi-Lab was positioned [close to]

Figure 3.5: Measurement procedure for the ITD individualization process (from Loh et al., 2024).(from Loh et al., 2022a).

Note. Right: head width; Center: head height; Left: head circumference.

> the cooperating [educational institutions] 's schoolyard [or playgrounds] and was not moved until data collection [with all participants] was finished. The [cooperating educational institutions were mostly] situated in a residential area with moderate traffic.

3.2.6 General procedure

The procedure was adapted slightly for every study to align with the expectations and requirements of each cooperating educational partner. However, a common procedure was carried out in every study, as described in the work by Loh et al. (2022a):

> [In general, every] participant was tested individually, and the procedure started with a preparational part that included the audiometry, the individualization process, and the headphone equalization. Preceding the experimental part, a recorded, spoken, and written instruction (shown on the monitor) including three practice blocks each was presented to every participant. In the first training block, the participant practiced the categorization task. In the second training block, the categorization task was combined with the spatial localization of the target. The distractor was still absent, meanwhile. Finally, in the third training block, the distractor was added; thus, the complete task was trained. Additionally, all participants were instructed to respond as fast and accurately as possible.All participants, including children and adults, had [generally] no problems understanding the task and during the

Figure 3.6: Mobile Hearing Laboratory "MobiLab" (from Loh et al., 2022a).

a)

b)

Note. a) A preschool participant together with *Hasi* (from Loh et al., 2024). b) Left: MobiLab from the outside on a primary school playground; Right: A primary school participant running the experiment (from Loh et al., 2022a).

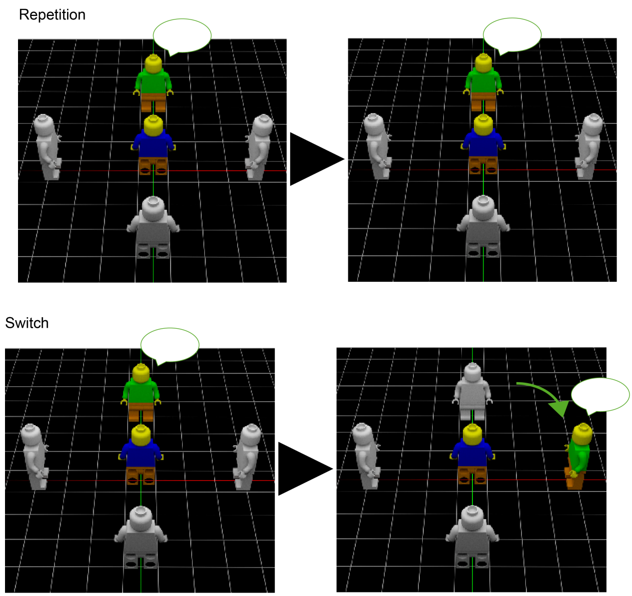
practice sessions[, otherwise it is reported specifically in the section of the study]. Individual feedback after the experiments revealed that the three-step introduction and training phase was beneficial. However, neither children nor adults had questions during the introductory session and continued immediately with the main experiment after a short break[, otherwise it is reported specifically in the section of the study].

3.2.7 General experiment design

Cognitive performance of intentional switching of auditory selective attention was measured according to the following factors and corresponding levels provided by the task and stimuli, as described by Loh et al. (2022a, 2024):

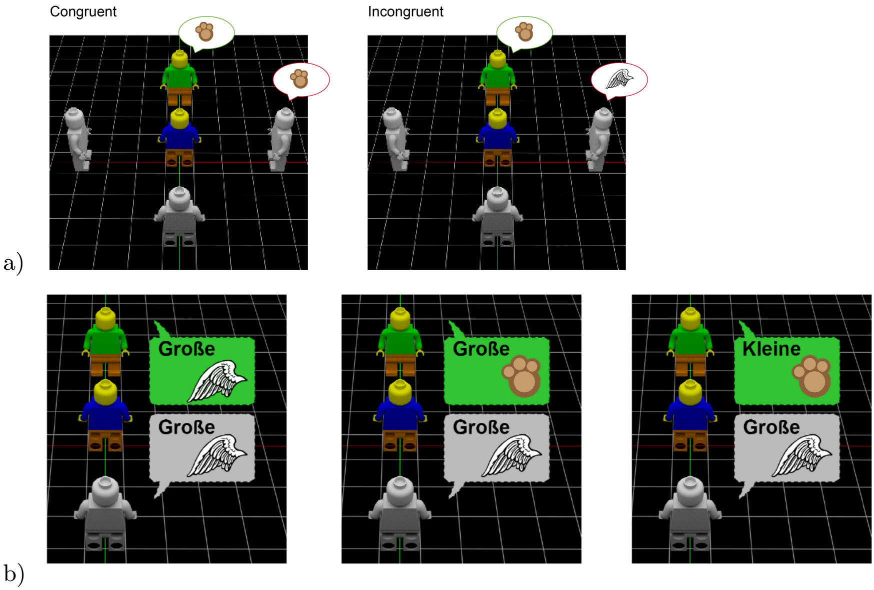
Attention transition (AT) The cues of two consecutive trials could either indicate the target on the same spatial position (e.g., left-left), i.e., the auditory attention focus is repeated (denoted as *repetition*), or on different spatial positions (e.g., left-right), i.e., the auditory attention focus is switched (denoted as *switch*).

Figure 3.7: Factor *Attention transition* (AT).



Congruency (C) Stimuli presented by the target and distractor speakers could be either from the same category (denoted as *congruent*, e.g., both animals could fly) or from different categories (denoted as *incongruent*, e.g., the target animal could not fly, and the distractor animal could fly).
Concerning the stimuli containing two words (big or *small* combined with an animal name), a third category, *semi-congruent*, was introduced. Both words from the target and distractor were from the same category (*congruent*), and both words from the target and distractor were from different categories (*incongruent*). If the category of one word from the target overlaps with the distractor and the category of the other does not, it is considered *semi-congruent*.

Figure 3.8: Factor *Congruency* (C).

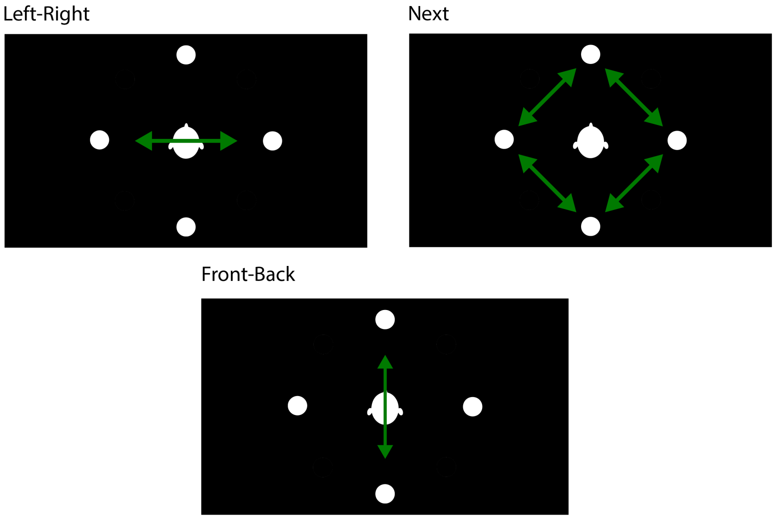


Note. a) Base categories. b) Extended categories concerning the double word stimuli: Left - congruent, center - semi-congruent, right - incongruent; *Große* = big; *Kleine* = small.

Target-distractor-position combination (TDPC) The four possible source positions resulted in three categories of how the positions of the target and distractor could be arranged to each other: 1. target and distractor were positioned on the left and right side of the listener (left-right, LR), 2. target and distractor were positioned on neighboring positions (Next), and 3. target and distractor were positioned in the front and the back of the listener (front-back, FB). From an acoustic perspective, the difficulty of locating sound sources within categories one to three could be considered incrementally increasing, e.g., the front-back combination represents the median plane as part of the cone of confusion, where sources are hard to localize and distinguish.

Noise (N) Noise was presented in a surrounding manner simultaneously from all four noise source positions in between the target and distractor source positions. The noise playback started at the beginning of a block and ended after completing

Figure 3.9: Factor *Target-distractopr-position combination* (TDPC).



a block in the first study, *ChildASA - Validation study*, i.e., the participants could stop the noise by responding and pressing the buttons. To avoid participants' control over the noise in the following studies, the noise was presented continuously during a noise block and stopped when the break started.

Four identical noise signals were played, and they were overall presented (in summation) at a signal-to-noise ratio (SNR) of +6 dB (equal loudness for noise and target) in the *ChildASA - Validation study*, which was chosen according to Stellmack et al. (1997). The resulting noise signal was separately calibrated with respect to the target and distractor SPL to ensure the defined SNR of the respective study. In the *ChildASA - Preschool study*, the SNR was 0 dB, and in the *ChildASA - Noise study*, the SNR was +6dB dB

3.3 Child-appropriate Paradigm Validation Study

The *ChildASA - Validation study*, as it is presented in this dissertation, was published in Loh et al. (2022a). This study aimed to validate a newly developed paradigm that appropriately investigates children's intentional switching of auditory selective attention as described in the work by Loh et al. (2022a):

It should incorporate realistic acoustic situations in educational buildings, such as classrooms, including room acoustics, spatially distributed sound sources, and background noises. For this purpose, a child-appropriate version of the binaural-listening paradigm was developed, which is comparable to the paradigm by Oberem et al. (2014).

In this child-appropriate paradigm, knowledge on numbers and the classification task (regarding the number *five*) were discarded to meet children's cognitive ability and knowledge. The categorization task was simplified to *flying*or *nonflying*using well-known animal names as stimulus material. This simplification was decided following the suggestion by Rueda et al. (2004), stating that children must more likely cope with problems in understanding the task and the response code.

An extended feedback system was designed to reflect the performance of a participant in every task directly. Children will, therefore, not be discouraged by too many errors, and participants are always aware of the result from a given input. This extended feedback system further integrated a progress visualization in the experiment in order to maintain an appropriate motivation level during the whole experiment, especially in children. Additionally, colorful pictures and childish game elements were used to meet the expectations of a computer game, which is more attractive and motivating for children. In consideration of realistic sound environments, this study addressed the topic of plausible sound reproduction in the field of listening experiments with children, in contrast to how listening experiments are conducted until nowadays. Different anthropometric sizes of children and adults were considered in the binaural sound reproduction so that a plausible perception of sound in space was guaranteed for all age groups. Furthermore, a noise condition was added in contrast to the no-noise condition to examine background noise effects on intentional attention control of adults and particularly in children. A stationary noise filtered with the frequency spectrum of children's speech was chosen. It was presented with low noise levels to obtain first insights.

To summarize, this work's objectives were threefold: First, the newly developed paradigm is validated as suited for children with results from previous studies that were designed for adults. The hypothesis tested is that the newly developed paradigm shows switch costs and congruency effects in a comparable manner to previous studies on intentional switching of auditory selective attention including spatial auditory aspects. Second, this cognitive ability and its extent is examined in children and compared to young adults. The hypothesis is that significant differences in switch costs and congruency effects exist between children and adults. Finally, the hypothesis that background noises affect these cognitive processes within a spatial acoustic setup is tested for young adults, children, and young adults compared to children. More specifically, it is assumed that there are significant differences in trials with noise versus without noise and that these differences vary between different spatial setups.

3.3.1 Participants

To validate the paradigm for children, primary school children (six to ten years old) were chosen as the test participant group, and young adults were recruited as the reference group as it was known from the previous work by Oberem et al. (2018) and as described by Loh et al. (2022a):

> The sample size was chosen in the same manner as previous studies (Lawo et al., 2014; Oberem et al., 2014, 2017) on auditory selective attention, to allow direct comparison and validation of this new paradigm suited for children. Twenty-four young adults (age: 18–26 years; M = 22 years, SD = 2 years, 12 female) and 24 primary school children (age: 6–10 years; M = 8 years, SD = 1 years, 12 female) participated in the experiment. Inclusion criteria were normal hearing abilities (within 25 dB[HL] defined as no impairment by the WHO (1991)), German-speaking, and no behavioral syndromes such as attention deficit hyperactivity disorder, which is assessed by consulting the responsible teachers in the case of children. To ensure normal hearing abilities, all listeners were screened by an ascending-pure-tone-audiometry procedure for frequencies between 125 Hz and 8 kHz using a diagnostic audiometer (ear3.0, AURITEC – Medizindiagnostische Systeme GmbH, Hamburg, Germany) complying with DIN 60645-1 (DIN e.V., 2018). All listeners could be considered novices regarding the task since they had never participated in a listening experiment on auditory selective attention before. For participating in the experiment, every adult received a financial compensation of 8€ and every child received a voucher of the same amount for a bookstore.
>
> Recruitment was performed after obtaining ethical approval by the Medical Ethics Committee at the RWTH Aachen University, with the study title *Studie zur selektiven auditiven Aufmerksamkeit bei Kindern im Vorschul— und Grundschulalter mit einem kindgerechten Paradigma*([English:] Study on auditory selective attention in preschool and primary school children using a child-appropriate paradigm) and the protocol number EK 036/18. Primary school children were recruited from the cooperating primary school through teacher–parent communication and were tested after the regular school day during all-day childcare. Informed consent was obtained from all participating

children and their families and adult participants prior to testing. Children gave verbal consent, with the possibility to revoke the consent and cancel at any time during the experiment session.

3.3.2 Experiment design in detail

The general experiment design was described in Section 3.2. In this section, experiment design specifics of the *ChildASA - Validation study* are highlighted as described in the work by Loh et al. (2022a):

> [Each trainings block contained 16 trials and the] main experiment contained twelve blocks with 48 trials each. The blocks were separated by short breaks, which could be extended if needed. The total duration of the experimental procedure did not exceed 90 minutes [...].
>
> The conditions target-distractor-position combination, auditory attention transition, and congruency were balanced, and all categories of every condition were presented in an equal number of trials with 24 repetitions each. Noise and no-noise conditions were blocked, and the order was counterbalanced across all participants. The mapping of the response buttons to the two categories (*flying*vs. *nonflying*) was counterbalanced over participants as well. [...]
>
> A repeated-measures ANOVA was conducted with a mixed-subject (within- and between-subject) design. [...] Independent variables were auditory attention transition (switch vs. repetition), congruency (incongruent or *Incong*vs. congruent or *Cong*), target-distractor-position combination (L-R vs. Next vs. F-B), and noise (no noise vs. noise) as within-subject variables. Age group (children vs. adults) was analyzed as a between-group variable. Dependent variables were reaction times and error rates. [...]
>
> For the response time (RT) and error rate (ER) analyses, practice trials, the first trial of each block, and trials following an error were removed from the data. RTs were Z-transformed for each participant separately, and then, values exceeding ± 2 SD were excluded from analyses as outliers (4.1%). Furthermore, error trials were discarded for the RT analyses.

3.3.3 Results and discussion

The complete ANOVA results for reaction time and error rates are summarized in Table 3.2. The following section will present the results and discussions relevant to the scope of this dissertation. More details and descriptions on the results and extended discussion can be found in Loh et al. (2022a).

Table 3.2: Complete ANOVA Results for RT and ER (adapted from Loh et al., 2022a).

	Reaction time (RT)				Error rate (ER)			
	df	*F*	*p*	η_p^2	*df*	*F*	*p*	η_p^2
N	**(1, 46)**	**43.4**	**<.001**	**.485**	**(1, 46)**	**9.8**	**.003**	**.176**
AT	**(1, 46)**	**6.6**	**.014**	**.125**	(1, 46)	0.5	.469	.011
C	**(1, 46)**	**8.4**	**.006**	**.154**	**(1, 46)**	**380.7**	**<.001**	**.892**
TD-PC	**(1.4, 62.3)**[a]	**32.4**	**<.001**	**.414**	**(2, 92)**	**196.5**	**<.001**	**.810**
N × AT	(1, 46)	0.8	.372	.017	(1, 46)	1.4	.241	.030
N × C	(1, 46)	0.1	.818	.001	(1, 46)	0.2	.655	.004
N × TD-PC	(1.4, 64.0)[a]	1.0	.361	.020	(2, 92)	0.7	.481	.016
AT × TD-PC	(2, 92)	1.0	.369	.021	(2, 92)	0.5	.602	.011
AT × C	(1, 46)	1.2	.284	.025	**(1, 46)**	**4.9**	**.032**	**.096**
C × TD-PC	**(2, 92)**	**5.4**	**.006**	**.104**	**(1.5, 70.6)**[a]	**96.4**	**<.001**	**.677**
N × AT × C	(1, 46)	0.1	.727	.003	(1, 46)	0.0	.920	.000
N × AT × TD-PC	(1.7, 76.0)[a]	0.6	.508	.013	(2, 92)	0.8	.458	.017
N × C × TD-PC	(1.7, 79.3)[a]	0.2	.825	.003	(2, 92)	4.0	.021	.080
AT × C × TD-PC	(2, 92)	1.0	.360	.022	(2, 92)	0.0	.972	.001
N × AT × C × TD-PC	(1.7, 79.2)[a]	0.0	.981	.000	(2, 92)	3.0	.055	.061
AG	**(1, 46)**	**20.7**	**<.001**	**.311**	**(1, 46)**	**45.8**	**<.001**	**.499**
N × AG	**(1, 46)**	**17.1**	**<.001**	**.271**	**(1, 46)**	**10.8**	**.002**	**.190**
AT × AG	(1, 46)	0.0	.975	.000	(1, 46)	0.4	.545	.008
C × AG	(1, 46)	0.2	.669	.004	(1, 46)	1.0	.326	.021
TD-PC × AG	(1.4, 62.3)[a]	3.4	.056	.069	(2, 92)	2.7	.075	.055
N × AT × AG	(1, 46)	0.1	.795	.001	(1, 46)	0.1	.799	.001
N × C × AG	(1, 46)	0.2	.675	.004	(1, 46)	0.0	.886	.000
N × TD-PC × AG	(1.4, 64.0)[a]	0.9	.368	.020	(2, 92)	1.0	.378	.021
AT × C × AG	(1, 46)	2.5	.120	.052	(1, 46)	0.8	.383	.017
AT × TD-PC × AG	(2, 92)	0.8	.463	.017	(2, 92)	1.1	.338	.023
C × TD-PC × AG	(2, 92)	5.4	.006	.105	(1.5, 70.6)[a]	11.5	<.001	.200
N × AT × C × AG	**(1, 46)**	**0.2**	**.694**	**.003**	**(1, 46)**	**0.3**	**.572**	**.007**
N × AT × TD-PC × AG	(1.7, 76.0)[a]	1.2	.289	.026	(2, 92)	0.7	.500	.015
N × C × TD-PC × AG	(1.7, 79.3)[a]	1.4	.260	.029	(2, 92)	0.2	.801	.005
AT × C × TD-PC × AG	(2, 92)	0.2	.794	.005	(2, 92)	0.7	.491	.015
N × AT × C × TD-PC × AG	(1.7, 79.2)[a]	0.6	.527	.013	(2, 92)	0.3	.760	.006

Note. AG = age group; N = noise; AT = attention transition; C = congruency; TD-PC = target-distractor-position combination. Significant effects are indicated in bold. [a] Greenhouse-Geisser correction applied due to violation of assumption.

Paradigm validity Results of the *ChildASA - Validation study* found the newly developed paradigm to be appropriate for children, and it reflected effects on intentional switching of auditory accurately (Loh et al., 2022a):

> Young adults yielded results comparable to those in previous research despite the reduced task complexity and an additionally implemented motivation system. This observation leads to the assumption that the specific categorization task is not essential for the paradigm. The categorization task must consist of at least two categories (e.g., *flying*vs. *nonflying*or *bigger than five*vs. *smaller than five*). This restriction for the task is enough to provide results to derive corresponding cognitive mechanisms of attention switch and the selection of relevant information. Moreover, a motivation system, including positive and negative feedback, does not affect intentional attention control. Even though we introduced the aspect of positive feedback not included in previous research (e.g., Koch et al., 2011; Oberem et al., 2014), attention switch costs and congruency effects remained consistent with previous results. This consistency in results confirms that the present paradigm is a good choice for examining children's intentional attention control and delivers robust results.

Age efffect As expected, Loh et al. (2022a) reported a significant age effect in reaction time and error rate:

> On average, children showed higher reaction times (-912 ms) and higher error rates (-11.7%) compared to young adults. However, the results of this work suggest that there is no significant difference in auditory attention flexibility between children and young adults. In general, attention switch costs and congruency effects were comparable in young adults and children[, ...] which is in line with the existing literature (A. Peng et al., 2018; Röer et al., 2018). Taken together, the findings of this work suggest that age-related differences in intentional switching of auditory selective attention are due to differences in processing resources or execution processes. In this context, it seems reasonable to assume that cognitive mechanisms of intentional attention control are comparable to young adults at the age of 6–10 years old. This age limit has also been found for different auditory processes in a developmental context (for review, see Litovsky, 2015). Further work should concentrate on age groups below 6 years old to explore whether abilities considering intentional attention switch and selection of target information are developed at an earlier stage of children's development.

Noise effect The *ChildASA - Validation study* further revealed an interesting noise effect on the intentional switching of auditory selective attention as described by Loh et al. (2022a):

> [... Results] revealed a negligible noise effect for young adults and a significantly reversed pattern for children. With the objective to investigate noise effects on intentional switching of auditory selective attention, a moderate signal-to-noise ratio of 6 dB[SPL] was chosen for children to provide first insight. The results of this study found that this signal-to-noise ratio is challenging for younger children but does not affect young adults, which is in line with previous research (e.g., Neuman et al., 2010).
>
> The technical implementation [...] allowed the participant a certain amount of control over the noise presentation. Noise playback was started shortly before the stimuli were present and stopped at the participant's response. Children could have realized this fact during noise conditions and responded faster while taking more errors into account. Hence, we can consider a speed-accuracy trade-off. One possible explanation for this phenomenon might be that children perceive the noise as unpleasant and choose to conduct an avoidance strategy. On that score, they conclude faster with the noise and can recover subsequently.
>
> The significant noise effect on children's auditory selective attention is noteworthy. [...] This might conclude in significant effects during children's development. Nevertheless, guidelines nowadays limiting noise levels in educational institutions are mainly developed based on experimental findings based on young adults. Taking this into consideration, results from this work offer indisputable evidence for the importance of child-appropriate guidelines.
>
> In future work, it would be interesting to change the technical implementation of the noise presentation and to examine further the corresponding noise effect on attention switch and relevant information selection.

Spatial processing Results on spatial sound processing in combination with auditory selective attention differed slightly between adults and children as observed by Loh et al. (2022a):

> The findings of this study on the spatial aspect of hearing correlate reasonably well with Oberem et al. (2014) and further support the concept of auditory processing benefits when using binaural cues. Depending on the spatial configurations of target and distractor, participants obtain different amounts of information due to interaural time and level differences. For example, differences between the left and right ears are higher than in the right-left-position

than in the front-back configuration, where differences are nearly zero. Thus, it is more complex to locate the target position accurately. It is explainable why performance decreases with the complexity of spatial configuration. Before attention selection happens, the target sound source must be identified using the provided acoustic information. Therefore, more cognitive load is needed to resolve the target's location if the spatial configuration becomes more complex. It becomes more difficult to select relevant information. This effect is reflected in the interaction of congruency and target-distractor position combination.

Remarkable insights into the age differences were revealed regarding the benefits of congruent configurations. Results from this study showed age differences regarding the selection of relevant information (congruency effect) but not in attentional flexibility. In other words, only the ability to select relevant information differed in young adults and children depending on the spatial configuration of target and distractor, as one can see in the error rates. Young adults seem to benefit more from easier spatial configurations. The response times revealed a weaker specification of congruency effects. This discrepancy of response times and error rates, especially regarding congruency, is in line with previous work on attentional flexibility (Nolden & Koch, 2017; Nolden et al., 2019; Oberem et al., 2017, 2018). Since the congruency effect does not change significantly for children over the different spatial configurations in RT, it can be assumed that they might not benefit from spatial auditory cues to the same extent as adults. It is likely that children have not developed the ability to make use of spatial auditory cues for auditory attention control in the present development stage and, therefore, cannot benefit from them as much as adults yet.

Litovsky (2015) indicated in her work that children at the age of 5 years old are already able to locate sound sources accurately. To be more precise, Litovsky and Godar (2010) found that children were able to distinguish single sound sources up to a minimum audible angle of approximately 10.2°. Since the spatial setup in this experiment included a separation of sound sources at an angle of 45°, it can be expected that children were able to utilize binaural cues for accurate perception when localizing the origin of sound sources. This leads to the conclusion that the worse performance compared to adults might have its origin in other auditory attention processes, at least for the horizontal plane and well-separated sound sources. However, it must be noted that the worse performance is especially noticeable within the hardest spatial setup: the front-back configuration. It might be explained by the smaller pinna sizes of children compared to adults, leading to a lack of spectral information to resolve this challenging spatial configuration.

3.4 Developmental Differences in Young Children

After validating the paradigm with primary school children (six to ten years old), the objective was to adapt the base paradigm for preschool children (three to six years old). Three major considerations arose considering the paradigm design when taking children younger than six years old into account:

- Duration of the experiment, since the current length was approx. 90 minutes and up to 120 minutes, including the preparation. Furthermore, it was unknown whether younger children would need even more time to solve the task.
- Maintenance of motivation in the course of the experiment, since younger children tend to be more easily distracted and lose motivation in continuously repeating tasks and events.
- Solvability of the task by younger children, since it is unknown whether the younger children understood and were able to solve the task.

The base paradigm was, therefore, restructured into shorter blocks, and a more extensive gamification system was added as described in Section 3.2.3. Additionally, a three-stage assessment of the task was added to the training to ensure that the participant understood the task sufficiently (see Section 3.4.1). Parts of this section were submitted to *Scientific Reports by Nature* in August 2023. In the following, these parts are referenced as Loh et al. (2024).

3.4.1 Participants

To achieve sufficient power, an a-priori power analysis was conducted based on the previous study by Loh et al. (2022a) to determine an appropriate sample size across the four age groups (three-, four-, five-, and six-years-old) (cf. Loh et al., 2024):

> To achieve a power of (1 - β) = .8 with an average effect size of $\eta_p^2 = .3$ at the standard $\alpha = .05$ error probability, the analysis via G*Power (Faul et al., 2007) resulted in a sample size of 20 participants per age group, thus a total of 80 participants.
>
> In total, 91 children (age: 3-6 years; full descriptives on age and gender can be found in Table 3.3) were recruited in cooperation with seven day-care centres in Aachen, Germany. Recruitment was performed after obtaining ethical approval from the Medical Ethics Committee at the RWTH Aachen University with the protocol number EK 476/21. All participants received

a 10 €-voucher as compensation. Informed consent was obtained from all participating children and their families prior to the experiment. Children gave verbal consent, with the possibility to cancel at any time during the experiment. The study was conducted in accordance to the rules of conduct stated in the Declaration of Helsinki.

Inclusion criteria were German mother tongue, normal hearing abilities, and never received attention deficit hyperactivity disorder diagnosis (all assessed by consulting the legal guardians via a questionnaire and educators).

[...] To make sure that the children understood the task and were able to continue with the main experiment, the three-stage training comprised an incremental assessment. In the first two training stages, including only the target speaker, a maximum of 20% error rate was accepted. In the third training stage, including both the target and distractor speaker, a maximum of 40% was tolerated. It was ensured that 50% of the trials were congruent and incongruent, respectively. The limits were chosen according to the guessing probability of 50% and the expected difficulty due to the number of present speakers. Each training stage could be repeated once if the maximum error rate was exceeded. If the participant failed any training stage twice, it was assumed they did not understand the task sufficiently and were excluded from the main experiment. This procedure was designed in reference to Jones et al. (2015).

Table 3.3: ChildASA-preschool: Overview of participants.

	Recruited			Completion rate[%]			Cancelled within training			
	N_T	N_f	N_m	T	f	m	N	N_A	N_B	N_C
$\sum$	85	39	46	58.8	59.0	58.7	35	1	3	31
3 y/o	15	9	6	46.7	55.6	33.3	8	1	3	4
4 y/o	23	11	12	56.5	36.4	75.0	10	-	-	10
5 y/o	30	12	18	60.0	75.0	50.0	12	-	-	12
6 y/o	17	7	10	70.6	71.4	70.0	5	-	-	5

Note. T = Total, f = female, m = male; N_A, N_B, N_C = cancelled after stage A, B, C, respectively.

3.4.2 Experiment design in detail

The experiment design comprised the independent variables attention transition (AT), congruency (C), and target-distractor-position combination (TD-PC) as within-variables and the age group (AG) as between-variable with four levels (three-, four-, five-, and six-years-old). A detailed description of the independent variables can be found in Section 3.2.7. To avoid sequential biases, balancing for the presentation of the independent variables was applied and the experiment design was setup as described in Loh et al. (2024):

> Attention transition, congruency, and target-distractor position were balanced within one block. At the same time, the factor noise was presented block-wise, and the order of blocks was counterbalanced over the participant ID. The first block was always noise-free so participants could adjust to the experimental conditions. The mapping of the response buttons of the two task categories was also counterbalanced over the participant ID. All factor combinations (conditions) were repeated 15 times and divided into six blocks with equally distributed numbers of each condition. Each block endured five minutes, and the total duration of the complete experiment (including all preparatory measurements, training and experimental part, and multiple breaks between the blocks) did not exceed 75 minutes. Participating children decided how many and how long the breaks between the blocks lasted to ensure enough recovery time.

Originally planned analyses were defined based on results of the previous study by Loh et al. (2022a) and preregistered on the Open Science Framework (access via https:/doi.org/10.17605/OSF.IO/EW4QR). Results according to the preregistration can be found in the supplementary information of the submitted manuscript by Loh et al. (2024). However, the preregistered analyses revealed that additional analyses were needed to understand the impact of random factors (Claesen et al., 2021; Nosek et al., 2019). Eventually, analyses, according to the factors of age group and noise, revealed insightful results presented in this dissertation. The analysis plan, therefore, comprised a four-step analysis (see Loh et al., 2024):

1. Investigation of success rate across age groups to complete the training and the whole experiment resulting in insights which age groups understood and were able to solve the categorization task.

2. Analyzing the development of response times and error rates during the experiment yielding explanations on maintaining motivation and attention within the duration of the experiment across age groups.

3. Focusing on age and noise effects, repeated-measures ANOVAs with mixed-subject design for the response times and error rates were conducted with the factors *age group* (AG, four levels: three-, four-, five-, and six-year-olds), *attention transition* (AT, two levels: repetition and switch), *congruency* (C, two levels: congruent and incongruent), and *noise* (N, two levels: noise and

no noise). To compensate for the lack of repetitions for each condition, the planned variable *target-distractor-position combination* (TD-PC) was not evaluated in the present analysis (see supplementary information in the submitted manuscript by Loh et al. (2024) for further details). Furthermore, the effect of gender was not considered in this study due to the small sample size in each age group.

The models were additionally controlled in linear mixed-effects models for the random effects *participant id*, *daycare*, and *trial number* to consider the within-participant variances, effects from the differences of daycare centers that participated in the experiment, and the duration of the experiment. However, the random effects did not change the models significantly from the results of the ANOVAs. Thus, only the results of the ANOVA were presented and evaluated.

4. Investigating the impact of trials with and without errors on the response time, the repeated-measures ANOVA for the response time was conducted with an additional factor *Error* with two levels *trials with error* versus *trials without errors*. The two ANOVA models were initially checked in comparison, revealing significant differences ($X^2(1) = 106.87, p < .0001$) between the models with a better fit using the new model, including the factor *Error*.

All data, models, and results were pre-processed and computed using the R version 4.2.2[1] using the packages *lm4* version 1.1-33[2] as reported in Loh et al. (2024) and IBM SPSS Statistics (Version 29).

3.4.3 Results and discussion

Completion rate across age groups

Before the main experiment, children had to complete a three-stage training, which comprised an incremental assessment of task understanding (stages A, B, and C). An overview of the participants finishing the training and assessment is added to Table 3.3. Thirty-five children did not pass the training stage, and one child did not complete the experiment after passing the training and assessment.

The completion rate reflects the percentage of children understanding the complete paradigm task and having the endurance to finish the whole experiment. In general, 58.8% of the participating children completed the experiment. Moreover,

[1] https:/www.r-project.org
[2] https:/cran.r-project.org/web/packages/lme4/index.html

it can be observed that the success rate increased with age from three- to six-year-olds (46.7%, three-year-olds, vs. 56.5%, four-year-olds, vs. 60.0%, five-year-olds, vs. 70.6%, six-year-olds).

Training stage A involved only the categorization task, and the maximum error rate was predefined to 20%. Training stage B comprised the categorization task while the target speaker switched the position from trial to trial. The maximum error rate of stage B was also defined as 20%. Training stage C eventually contained the final task, which included the distractor speaker. At this stage, a maximum of 40% error rate was tolerated. If the maximum error rate was exceeded, the run of a training stage was noted as failed and could be repeated once. If both runs failed, the training was canceled, and the child did not move on to the main experiment.

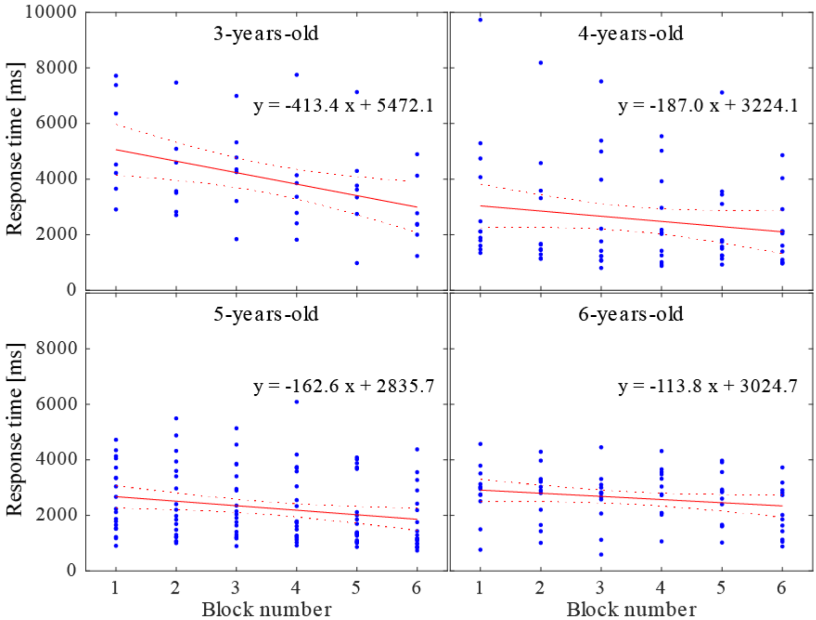
In this study, children mainly canceled after training stage C, unable to distinguish the distractor speaker from the target speaker. However, four three-year-olds canceled in stages A and B, showing that these children could not solve the categorization task. This result indicates that the paradigm task might be too complex for the youngest children, though 46.7% of three-year-olds were able to complete the whole experiment. In this case, it is necessary to consider differences in children's development. At the age between three to six years, children's development can strongly vary and depend on external factors, such as the parent's socioeconomic status and their influences on the development of their children. With reference to the supplementary information in the submitted manuscript by Loh et al. (2024), it is noticeable that the parents of the three-year-olds participating in this experiment reported a significantly higher socioeconomic status than the average of the parents within the other age groups.

Performance over time

To investigate possible attention declines in the course of the experiment, linear regression analyses (with the model $y = c_{\text{reg}} * x + c_{\text{int}}$) of the response times (cf. Figure 3.10) and error rates (cf. Figure 3.11) over time within the main experiment was conducted. For this, the blocks were numbered from one to six, and the block number was treated as an independent variable representing the course of time.

Furthermore, the data was pre-processed to prevent biases in the results: the first trial of each block, trials following an error, and values where the response times exceeded $\pm 2SD$ were filtered as outliers, assuming possible irritation of the participants in these trials. Additionally, error trials were removed from the response time analyses to avoid overlaying trade-off.

Figure 3.10: ChildASA-preschool: RT over block number per age group.

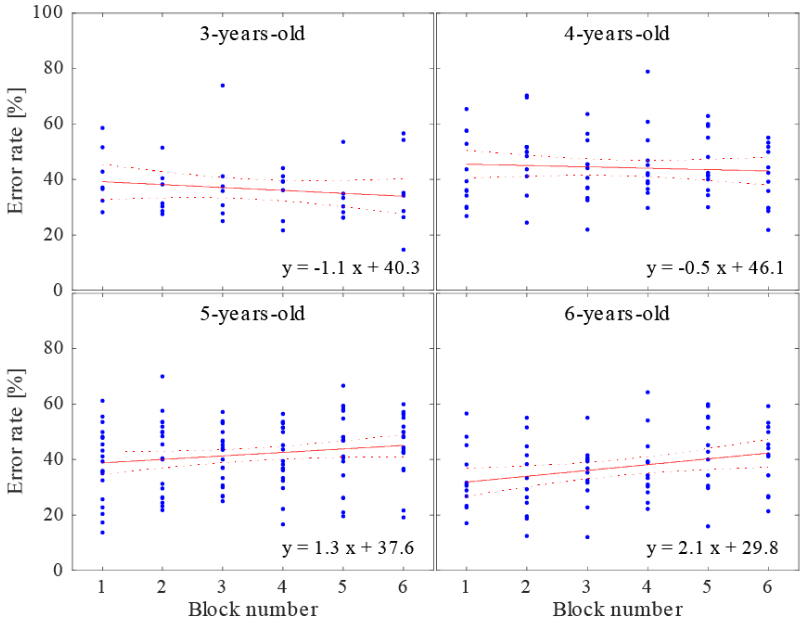


In response times, a steeper decrease of response times with increasing block number was observed for three-year-olds ($c_{\text{reg}} = -414.4\,\text{ms}$, with $p = .002$ and adj. $R^2 = .192$) compared to the older children ($c_{\text{reg}} = [-187.0, -113.8] ms$). The decrease in response time with the course of time reduced with increasing age: $c_{\text{reg}} = -187.0ms$ for four-year-olds ($p > .05$) vs $c_{\text{reg}} = -162.6ms$ for five-year-olds ($p = .010$ and adj. $R^2 = .051$) vs $c_{\text{reg}} = -113.8ms$ for six-year-olds ($p = .033$ and adj. $R^2 = .051$). Additionally, the response time results indicated a lower variation of results for the older (five- and six-year-olds) than the younger children.

In error rates, a split was found between the younger (three- and four-year-olds) and the older (five- and six-year-olds) children. While error rates in the course of time decreased insignificantly for the three- ($c_{\text{reg}} = -1.1\%$, with $p > .05$) and four-year-olds ($c_{\text{reg}} = -0.5\%$, with $p > .05$), it increased for five- ($c_{\text{reg}} = 1.3\%$, with $p = .001$ and adj. $R^2 = .085$) and six-year-olds ($c_{\text{reg}} = 2.1\%$, with $p = .005$ and adj. $R^2 = .192$).

Taking the analyses of error rates and response times together, it can be assumed that the five- and six-year-olds revealed a slight decline in attention in the course

Figure 3.11: ChildASA-preschool: ER over block number per age group.



of time within the main experiment. It seems they took a speed-accuracy trade-off into account the longer they participated in the experiment. This effect was not observable for the younger children (three- and four-year-olds).

Age group and noise effects

The paradigm differentiates between three within variables, *attention transition* (AT), *congruency* (C), and *noise* (N), and one between variable, *age group* (AG). For the model, practice trials, the first trial of each block, and trials following an error were filtered from the data. Additionally, response times were z-transformed for each participant, values exceeding $\pm 2SD$ were excluded as outliers, assuming possible irritation of the participants after negative feedback, and all error trials were excluded to avoid overlayed trade-offs. In case of violation of the assumptions for running ANOVAs, Greenhouse-Geisser correction was applied. The complete ANOVA results for response times and error rates are presented in Table 3.4, and corresponding significant post-hoc tests are summarized in Table 3.5.

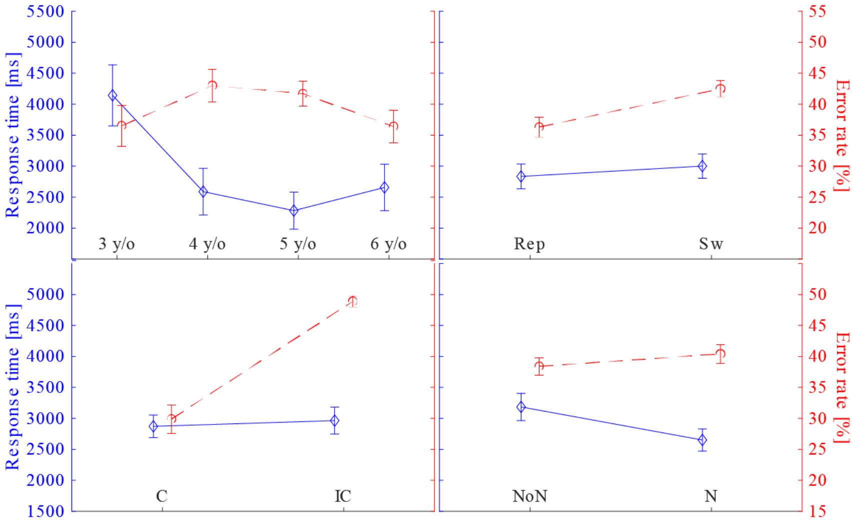
Table 3.4: ChildASA-preschool: ANOVA results for RT and ER ($N = 50$).

	Response time				Error rate			
	df	***F***	***p***	η_p^2	***df***	***F***	***p***	η_p^2
AT	**(1, 46)**	**9.2**	**.004**	**.166**	**(1, 46)**	**27.4**	**<.001**	**.373**
C	(1, 46)	1.4	.239	.030	**(1, 46)**	**68.4**	**<.001**	**.598**
N	**(1, 46)**	**39.0**	**<.001**	**.459**	(1, 46)	3.7	.060	.075
AT × C	(1, 46)	2.3	.134	.048	**(1, 46)**	**6.9**	**.012**	**.130**
AT × N	**(1, 46)**	**4.1**	**.049**	**.082**	(1, 46)	0.3	.615	.006
C × N	**(1, 46)**	**7.9**	**.007**	**.147**	(1, 46)	2.4	.126	.050
AT × C × N	(1, 46)	3.1	.083	.064	(1, 46)	3.1	.085	.063
AG	**(3, 46)**	**3.6**	**.021**	**.188**	(3, 46)	1.7	.183	.099
AG × AT	(3, 46)	2.4	.082	.134	(3, 46)	1.5	.215	.092
AG × C	(3, 46)	2.2	.104	.124	(3, 46)	0.7	.542	.045
AG × N	(3, 46)	2.6	.061	.147	(3, 46)	2.4	.080	.135
AG × AT × C	**(3, 46)**	**2.8**	**.049**	**.156**	(3, 46)	1.5	.236	.087
AG × AT × N	(3, 46)	1.3	.283	.079	(3, 46)	0.1	.937	.009
AG × C × N	**(3, 46)**	**2.8**	**.048**	**.156**	(3, 46)	0.4	.754	.025
AG × AT × C × N	(3, 46)	1.3	.300	.076	(3, 46)	0.5	.697	.030

Note. AG = age group; N = noise; AT = attention transition; C = congruency. Significant effects are indicated in bold.

Descriptives of the main effects are shown in Figure 3.12. Attention transition and noise yielded significant main effects in response times, while in error rates, the main effects of attention transition and congruency were found significant. Attention transition and congruency in interaction with noise yielded a significant interaction effect in response times. The interaction effect of attention transition and congruency was significant in error rates. It was observable that age group effects were mainly found in the response time, especially the three-way interaction of age group and attention transition with congruency and noise, respectively. In general, the effects of the cognitive flexibility of intentional switching of auditory selective attention could be found for children younger than six years old and were mainly observed in error rate. Interestingly, the speed-accuracy trade-off effect, as found in Loh et al. (2022a), was also observed in this study, though not to the same extent.

Figure 3.12: ChildASA-preschool: Main effects age group, attention transition, congruency, and noise in reaction time and error rate.



Note. The error bars represent standard errors. AG = age group; N = Noise; NoN = No Noise; C = Congruent; IC = Incongruent; Rep = Repetition; Sw = Switch.

Descriptives of the two-way interaction effects with age group are presented in Figure 3.13 and three-way interaction of age group and noise in Figure 3.14.

Table 3.5: ChildASA-preschool: ANOVA - significant post-hoc tests.

		G_1	G_2	p (G_1 - G_2)	M_{diff} [ms]	[95 %-CI] [ms]
RT: AG						
		3 y/o	5 y/o	0.014	1862	[274, 3449]
RT: AT × N						
	NoN	Rep	Sw	0.003	-298	[-487, -109]
	Rep	NoN	N	0.000	403	[239, 566]
	Sw	NoN	N	0.000	665	[407, 923]
RT: C × N						
	NoN	C	IC	0.023	-237	[-439, -35]
	C	NoN	N	0.000	389	[207, 572]
	IC	NoN	N	0.000	678	[461, 896]
RT: AG × AT × C						
Rep	C	3 y/o	5 y/o	0.028	1623	[120, 3126]
Sw	C	3 y/o	5 y/o	0.017	1667	[208, 3126]
	IC	3 y/o	4 y/o	0.028	2112	[152, 4073]
	IC	3 y/o	5 y/o	0.003	2480	[658, 4303]
	IC	3 y/o	6 y/o	0.021	2184	[224, 4144]
3 y/o	IC	Rep	Sw	0.001	-795	[-1257, -333]
	Sw	C	IC	0.004	-815	[-1351, -280]
RT: AG × C × N						
C	NoN	3 y/o	5 y/o	0.030	1761	[114, 3409]
	N	3 y/o	5 y/o	0.020	1528	[163, 2893]
IC	NoN	3 y/o	4 y/o	0.037	2226	[89, 4362]
	NoN	3 y/o	5 y/o	0.004	2614	[628, 4600]
	NoN	3 y/o	6 y/o	0.032	2269	[132, 4405]
3 y/o	NoN	C	IC	0.001	-893	[-1401, -385]
	C	NoN	N	0.017	564	[106, 1022]
3 y/o	IC	NoN	N	0.000	1398	[852, 1945]
4 y/o	C	NoN	N	0.006	498	[148, 848]
	IC	NoN	N	0.012	541	[123, 958]
5 y/o	C	NoN	N	0.021	330	[52, 608]
	IC	NoN	N	0.036	448	[31, 866]
		G_1	G_2	p (G_1 - G_2)	M_{diff} [%]	[95 %-CI] [%]
ER: AT × C						
	C	Rep	Sw	0.023	-3.3	[-6.1, -0.5]
	IC	Sw	Rep	0.000	-9.3	[-13.1, -5.5]
	Rep	C	IC	0.000	-16.1	[-20.4, -11.7]
	Sw	C	IC	0.000	-22.1	[-28.0, -16.2]

Note. All results were Bonferroni corrected when required. N = Noise; NoN = No Noise; C = Congruent IC = Incongruent; Rep = Repetition; Sw = Switch.

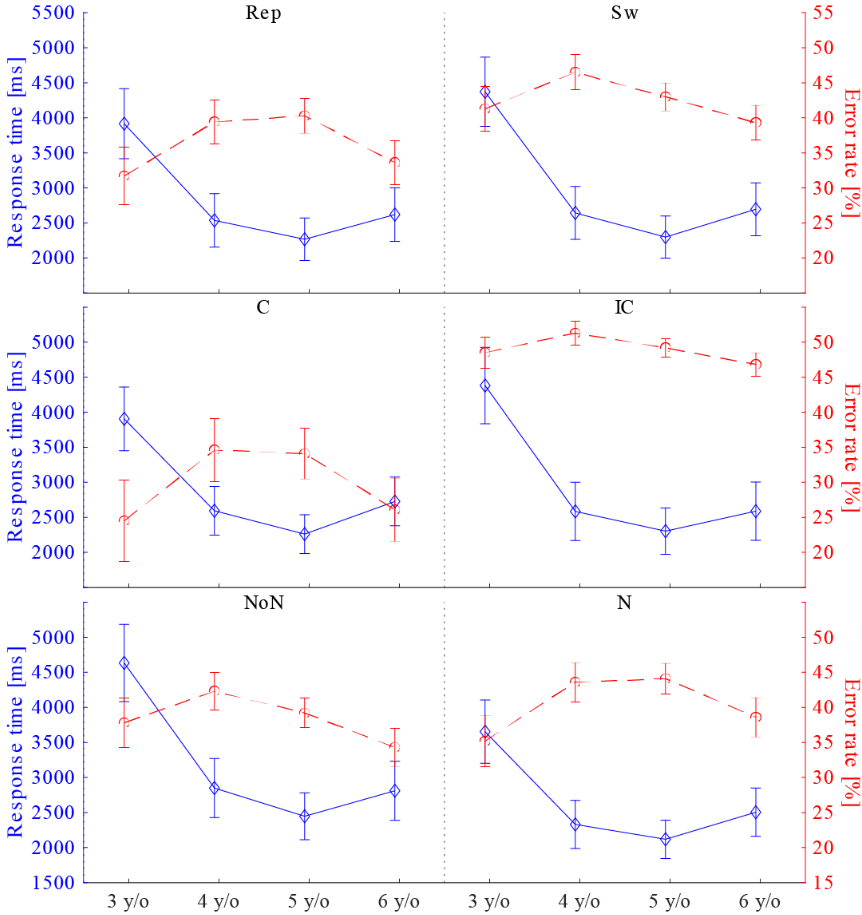
Results indicated significant differences between three-year-olds and the other age groups. In contrast, the other age groups did not reveal significant differences, leading to the assumption that children's cognitive flexibility could be fully developed starting at the age of four. Furthermore, a tendency to decrease error rates with increasing age was found while response times remained constant across age groups starting from the age of four. However, this trend was not significant and must be interpreted with care.

The results of the three-year-olds must be interpreted with care due to the small sample size. A tendency towards a higher impairment of intentional switching of auditory selective attention was observed in more complex situations, such as incongruent and noise conditions. Three-year-olds' ability to correctly select target speech and to inhibit the distractor speaker was similar to that of older children. Interestingly, response times and error rates in noise and no noise conditions were inversed to expectations. Response times in noise conditions were lower than in noise-free conditions, while error rates were higher in noise-free conditions than in noise conditions. It could be assumed that the three-year-olds tried to put more effort into solving the task but could still not solve it. This could indicate mental overload and not having a coping strategy developed for the three-year-olds, but still, it was to their advantage since their error rates were comparatively low.

The present findings suggest similar cognitive flexibility in children starting at the age of four years and children aged six to ten, as well as young adults older than 18. No specific developmental differences were found above the age of four, which assumes that intentional auditory selective attention switching is developed to a reasonable extent by age four. Though not significantly different, performance tended to increase with age, as indicated by decreasing error rate and relative constant reaction times. This result aligned with the previous study by Jones et al. (2015) and added insights to the findings by Stellmack et al. (1997) and Doyle (1973). It was further in line with previous studies on the intentional switch of auditory selective attention (Koch et al., 2011; Loh et al., 2022a; Nolden et al., 2019; Oberem et al., 2014), including adults and older children.

Summarizing the results from the four age groups, the three-year-olds represented the group of children not reliably solving the paradigm (i.e., non-solvers). At the same time, the four-year-olds posed to be the age of transition, and the pooled group of five- and six-year-olds as solvers where the abilities to switch the auditory selective attention intentionally were settled. Therefore, it can be assumed that this cognitive ability develops around the age of three to four. However, this result is preliminary and requires further studies, including bigger sample sizes and focusing on the age groups around three- and four-year-olds to validate this

Figure 3.13: ChildASA-preschool: Interaction with AG in RT and ER.

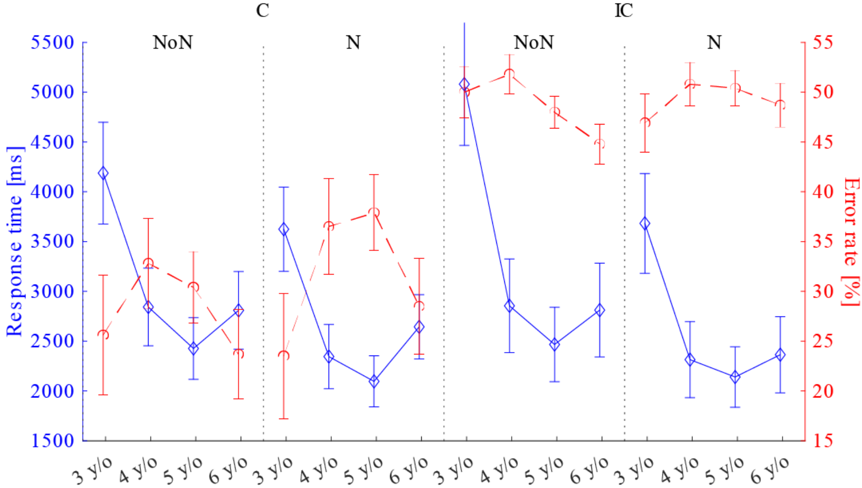


Note. The error bars represent standard errors. N = Noise; NoN = No Noise; C = Congruent; IC = Incongruent; Rep = Repetition; Sw = Switch.

finding.

Noise effects on attention transition and the ability to select relevant information were observed in this study. The average difference in response times between noise and noise-free conditions was increased in switch compared to repetition

Figure 3.14: ChildASA-preschool: Interaction AG × C × N in RT and ER.



Note. The error bars represent standard errors. N = Noise; NoN = No Noise; C = Congruent; IC = Incongruent; Rep = Repetition; Sw = Switch.

conditions. A similar effect was found for incongruent versus congruent trials (cf. Table 3.5). These observations suggested an increase in cognitive load within noisy conditions. Typically, higher response times in noise conditions and lower response times in noise-free conditions (and error rates, respectively) were expected, reflecting the benefit of noise-free hearing situations. However, the opposite was observed for the main effect of noise in response times, yielding higher response times in no-noise trials versus noise trials. This finding indicated a trend towards a speed-accuracy trade-off as a noise avoidance strategy to escape noisy conditions faster. Especially when further taking the different age groups into account, it became clear that each age group had different strategies to deal with the noise conditions. These observed noise effects impairing cognitive processes aligned with previous studies (Klatte et al., 2013), and added to the findings by Loh et al. (2022a) where a similar noise avoidance strategy was found for children aged six to ten years.

Response times of trials with and without errors

To investigate the effect of trials with error in comparison to trials without errors, the difference response time ($\Delta RT = RT_{\text{no error trials}} - RT_{\text{error trials}}$) was calculated and investigated in an ANOVA with three within variables,*attention transition* (AT), *congruency* (C), and *noise* (N), and one between variable, *age group* (AG). The first trial of each block, trials following an error, and values where the response times exceeded $\pm 2SD$ were filtered as outliers, assuming possible irritation of the participants in these trials. Error trials were excluded from the response time analysis to avoid overlayed trade-offs. The complete ANOVA results for response times and error rates are presented in Table 3.6. The descriptives of the significant main effect and the significant post-hoc tests are presented in Figure 3.15.

Table 3.6: ChildASA-preschool: ANOVA results for ΔRT ($N = 50$).

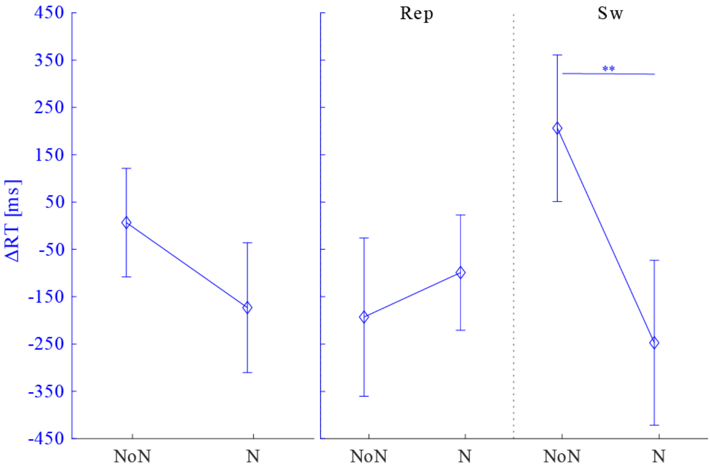
	Response time			
	df	***F***	***p***	η_p^2
AT	(1, 43)	0.8	.373	.019
C	(1, 43)	1.2	.286	.026
N	**(1, 43)**	**4.2**	**.046**	**.089**
AT × C	(1, 43)	0.3	.612	.006
AT × N	**(1, 43)**	**5.4**	**.025**	**.112**
C × N	(1, 43)	0.3	.603	.006
AT × C × N	(1, 43)	0.5	.497	.011
AG	(3, 43)	1.1	.348	.073
AG × AT	(3, 43)	1.1	.348	.073
AG × C	(3, 43)	1.8	.162	.111
AG × N	(3, 43)	1.4	.269	.086
AG × AT × C	(3, 43)	0.5	.673	.035
AG × AT × N	(3, 43)	2.1	.119	.126
AG × C × N	(3, 43)	0.5	.651	.037
AG × AT × C × N	(3, 43)	0.3	.829	.020

Note. AG = age group; N = noise; AT = attention transition; C = congruency. Significant effects are indicated in bold.

A significant main effect was only found for the noise conditions. Furthermore, the interaction effect between attention transition and noise yielded a signifcant effect. The post-hoc analyses revealed only one significant difference between the noise-free and noise conditions within switch trials. In this case, no noise trials yielded a higher difference in response time ($\Delta RT = 7$ ms) compared to noise condition($\Delta RT = -173$ ms). A negative difference indicated a higher reaction time in error trials than in no-error trials.

The present results revealed that noise generally affected the cognitive processes of attention transition, especially in the complex situation of spatial reorientation

Figure 3.15: ChildASA-preschool: Noise effects in ΔRT.



Note. The error bars represent standard errors. N = Noise; NoN = No Noise; Rep = Repetition; Sw = Switch.

when the target changed position.

Concluding remarks

This study examined age and noise effects on intentional switching of auditory selective attention, an essential auditory cognitive process to locate and identify relevant information in complex spatial acoustic situations. The child-friendly paradigm from previous work by Loh et al. (2022a) was validated for children younger than six years old starting at the age of four years. Additional insights were gained on developmental changes within children's cognitive abilities between three- to six-year-olds. With increasing age, children were more capable of understanding and solving a categorization task presented in spatial and relocating acoustic scenes, primarily reflected in the completion rate across the age groups (cf. Table 3.3).

The general main effects of attention transition, congruency, and noise were not observed to the same extent as in the previous studies (Koch et al., 2011; Loh et al., 2022a; Nolden et al., 2019; Oberem et al., 2014). Switch costs in attention transition were observed in tendencies, while the congruency effect was found to be significant only in the error rates and the noise effect only in the response times. An explanation could be that the age effects (i.e., different behavioral patterns in the different age groups) overlayed the main effects, primarily reflected in the complexity of the conditions (switch cost, performance decrease due to incongruency, and noise effects).

Results of this study indicated that younger children (three-year-olds) did not manage to solve the paradigm reliably or only for easier conditions. Three-year-olds showed a reversed pattern of response times and error rates compared to older children. It could be assumed that the three-year-olds tried to put more effort into solving the problem but could still not reliably solve it. This could indicate mental overload and not having a coping strategy developed for the three-year-olds, but still, it was to their advantage since their error rates were comparatively low. However, this result needs further investigation since the number of three-year-olds who completed the whole experiment was significantly lower than in the other age groups.

Tendencies of different strategies during the experiment were observed across age groups, not only for conditions with higher complexity, e.g., noise in combination with the switch or incongruent trials but also in the course of the experiment (cf. Figure 3.10 and Figure 3.11).
Younger age groups (three- and four-year-olds) primarily showed a decrease in response times and constant error rates in the course of the experiment, while the older age groups (five- and six-year-olds) tended to yield higher error rates, additionally to the decrease in response times in the course of the experiment. This supports the previous findings regarding the speed-accuracy trade-off concerning the course of the experiment. It leads to the assumption that children of different age stages applied different listening strategies to solve the task, which also depends on the experiment's duration. The youngest age group investigated in this study (three-year-olds) showed a more substantial decline in response times in the course of the experiment than the older children.

These analyses resulted from those preregistered on the Open Science Framework (https:/doi.org/10.18154/RWTH-2023-00740), which were added to the submitted manuscript (Loh et al., 2024). The preregistered analyses focused on the paradigm's spatial aspects, suggesting an important role of spatial distribution and combination of sound sources. For this study, it was chosen to focus on the age and noise effects that appeared during the analyses. However, the results from the preregistered analyses and the previous work by Loh et al. (2022a) suggest that children can distinguish sound sources and information presented from different locations starting from a very young age of four. Still, they might be unable to use the natural binaural cues within the cognitive process of auditory selective attention switching to their advantage. This is consistent with the theory that children's poorer performance in a selective listening task may be caused by their immature listening strategies (Werner, 2019). Physiologically, this finding could be supported by the ongoing myelination of the connection between the two cerebral hemispheres, which is mainly responsible for interpreting binaural cues (Boothroyd, 1997; Chevalier et al., 2015). Future work might focus on the abilities and strategies of children to resolve spatial cues provided by the natural hearing processes. Since cognitive flexibility was similar across the young age of three years until reaching adulthood, it is assumed that children might have different listening strategies than adults. This study provided the first substantial information on how children locate and attend to sound and information in spatial setups.

3.5 Comparing Physiological and Cognitive Noise Effects

This chapter examines health effects in children compared to adults reflected in heart rate variability parameters induced by noise exposure within a listening experiment on auditory cognition. This study aimed to investigate the relation between physiological and cognitive noise effects in children.

3.5.1 Participants

Twenty-seven young adults (age: 20 - 26 years; M = 23.1 years, SD = 1.6 years, 12 female) and 25 children (age: 6 - 9 years; M = 7.36 years, SD = 0.81 years, 14 female) participated this study after obtaining ethical approval by the Medical Ethics Committee of the RWTH Aachen University (protocol number EK 231/22). Recruiting was performed with a cooperating primary school in Aachen, Germany. Children and their legal guardians were approached via information material distributed by the school's teachers. All participants (children represented by their legal guardians) gave signed consent before the start of the experiment and were instructed that they could cancel the experiment at any time. Children further gave verbal consent. All participants completing the experiment were compensated with a 10 €voucher for a local book store in Aachen, Germany.

Inclusion criteria were German-speaking, not diagnosed with any attention disorder, and normal hearing abilities within 25 dB[HL] as defined by WHO (1991). To ensure normal hearing, a pure-tone-audiometry for frequencies between 125 Hz and 8 kHz was conducted using a diagnostic audiometer (ear3.0, AURITEC – Medizindiagnostische Systeme GmbH, Hamburg, Germany) complying with DIN 60645-1 (DIN e.V., 2018).

3.5.2 Evaluation of heart rate variability

Heart rate variability was assessed via a Polar H7 sensor mounted on a chest band (cf. Figure 3.16) for each participant individually, and RR intervals were measured for the evaluation. The H7 sensor was connected via Bluetooth to the computer, and thus, connection issues could not be reliably avoided. To filter unusual measurements and missing RR intervals, unplausible RR intervals that exceeded a standard range of children (Paul, 2015) and adults (Gambi et al., 2017; Ostchega et al., 2011) at rest were discarded before the HRV parameters were derived.

Figure 3.16: ChildASA-noise: HRV measurement using Polar H7.

This study focused on the independent variables evaluated blockwise, i.e., noise (with two levels: no noise vs. noise) and age group (with two levels: children vs. adults). Variables requiring a trialwise evaluation were not explicitly examined here except for validating the adapted paradigm. HRV parameters were calculated for each trial over an interval of 30 s using a *moving window*, i.e., the first window would start with the beginning of the first trial and end 30 s later, and the second window then began with the beginning of the second trial. The *moving window* ended with the last 30 s window of a block. Finally, all HRV parameters of each trial were averaged per block to achieve one value for each HRV parameter per participant and block. To ensure the highest reaction time and error rate comparability, the error rates were calculated for each block, and reaction times per trial were averaged over each block.

3.5.3 Experiment design in detail

Each participant was introduced to the task individually in three training rounds with 16 trials each. The main experiment comprised a total of twelve blocks. The experiment for adults was slightly different from the experiment for children since the experiment for children had to be reduced to ensure reasonable stress and attention levels throughout the experiment with a planned duration of 90 minutes. The adults' study included three noise conditions (no noise versus SNR of +6 dB[SPL] versus SNR of 0 dB[SPL]). As a result, the noise conditions were chosen to represent a condition with general background noise and a condition where the noise could also impair the listening process. Each condition was repeated 16 times, leading to 72 trials per block. The noise condition with an SNR of +6 dB[SPL] was omitted for children. Therefore, the results of this condition from the adults' experiment were not included in the final evaluation of this study. The first 50% of children participants were presented with 16 repetitions, ending up with 48 trials per block. Nevertheless, the duration of children's participation still exceeded 2 hours and required a further adaptation of the repetition number from 16 to 12 presented, leading to 36 trials per block.

All trials from the training rounds prior to the main experiment were excluded from the analyses. Further decisions on excluding trials for statistical analyses are shown in Table 3.7.

This study includes objectives to examine the relation between physiological and cognitive noise effects within an auditory cognition paradigm on intentional switching of auditory selective attention:

1. The general paradigm was validated after adapting the number of repetitions within the children's experiment and the double-word stimuli, including the three-level congruency variable.
2. HRV parameters were examined by comparing age effects between adults and children and noise effects regarding noise versus noise-free conditions.
3. Performance over time was examined for all dependent variables to reflect increasing stress levels with the course of the experiment since participants were exposed to noise for a long duration but with interleaved exposure breaks.
4. Linear regression models between the standard dependent variables (reaction times and error rates) and the HRV parameters were analyzed to derive dependencies between HRV and behavioral measures of cognitive paradigms.

Table 3.7: ChildASA-noise: Trial exclusion criteria and outlier treatment.

Exclusion Criteria	**Excl. trials [%]**	**ANOVA**			**Lin. Reg.**
		RT	ER	HRV	$RT_{\mathbf{block}}$ & $ER_{\mathbf{block}}$
$RT < 1200$ ms	1.1	X	X		
$RT > 7227$ ms	4.7	X	X		
$RT_{\text{block}} < 1200$ ms	0			X	X
$RT_{\text{block}} > 7227$ ms	2.7			X	X
$\lvert RT \rvert > 2SD(RT)$	3.7	X	X		
Error trials	21.5	X			
Trials following an error trial	14.4	X	X		
First trial of a block	1.9	X	X		
$62 < HR_{\text{children}} < 133$ [bpm]	1.9			X	
$40 < HR_{\text{adults}} < 100$ [bpm]	8.2			X	
HRV parameter=0				X	
Total [%]		40.5	28.1	2.7	2.7

Note. Exclusion of unsual heart rates (HR) occured before the statistical removal of outliers. All other parameters were calculated afterwards. The percentage of HRV parameters equal to zero was not specified here as it was different for every HRV parameter. The percentage of excluded trials for the criterion *Trial following an error trial* excluded error trials.

3.5.4 Results and discussion

Paradigm validity

Due to missing data in the datasets, not all participants could be included in the evaluation. For the RT analysis, 14 children and 27 adult datasets were analyzed, and 24 children and 27 adult datasets were included in the ER analysis. Results of the repeated-measures ANOVA for RT and ER are summarized in Table 3.8 and descriptives of the main effects are shown in Figure 3.17. For the sake of clarity and brevity of this thesis, post-hoc analyses are added to the appendix since they did not add significantly to the insights of this dissertation (cf. Appendix B.1).

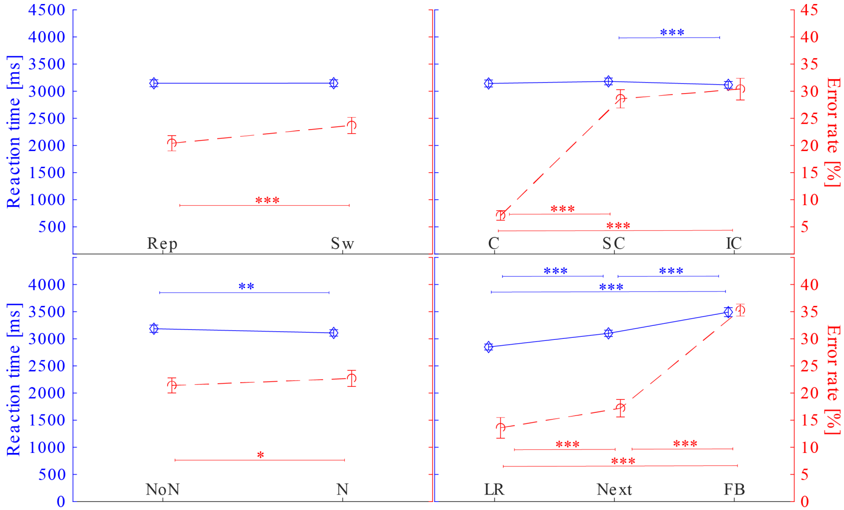
Based on the significance found within the main interaction effects, it can be generally stated that the adapted paradigm, including more prolonged stimuli and reduced condition repetition number in the presentation, was valid. The cognitive flexibility of intentional switching of auditory selective attention was reflected in transition, congruency, and target-distractor position combination. It was even more pronounced in error rates than in reaction time, which aligned with previous studies (Loh et al., 2022a, 2024; Nolden et al., 2019). Also, a notable noise

Table 3.8: ChildASA-noise: ANOVA results for RT and ER.

	Reaction time (RT)				Error rate (ER)			
	df	*F*	*p*	η_p^2	*df*	*F*	*p*	η_p^2
AT	(1, 39)	0.0	.945	.000	**(1, 49)**	**18.4**	**< .001**	**.273**
C	**(1.5, 56.6)**$^{\mathbf{a}}$	**4.1**	**.033**	**.095**	**(1.7, 83.1)**$^{\mathbf{a}}$	**186.1**	**< .001**	**.792**
N	**(1, 39)**	**7.3**	**.010**	**.158**	**(1, 49)**	**4.2**	**.047**	**.078**
TD-PC	**(1.4, 54.4)**$^{\mathbf{a}}$	**121.2**	**< .001**	**.757**	**(1.6, 77.1)**$^{\mathbf{a}}$	**143.8**	**< .001**	**.746**
AT × C	(1.7, 64.5)^a	0.2	.795	.005	**(2, 98)**	**13.3**	**< .001**	**.213**
AT × N	(1, 39)	0.0	.953	.000	(1, 49)	1.9	.176	.037
AT × TD-PC	(1.7, 64.4)^a	0.9	.389	.023	**(1.5, 73.5)**$^{\mathbf{a}}$	**7.9**	**.002**	**.139**
C × N	**(1.7, 66.1)**$^{\mathbf{a}}$	**3.4**	**.045**	**.081**	(2, 98)	0.8	.452	.016
C × TD-PC	(3.0, 115.5)^a	1.9	.136	.046	**(3.0, 145.0)**$^{\mathbf{a}}$	**74.4**	**< .001**	**.603**
N × TD-PC	(1.5, 58.9)^a	3.1	.066	.074	(2, 98)	0.9	.429	.017
AT × C × N	(2, 78)	1.6	.209	.039	(2, 98)	1.4	.263	.027
AT × C × TD-PC	(2.9, 113.7)^a	0.2	.918	.004	**(4, 196)**	**7.6**	**< .001**	**.134**
AT × N × TD-PC	(2, 78)	1.6	.217	.038	(2, 98)	1.1	.346	.021
C × N × TD-PC	(2.9, 114.9)^a	0.6	.617	.015	(3.3, 163.9)^a	0.5	.670	.011
AT × C × N × TD-PC	**(3.2, 123.7)**$^{\mathbf{a}}$	**3.4**	**.017**	**.081**	(3.3, 164.0)^a	0.6	.631	.012
AG	**(1, 39)**	**102.3**	**< .001**	**.724**	**1, 49)**	**39.4**	**< .001**	**.446**
AG × AT	(1, 39)	3.9	.055	.091	**(1, 49)**	**11.8**	**.001**	**.194**
AG × C	(1.4, 56.6)^a	1.6	.218	.039	**(1.7, 83.1)**$^{\mathbf{a}}$	**14.2**	**< .001**	**.225**
AG × N	(1, 39)	1.5	.229	.037	(1, 49)	0.2	.663	.004
AG × TD-PC	**(1.4, 54.4)**$^{\mathbf{a}}$	**3.7**	**.046**	**.087**	**(1.6, 77.1)**$^{\mathbf{a}}$	**18.8**	**< .001**	**.277**
AG × AT × C	(1.7, 64.5)^a	0.3	.666	.009	**(2, 98)**	**20.7**	**< .001**	**.297**
AG × AT × N	(1, 39)	1.1	.311	.026	(1, 49)	0.9	.340	.019
AG × AT × TD-PC	(1.7, 64.4)^a	0.5	.592	.012	**(1.5, 73.5)**$^{\mathbf{a}}$	**7.3**	**.003**	**.130**
AG × C × N	(1.7, 66.1)^a	1.1	.332	.027	(2, 98)	0.5	.626	.010
AG × C × TD-PC	(3.0, 115.5)^a	1.6	.183	.041	**(3.0, 145.0)**$^{\mathbf{a}}$	**15.1**	**< .001**	**.236**
AG × N × TD-PC	(1.5, 58.9)^a	0.8	.436	.019	(2, 98)	0.3	.716	.007
AG × AT × C × N	(2, 78)	1.5	.228	.037	(2, 98)	0.6	.567	.011
AG × AT × C × TD-PC	(2.9, 113.7)^a	0.5	.679	.013	**(4, 196)**	**8.6**	**< .001**	**.149**
AG × AT × N × TD-PC	(2, 78)	1.1	.333	.028	(2, 98)	2.0	.143	.039
AG × C × N × TD-PC	(2.9, 114.9)^a	0.3	.787	.009	(3.3, 163.9)^a	0.7	.555	.014
AG × AT × C × N × TD-PC	(3.2, 123.7)^a	2.1	.103	.051	(3.3, 164.0)^a	0.6	.667	.011

Note. AG = age group; N = noise; AT = attention transition; C = congruency; TD-PC = target-distractor-position combination. Significant effects are indicated in bold. ^{a} Greenhouse-Geisser correction applied due to violation of assumption.

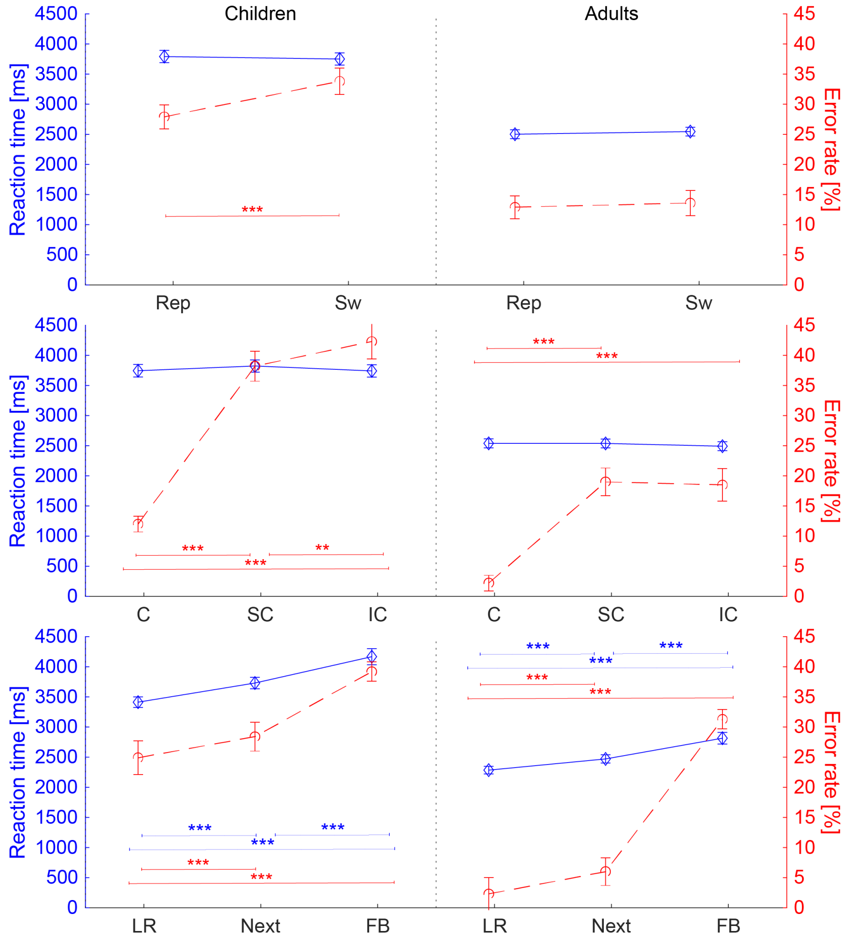
Figure 3.17: ChildASA-noise: Main effects for RT and ER.



effect, comprising lower reaction times and higher error rates in noisy conditions than in noise-free conditions, was observed. Regarding congruency, the error rate yielded significant differences between congruent and semi-congruent, as well as incongruent trials. On the contrary, semi-congruent and incongruent trials were not significantly different. It is noteworthy that the effect in reaction times was observed in a reversed manner, indicating a difference in cognitive processing, namely that more complex configurations of the stimuli required more cognitive capacity but at the same time led to a higher decisiveness.

Comparing the age groups adults and children (cf. Figure 3.18), no interaction effect with noise was found, the interaction effects with attention transition were less pronounced, and the interaction effects with congruency was primarily observed in error rates, as in previous studies including children participants(cf. Loh & Fels, 2021; Loh et al., 2024). The interaction effects of age group and target-distractor position combination revealed significant effects between the categories as expected and found in previous studies including adult participants (cf. Koch et al., 2011; Oberem et al., 2014). Contrary to the previous study with children participants within the same age range (Loh et al., 2022a), the children's effect introduced by the target-distractor position combination was comparable to adults' effects. The previous assumption that children could not process spatial auditory cues to the same extent as adults could not be verified in this study

Figure 3.18: ChildASA-noise: Interaction effects of age groups for RT and ER.



and should be further examined in future studies. The possibility of controlling noise presentation in the paradigm, which was the main difference between the two paradigms, might mask the spatial benefits.

The adapted paradigm could reflect the cognitive performances of intentional auditory selective attention switching for children and adults. Even though the stimuli consisted of two words, congruency effects did not change significantly.

It still revealed a two-level complexity (congruent vs incongruent) instead of three levels (congruent, semi-congruent, and incongruent). However, using more prolonged stimuli will benefit future studies on acoustic properties that reflect more prolonged effects than the usual stimuli length of 500 ms.

Analysis of Heart Rate Variability

For the ANOVAs on HRV parameters, 25 children and 27 adults' datasets were evaluated. Datasets were excluded due to a significant amount of missing data (more than 25%). Repeated measures ANOVA results for all HRV parameters are shown in Table 3.9.

Table 3.9: ChildASA-noise: ANOVA results for HRV parameters.

HRV parameter		*df*	*F*	*p*	η_p^2
meanHR	N	(1, 50)	0.8	.366	.016
	AG	**(1, 50)**	**38.7**	**< .001**	**.436**
	N × AG	**(1, 50)**	**5.7**	**.021**	**.102**
meanNN	N	(1, 50)	0.3	.573	.006
	AG	**(1, 50)**	**31.8**	**< .001**	**.389**
	N × AG	**(1, 50)**	**4.4**	**.040**	**.081**
StdHR	N	(1, 50)	0.3	.600	.006
	AG	**(1, 50)**	**44.0**	**< .001**	**.468**
	N × AG	(1, 50)	1.9	.176	.036
StdNN	N	(1, 50)	0.0	.914	.000
	AG	(1, 50)	1.7	.197	.033
	N × AG	**(1, 50)**	**4.9**	**.031**	**.090**
RMSSD	N	(1, 50)	0.6	.437	.012
	AG	(1, 50)	0.0	.918	.000
	N × AG	(1, 50)	0.3	.593	.006
SD1	N	(1, 50)	0.2	.648	.004
	AG	(1, 50)	0.5	.505	.009
	N × AG	(1, 50)	3.6	.064	.067
SD2	N	(1, 50)	0.3	.567	.007
	AG	(1, 50)	2.9	.096	.054
	N × AG	**(1, 50)**	**6.6**	**.014**	**.116**
SD1/SD2	N	(1, 50)	1.8	.189	.034
	AG	(1, 50)	1.5	.222	.030
	N × AG	(1, 50)	0.5	.467	.011

Note. AG = age group; N = noise. Significant effects are indicated in bold.

The results revealed no significant noise effect for all HRV parameters. Significant age group effects were found for meanHR, meanNN, and StdHR. Since the measure meanNN is considered as the inverse value of meanHR (i.e., $\frac{1}{meanHR}$), it was expected that results would be similar and, thus, redundant. Significant interaction effects of age group and noise were found for meanHR, meanNN, StdNN, and SD2. Post-hoc analyses revealed mainly differences between adults and children in each (noise and noise-free) condition for meanHR and meanNN. For StdNN, no significant difference was found in the post-hoc analyses. Regarding SD2, differences between noise and noise-free conditions were found for adults, and an age difference was found in the noisy condition.

Table 3.10: ChildASA-noise: Post-hoc test results for HRV ANOVA analyses.

	G_1	G_2	p (G_1 - G_2)	M_{diff}	**[95 %-CI]**
meanHR				[bpm]	[bpm]
Children	NoN	N	.026	-1076	[-2.020, -0.132]
Adults	NoN	N	.292	0	[-0.427, 1.390]
NoN	**Children**	**Adults**	**<.001**	**14**	**[9.315, 19.257]**
N	**Children**	**Adults**	**<.001**	**16**	**[10.999, 20.687]**
meanNN				[ms]	[ms]
Children	NoN	N	.069	0.008	[-0.001, 0.018]
Adults	NoN	N	.273	-0.005	[-0.014, 0.004]
NoN	**Children**	**Adults**	**<.001**	**-0.122**	**[-0.169, -0.075]**
N	**Children**	**Adults**	**<.001**	**-0.135**	**[-0.181, -0.090]**
StdNN					
Children	NoN	N	.149	0.003	[-0.001, 0.006]
Adults	NoN	N	.099	-0.003	[-0.006, 0.001]
NoN	Children	Adults	.099	0.009	[-0.002, 0.021]
N	Children	Adults	.429	0.004	[-0.006, 0.013]
SD2				[ms]	[ms]
Children	NoN	N	.175	0.003	[-0.001, 0.007]
Adults	**NoN**	**N**	**.028**	**-0.004**	**[-0.008, 0.000]**
NoN	Children	Adults	.175	0.003	[-0.001, 0.007]
N	**Children**	**Adults**	**.028**	**-0.004**	**[-0.008, 0.000]**

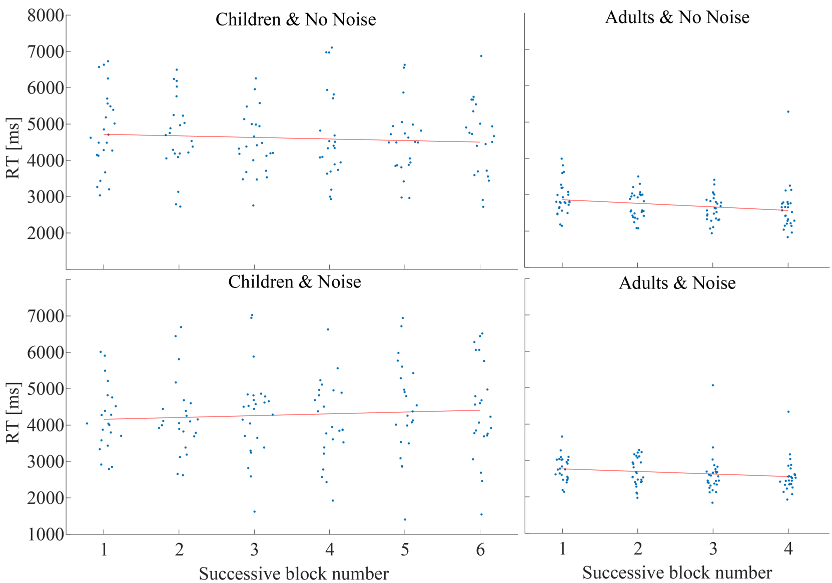
Note. Significant effects are indicated in bold. N = Noise; NoN = No Noise.

These findings pointed towards an age effect between children of six to 10 years and young adults within the physiological stress parameters, which aligned with previous research indicating that heart rates generally differ between children and adults (cf. Gambi et al., 2017; Ostchega et al., 2011; Paul, 2015) with higher meanHR for children than adults. Furthermore, results indicated that SD2 was sensitive to measure differences within noise specifically.

Performance over Time

Linear regressions were conducted for each dependent variable (RT, ER, and all HRV parameters) to investigate children's and adults' performance over time within noise and noise-free conditions. Age group (adult vs. children) and noise conditions were split, so four conditions were investigated for each dependent variable and the independent variable of block number. The increasing block number represented the progression of time. The linear regression results are presented in Table 3.11. An exemplary representation of the regression of meanHR in dependncy of successive block number is shown in Figure 3.19.

Figure 3.19: ChildASA-noise: RT over time in the course of the experiment.



Significant regressions were mainly found for children in the dependent variables

Table 3.11: ChildASA-noise: Linear regression results considering the course of time.

Dependent variable	**Factors**	*df*	*F*	*p*	R^2	c_{reg}
RT_{block}	Children & No Noise	(1, 137)	0.7	.402		
	Children & Noise	(1, 139)	0.8	.386		
	Adults & No Noise	**(1, 105)**	**5.7**	**.018**	**.052**	**-98 ms**
	Adults & Noise	(1, 104)	3.3	.074		
ER_{block}	Children & No Noise	(1, 137)	0.3	.566		
	Children & Noise	(1, 139)	2.9	.090		
	Adults & No Noise	(1, 105)	2.3	.129		
	Adults & Noise	(1, 104)	0.0	.955		
$meanHR$	Children & No Noise	**(1, 120)**	**5.5**	**.021**	**.044**	**-1 bpm**
	Children & Noise	(1, 127)	2.2	.140		
	Adults & No Noise	(1, 97)	0.9	.344		
	Adults & Noise	(1, 96)	1.1	.294		
$meanNN$	Children & No Noise	**(1, 120)**	**7.2**	**.008**	**.056**	**0.008**
	Children & Noise	(1, 127)	3.0	.088		
	Adults & No Noise	(1, 97)	0.8	.373		
	Adults & Noise	(1, 96)	1.1	.301		
$StdHR$	Children & No Noise	(1, 120)	3.6	.059		
	Children & Noise	(1, 127)	3.1	.080		
	Adults & No Noise	(1, 97)	0.0	.877		
	Adults & Noise	(1, 96)	0.7	.396		
$StdNN$	Children & No Noise	**(1, 120)**	**8.5**	**.004**	**.066**	**0.003**
	Children & Noise	**(1, 127)**	**5.7**	**018**	**.043**	**0.002**
	Adults & No Noise	(1, 97)	0.1	.743		
	Adults & Noise	(1, 96)	1.1	.299		
$RMSSD$	Children & No Noise	**(1, 120)**	**8.4**	**.004**	**.066**	**0.001**
	Children & Noise	(1, 127)	3.4	.069		
	Adults & No Noise	(1, 97)	1.4	.247		
	Adults & Noise	(1, 96)	0.1	.807		
$SD1$	Children & No Noise	**(1, 120)**	**8.4**	**.004**	**.066**	**0.003**
	Children & Noise	(1, 127)	3.2	.075		
	Adults & No Noise	(1, 97)	0.3	.600		
	Adults & Noise	(1, 96)	0.5	.495		
$SD2$	Children & No Noise	**(1, 120)**	**7.6**	**.007**	**.060**	**0.004**
	Children & Noise	**(1, 127)**	**6.1**	**.015**	**.046**	**0.003**
	Adults & No Noise	(1, 97)	1.1	.304		
	Adults & Noise	(1, 96)	1.1	.301		
$SD1/SD2$	Children & No Noise	(1, 120)	2.4	.123		
	Children & Noise	(1, 127)	0.0	.926		
	Adults & No Noise	(1, 97)	1.5	.222		
	Adults & Noise	(1, 96)	0.1	.734		

Note. Significant effects are indicated in bold.

meanHR, meanNN, StdNN, RMSSD, SD1, and SD2. Significant relations were found for no noise conditions, but for the parameters StdNN and SD2, significant relations were also found for noisy conditions. In RT, a significant relation was found for adults within noise-free conditions. Investigating the regression coefficients, it was questionable whether the found relations were reasonable since the declines were relatively small. As shown in Figure 3.19, it was observed that the variation of children's results was relatively high so that the decline of -1 bpm was neglectable in terms of reasonable numbers for heart rate changes though it was statistically significant. However, a tendency to increase values for the HRV parameters was observed for the children, which could be interpreted as relaxation with the progression of the experiment. An explanation for this could be the increasing familiarity with the task, resulting in less tension with the progression of time.

Summarizing the findings regarding the stress level reflected in the cognitive and physiological parameters, it was found that a tendency of relaxation was reflected in the HRV parameters. However, this observation must be treated with caution since the changes in HRV were in the range of regular variations of heart rate that cannot be interpreted as stress reactions without limitations.

Relation between cognitive and HRV parameters

For the linear regression analyses on the cognitive parameters (RT and ER), as indicators for cognitive flexibility, and the HRV parameters, 25 children and 27 adults datasets were included in the evaluation. Datasets were excluded in case the amount of missing data exceeded 25%. Results of the linear regression analyses are summarized in Table 3.12.

Linear regression analyses between ER and the HRV parameters were all found insignificant. Results revealed mainly significant regressions between RT and HRV parameters. In RMSSD and SD1, significant effects were found for adults. In RMSSD, the effect was found for noisy and noise-free conditions, indicating an increase in RT was related to an increase in RMSSD. However, the same effect for SD1 was only found for noise-free conditions. Notably, the SD1/SD2 not only revealed a significant relation to RT for adults in both noisy and noise-free conditions but also yielded a significant effect for children in the noise-free conditions.

These observations indicate a relation between RT and specific HRV, leading to the assumption that changes in the HRV were connected to RT. As cognitive functions require time to process and execute, it can be assumed that increasing RT indicates a higher cognitive load, as it was explained in previous studies (cf., Loh et al., 2022a; Nolden et al., 2019; Oberem et al., 2018). An increase

Table 3.12: ChildASA-noise: Linear regression results on cognitive parameters in relation to HRV parameters.

	AG	N	Reaction time (RT)					Error rate (ER)				
			df	*F*	*p*	R^2	**Stand. Coeff**	*df*	*F*	*p*	R^2	**Stand. Coeff**
meanNN	Children	NoN	(1, 23)	0.1	.734	.005	-0.072	(1, 23)	2.9	.100	.114	-0.337
	Children	N	(1, 23)	0.0	.845	.002	0.041	(1, 23)	2.8	.108	.109	-0.329
	Adults	NoN	(1, 25)	1.5	.228	.058	0.240	(1, 25)	0.8	.389	.030	-0.173
	Adults	N	(1, 25)	2.1	.156	.079	0.281	(1, 25)	0.0	.949	.000	-0.013
meanHR	Children	NoN	(1, 23)	0.2	.669	.008	0.090	(1, 23)	2.5	.129	.097	0.312
	Children	N	(1, 23)	0.0	.906	.001	-0.025	(1, 23)	2.5	.128	.098	0.312
	Adults	NoN	(1, 25)	1.0	.333	.038	-0.194	(1, 25)	0.7	.404	.028	0.167
	Adults	N	(1, 25)	1.6	.211	.062	-0.249	(1, 25)	0.0	.928	.000	-0.018
StdNN	Children	NoN	(1, 23)	0.1	.778	.004	-0.059	(1, 23)	0.6	.431	.027	-0.165
	Children	N	(1, 23)	0.9	.360	.037	-0.191	(1, 23)	0.2	.635	.010	-0.100
	Adults	NoN	(1, 25)	3.0	.096	.107	0.327	(1, 25)	0.0	.828	.002	0.044
	Adults	N	(1, 25)	0.8	.369	.032	0.180	(1, 25)	0.3	.618	.010	0.101
StdHR	Children	NoN	(1, 23)	0.2	.669	.008	0.090	(1, 23)	2.5	.129	.097	0.312
	Children	N	(1, 23)	0.0	.906	.001	-0.025	(1, 23)	2.5	.128	.098	0.312
	Adults	NoN	(1, 25)	1.0	.333	.038	-0.194	(1, 25)	0.7	.404	.028	0.167
	Adults	N	(1, 25)	1.6	.211	.062	-0.249	(1, 25)	0.0	.928	.000	-0.018
RMSSD	Children	NoN	(1, 23)	1.6	.220	.065	-0.254	(1, 23)	1.4	.250	.057	-0.239
	Children	N	(1, 23)	1.5	.235	.061	-0.247	(1, 23)	0.3	.578	.014	-0.117
	Adults	NoN	**(1, 25)**	**14.1**	**.001**	**.361**	**0.601**	(1, 25)	1.2	0.310	.041	0.203
	Adults	N	**(1, 25)**	**6.6**	**.016**	**.209**	**0.457**	(1, 25)	0.9	0.4	.033	0.182
SD1	Children	NoN	(1, 23)	1.2	.285	.049	-0.222	(1, 23)	1.7	.203	.069	-0.263
	Children	N	(1, 23)	1.4	.246	.058	-0.241	(1, 23)	1.1	.306	.046	-0.213
	Adults	NoN	**(1, 25)**	**8.2**	**.008**	**.248**	**0.498**	(1, 25)	0.2	.679	.007	0.083
	Adults	N	(1, 25)	4.1	.055	.140	0.374	(1, 25)	0.3	.621	.010	0.100
SD2	Children	NoN	(1, 23)	0.0	.984	.000	0.004	(1, 23)	0.4	.556	.015	-0.124
	Children	N	(1, 23)	0.7	.426	.028	-0.167	(1, 23)	0.1	.792	.003	-0.056
	Adults	NoN	(1, 25)	0.4	.533	.016	0.125	(1, 25)	0.0	.922	.000	-0.020
	Adults	N	(1, 25)	0.0	.898	.001	0.026	(1, 25)	0.2	.677	.007	0.084
SD1/SD2	Children	NoN	**(1, 23)**	**8.0**	**.010**	**.258**	**-0.508**	(1, 23)	3.3	.084	.124	-0.353
	Children	N	(1, 23)	1.5	.229	.062	-0.250	(1, 23)	2.8	.105	.110	-0.332
	Adults	NoN	**(1, 25)**	**17.7**	**<.001**	**.414**	**0.644**	(1, 25)	1.4	.254	.052	0.227
	Adults	N	**(1, 25)**	**9.7**	**.005**	**.279**	**0.528**	(1, 25)	0.5	.470	.021	0.145

Note. AG = age group; N = noise; NoN = no noise; Significant effects are indicated in bold.

in HRV indicates relaxation. These two facts, combined with higher RT and higher cognitive load in relation to higher HRV and higher relaxation, appeared unintuitive. It was expected that a higher cognitive load would lead to stress reactions rather than relaxation.

Taken together, it is questionable whether stress in a cognitive paradigm with continuously changing conditions, though the noise is continuously presented, can be measured using HRV parameters. Present findings indicated that noise effects in auditory cognitive functions, such as the intentional switching of auditory selective attention, did not translate directly into bodily reactions to noise, interpreted as stress reactions. HRV parameters were intended for long-term stress reactions assessed over long-term measurements (more than 7-8 hours) (cf. Baek et al., 2015; Kim et al., 2021; Lackner et al., 2020). Though validated for short time frames, it appeared not to be sensitive to short-term noise exposure as it was applied in this study with the noise within blocks (less than 15 minutes).

Concluding remarks

Overall, the adapted paradigm on intentional auditory selective attention switching with more prolonged (two-word) stimuli and less repetition of conditions was validated, and the task was found to be not more complex than before.

Assessing HRV parameters might add insights into children's susceptibility to noise within paradigms on auditory cognitive functions. A higher susceptibility to noise is hereby regarded as a higher risk of being affected by noise. Previous research (Cohen et al., 1986; Stansfeld & Clark, 2015) showed that children may be more susceptible to environmental exposures, such as noise, than adults because of several factors, e.g., less cognitive capacity to understand the environmental exposures or having a less well-developed repertoire to cope with the challenges coming along with the exposures. However, a negative impact of noise could not be shown to the same extent in this work compared to previous reviews provided by, for example, Stansfeld and Clark (2015). In this work, differences of HRV parameters were found between children and adults in noisy compared to noise-free conditions. The variation of differences indicated different levels of susceptibility to noise by children compared to adults. Still, it cannot be applied interchangeably with the behavioral parameters RT and ER, as they were traditionally used in the paradigm to examine intentional switching of auditory selective attention. The paradigm task and setup used in this study were not sufficiently suitable for assessing the noise effect on HRV parameters to gain insights into stress levels. Specific HRV parameters, such as SD2, can reveal a certain level of noise susceptibility to noise when examining the change of this parameter, for example, over time. A higher value of the HRV parameter SD2 indicated a better ability to adapt to the situation, which represents a stronger relaxation process. In general, the results of the linear regression analyses indicated a tendency for both age groups to adapt better to the task when there was no noise present in the background. Significant linear regressions were, however, primarily found for children, so that a stronger attenuation of the stress response on noise by the children than adults could be assumed, which further indicates a higher noise susceptibility of children compared to adult, i.e., children being more reactive noise exposure. It remains to future work, to investigated this relationship with a better suited paradigm.

3.6 Insights on Children's Auditory Cognition in Noise

A child-appropriate paradigm was developed to examine the intentional switching of auditory selective attention for children. The requirements for a child-appropriate paradigm comprised a categorization consisting of at least two categories to provide results on attention switch, relevant information selection, and a motivation system including positive and negative feedback, which yielded consistent results with previous studies on intentional auditory selective attention switching. The stimuli used in the paradigm can either be one-word stimuli, i.e., an animal name, or two-word stimuli, i.e., the words *big* or *small* and an animal name. The extension of the stimuli did not raise the complexity of the categorization task, and the congruency (i.e., representation of a selection of relevant information) remained consistent.

The newly developed paradigm for children was validated for ages four to ten. Starting from age five, comparable cognitive flexibility of intentional auditory selective attention switching as adults was observed, leading to the assumption that this cognitive ability is generally fully developed at this point. The turning point of the development of intentional switching of auditory selective attention tended to be at the age of four, revealing increasing performance with increasing age. However, the sample size of three years old was too small to state this without limitations.

Significant noise effects on intentional switching of auditory selective attention were observed for children and adults alike. However, children revealed a higher susceptibility to noise than adults. Higher noise susceptibility is hereby interpreted as a stronger influence on cognitive ability induced by noise, revealing a higher reactivity when noise is present. Tendencies of a speed-accuracy trade-off were observed for children of all ages in all studies conducted within the scope of this dissertation. Additional investigations on attention retention throughout the listening experiments and the reaction time in error versus non-error trials were found to depend on the presence of noise and the progression of time. Children starting at the age of five especially revealed an increasing tendency toward speed-accuracy trade-off with the increasing amount of time spent in the listening experiment. Furthermore, different strategies to deal with noisy conditions were found between the age groups three to six.

The ability to process spatial information in a paradigm on intentional switching of auditory selective attention was generally in line with previous research. The cognitive performances depended on the complexity of the spatial configuration of the target and distractor and the corresponding availability of information provided through binaural cues, such as interaural time and level differences. However, age differences between adults and children were inconclusive regarding spatial abilities across the studies conducted in the scope of this dissertation. The first study, *ChildASA - validation study*, revealed that children did not benefit from the spatial cues to the same extent as adults. The *ChildASA - noise study*, on the contrary, found that children were able to process spatial cues in the same manner as adults. Their performance increased, as did adults', with the decrease in the complexity of spatial configuration of target and distractor. Generally, there are two explanations for this phenomenon: On the one hand, the possibility of controlling noise conditions might have masked the spatial effect as given in the *ChildASA - validation study*. On the other hand, developmental differences in spatial processing of children at the same age could be the reason since the differences in the benefit of available spatial cues were also observed to a certain extent for very young children between three and six years old. Though children can be categorized into age groups, their cognitive-developmental state can still vary a lot based on other factors, e.g., socio-economic status, etc., as it was found in the *ChildASA - preschool study* and the *ChildASA - noise study* and the *ChildASA - validation study* were conducted in different areas of Aachen with different general socio-economic status. Since the sample size of the age groups with regard to the variable of socio-economic status in the *ChildASA - preschool study* was too small to derive final conclusions, this observation had to be interpreted with caution.

Dependencies between cognitive behavior in terms of intentional switching of auditory selective attention and physiological reactions measured via heart rate variability as a result of noise-induced stress could not be observed. Though the heart rate variability measures were promising for assessing physiological reactions easily and unobtrusively, results indicated that HRV parameters and cognitive behavior parameters (here RT and ER) could not be used interchangeably. However, HRV parameters, such as SD2, can add insights into noise susceptibility, i.e., degree of reactivity to noise, which was observed as bodily reaction,s e.g., in terms of a stronger parasympatic response, especially when investigating differences between adults and children.

4 Summary and Outlook

This dissertation aimed to assess noise effects on children's auditory cognition and health by bringing close-to-real-life acoustic scenes into listening experiments on intentional auditory selective attention switching for children between three and ten years old in a controlled and reproducible manner with high ecological validity.

The first objective was to determine the general noise situation in German pre- and primary schools by adding children's perspectives through measurements using head and torso simulators (HATS), considering children's environment and activities, as well as child-appropriate noise ratings. Room acoustic and long-term in-situ noise measurements using HATSs of varying anthropometric sizes were examined to assess differences in the physical signal and their influences on children's noise perception.

In general, acoustic conditions in German classrooms and playrooms were found to be improvable by focusing on high-frequency sound absorption and optimizing reverberation times and speech intelligibility, enhancing the auditory environment for children as an essential factor for learning and communication.
In long-term noise measurements, psychoacoustic parameters provided deeper insights into noise assessment than traditional sound pressure level (SPL)-based parameters. Differences between adult and child HATS were particularly pronounced in higher frequencies, where psychoacoustic parameters like sharpness played a crucial role. Furthermore, this observation was strengthened within the room acoustic analyses, focusing on the frequency bands from 250 Hz to 16 kHz. Differences were mainly found in higher frequencies and parameters such as C50, D50, and TS, which tended to be more responsive to anthropometric differences between HATSs than reverberation parameters like T20, T30, and EDT.
The STI analysis on HATS measurements revealed that single-value measures might obscure significant differences in the higher frequency spectrum when averaging across frequency bands to achieve the final single-value measure. The binaural version of STI displayed differences between adult and child HATS, which were further explored through modulation transfer indices (MTIs), offering a more detailed frequency-dependent analysis. This approach helped capture

distinctions between HATS of varying anthropometric sizes primarily found in the higher frequency bands.
These results were also reflected in the analyses of psychoacoustic and SPL-based parameters across frequency bands derived from long-term measurements, showing how noise parameters vary between adults and children, particularly in higher frequencies. Frequency response curves obtained from measurements with adult and child HATS, alongside omnidirectional reference microphones, highlighted these differences, eventually reflected in the psychoacoustic parameters.

The advantages of psychoacoustic parameters over level-based ones were further supported by their stronger correlation with subjective noise ratings by children. Using the INCH questionnaire, this dissertation demonstrated that psychoacoustic measures aligned better with children's subjective noise ratings than SPL-based metrics. Furthermore, differences between adults and children's noise ratings were observed, with children experiencing or expressing stronger emotional reactions towards the assessed sound types than adults. Differences between preschool (three to six years old) and primary school (six to ten years old) children were further found with young children, i.e., preschool-aged children, to be more susceptible to high-frequency sounds reflected in stronger negative emotional reactions reported in the INCH questionnaire, which aligned with previous research.

Different educational activities in pre-primary and primary schools were considered influential for existing noise levels and were, therefore, systematically explored in this dissertation to understand their impact. By analyzing these activities from an acoustic perspective, focusing on adult-child interactions and communication pathways, the study found that activities with more structured communication patterns tended to lower noise levels. However, differences in the higher frequency bands had a relatively minor impact when addressing activities, though differences between adult and child HATS were still observed. These findings add valuable insights into how children might perceive noise compared to adults.

This dissertation showed that anthropometric differences between child and adult HATS can significantly impact measured acoustic parameters, especially in educational buildings and contexts. While this dissertation offers essential findings on the differences within the physical sound signals measured using HATSs of varying anthropometric sizes and corresponding acoustic parameters, it emphasizes the need for further research to understand how children perceive acoustic and noise parameters fully. Since most acoustic parameters were developed for adults, their validity for assessing children's noise perception remains uncertain and is subject to future research.

The second objective was to investigate the noise effects on children's auditory cognition and health based on insights from the previously assessed acoustic conditions in German pre- and primary schools. This dissertation developed a child-appropriate paradigm for studying auditory selective attention switching in spatial and noisy acoustic environments using close-to-real-life sound reproduction methods, a motivation system, and gamification of the paradigm.

The newly developed paradigm was validated for children aged four to ten. It required them to solve a categorization task to assess intentional auditory selective attention switching and incorporate a motivation system with immediate feedback. By age five, children displayed similar cognitive flexibility in intentional auditory selective attention switching to adults. The development of this ability appears to begin around age four, with performance improving as children age. However, the sample size for three-year-olds was too small in the study conducted within this dissertation to draw definitive conclusions.

Significant noise effects on intentional auditory selective attention switching were observed in children and adults, with children being more sensitive to noise, i.e., the noise will lead to stronger reactions regardless of whether beneficial or detrimental. A speed-accuracy trade-off was noted, especially in children aged five and older, with decreased reaction times and increased error rates. Additionally, children employed different strategies to cope with noisy conditions, varying by age group, particularly between ages three and six.

This dissertation further examined how children processed spatial information in auditory cognition tasks, especially in intentional switching of auditory selective attention. Results, aligning with previous research, showed that performance depended on the complexity of spatial configurations and the availability of binaural cues, such as interaural time differences. However, age-related differences were inconclusive. In the initial validation study, children benefitted less from spatial cues than adults, but later noise studies showed that children's ability to process spatial cues improved as task complexity decreased. Possible explanations for the mixed results include either the potential to control noise presentation within the experiments by the participants that might have introduced masking of spatial effects in the validation study or developmental variability in spatial processing among children, which leads to the conclusion that categorizing children in terms of the actual age cannot sufficiently reflect their developmental stage. Developmental differences in auditory processing were found to show tendencies towards linkages to socio-economic status, as shown in the preschool studies conducted in different regions with varying socio-economic backgrounds. Nevertheless, these findings should be treated with caution due to the small sample sizes.

Lastly, no clear relationship was found between behavioral response measures (reaction times and error rates) in auditory selective attention tasks and physiological responses (e.g., heart rate variability (HRV)) under noise-induced stress. While HRV measures were promising as non-intrusive indicators of noise susceptibility, meaning reflecting different levels of reactivity to noise, e.g., in terms of cognitive effects representing noise-induced stress levels, they could not fully substitute or represent behavioral response metrics from cognitive paradigm design. Nonetheless, particular HRV parameters, such as SD2, may offer additional insights into noise susceptibility, i.e., as distinct levels of bodily reactivity to noise, especially when comparing adults and children.

In conclusion, this dissertation provides significant insights into children's auditory cognition under noise exposure, especially in educational environments. Furthermore, it revealed significant differences in measured signals when applying head and torso simulators of varying anthropometric sizes. It highlights the need for child-specific acoustic standards and further research into developmental and socio-economic factors influencing auditory perception.

A

Appendix - Towards Child-Appropriate Noise Assessment Methods

A.1 Complete overview on RA parameters

Table A.1: Complete RA parameters derived from the reference microphone per measured room.

Room	N_{pos}	T_{20} [s]			T_{30} [s]			EDT [s]			C_{50} [dB]			D_{50} [%]			T_S [s]		
		BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.
CR01	11	0.479	0.485	1.38	0.499	0.490	1.40	0.424	0.444	1.51	7.8	6.7	0.51	83.2	81.8	0.85	0.027	0.028	1.89
CR02	11	0.567	0.601	1.08	0.588	0.601	1.15	0.528	0.579	1.04	4.9	4.0	0.87	74.3	70.8	0.97	0.037	0.040	1.07
CR03	12	0.928	1.035	1.37	0.968	1.047	1.43	0.884	1.021	1.34	1.6	0.3	0.08	58.1	51.9	0.79	0.062	0.071	1.39
CR04	10	0.954	1.095	1.52	0.970	1.103	1.52	0.900	1.059	1.52	1.9	0.5	-0.05	58.6	52.7	0.71	0.062	0.071	1.61
CR05	10	0.522	0.580	1.04	0.544	0.592	0.98	0.422	0.456	1.45	7.5	6.3	0.53	82.9	80.1	0.85	0.027	0.030	1.73
CR06	11	0.529	0.618	0.94	0.536	0.620	0.95	0.524	0.628	0.91	5.6	3.9	0.99	76.5	70.4	1.01	0.034	0.040	0.96
CR07	10	0.531	0.584	1.06	0.612	0.579	1.37	0.538	0.584	0.99	4.5	3.7	0.95	72.0	69.2	1.00	0.039	0.043	1.03
CR08	8	0.503	0.557	0.90	0.511	0.564	0.91	0.486	0.543	0.95	6.6	5.5	0.91	80.3	77.4	1.00	0.030	0.033	1.06
PR09	12	0.459	0.502	1.40	0.466	0.503	1.41	0.443	0.504	1.33	6.5	4.7	0.56	78.5	74.0	0.87	0.032	0.037	1.48
PR10	6	0.391	0.397	1.33	0.415	0.412	1.40	0.374	0.363	1.46	8.8	8.5	0.61	84.8	86.5	0.88	0.026	0.024	1.51
PR11	10	0.340	0.338	0.94	0.351	0.336	0.94	0.306	0.325	1.05	10.4	9.5	0.91	89.7	88.5	0.99	0.020	0.021	1.17
PR12	8	0.464	0.483	1.46	0.475	0.490	1.44	0.458	0.482	1.44	6.4	5.0	0.50	78.1	75.2	0.85	0.033	0.037	1.58
PR13	11	0.543	0.578	1.13	0.562	0.584	1.06	0.479	0.544	1.46	6.6	4.9	0.47	78.3	74.3	0.83	0.032	0.035	1.66
PR14	10	0.439	0.412	1.97	0.448	0.416	1.95	0.440	0.449	1.94	7.3	5.9	0.33	78.5	77.4	0.76	0.034	0.034	2.15
PR15	10	0.465	0.483	1.22	0.477	0.485	1.18	0.419	0.451	1.46	7.6	6.2	0.54	82.5	80.2	0.85	0.027	0.031	1.71
PR16	4	0.462	0.462	1.36	0.469	0.465	1.37	0.440	0.437	1.29	7.0	6.7	0.68	78.3	78.4	0.93	0.032	0.032	1.27
PR17	12	0.524	0.460	1.41	0.532	0.464	1.37	0.483	0.458	1.46	6.7	6.4	0.53	78.9	79.8	0.84	0.032	0.030	1.72
PR18	12	0.658	0.754	1.45	0.678	0.776	1.39	0.622	0.725	1.52	4.6	2.8	0.32	69.9	64.9	0.79	0.044	0.049	1.65

Table A.2: Complete RA parameters derived from the A-$HATS_{av}$ per measured room.

		T_{20} [s]			T_{30} [s]			EDT [s]			C_{50} [dB]			D_{50} [%]			T_S [s]		
Room	N_{pos}	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.
CR01	11	0.480	0.477	1.35	0.554	0.490	1.49	0.433	0.471	1.44	7.9	6.8	0.53	83.6	81.8	0.86	0.025	0.026	1.84
CR02	11	0.569	0.606	1.10	1.072	0.606	2.93	0.526	0.588	1.08	5.6	4.2	0.83	76.3	71.5	0.97	0.035	0.040	1.10
CR03	12	0.926	1.044	1.37	0.987	1.052	1.49	0.880	1.004	1.43	2.0	0.5	-0.06	58.6	52.7	0.71	0.061	0.069	1.60
CR04	10	0.963	1.094	1.52	0.990	1.113	1.55	0.902	1.049	1.58	2.4	0.5	-0.03	59.2	52.9	0.71	0.062	0.071	1.69
CR05	10	2.336	0.603	7.71	3.947	0.997	8.16	0.449	0.475	1.41	7.7	6.9	0.55	83.1	82.3	0.87	0.029	0.028	1.95
CR06	11	0.578	0.612	1.11	2.601	0.619	8.36	0.515	0.611	0.96	6.3	4.6	0.93	78.5	73.6	1.01	0.031	0.036	0.99
CR07	10	0.524	0.570	1.06	0.636	0.576	1.49	0.517	0.574	1.08	5.5	3.9	0.79	75.5	70.2	0.96	0.034	0.040	1.10
CR08	8	0.504	0.550	0.90	1.571	0.562	4.84	0.476	0.541	0.92	6.6	5.2	0.95	79.8	76.1	1.02	0.030	0.033	0.99
PR09	12	0.455	0.501	1.44	0.468	0.506	1.47	0.410	0.486	1.36	7.8	5.8	0.59	82.7	78.4	0.89	0.027	0.032	1.54
PR10	6	0.421	0.383	1.57	4.316	0.416	24.34	0.339	0.348	1.34	9.8	8.8	0.66	87.6	86.8	0.91	0.023	0.023	1.61
PR11	10	0.340	0.339	0.92	0.946	1.618	4.20	0.289	0.307	0.95	11.2	10.4	0.96	91.8	91.3	1.00	0.018	0.018	1.13
PR12	8	0.452	0.459	1.38	0.474	0.487	1.43	0.446	0.432	1.49	7.5	6.7	0.50	81.4	81.5	0.84	0.028	0.029	1.80
PR13	11	0.579	0.620	1.22	4.064	2.683	13.69	0.457	0.517	1.39	7.5	5.9	0.57	82.6	79.1	0.87	0.027	0.031	1.77
PR14	10	0.419	0.403	2.11	0.461	0.407	2.32	0.409	0.429	2.27	9.5	7.3	0.36	84.4	82.4	0.79	0.026	0.026	2.79
PR15	10	0.481	0.495	1.22	0.501	0.498	1.24	0.418	0.468	1.31	7.8	6.4	0.60	83.2	80.7	0.88	0.026	0.028	1.60
PR16	4	0.460	0.493	1.53	0.463	0.503	1.51	0.364	0.392	1.17	9.2	7.9	0.78	86.2	83.9	0.96	0.025	0.028	1.29
PR17	12	0.514	0.446	1.39	1.038	0.453	3.66	0.469	0.429	1.39	7.5	7.3	0.60	81.8	83.4	0.88	0.028	0.026	1.69
PR18	12	0.659	0.766	1.42	0.693	0.790	1.42	0.589	0.750	1.56	5.5	2.8	0.36	72.8	64.3	0.81	0.039	0.050	1.64

Table A.3: Complete RA parameters derived from the C-HATS$_{av}$ per measured room.

		T_{20} [s]			T_{30} [s]			EDT [s]			C_{50} [dB]			D_{50} [%]			T_S [s]		
Room	N_{pos}	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.
CR01	11	0.477	0.474	1.39	1.086	0.481	4.21	0.432	0.453	1.53	8.0	6.7	0.49	83.7	81.4	0.84	0.026	0.028	2.06
CR02	11	0.571	0.609	1.12	0.941	0.611	2.45	0.527	0.580	1.11	5.6	4.2	0.62	76.4	72.1	0.90	0.034	0.040	1.35
CR03	12	0.935	1.052	1.39	1.071	1.058	1.70	0.875	1.009	1.39	2.2	0.3	0.04	59.7	51.7	0.74	0.060	0.071	1.58
CR04	10	0.962	1.086	1.54	1.149	1.113	1.97	0.897	1.046	1.58	2.3	0.0	-0.10	58.9	49.9	0.67	0.062	0.074	1.78
CR05	10	0.828	0.559	2.34	3.701	1.119	6.97	0.425	0.450	1.56	7.9	6.5	0.49	83.2	80.4	0.83	0.028	0.030	2.01
CR06	11	0.800	0.612	1.95	3.913	0.630	3.35	0.507	0.610	0.97	6.4	3.7	0.73	78.5	69.9	0.95	0.032	0.041	1.24
CR07	10	0.534	0.571	1.11	1.363	0.572	4.34	0.512	0.541	1.10	5.7	4.6	0.68	76.5	73.7	0.92	0.033	0.037	1.23
CR08	8	0.503	0.548	0.90	2.056	0.566	6.72	0.491	0.542	0.95	6.7	5.2	0.84	80.8	76.2	0.98	0.029	0.033	1.12
PR09	12	0.453	0.493	1.42	0.936	0.501	3.90	0.416	0.476	1.34	7.3	5.7	0.55	81.3	78.0	0.86	0.029	0.033	1.63
PR10	6	0.842	0.364	4.26	5.980	1.091	6.26	0.352	0.326	1.55	9.9	9.7	0.61	87.2	89.0	0.89	0.023	0.022	1.78
PR11	10	0.972	0.328	4.62	1.367	0.336	2.93	0.294	0.291	1.04	11.1	10.7	0.91	91.6	91.8	0.99	0.019	0.018	1.36
PR12	8	0.463	0.479	1.49	1.877	0.510	8.70	0.419	0.427	1.53	8.1	6.5	0.48	82.7	80.8	0.83	0.027	0.030	2.10
PR13	11	1.059	0.720	3.26	3.763	2.915	4.02	0.461	0.512	1.59	7.8	5.6	0.52	82.4	77.2	0.86	0.029	0.034	1.87
PR14	10	0.422	0.369	2.15	3.039	0.385	3.52	0.388	0.367	2.14	8.9	7.9	0.33	82.5	84.8	0.76	0.030	0.026	2.78
PR15	10	0.465	0.475	1.22	0.792	0.482	2.62	0.430	0.475	1.58	7.8	5.7	0.47	82.6	78.0	0.83	0.026	0.032	1.95
PR16	4	0.363	0.358	1.25	1.674	0.438	8.07	0.274	0.273	1.46	11.8	10.8	0.67	92.0	91.7	0.94	0.018	0.019	1.65
PR17	12	0.713	0.454	2.29	1.718	1.286	3.53	0.482	0.445	1.45	7.1	6.4	0.48	79.4	79.9	0.82	0.032	0.030	1.96
PR18	12	0.650	0.740	1.49	2.010	0.790	5.61	0.609	0.700	1.80	5.4	3.3	0.23	72.2	67.3	0.72	0.040	0.046	2.10

Table A.4: Complete RA parameters derived from the A-$HATS_{prom}$ per measured room.

		T_{20} [s]			T_{30} [s]			EDT [s]			C_{50} [dB]			D_{50} [%]			T_S [s]		
Room	N_{pos}	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.
CR01	11	0.485	0.481	1.36	0.587	0.493	1.62	0.458	0.493	1.41	8.3	7.1	0.54	84.7	83.1	0.87	0.026	0.028	1.81
CR02	11	0.581	0.619	1.13	1.080	0.615	2.93	0.553	0.616	1.08	6.2	4.8	0.86	78.5	74.3	0.98	0.037	0.042	1.10
CR03	12	0.942	1.062	1.39	1.000	1.063	1.51	0.917	1.043	1.47	2.5	1.0	0.07	61.3	55.5	0.73	0.064	0.072	1.59
CR04	10	0.981	1.112	1.54	1.002	1.126	1.56	0.942	1.089	1.57	3.0	1.1	0.07	62.2	56.2	0.72	0.065	0.074	1.65
CR05	10	2.358	0.617	7.63	4.071	1.235	6.83	0.480	0.506	1.34	8.3	7.4	0.55	84.5	84.1	0.87	0.031	0.030	1.89
CR06	11	0.634	0.621	1.29	2.665	0.626	8.50	0.543	0.638	0.95	6.8	5.1	0.92	80.5	75.9	1.00	0.034	0.039	0.95
CR07	10	0.535	0.581	1.08	0.652	0.583	1.53	0.542	0.599	1.10	6.0	4.3	0.83	77.3	72.1	0.97	0.036	0.041	1.12
CR08	8	0.514	0.564	0.91	1.617	0.567	4.95	0.503	0.577	0.92	7.2	5.8	0.93	81.8	78.5	1.01	0.032	0.036	0.96
PR09	12	0.463	0.511	1.45	0.473	0.510	1.47	0.424	0.499	1.35	8.2	6.1	0.59	83.6	79.4	0.89	0.028	0.033	1.52
PR10	6	0.429	0.391	1.60	4.322	0.424	24.17	0.357	0.366	1.29	10.3	9.4	0.66	88.7	88.5	0.92	0.024	0.026	1.56
PR11	10	0.349	0.345	0.92	0.996	1.647	4.39	0.309	0.320	0.93	11.8	10.7	0.95	92.7	91.9	1.00	0.019	0.019	1.09
PR12	8	0.456	0.467	1.38	0.479	0.495	1.44	0.462	0.454	1.52	7.8	7.2	0.54	82.3	83.1	0.86	0.029	0.030	1.84
PR13	11	0.591	0.629	1.22	4.479	2.689	14.98	0.481	0.533	1.36	8.0	6.2	0.57	83.8	79.9	0.87	0.029	0.032	1.66
PR14	10	0.427	0.411	2.14	0.470	0.413	2.34	0.425	0.448	2.22	10.0	7.9	0.37	85.4	83.8	0.80	0.027	0.028	2.70
PR15	10	0.488	0.504	1.23	0.506	0.501	1.25	0.434	0.485	1.31	8.0	6.7	0.61	83.9	81.7	0.88	0.027	0.030	1.59
PR16	4	0.474	0.512	1.53	0.471	0.513	1.52	0.400	0.421	1.20	9.8	8.6	0.80	87.8	85.5	0.97	0.027	0.029	1.26
PR17	12	0.526	0.454	1.41	1.062	0.458	3.73	0.500	0.445	1.45	7.9	7.6	0.63	83.1	84.5	0.89	0.030	0.027	1.70
PR18	12	0.671	0.781	1.43	0.700	0.801	1.41	0.614	0.779	1.51	6.1	3.4	0.40	75.1	67.6	0.83	0.042	0.052	1.66

Table A.5: Complete RA parameters derived from the C-HATS$_{\text{prom}}$ per measured room.

Room	N_{pos}	T_{20} [s]			T_{30} [s]			EDT [s]			C_{50} [dB]			D_{50} [%]			T_S [s]		
		BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.	BB	Mid	L.v.H.
CR01	11	0.483	0.481	1.39	1.148	0.486	4.47	0.461	0.480	1.44	8.3	7.1	0.50	84.6	82.6	0.85	0.027	0.029	1.99
CR02	11	0.580	0.619	1.13	0.966	0.616	2.52	0.550	0.597	1.07	6.0	4.5	0.62	77.7	73.6	0.90	0.035	0.041	1.32
CR03	12	0.949	1.070	1.41	1.098	1.068	1.75	0.911	1.049	1.42	2.7	0.9	0.16	62.2	55.2	0.77	0.063	0.075	1.58
CR04	10	0.978	1.100	1.57	1.258	1.122	2.23	0.933	1.079	1.61	2.8	0.6	0.03	61.7	53.5	0.71	0.065	0.078	1.80
CR05	10	0.839	0.569	2.33	4.119	1.594	5.42	0.446	0.470	1.51	8.2	6.9	0.52	84.3	81.9	0.85	0.029	0.032	1.98
CR06	11	0.809	0.619	1.95	4.045	0.635	3.41	0.529	0.625	0.96	6.9	4.1	0.71	80.1	71.5	0.94	0.033	0.043	1.17
CR07	10	0.544	0.578	1.14	1.403	0.578	4.46	0.533	0.561	1.12	6.1	4.9	0.70	77.8	75.0	0.93	0.035	0.038	1.23
CR08	8	0.509	0.555	0.91	2.143	0.571	6.99	0.519	0.561	0.89	7.2	5.5	0.80	82.3	77.7	0.97	0.030	0.035	1.06
PR09	12	0.458	0.496	1.44	1.007	0.503	4.25	0.429	0.487	1.34	7.6	6.0	0.57	82.2	78.8	0.87	0.030	0.033	1.62
PR10	6	0.958	0.374	4.89	7.372	1.096	5.05	0.376	0.346	1.48	10.4	10.1	0.62	88.0	89.8	0.89	0.024	0.022	1.71
PR11	10	0.983	0.330	4.63	1.413	0.339	3.04	0.307	0.299	1.03	11.6	10.9	0.90	92.2	92.1	0.98	0.020	0.018	1.30
PR12	8	0.473	0.487	1.51	1.926	0.518	8.83	0.436	0.439	1.50	8.5	6.7	0.48	83.8	81.2	0.84	0.028	0.030	2.03
PR13	11	1.138	0.875	3.49	3.911	2.933	4.20	0.489	0.535	1.56	8.3	6.0	0.53	83.8	78.7	0.87	0.030	0.036	1.81
PR14	10	0.432	0.378	2.16	4.058	0.393	3.23	0.404	0.385	2.12	9.3	8.3	0.36	83.6	85.9	0.77	0.031	0.027	2.65
PR15	10	0.473	0.483	1.23	0.810	0.485	2.67	0.447	0.491	1.55	8.2	6.0	0.48	83.6	79.0	0.84	0.027	0.033	1.87
PR16	4	0.377	0.371	1.27	1.827	0.484	8.21	0.295	0.296	1.45	12.4	11.3	0.68	92.9	92.7	0.94	0.019	0.020	1.59
PR17	12	0.720	0.463	2.27	2.095	2.092	2.95	0.503	0.465	1.44	7.5	6.9	0.50	80.6	81.8	0.83	0.033	0.032	1.93
PR18	12	0.658	0.750	1.50	2.027	0.801	5.62	0.630	0.723	1.75	5.9	3.7	0.25	73.8	68.9	0.73	0.042	0.047	2.05

A.2 Complete overview on in-situ parameters

Table A.6: In-situ level parameters derived from the reference microphone per measured room.

Room	$N_{\mathbf{pos}}$	$L_{\mathbf{Z,eq}}$ [dB]	$L_{\mathbf{ZF,mean}}$ [dB]	$L_{\mathbf{Z,10}}$ [dB]	$L_{\mathbf{Z,90}}$ [dB]	$L_{\mathbf{Z,eq}}$ (L.v.H)	$L_{\mathbf{A,eq}}$ [dB]	$L_{\mathbf{AF,mean}}$ [dB]	$L_{\mathbf{A,10}}$ [dB]	$L_{\mathbf{A,90}}$ [dB]	$L_{\mathbf{A,eq}}$ (L.v.H)
CR01	11	69.5	59.0	69.3	48.9	1.10	62.1	53.8	65.2	42.4	0.81
CR02	11	65.4	55.9	63.6	48.9	1.08	61.1	42.8	56.9	31.8	0.81
CR03	12	68.9	61.6	70.3	52.8	1.07	63.9	56.6	67.6	44.5	0.79
CR04	10	69.5	58.6	69.6	48.6	1.13	62.6	49.8	65.1	35.8	0.83
CR05	10	69.1	62.4	70.8	54.4	1.12	61.1	52.2	64.0	41.2	0.84
CR06	11	65.3	59.9	68.1	52.6	1.11	61.9	54.3	65.1	43.6	0.83
CR07	10	61.2	54.0	62.9	46.1	1.06	57.9	45.0	59.9	28.9	0.77
CR08	8	67.5	59.2	68.9	49.8	1.08	64.2	53.5	65.7	41.4	0.81
PR09	12	69.0	61.9	70.1	53.5	1.10	63.5	56.3	66.6	45.6	0.83
PR10	6	67.3	58.5	66.6	50.4	1.13	59.2	50.7	62.0	38.0	0.84
PR11	10	65.2	59.0	67.3	51.1	1.05	63.0	55.5	65.8	44.6	0.79
PR12	8	65.8	60.5	67.8	53.5	1.11	61.3	55.2	64.3	46.0	0.84
PR13	11	68.2	57.0	66.8	48.4	1.09	59.0	48.0	60.8	35.1	0.81
PR14	10	69.7	63.1	71.1	55.4	1.08	66.2	58.8	68.6	49.2	0.82
PR15	10	68.7	58.6	68.1	49.2	1.07	61.7	50.8	64.2	36.7	0.81
PR16	4	67.0	57.1	65.9	48.0	1.11	59.9	50.4	62.6	37.6	0.82
PR17	12	67.1	57.8	68.2	48.1	1.13	59.2	48.4	61.7	34.6	0.84
PR18	12	67.0	59.9	69.0	50.7	1.07	62.7	53.3	64.8	41.8	0.81

Table A.7: In-situ level parameters derived from the A-HATS$_{av}$ per measured room.

Room	N_{pos}	$L_{Z,eq}$ [dB]	$L_{ZF,mean}$ [dB]	$L_{Z,10}$ [dB]	$L_{Z,90}$ [dB]	$L_{Z,eq}$ (L.v.H)	$L_{A,eq}$ [dB]	$L_{AF,mean}$ [dB]	$L_{A,10}$ [dB]	$L_{A,90}$ [dB]	$L_{A,eq}$ (L.v.H)
CR01	11	72.7	65.9	76.2	54.9	1.16	70.2	62.2	73.6	50.1	0.90
CR02	11	69.1	59.0	71.1	48.2	1.19	66.6	51.9	67.2	38.9	0.92
CR03	12	76.5	71.4	80.5	61.7	1.15	74.2	67.4	78.2	55.4	0.89
CR04	10	73.3	64.9	77.2	53.4	1.14	71.2	59.9	74.7	46.1	0.88
CR05	10	72.7	65.5	76.2	56.1	1.18	69.5	59.9	72.7	48.2	0.93
CR06	11	70.8	65.0	74.2	56.6	1.17	68.2	60.4	71.7	49.1	0.91
CR07	10	68.5	59.9	71.5	47.8	1.13	66.4	54.0	68.9	36.9	0.88
CR08	8	73.4	67.1	76.6	57.6	1.14	71.5	62.5	74.5	49.3	0.89
PR09	12	73.0	67.7	76.3	58.5	1.17	70.8	64.0	74.0	53.7	0.92
PR10	6	69.6	63.5	72.6	53.6	1.16	67.6	60.0	70.7	48.2	0.90
PR11	10	71.3	65.4	74.5	56.3	1.10	69.8	62.7	73.0	51.9	0.86
PR12	8	71.9	67.9	75.3	60.2	1.15	69.6	64.3	73.1	54.8	0.91
PR13	11	68.7	60.4	71.1	49.8	1.13	66.8	55.9	68.9	42.6	0.87
PR14	10	74.9	70.0	78.4	62.0	1.15	73.1	66.6	76.3	57.1	0.91
PR15	10	69.9	63.0	73.2	51.9	1.14	68.0	58.5	71.2	44.0	0.89
PR16	4	70.7	64.0	73.8	54.2	1.14	69.1	60.4	72.1	48.3	0.89
PR17	12	70.8	63.7	73.9	53.9	1.14	68.3	59.1	71.4	46.5	0.89
PR18	12	71.9	64.4	74.3	54.9	1.12	70.2	60.7	72.1	49.7	0.88

Table A.8: In-situ level parameters derived from the C-HATS$_{\mathrm{av}}$ per measured room.

Room	$N_{\mathbf{pos}}$	$L_{\mathbf{Z,eq}}$ [dB]	$L_{\mathbf{ZF,mean}}$ [dB]	$L_{\mathbf{Z,10}}$ [dB]	$L_{\mathbf{Z,90}}$ [dB]	$L_{\mathbf{Z,eq}}$ (L.v.H)	$L_{\mathbf{A,eq}}$ [dB]	$L_{\mathbf{AF,mean}}$ [dB]	$L_{\mathbf{A,10}}$ [dB]	$L_{\mathbf{A,90}}$ [dB]	$L_{\mathbf{A,eq}}$ (L.v.H)
CR01	11	70.4	63.5	73.5	53.4	1.06	68.3	60.1	71.4	48.8	0.82
CR02	11	66.9	57.0	68.4	46.9	1.11	64.0	49.2	63.3	36.3	0.85
CR03	12	72.5	67.7	76.0	59.2	1.07	70.7	63.9	74.1	52.8	0.82
CR04	10	70.5	61.5	73.6	50.6	1.06	69.2	56.6	71.7	43.0	0.81
CR05	10	69.9	63.5	73.3	54.9	1.09	67.0	57.8	70.0	46.7	0.85
CR06	11	69.4	64.4	72.7	57.0	1.08	67.1	59.4	70.4	48.9	0.83
CR07	10	66.4	57.9	68.9	47.0	1.06	65.0	51.7	66.6	35.2	0.81
CR08	8	70.8	64.4	73.8	55.9	1.05	69.4	59.8	71.9	47.2	0.81
PR09	12	72.9	67.5	76.3	58.3	1.10	70.9	63.6	74.0	53.0	0.87
PR10	6	68.8	63.0	71.8	53.5	1.07	67.3	59.5	70.0	48.0	0.83
PR11	10	68.0	62.2	71.0	53.2	1.03	67.0	59.4	69.9	48.4	0.79
PR12	8	70.8	66.3	73.9	58.7	1.09	68.9	62.6	71.9	53.2	0.86
PR13	11	68.6	60.8	70.9	51.1	1.05	67.2	56.0	68.8	43.2	0.81
PR14	10	73.8	68.6	77.1	60.6	1.07	72.6	65.4	75.5	55.6	0.84
PR15	10	70.9	63.9	73.7	53.8	1.05	69.4	59.1	72.2	44.9	0.82
PR16	4	70.0	62.8	72.5	53.3	1.03	69.2	59.3	71.3	47.3	0.80
PR17	12	70.3	63.5	73.0	54.6	1.08	68.0	58.5	70.7	46.3	0.83
PR18	12	70.5	63.6	73.0	54.7	1.05	69.3	59.9	71.2	49.2	0.81

Table A.9: In-situ level parameters derived from the A-$HATS_{prom}$ per measured room.

Room	N_{pos}	$L_{Z,eq}$ [dB]	$L_{ZF,mean}$ [dB]	$L_{Z,10}$ [dB]	$L_{Z,90}$ [dB]	$L_{Z,eq}$ (L.v.H)	$L_{A,eq}$ [dB]	$L_{AF,mean}$ [dB]	$L_{A,10}$ [dB]	$L_{A,90}$ [dB]	$L_{A,eq}$ (L.v.H)
CR01	11	73.2	61.1	71.8	50.0	1.16	83.4	57.5	69.2	45.2	0.88
CR02	11	64.9	54.5	66.6	43.5	1.19	63.3	47.3	62.8	34.0	0.91
CR03	12	91.3	66.5	75.7	56.7	1.21	93.2	62.6	73.4	50.3	0.88
CR04	10	69.4	60.5	72.9	48.6	1.14	67.1	55.6	70.5	41.3	0.86
CR05	10	88.3	60.7	71.4	51.1	1.17	68.9	55.2	68.0	43.3	0.91
CR06	11	72.5	60.2	69.6	51.8	1.16	99.4	55.7	67.2	44.3	0.89
CR07	10	72.2	55.4	67.0	43.5	1.12	62.5	49.7	64.5	32.9	0.85
CR08	8	90.6	62.3	72.0	53.0	1.14	73.0	57.9	69.9	44.6	0.86
PR09	12	71.2	62.9	71.8	53.5	1.18	72.9	59.3	69.5	48.7	0.86
PR10	6	76.5	59.0	68.5	49.0	1.15	67.0	55.6	66.6	43.7	0.88
PR11	10	83.9	60.7	70.0	51.4	1.10	73.1	58.0	68.6	47.1	0.83
PR12	8	74.4	63.4	70.9	55.6	1.14	73.4	59.8	68.8	50.4	0.88
PR13	11	78.4	55.4	66.3	44.9	1.15	65.7	51.1	64.1	37.7	0.85
PR14	10	86.9	65.5	74.0	57.6	1.15	84.5	62.1	71.9	52.6	0.87
PR15	10	70.1	58.3	68.6	46.9	1.15	69.0	53.9	66.7	39.2	0.88
PR16	4	70.1	59.4	69.4	49.5	1.13	67.3	55.9	67.8	43.7	0.84
PR17	12	82.3	58.8	69.2	48.6	1.14	86.7	54.3	66.8	41.5	0.87
PR18	12	86.0	59.4	69.6	49.8	1.14	87.6	55.7	67.4	44.6	0.85

Table A.10: In-situ level parameters derived from the C-HATS$_{\mathrm{prom}}$ per measured room.

Room	$N_{\mathbf{pos}}$	$L_{\mathbf{Z,eq}}$ [dB]	$L_{\mathbf{ZF,mean}}$ [dB]	$L_{\mathbf{Z,10}}$ [dB]	$L_{\mathbf{Z,90}}$ [dB]	$L_{\mathbf{Z,eq}}$ (L.v.H)	$L_{\mathbf{A,eq}}$ [dB]	$L_{\mathbf{AF,mean}}$ [dB]	$L_{\mathbf{A,10}}$ [dB]	$L_{\mathbf{A,90}}$ [dB]	$L_{\mathbf{A,eq}}$ (L.v.H)
CR01	11	91.3	58.7	68.9	48.4	1.06	73.2	55.3	66.8	43.9	0.80
CR02	11	89.6	52.0	63.5	41.6	1.11	59.5	44.2	58.5	30.9	0.82
CR03	12	73.0	63.4	72.1	54.3	1.10	67.7	59.7	70.3	48.1	0.80
CR04	10	68.9	56.7	68.9	45.7	1.05	69.8	52.0	67.1	38.2	0.79
CR05	10	70.8	58.7	68.6	50.2	1.08	74.4	53.2	65.5	42.3	0.82
CR06	11	75.3	59.5	68.0	51.9	1.07	63.5	54.6	65.8	44.0	0.81
CR07	10	69.3	52.6	63.8	41.4	1.05	88.5	46.7	61.7	30.0	0.78
CR08	8	72.3	59.8	69.3	51.3	1.04	75.3	55.2	67.5	42.6	0.78
PR09	12	86.2	62.1	71.0	52.9	1.10	72.3	58.4	68.8	47.7	0.85
PR10	6	71.1	58.8	68.1	48.9	1.06	65.9	55.6	66.7	43.8	0.81
PR11	10	69.2	57.2	66.3	48.3	1.01	69.6	54.6	65.3	43.5	0.76
PR12	8	80.8	61.6	69.3	53.9	1.05	64.9	58.0	67.4	48.3	0.86
PR13	11	75.6	55.5	65.9	45.7	1.06	79.6	51.0	63.9	38.0	0.79
PR14	10	83.4	63.8	72.4	55.6	1.05	100.0	60.7	70.9	50.7	0.81
PR15	10	87.0	58.8	68.8	48.4	1.04	69.6	54.1	67.3	39.8	0.79
PR16	4	67.0	58.0	68.0	48.4	1.03	69.8	54.7	67.0	42.5	0.77
PR17	12	93.2	58.9	68.5	50.1	1.08	70.2	54.0	66.2	41.9	0.82
PR18	12	79.0	58.4	68.0	49.4	0.98	83.6	54.8	66.4	44.0	0.78

Table A.11: In-situ psychoacoustic parameters derived from the reference microphone per measured room.

Room	$N_{\mathbf{pos}}$	$N_{\mathbf{mean}}$	$N_{\mathbf{5}}$	$N_{\mathbf{90}}$	$S_{\mathbf{mean}}$	$S_{\mathbf{5}}$	$S_{\mathbf{90}}$	$R_{\mathbf{mean}}$	$R_{\mathbf{5}}$	$R_{\mathbf{90}}$	$FS_{\mathbf{mean}}$	$FS_{\mathbf{5}}$	$FS_{\mathbf{90}}$
		(sone)			(acum)			(asper)			(vacil)		
CR01	11	8.3	18.0	3.0	1.56	2.14	1.23	0.08	0.16	0.03	0.08	0.17	0.04
CR02	11	4.5	12.1	1.4	1.29	1.88	0.80	0.09	0.24	0.02	0.03	0.10	0.00
CR03	12	10.4	21.3	3.9	1.53	2.01	1.25	0.07	0.13	0.03	0.07	0.14	0.03
CR04	10	7.4	19.1	1.9	1.45	1.95	1.12	0.06	0.13	0.01	0.05	0.11	0.02
CR05	10	8.0	18.2	3.1	1.52	2.05	1.20	0.08	0.17	0.03	0.06	0.13	0.03
CR06	11	9.0	19.0	3.6	1.61	2.16	1.29	0.07	0.14	0.02	0.08	0.16	0.03
CR07	10	5.0	13.3	0.9	1.53	2.11	1.13	0.06	0.14	0.01	0.05	0.13	0.01
CR08	8	8.7	20.0	2.9	1.62	2.17	1.29	0.07	0.15	0.02	0.08	0.16	0.03
PR09	12	9.6	19.5	3.9	1.45	1.93	1.14	0.08	0.15	0.03	0.09	0.18	0.04
PR10	6	6.1	13.7	1.7	1.37	1.88	1.04	0.08	0.15	0.03	0.09	0.18	0.04
PR11	10	8.5	17.9	3.4	1.50	1.99	1.20	0.09	0.17	0.03	0.10	0.20	0.04
PR12	8	8.7	16.6	4.0	1.45	1.90	1.15	0.08	0.15	0.04	0.08	0.15	0.04
PR13	11	5.4	13.4	1.5	1.43	2.08	1.05	0.07	0.15	0.02	0.07	0.16	0.03
PR14	10	10.7	20.9	4.9	1.42	1.86	1.13	0.10	0.18	0.05	0.09	0.18	0.04
PR15	10	6.6	15.9	1.7	1.46	2.04	1.11	0.07	0.16	0.02	0.09	0.18	0.03
PR16	4	6.0	14.2	1.7	1.37	1.89	1.02	0.08	0.15	0.03	0.08	0.17	0.04
PR17	12	5.8	14.6	1.3	1.35	1.88	1.00	0.07	0.15	0.02	0.08	0.17	0.03
PR18	12	8.0	17.6	2.8	1.41	1.92	1.10	0.07	0.15	0.03	0.07	0.14	0.03

Table A.12: In-situ psychoacoustic parameters derived from the A-HATS$_{\mathrm{av}}$ per measured room.

Room	$N_{\mathbf{pos}}$	$N_{\mathbf{mean}}$	$N_{\mathbf{5}}$	$N_{\mathbf{90}}$	$S_{\mathbf{mean}}$	$S_{\mathbf{5}}$	$S_{\mathbf{90}}$	$R_{\mathbf{mean}}$	$R_{\mathbf{5}}$	$R_{\mathbf{90}}$	$FS_{\mathbf{mean}}$	$FS_{\mathbf{5}}$	$FS_{\mathbf{90}}$
			(sone)			(acum)			(asper)			(vacil)	
CR01	11	8.8	18.8	3.0	1.27	1.73	1.01	0.12	0.20	0.06	0.09	0.19	0.04
CR02	11	5.2	14.5	1.3	1.22	1.74	0.86	0.08	0.16	0.04	0.05	0.13	0.01
CR03	12	13.3	26.2	5.3	1.28	1.65	1.05	0.12	0.18	0.07	0.08	0.16	0.04
CR04	10	9.1	21.8	2.7	1.27	1.69	0.99	0.10	0.17	0.05	0.06	0.12	0.02
CR05	10	8.5	19.6	3.1	1.25	1.68	0.98	0.12	0.20	0.07	0.06	0.13	0.03
CR06	11	8.3	18.0	3.2	1.34	1.82	1.07	0.10	0.17	0.05	0.08	0.17	0.04
CR07	10	5.9	14.9	1.0	1.26	1.75	0.93	0.09	0.17	0.04	0.06	0.14	0.01
CR08	8	9.8	21.3	3.3	1.37	1.82	1.10	0.10	0.18	0.05	0.09	0.17	0.04
PR09	12	9.9	19.7	4.2	1.21	1.61	0.95	0.11	0.18	0.07	0.09	0.18	0.04
PR10	6	7.0	15.0	2.3	1.22	1.66	0.92	0.10	0.17	0.06	0.10	0.19	0.04
PR11	10	8.5	18.0	3.3	1.26	1.67	0.99	0.11	0.19	0.06	0.10	0.19	0.05
PR12	8	9.9	18.4	4.6	1.24	1.62	0.98	0.12	0.19	0.07	0.09	0.16	0.05
PR13	11	5.7	13.8	1.6	1.16	1.70	0.84	0.10	0.18	0.05	0.08	0.17	0.03
PR14	10	11.2	21.8	5.2	1.19	1.57	0.94	0.13	0.20	0.08	0.09	0.18	0.04
PR15	10	6.7	15.4	1.7	1.19	1.68	0.86	0.10	0.18	0.05	0.09	0.19	0.04
PR16	4	7.3	16.4	2.4	1.22	1.66	0.92	0.10	0.17	0.06	0.09	0.19	0.04
PR17	12	7.4	17.2	2.3	1.20	1.65	0.90	0.10	0.18	0.05	0.09	0.19	0.04
PR18	12	8.1	17.9	3.0	1.21	1.64	0.93	0.11	0.18	0.06	0.07	0.15	0.03

Table A.13: In-situ psychoacoustic parameters derived from the C-HATS$_{av}$ per measured room.

Room	$N_{\mathbf{pos}}$	$N_{\mathbf{mean}}$	$N_{\mathbf{5}}$	$N_{\mathbf{90}}$	$S_{\mathbf{mean}}$	$S_{\mathbf{5}}$	$S_{\mathbf{90}}$	$R_{\mathbf{mean}}$	$R_{\mathbf{5}}$	$R_{\mathbf{90}}$	$FS_{\mathbf{mean}}$	$FS_{\mathbf{5}}$	$FS_{\mathbf{90}}$
		(sone)			(acum)			(asper)			(vacil)		
CR01	11	8.3	17.7	3.1	1.50	2.04	1.20	0.09	0.16	0.04	0.08	0.17	0.04
CR02	11	4.5	12.1	1.0	1.43	2.00	1.06	0.05	0.13	0.01	0.04	0.11	0.01
CR03	12	11.0	21.6	4.6	1.45	1.88	1.20	0.08	0.14	0.04	0.07	0.14	0.03
CR04	10	7.8	19.6	2.1	1.46	1.94	1.14	0.06	0.13	0.02	0.05	0.11	0.02
CR05	10	7.7	17.6	2.9	1.45	1.94	1.16	0.09	0.17	0.04	0.06	0.13	0.03
CR06	11	8.4	17.7	3.4	1.56	2.07	1.26	0.07	0.14	0.02	0.07	0.15	0.03
CR07	10	5.3	13.8	0.8	1.47	1.99	1.12	0.06	0.15	0.01	0.05	0.13	0.01
CR08	8	8.6	19.3	2.9	1.58	2.09	1.28	0.08	0.16	0.03	0.08	0.15	0.03
PR09	12	10.5	21.3	4.4	1.40	1.85	1.11	0.09	0.15	0.04	0.09	0.18	0.04
PR10	6	7.0	14.6	2.5	1.42	1.84	1.15	0.08	0.15	0.04	0.09	0.19	0.04
PR11	10	7.3	15.3	2.8	1.44	1.92	1.14	0.09	0.18	0.04	0.09	0.19	0.04
PR12	8	9.5	18.1	4.4	1.41	1.83	1.14	0.09	0.15	0.05	0.08	0.16	0.04
PR13	11	6.2	15.1	1.9	1.40	1.98	1.05	0.08	0.16	0.02	0.08	0.18	0.03
PR14	10	11.1	21.8	5.1	1.38	1.79	1.12	0.11	0.18	0.06	0.09	0.19	0.04
PR15	10	7.6	17.8	2.1	1.43	1.96	1.10	0.08	0.16	0.03	0.09	0.19	0.03
PR16	4	7.2	16.4	2.5	1.41	1.90	1.09	0.09	0.16	0.04	0.09	0.19	0.04
PR17	12	7.5	17.4	2.4	1.39	1.88	1.07	0.07	0.15	0.03	0.09	0.18	0.04
PR18	12	8.1	17.4	3.2	1.41	1.88	1.12	0.08	0.15	0.03	0.07	0.15	0.03

Table A.14: In-situ psychoacoustic parameters derived from the A-HATS$_{\text{prom}}$ per measured room.

Room	N_{pos}	N_{mean}	N_5	N_{90}	S_{mean}	S_5	S_{90}	R_{mean}	R_5	R_{90}	FS_{mean}	FS_5	FS_{90}
		(sone)			(acum)			(asper)			(vacil)		
CR01	11	9.7	20.9	3.4	1.33	1.84	1.05	0.13	0.22	0.07	0.10	0.21	0.04
CR02	11	5.8	16.2	1.4	1.26	1.83	0.89	0.10	0.18	0.04	0.05	0.14	0.01
CR03	12	14.4	28.6	5.7	1.33	1.73	1.08	0.13	0.20	0.08	0.08	0.17	0.04
CR04	10	10.5	24.8	3.0	1.32	1.78	1.03	0.11	0.19	0.06	0.07	0.14	0.03
CR05	10	9.4	21.4	3.5	1.31	1.79	1.02	0.13	0.24	0.08	0.07	0.15	0.03
CR06	11	9.1	19.9	3.5	1.39	1.89	1.10	0.11	0.20	0.06	0.09	0.18	0.04
CR07	10	6.7	16.6	1.3	1.31	1.83	0.96	0.10	0.20	0.04	0.06	0.15	0.01
CR08	8	10.7	23.2	3.7	1.42	1.91	1.13	0.12	0.21	0.06	0.09	0.19	0.04
PR09	12	10.8	21.7	4.6	1.24	1.67	0.97	0.13	0.20	0.08	0.10	0.19	0.04
PR10	6	7.9	17.2	2.6	1.28	1.76	0.96	0.11	0.20	0.06	0.11	0.22	0.05
PR11	10	9.4	20.0	3.6	1.31	1.78	1.02	0.13	0.22	0.07	0.11	0.22	0.05
PR12	8	11.1	20.6	5.2	1.30	1.72	1.03	0.13	0.21	0.08	0.10	0.18	0.05
PR13	11	6.2	15.0	1.8	1.23	1.83	0.88	0.11	0.20	0.05	0.09	0.19	0.04
PR14	10	12.5	24.4	5.9	1.24	1.65	0.97	0.14	0.23	0.09	0.10	0.20	0.05
PR15	10	7.4	17.1	1.9	1.25	1.79	0.91	0.11	0.20	0.06	0.10	0.21	0.04
PR16	4	8.1	18.3	2.8	1.28	1.75	0.96	0.11	0.19	0.06	0.11	0.21	0.05
PR17	12	8.1	18.8	2.5	1.26	1.74	0.95	0.11	0.20	0.05	0.10	0.20	0.04
PR18	12	8.8	19.5	3.3	1.25	1.72	0.95	0.12	0.20	0.07	0.08	0.16	0.03

Table A.15: In-situ psychoacoustic parameters derived from the C-HATS$_{prom}$ per measured room.

Room	N_{pos}	N_{mean}	N_5	N_{90}	S_{mean}	S_5	S_{90}	R_{mean}	R_5	R_{90}	FS_{mean}	FS_5	FS_{90}
		(sone)			(acum)			(asper)			(vacil)		
CR01	11	9.2	19.6	3.4	1.56	2.14	1.24	0.10	0.19	0.04	0.09	0.19	0.04
CR02	11	4.9	13.1	1.1	1.49	2.08	1.10	0.06	0.15	0.01	0.05	0.13	0.01
CR03	12	12.7	25.2	5.1	1.50	1.96	1.23	0.09	0.16	0.04	0.08	0.16	0.04
CR04	10	8.6	21.6	2.4	1.52	2.01	1.19	0.07	0.15	0.02	0.06	0.12	0.02
CR05	10	8.6	19.4	3.4	1.52	2.05	1.20	0.10	0.20	0.04	0.07	0.14	0.03
CR06	11	9.2	19.5	3.7	1.61	2.15	1.29	0.08	0.17	0.03	0.08	0.16	0.04
CR07	10	5.7	14.7	0.9	1.53	2.08	1.17	0.07	0.18	0.01	0.06	0.15	0.01
CR08	8	9.6	21.5	3.3	1.65	2.20	1.32	0.09	0.19	0.03	0.08	0.17	0.04
PR09	12	11.1	22.3	4.7	1.43	1.90	1.14	0.10	0.17	0.05	0.10	0.19	0.04
PR10	6	8.5	18.0	3.0	1.54	2.07	1.22	0.10	0.19	0.05	0.11	0.24	0.05
PR11	10	7.9	16.8	3.1	1.49	2.01	1.18	0.10	0.20	0.04	0.10	0.21	0.04
PR12	8	10.5	20.0	4.9	1.45	1.89	1.16	0.10	0.18	0.05	0.09	0.17	0.05
PR13	11	6.7	16.1	2.0	1.45	2.07	1.08	0.09	0.18	0.03	0.09	0.19	0.03
PR14	10	12.2	24.0	5.6	1.44	1.89	1.16	0.12	0.21	0.07	0.10	0.21	0.05
PR15	10	8.2	19.1	2.3	1.48	2.06	1.14	0.10	0.19	0.03	0.10	0.21	0.04
PR16	4	8.0	18.4	2.7	1.47	2.00	1.13	0.10	0.18	0.05	0.11	0.22	0.04
PR17	12	8.4	19.1	2.8	1.43	1.95	1.10	0.09	0.17	0.03	0.09	0.20	0.04
PR18	12	8.7	18.7	3.4	1.45	1.96	1.15	0.09	0.18	0.04	0.07	0.16	0.03

A.3 INCH results per age group

Table A.16: INCH survey descriptives (adults participants).

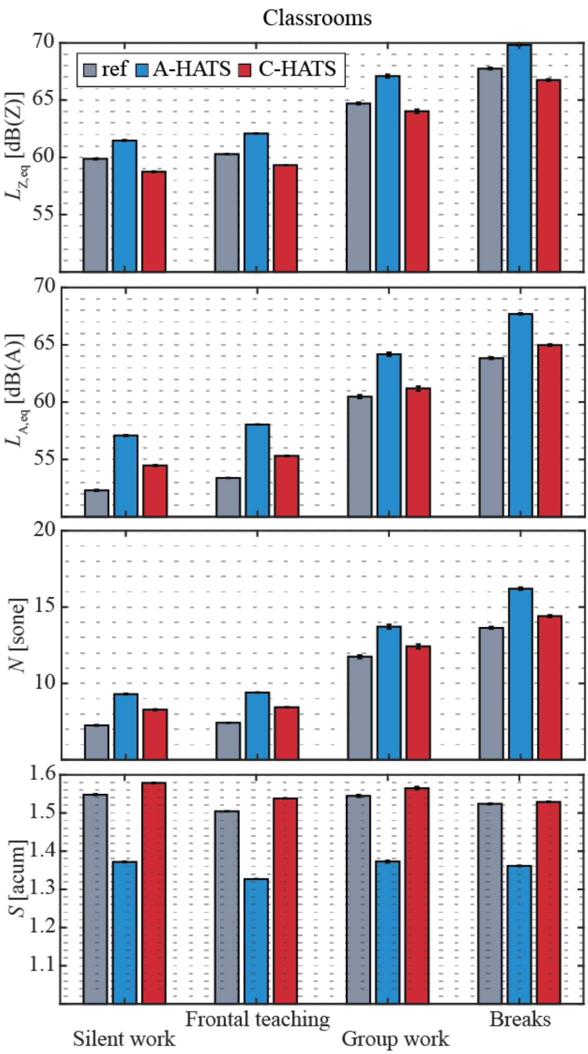
	Playrooms					
Age category	**< 20 y/o**	**20-30 y/o**	**31-40 y/o**	**41-50 y/o**	**51-60 y/o**	**>61 y/o**
N	3	16	7	5	2	-
Frequency						
Angry and yelling	33.3%	68.8%	14.3%	50.0%	50.0%	-
Loud and strong	66.7%	81.3%	85.7%	80.0%	100.0%	-
Scraping and screeching	33.3%	31.3%	28.6%	20.0%	50.0%	-
Sadness						
Angry and yelling	0.0%	25.0%	14.3%	0.0%	0.0%	-
Loud and strong	0.0%	37.5%	14.3%	40.0%	0.0%	-
Scraping and screeching	0.0%	26.7%	14.3%	20.0%	0.0%	-
Angriness						
Angry and yelling	33.3%	31.3%	42.9%	40.0%	50.0%	-
Loud and strong	50.0%	50.0%	42.9%	40.0%	50.0%	-
Scraping and screeching	33.3%	31.3%	42.9%	0.0%	0.0%	-
	Classrooms					
Age category	**< 20 y/o**	**20-30 y/o**	**31-40 y/o**	**41-50 y/o**	**51-60 y/o**	**>61 y/o**
N	1	4	8	8	5	3
Frequency						
Angry and yelling	0.0%	0.0%	12.5%	37.5%	40.0%	0.0%
Loud and strong	0.0%	50.0%	37.5%	62.5%	100.0%	0.0%
Scraping and screeching	0.0%	0.0%	0.0%	0.0%	60.0%	0.0%
Sadness						
Angry and yelling	0.0%	0.0%	0.0%	0.0%	60.0%	0.0%
Loud and strong	0.0%	0.0%	0.0%	0.0%	20.0%	33.3%
Scraping and screeching	0.0%	50.0%	12.5%	37.5%	0.0%	50.0%
Angriness						
Angry and yelling	0.0%	75.0%	50.0%	25.0%	60.0%	0.0%
Loud and strong	0.0%	50.0%	50.0%	25.0%	60.0%	33.3%
Scraping and screeching	100.0%	50.0%	62.5%	25.0%	20.0%	66.7%

Table A.17: INCH survey descriptives (children participants).

	Playrooms				
Age category	**3 y/o**	**4 y/o**	**5 y/o**	**6 y/o**	
N	13	25	29	16	-
Frequency					
Angry and yelling	30.8%	64.0%	62.1%	43.8%	-
Loud and strong	53.8%	68.0%	72.4%	56.3%	-
Scraping and screeching	30.8%	36.0%	10.3%	12.5%	-
Sadness					
Angry and yelling	61.5%	36.0%	55.2%	37.5%	-
Loud and strong	61.5%	44.0%	37.9%	37.5%	-
Scraping and screeching	15.4%	32.0%	32.1%	28.6%	-
Angriness					
Angry and yelling	30.8%	47.8%	48.3%	25.0%	-
Loud and strong	41.7%	47.8%	51.7%	40.0%	-
Scraping and screeching	38.5%	21.7%	24.1%	37.5%	-
	Classrooms				
Age category	**6 y/o**	**7 y/o**	**8 y/o**	**9 y/o**	**10 y/o**
N	20	49	24	7	1
Frequency					
Angry and yelling	55.0%	38.8%	41.7%	85.7%	100.0%
Loud and strong	55.0%	65.3%	58.3%	71.4%	100.0%
Scraping and screeching	5.0%	14.3%	8.3%	0.0%	100.0%
Sadness					
Angry and yelling	50.0%	42.9%	33.3%	28.6%	100.0%
Loud and strong	60.0%	49.0%	50.0%	57.1%	100.0%
Scraping and screeching	40.0%	16.3%	58.3%	14.3%	100.0%
Angriness					
Angry and yelling	65.0%	44.9%	58.3%	42.9%	100.0%
Loud and strong	60.0%	44.9%	58.3%	28.6%	100.0%
Scraping and screeching	55.0%	22.4%	37.5%	57.1%	0.0%

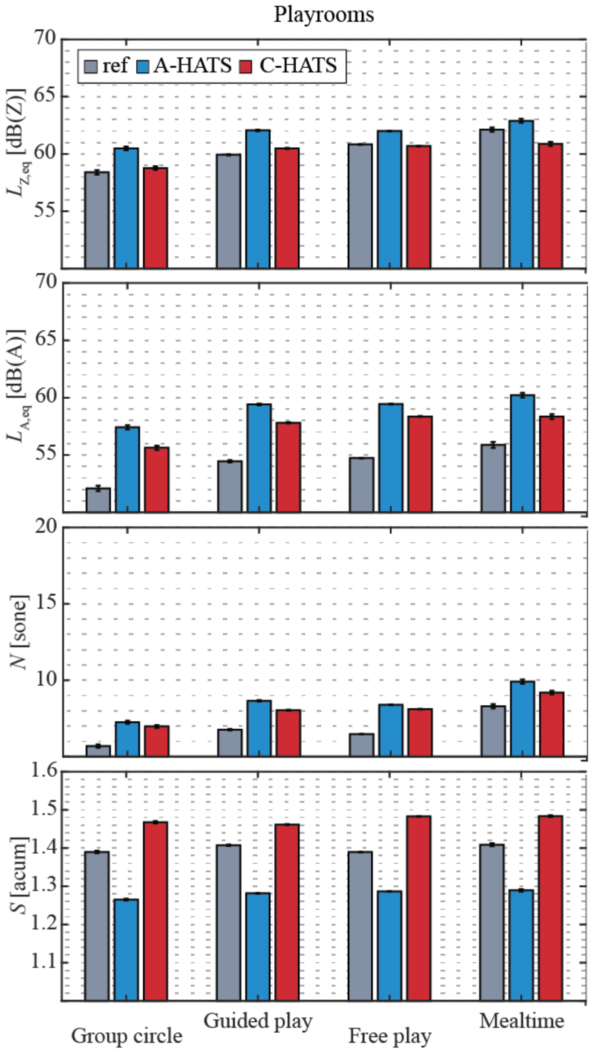
A.4 ABAS results (prominent-ear method)

Figure A.1: ABAS in noise parameters (prominent-ear method) with regard to classroom activities.



Note. Classrooms in primary schools. Comparing results of the four activities, respectively, obtained from the reference microphone (Ref), adult (aHATS) and child head and torso simulator (cHATS) evaluated using the prominent-ear method.

Figure A.2: ABAS in noise parameters (prominent-ear method) with regard to playroom activities.



Note. Playrooms in daycare centers . Comparing results of the four activities, respectively, obtained from the reference microphone (Ref), adult (A-HATS) and child head and torso simulator (C-HATS) evaluated using the prominent-ear method.

B

Appendix - Children's Auditory Cognition in Noise

B.1 ChildASA noise study: Post-hoc analyses for RT and ER

Table B.1: ChildASA-noise: Post-hoc test results according to Bonferroni for the reaction time (RT).

			G_1	G_2	p (G_1 - G_2)	M_{diff} [ms]	[95 %-CI] [ms]
C	NoN	L-R	Rep	Sw	.031	132	[13 , 251]
		F-B	Rep	Sw	.006	171	[52, 290]
Rep	NoN	F-B	C	IC	< .001	236	[131, 341]
			SC	IC	.001	245	[84, 407]
Sw	NoN	L-R	C	SC	.017	-260	[-482, -37]
	N	F-B	C	IC	.031	232	[17, 446]
Rep	C	L-R	NoN	N	< .001	221	[92, 351]
		F-B	NoN	N	.034	136	[11, 261]
Sw	C	Next	NoN	N	.006	161	[48, 273]
	SC	L-R	NoN	N	.011	315	[77, 552]
		Next	NoN	N	.039	150	[8, 293]
	IC	F-B	NoN	N	.029	129	[14, 244]
Rep	C	NoN	L-R	F-B	< .001	594	[405, 783]
			L-R	Next	< .001	489	[335, 644]
		N	L-R	F-B	< .001	509	[337, 681]
			L-R	Next	.005	259	[67, 450]
			Next	F-B	.008	250	[56, 444]
	SC	NoN	L-R	F-B	< .001	592	[333, 850]
			L-R	Next	< .001	491	[242, 741]
		N	L-R	F-B	< .001	603	[386, 821]
			L-R	Next	.003	311	[96, 527]
			Next	F-B	.001	292	[114, 470]
	IC	NoN	L-R	F-B	< .001	831	[560, 1102]
			L-R	Next	< .001	464	[224, 703]
			Next	F-B	< .001	367	[230, 504]
		N	L-R	F-B	< .001	612	[282, 941]
			L-R	Next	< .001	419	[172, 666]
Sw	C	NoN	L-R	F-B	< .001	633	[471, 795]
			L-R	Next	.001	263	[103, 423]
			Next	F-B	< .001	370	[222, 518]
		N	L-R	F-B	< .001	537	[339, 736]
			L-R	Next	< .001	456	[284, 629]
	SC	NoN	L-R	F-B	< .001	788	[485, 1091]
			L-R	Next	.001	450	[177, 723]
			Next	F-B	< .001	338	[150, 525]
		N	L-R	F-B	< .001	557	[264, 849]
			L-R	Next	.029	286	[24, 548]
			Next	F-B	.001	271	[98, 444]
	IC	NoN	L-R	F-B	< .001	678	[502, 853]
			L-R	Next	.003	423	[121, 724]
			Next	F-B	.030	255	[19, 490]
		N	L-R	F-B	< .001	779	[577, 980]
			L-R	Next	< .001	378	[178, 577]
			Next	F-B	< .001	401	[196, 606]

Note. N = Noise; NoN = No Noise; C = Congruent; SC = Semi-Congruent; IC = Incongruent; F-B = Front-Back; L-R = Left-Right; Rep = Repetition; Sw = Switch.

Table B.2: ChildASA-noise: Post-hoc test results according to Bonferroni for the reaction time (ER) with regard to age group.

			G_1	G_2	p (G_1 - G_2)	M_{diff} [%]	[95 %-CI] [%]
Children	C	FB	Rep	Sw	.002	59	[2.3 , 9.4]
Children	SC	FB	Rep	Sw	< .001	-29.2	[-37.3 , -21]
Children	SC	LR	Rep	Sw	.027	9	[1.1 , 16.9]
Children	SC	Next	Rep	Sw	< .001	-23.7	[-30 , -17.4]
Children	IC	FB	Rep	Sw	.005	-15.2	[-25.6 , -4.7]
Children	IC	LR	Rep	Sw	.002	10.1	[3.8 , 16.5]
Children	IC	Next	Rep	Sw	.003	-12.2	[-20 , -4.5]
Children	Rep	LR	C	SC	< .001	-18.3	[-26.5 , -10]
Children	Rep	LR	C	IC	< .001	-29.1	[-37.5 , -20.7]
Children	Rep	LR	IC	SC	.02	10.9	[1.4 , 20.3]
Children	Rep	C	LR	Next	.011	5.8	[1.1 , 10.5]
Children	Rep	SC	LR	Next	.028	10.9	[0.9 , 20.9]
Children	Rep	IC	LR	Next	.001	14.3	[4.8 , 23.8]
Children	Rep	FB	C	SC	< .001	-21	[-30.3 , -11.8]
Children	Rep	FB	C	IC	< .001	-28.2	[-38.4 , -18]
Children	Rep	FB	IC	SC	.012	7.2	[1.3 , 13.1]
Children	Rep	Next	C	SC	< .001	-13.2	[-19.4 , -7]
Children	Rep	Next	C	IC	< .001	-20.6	[-28.1 , -13.2]
Children	Sw	LR	C	SC	< .001	-53.3	[-63.4 , -43.3]
Children	Sw	LR	C	IC	< .001	-50.2	[-57.7 , -42.7]
Children	Sw	FB	C	SC	< .001	-14	[-20.8 , -7.2]
Children	Sw	FB	C	IC	< .001	-20	[-27 , -13.1]
Children	Sw	FB	IC	SC	.025	6	[0.6 , 11.4]
Children	Sw	Next	C	SC	< .001	-37.6	[-45.8 , -29.4]
Children	Sw	Next	C	IC	< .001	-33.6	[-43.4 , -23.7]
Children	Sw	SC	LR	FB	< .001	40.4	[30.8 , 50]
Children	Sw	SC	LR	Next	< .001	16.4	[6.8 , 26]
Children	Sw	SC	Next	FB	< .001	24	[15.1 , 32.8]
Children	Sw	IC	LR	FB	< .001	31.2	[20.5 , 41.9]
Children	Sw	IC	LR	Next	< .001	17.2	[7.5 , 27]
Children	Sw	IC	Next	FB	< .001	14	[5.4 , 22.6]
Adults	Rep	LR	C	SC	< .001	-44.7	[-52.2 , -37]
Adults	Rep	LR	C	IC	< .001	-40.7	[-48.6 , -32.7]
Adults	Rep	Next	C	SC	.034	-6.2	[-12.1 , -0.4]
Adults	Rep	SC	LR	FB	< .001	44.4	[33 , 55.8]
Adults	Rep	SC	LR	Next	< .001	40	[30.6 , 49.4]
Adults	Rep	IC	LR	FB	< .001	41.2	[27 , 55.3]
Adults	Rep	IC	LR	Next	< .001	36.8	[27.9 , 45.8]
Adults	Sw	LR	C	SC	< .001	-39.6	[-49 , -30.1]
Adults	Sw	LR	C	IC	< .001	-41.8	[-48.8 , -34.7]
Adults	Sw	SC	LR	FB	< .001	41.2	[32.1 , 50.2]
Adults	Sw	SC	LR	Next	< .001	34.6	[25.6 , 43.6]
Adults	Sw	IC	LR	FB	< .001	42.9	[32.8 , 53]
Adults	Sw	IC	LR	Next	< .001	36.7	[27.5 , 45.9]

Note. C = Congruent; SC = Semi-congruent; IC = Incongruent; FB = Front-Back; LR = Left-Right; Rep = Repetition; Sw = Switch.

Bibliography

Åhlander, V. L., Rydell, R., & Löfqvist, A. (2011). Speaker's Comfort in Teaching Environments: Voice Problems in Swedish Teaching Staff. *Journal of Voice*, *25*(4), 430–440. https://doi.org/10.1016/j.jvoice.2009.12.006

ANSI Inc. (2002). ANSI-S12.60-2002 - Acoustical performance criteria, design requirements, and guideslines for schools.

Astolfi, A., Parati, L., D'Orazio, D., & Garai, M. (2019a). The new Italian standard UNI 11532 on acoustics for schools. *Proceedings of the 23rd International Congress on Acoustics*, 7004–7011.

Astolfi, A., Puglisi, G. E., Murgia, S., Minelli, G., Pellerey, F., Prato, A., & Sacco, T. (2019b). Influence of Classroom Acoustics on Noise Disturbance and Well-Being for First Graders. *Frontiers in Psychology*, *10*, 2736. https://doi.org/10.3389/fpsyg.2019.02736

Baek, H. J., Cho, C.-H., Cho, J., & Woo, J.-M. (2015). Reliability of Ultra-Short-Term Analysis as a Surrogate of Standard 5-Min Analysis of Heart Rate Variability. *Telemedicine and e-Health*, *21*(5), 404–414. https://doi.org/10.1089/tmj.2014.0104

Barron, M. (1971). The subjective effects of first reflections in concert halls - the need for lateral relections. *J. Sound Vib*, *15*(4), 475–494. https://doi.org/10.1016/0022-460x(71)90406-8

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., & Stansfeld, S. (2014). Auditory and non-auditory effects of noise on health. *Lancet (London, England)*, *383*(9925), 1325–1332. https://doi.org/10.1016/S0140-6736(13)61613-X

Berggren, M., Lindström, F., Persson Waye, K., & Claesson, I. (2008). Analysis of How the Noise Level Depends on Different Activities in a Child Day-Care Center. *Proceedings of the 37th International Congress on Noise Control Engineering, INTER-NOISE 2008, Shanghai*. Retrieved May 14, 2024, from https://urn.kb.se/resolve?urn=urn:nbn:se:bth-8440

Berzborn, M., Bomhardt, R., Klein, J. C., Richter, J.-G., & Vorländer, M. (2017). The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic

Measurements and Signal Processing. *Fortschritte Der Akustik - DAGA 2017*. https://publications.rwth-aachen.de/record/687308

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press.

Bomhardt, R., & Fels, J. (2014). Analytical Interaural Time Difference Model for the Individualization of Arbitrary Head-Related Impulse Responses. *Audio Engineering Society Convention 137*. http://www.aes.org/e-lib/browse.cfm?elib=17454

Boothroyd, A. (1997). Auditory development of the hearing child. *Scandinavian Audiology-Supplement Only*, *46*, 9–16.

Bradley, J. S., Reich, R., & Norcross, S. G. (1999). A just noticeable difference in C50 for speech. *Applied Acoustics*, *58*(2), 99–108. https://doi.org/10.1016/s0003-682x(98)00075-9

Broadbent, D. E. (2013, October 22). *Perception and Communication*. Elsevier.

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*(5), 1465–1487. https://doi.org/10.3758/s13414-015-0882-9

Building Bulletin 93. (2015). Acoustic Design of Schools—Performance Standards.

Chemla, D., Young, J., Badilini, F., Maison-Blanche, P., Affres, H., Lecarpentier, Y., & Chanson, P. (2005). Comparison of fast Fourier transform and autoregressive spectral analysis for the study of heart rate variability in diabetic patients. *International Journal of Cardiology*, *104*(3), 307–313. https://doi.org/10.1016/j.ijcard.2004.12.018

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. https://doi.org/10.1121/1.1907229

Chevalier, N., Kurth, S., Doucette, M. R., Wiseheart, M., Deoni, S. C. L., Iii, D. C. D., O'Muircheartaigh, J., Blackwell, K. A., Munakata, Y., & LeBourgeois, M. K. (2015). Myelination Is Associated with Processing Speed in Early Childhood: Preliminary Insights. *PLOS ONE*, *10*(10), e0139897. https://doi.org/10.1371/journal.pone.0139897

Choi, Y.-J. (2016). Effect of occupancy on acoustical conditions in university classrooms. *Applied Acoustics*, *114*, 36–43. https://doi.org/10.1016/j.apacoust.2016.07.010

Cipresso, P., Colombo, D., & Riva, G. (2019). Computational Psychometrics Using Psychophysiological Measures for the Assessment of Acute Mental Stress. *Sensors*, *19*(4), 781. https://doi.org/10.3390/s19040781

Claesen, A., Gomes, S., Tuerlinckx, F., & Vanpaemel, W. (2021). Comparing dream to reality: An assessment of adherence of the first generation of preregistered studies. *Royal Society Open Science*, *8*(10), 211037. https://doi.org/10.1098/rsos.211037

Cohen, S., Evans, G. W., Stokols, D., Krantz, D. S., Cohen, S., Evans, G. W., Stokols, D., & Krantz, D. S. (1986). Environmental stress and health. *Behavior, health, and environmental stress*, 103–142. https://doi.org/10.1007/978-1-4757-9380-2_4

Crandell, C. C., & Smaldino, J. J. (2000). Classroom Acoustics for Children With Normal Hearing and With Hearing Impairment. *Language, Speech, and Hearing Services in Schools*, *31*(4), 362–370. https://doi.org/10.1044/0161-1461.3104.362

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80–90. https://doi.org/10.1037/h0039515

Dibbets, P., & Jolles, J. (2006). The Switch Task for Children: Measuring mental flexibility in young children. *Cognitive Development*, *21*(1), 60–71. https://doi.org/10.1016/j.cogdev.2005.09.004

DIN e.V. (2009). DIN 45692 - Measurement technique for the simulation of the auditory sensation of sharpness.

DIN e.V. (2010). DIN 45631/A1 - Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of the loudness of time-variant sound.

DIN e.V. (2016). DIN 18041 - Acoustic quality in rooms - Specifications and instructions for the room acoustic design.

DIN e.V. (2018). DIN 60645-1 - Akustik – Audiometer – Teil 1: ReintonAudiometer.

Doyle, A.-B. (1973). Listening to distraction: A developmental study of selective attention. *Journal of Experimental Child Psychology*, *15*(1), 100–115. https://doi.org/10.1016/0022-0965(73)90134-3

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fels, J., Buthmann, P., & Vorländer, M. (2004). Head-Related Transfer Functions of Children. *Acta Acustica united with Acustica*, *90*(5), 918–927. https://www.ingentaconnect.com/content/dav/aaua/2004/00000090/00000005/art00012

Fels, J. (2008). From children to adults: How binaural cues and ear canal impedances grow. *The Journal of the Acoustical Society of America*, *124*(6), 3359–3359. https://doi.org/10.1121/1.3020603

Fels, J., & Vorländer, M. (2009). Anthropometric Parameters Influencing Head-Related Transfer Functions. *Acta Acustica united with Acustica*, *95*(2), 331–342. https://doi.org/10.3813/aaa.918156

Finnish Standards Association. (2004). SFS 5907 - Acoustic classification of spaces of buildings.

Gambi, E., Agostinelli, A., Belli, A., Burattini, L., Cippitelli, E., Fioretti, S., Pierleoni, P., Ricciuti, M., Sbrollini, A., & Spinsante, S. (2017). Heart Rate Detection Using Microsoft Kinect: Validation and Comparison to Wearable Devices. *Sensors*, *17*(8), 1776. https://doi.org/10.3390/s17081776

Genuit, K., & Fiebig, A. (2014). A critical view on measurements of noise: Improvements on noise annoyance predictions by psychoacoustics. *Lärmbekämpfung*, *9*(1).

Huang-Pollock, C. L., Carr, T. H., & Nigg, J. T. (2002). Development of selective attention: Perceptual load influences early versus late attentional selection in children and adults. *Developmental Psychology*, *38*(3), 363–375. https://doi.org/10.1037/0012-1649.38.3.363

IEC. (2012). IEC 60268-16 - Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.

ISO. (1993). ISO 9568 - Cinematography - Background acoustic noise levels in theatre, review rooms and dubbing rooms.

ISO. (2008). ISO 3382-2 - Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms.

ISO. (2009). ISO 3382-1 - Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces.

ISO. (2017). ISO 532-1 - Acoustics - Methods for calculating loudness - Part 1: Zwicker method.

Jones, P. R., Moore, D. R., & Amitay, S. (2015). Development of auditory selective attention: Why children struggle to hear in noisy environments. *Developmental Psychology*, *51*(3), 353–369. https://doi.org/10.1037/a0038570

Kim, J. W., Seok, H. S., & Shin, H. (2021). Is Ultra-Short-Term Heart Rate Variability Valid in Non-static Conditions? *Frontiers in Physiology*, *12*. https://doi.org/10.3389/fphys.2021.596060

Klatte, M., Bergstroem, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, *4*, 578. https://doi.org/10.3389/fpsyg.2013.00578

Klatte, M., Hellbrück, J., Seidel, J., & Leistner, P. (2010a). Effects of Classroom Acoustics on Performance and Well-Being in Elementary School

Children: A Field Study. *Environment and Behavior*, *42*(5), 659–692. https://doi.org/10.1177/0013916509336813

Klatte, M., Lachmann, T., & Meis, M. (2010b). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise and Health: Special issue on Noise, Memory and Learning*, *12*(49), 270. https://doi.org/10.4103/1463-1741.70506

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010c). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology*, *22*(8), 1168–1191. https://doi.org/10.1080/09541440903378250

Koch, I., Lawo, V., Fels, J., & Vorländer, M. (2011). Switching in the cocktail party: Exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(4), 1140–1147. https://doi.org/10.1037/a0022189

Lackner, H. K., Eglmaier, M. T. W., Hackl-Wimmer, S., Paechter, M., Rominger, C., Eichen, L., Rettenbacher, K., Walter-Laager, C., & Papousek, I. (2020). How to Use Heart Rate Variability: Quantification of Vagal Activity in Toddlers and Adults in Long-Term ECG. *Sensors*, *20*(20), 5959. https://doi.org/10.3390/s20205959

Lavandier, M., & Best, V. (2020). Modeling Binaural Speech Understanding in Complex Situations. In J. Blauert & J. Braasch (Eds.), *The Technology of Binaural Understanding* (pp. 547–578). Springer International Publishing. https://doi.org/10.1007/978-3-030-00386-9_19

Lawo, V., Fels, J., Oberem, J., & Koch, I. (2014). Intentional attention switching in dichotic listening: Exploring the efficiency of nonspatial and spatial selection. *Quarterly Journal of Experimental Psychology*, *67*(10), 2010–2024. https://doi.org/10.1080/17470218.2014.898079

Leccese, F., Rocca, M., & Salvadori, G. (2018). Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes. *Applied Acoustics*, *140*, 143–149. https://doi.org/10.1016/j.apacoust.2018.05.019

Liang, L., & Yu, G. (2020). Binaural speech transmission index with spatialized virtual speaker in near field: Distance and direction dependence. *The Journal of the Acoustical Society of America*, *148*(2), EL202–EL207. https://doi.org/10.1121/10.0001808

Lindau, A., & Weinzierl, S. (2012). Assessing the Plausibility of Virtual Acoustic Environments. *Acta Acustica united with Acustica*, *98*(5), 804–810. https://doi.org/10.3813/AAA.918562

Litovsky, R. Y. (2015). Development of the auditory system. *Handbook of clinical neurology*, *129*, 55–72. https://doi.org/10.1016/B978-0-444-62630-1.00003-2

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *The Journal of the Acoustical Society of America*, *106*(4), 1633–1654. https://doi.org/10.1121/1.427914

Litovsky, R. Y., & Godar, S. P. (2010). Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. *The Journal of the Acoustical Society of America*, *128*(4), 1979–1991. https://doi.org/10.1121/1.3478849

Loh, K., & Fels, J. (2021). A concept to evaluate activity-based acoustic settings in primary schools appropriate for children's hearing.

Loh, K., Fintor, E., Nolden, S., & Fels, J. (2022a). Children's intentional switching of auditory selective attention in spatial and noisy acoustic environments in comparison to adults. *Developmental Psychology*, *58*(1), 69–82. https://doi.org/10.1037/dev0001239

Loh, K., Hoog Antink, C., Nolden, S., & Fels, J. (2021). Combined assessment of cognitive and physiological parameters in child-appropriate listening experiments.

Loh, K., Seitz, J., Braren, H., & Fels, J. (2023a). Speech Transmission Index Measured using Adult and Children Head and Torso Simulators. *Proceeding of Forum Acusticum 2023*.

Loh, K., Seitz, J., Nolden, S., Gudi-Mindermann, H., Dohmen, M., van Kamp, I., Botteldooren, D., Hornikx, M., Bolte, G., & Fels, J. (2024). *Cognitive milestone at the age of four: Intentional switching of auditory selective attention in preschool children* [Manuscript submitted for publication].

Loh, K., Seitz, J., Rust, F., & Fels, J. (2023b). A concept to evaluate activity-based acoustic settings in preschools for children aged three to six. *Proceedings of Inter-noise 2023*. https://doi.org/10.3397/IN_2023_0824

Loh, K., Yadav, M., Persson Waye, K., Klatte, M., & Fels, J. (2022b). Toward Child-Appropriate Acoustic Measurement Methods in Primary Schools and Daycare Centers. *Frontiers in Built Environment*, *8*. https://doi.org/10.3389/fbuil.2022.688847

Masiero, B., & Fels, J. (2011). Perceptually Robust Headphone Equalization for Binaural Reproduction. *Audio Engineering Society Convention 130*. http://www.aes.org/e-lib/browse.cfm?elib=15855

McAllister, A. M., Granqvist, S., Sjölander, P., & Sundberg, J. (2009). Child Voice and Noise: A Pilot Study of Noise in Day Cares and the Effects on 10 Children's Voice Quality According to Perceptual Evaluation. *Journal of Voice*, *23*(5), 587–593. https://doi.org/10.1016/j.jvoice.2007.10.017

McCarty, R. (2016). Chapter 4 - The Fight-or-Flight Response: A Cornerstone of Stress Research. In G. Fink (Ed.), *Stress: Concepts, Cognition, Emotion, and Behavior* (pp. 33–37). Academic Press. https://doi.org/10.1016/B978-0-12-800951-2.00004-2

Middlebrooks, J. C., & Green, D. M. (1990). Directional dependence of interaural envelope delays. *The Journal of the Acoustical Society of America*, *87*(5), 2149–2162. https://doi.org/10.1121/1.399183

Møller, H., Sørensen, M. F., Hammershøi, D., & Jensen, C. B. (1995). Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society*, *43*(5), 300–321.

Møller, H., Sørensen, M. F., Jensen, C. B., & Hammershøi, D. (1996). Binaural technique: Do we need individual recordings? *Journal of the Audio Engineering Society*, *44*(6), 451–469.

Moore, B. C. J., & Glasberg, B. R. (2007). Modeling binaural loudness. *The Journal of the Acoustical Society of America*, *121*(3), 1604–1612. https://doi.org/10.1121/1.2431331

Munoz, M. L., Roon, A. van, Riese, H., Thio, C., Oostenbroek, E., Westrik, I., Geus, E. J. C. de, Gansevoort, R., Lefrandt, J., Nolte, I. M., & Snieder, H. (2015). Validity of (Ultra-)Short Recordings for Heart Rate Variability Measurements. *PLOS ONE*, *10*(9), e0138921. https://doi.org/10.1371/journal.pone.0138921

Neuman, A. C., & Hochberg, I. (1983). Children's perception of speech in reverberation. *The Journal of the Acoustical Society of America*, *73*(6), 2145–2149. https://doi.org/10.1121/1.389538

Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010). Combined Effects of Noise and Reverberation on Speech Recognition Performance of Normal-Hearing Children and Adults. *Ear and Hearing*, *31*(3), 336–344. https://doi.org/10.1097/AUD.0b013e3181d3d514

Noguchi, S., Yoshizawa, H., Sato, M., & Konishi, T. (2019). The sound environment of German preschools and preschool teachers' thoughts about sound generated by children. *Proceedings of the ICA 2019 and EAA Euroregio : 23rd International Congress on Acoustics*, 5918–5925. https://doi.org/10.18154/RWTH-CONV-239560

Nolden, S., Ibrahim, C. N., & Koch, I. (2019). Cognitive control in the cocktail party: Preparing selective attention to dichotically presented voices supports distractor suppression. *Attention, Perception, & Psychophysics*, *81*(3), 727–737. https://doi.org/10.3758/s13414-018-1620-x

Nolden, S., & Koch, I. (2017). Intentional switching of auditory attention between long and short sequential tone patterns. *Attention, Perception, &*

Psychophysics, *79*(4), 1132–1146. https://doi.org/10.3758/s13414-017-1298-5

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van 't Veer, A. E., & Vazire, S. (2019). Preregistration Is Hard, And Worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818. https://doi.org/10.1016/j.tics.2019.07.009

Oberem, J., Koch, I., & Fels, J. (2017). Intentional switching in auditory selective attention: Exploring age-related effects in a spatial setup requiring speech perception. *Acta Psychologica*, *177*, 36–43. https://doi.org/10.1016/j.actpsy.2017.04.008

Oberem, J., Lawo, V., Koch, I., & Fels, J. (2014). Intentional Switching in Auditory Selective Attention: Exploring Different Binaural Reproduction Methods in an Anechoic Chamber. *Acta Acustica united with Acustica*, *100*(6), 1139–1148. https://doi.org/10.3813/AAA.918793

Oberem, J., Masiero, B., & Fels, J. (2016). Experiments on authenticity and plausibility of binaural reproduction via headphones employing different recording methods. *Applied Acoustics*, *114*, 71–78. https://doi.org/10.1016/j.apacoust.2016.07.009

Oberem, J., Seibold, J., Koch, I., & Fels, J. (2018). Intentional switching in auditory selective attention: Exploring attention shifts with different reverberation times. *Hearing Research*, *359*, 32–39. https://doi.org/10.1016/j.heares.2017.12.013

Okano, T. (2002). Judgments of noticeable differences in sound fields of concert halls caused by intensity variations in early reflections. *The Journal of the Acoustical Society of America*, *111*(1), 217–229. https://doi.org/10.1121/1.1426374

Ostchega, Y., Porter, K. S., Hughes, J., Dillon, C. F., & Nwankwo, T. (2011). Resting pulse rate reference data for children, adolescents, and adults: United States, 1999-2008. *National Health Statistics Reports*, (41), 1–16.

Paul, T. (2015). Herzrhythmusstörungen bei Kindern und Jugendlichen. In G. F. Hoffmann, M. J. Lentze, J. Spranger, & F. Zepp (Eds.), *Pädiatrie: Grundlagen und Praxis* (pp. 1–10). Springer. https://doi.org/10.1007/978-3-642-54671-6_205-1

Pausch, F., & Fels, J. (2019). MobiLab – A Mobile Laboratory for On-Site Listening Experiments in Virtual Acoustic Environments. *Acta Acustica united with Acustica*, *105*(5), 875–887. https://doi.org/10.3813/AAA.919367

Peng, A., Kirkham, N. Z., & Mareschal, D. (2018). Task switching costs in preschool children and adults. *Journal of Experimental Child Psychology*, *172*, 59–72. https://doi.org/10.1016/j.jecp.2018.01.019

Peng, Z., Lau, S.-K., Wang, L. M., Browne, S. D., & Roy, K. P. (2013). Binaural room impulse response database acquired from a variable acoustics classroom. *Proceedings of Meetings on Acoustics*, *15*(1), 015003. https://doi.org/10.1121/1.4793567

Persson Waye, K., & Karlberg, J. (2021). Sound Quality Characteristics of Importance for Preschool Children's Perception and Wellbeing After an Acoustic Intervention. *Frontiers in Built Environment*, *7*. https://doi.org/10.3389/fbuil.2021.688836

Persson Waye, K., van Kamp, I., & Dellve, L. (2013). Validation of a questionnaire measuring preschool children's reactions to and coping with noise in a repeated measurement design. *BMJ open*, *3*(5). https://doi.org/10.1136/bmjopen-2012-002408

Persson Waye, K., van Kamp, I., & Kim, J.-L. (2016). Acoustic intervention at preschools impact on children´s perception and response to high frequency sound qualities. *Proceedings of the 22nd International Congress on Acoustics*.

Picard, M., & Boudreau, C. (1999). Characteristics of the noise found in day-care centers. *The Journal of the Acoustical Society of America*, *105*(2_Supplement), 1127. https://doi.org/10.1121/1.425260

Picard, M., & Bradley, J. S. (2001). Revisiting Speech Interference in Classrooms:Revisando la interferencia en el habla dentro del salón de clases. *Audiology*. Retrieved May 14, 2024, from https://www.tandfonline.com/doi/abs/10.3109/00206090109073117

Prodi, N., Farnetani, A., Yiluya, S., & Fels, J. (2007). Investigating Classroom Acoustics by means of Advanced Reproduction Techniques. *Audio Engineering Society Convention 122*.

Röer, J. P., Bell, R., Körner, U., & Buchner, A. (2018). Equivalent auditory distraction in children and adults. *Journal of Experimental Child Psychology*, *172*, 41–58. https://doi.org/10.1016/j.jecp.2018.02.005

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Sala, E., & Rantala, L. (2016). Acoustics and activity noise in school classrooms in Finland. *Applied Acoustics*, *114*, 252–259. https://doi.org/10.1016/j.apacoust.2016.08.009

Salahuddin, L., Cho, J., Jeong, M. G., & Kim, D. (2007). Ultra Short Term Analysis of Heart Rate Variability for Monitoring Mental Stress in Mobile Settings. *2007 29th Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society, 4656–4659. https://doi.org/10.1109/IEMBS.2007.4353378

Schaaff, K., & Adam, M. T. (2013). Measuring Emotional Arousal for Online Applications: Evaluation of Ultra-short Term Heart Rate Variability Measures. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 362–368. https://doi.org/10.1109/ACII.2013.66

Schlesinger, A., Ramirez, J.-P., Schuitman, J. v. D., & Boone, M. M. (2009). Report on a binaural extension of the Speech Transmission Index method for nonlinear systems and narrowband interference. *Proceedings of the International Symposium on Auditory and Audiological Research*, *2*, 353–362. Retrieved May 15, 2023, from https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-36

Schlittmeier, S. J., Weißgerber, T., Kerber, S., Fastl, H., & Hellbrück, J. (2012). Algorithmic modeling of the irrelevant sound effect (ISE) by the hearing sensation fluctuation strength. *Attention, Perception, & Psychophysics*, *74*(1), 194–203. https://doi.org/10.3758/s13414-011-0230-7

Schmitz, A. (1995). A New Digital Artificial Head Measuring System. *Acta Acustica united with Acustica*, *81*(4), 416–420.

Shaffer, F., Shearman, S., & Meehan, Z. M. (2016). The Promise of Ultra-Short-Term (UST) Heart Rate Variability Measurements. *Biofeedback*, *44*(4), 229–233. https://doi.org/10.5298/1081-5937-44.3.09

Shield, B., Conetta, R., Dockrell, J., Connolly, D., Cox, T., & Mydlarz, C. (2015). A survey of acoustic conditions and noise levels in secondary school classrooms in England. *The Journal of the Acoustical Society of America*, *137*(1), 177–188. https://doi.org/10.1121/1.4904528

Shield, B., & Dockrell, J. (2003). The Effects of Noise on Children at School: A Review. *Building Acoustics*, *10*(2), 97–116. https://doi.org/10.1260/135101003768965960

Shield, B., & Dockrell, J. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *The Journal of the Acoustical Society of America*, *123*(1), 133–144. https://doi.org/10.1121/1.2812596

Shinn-Cunningham, B. G., Kopco, N., & Martin, T. J. (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses. *The Journal of the Acoustical Society of America*, *117*(5), 3100–3115. https://doi.org/10.1121/1.1872572

Södersten, M., Granqvist, S., Hammarberg, B., & Szabo, A. (2002). Vocal Behavior and Vocal Loading Factors for Preschool Teachers at Work Studied with Binaural DAT Recordings. *Journal of Voice*, *16*(3), 356–371. https://doi.org/10.1016/S0892-1997(02)00107-8

Sottek, R. (1993). *Models for signal processing in the human auditory system* [Doctoral thesis]. RWTH Aachen University.

Stansfeld, S., & Clark, C. (2015). Health effects of noise exposure in children. *Current environmental health reports*, *2*, 171–178. https://doi.org/10.1007/s40572-015-0044-1

Stellmack, M. A., Willihnganz, M. S., Wightman, F. L., & Lutfi, R. A. (1997). Spectral weights in level discrimination by preschool children: Analytic listening conditions. *The Journal of the Acoustical Society of America*, *101*(5), 2811–2821. https://doi.org/10.1121/1.419479

Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of Mental Stress on Heart Rate and Heart Rate Variability. In J. Vander Sloten, P. Verdonck, M. Nyssen, & J. Haueisen (Eds.), *4th European Conference of the International Federation for Medical and Biological Engineering* (pp. 1366–1369). Springer. https://doi.org/10.1007/978-3-540-89208-3_324

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299. https://doi.org/10.1037/h0027242

Valkenburg, P. M., & Piotrowski, J. T. (2017). *Plugged in: How media attract and affect youth*. Yale University Press. https://doi.org/10.12987/yale/9780300218879.001.0001
10.12987/yale/9780300218879.001.0001.

van Wijngaarden, S. J., & Drullman, R. (2008). Binaural intelligibility prediction based on the speech transmission index. *The Journal of the Acoustical Society of America*, *123*(6), 4514–4523. https://doi.org/10.1121/1.2905245

Wang, L. M., & Brill, L. C. (2021). Speech and noise levels measured in occupied K–12 classrooms. *The Journal of the Acoustical Society of America*, *150*(2), 864–877. https://doi.org/10.1121/10.0005815

Werner, L. A. (2019). *Human auditory development*. Routledge.

WHO. (1991). *Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment Programme Planning*. Geneva: World Health Organization. Geneva.

Wu, L., Shi, P., Yu, H., & Liu, Y. (2020). An optimization study of the ultra-short period for HRV analysis at rest and post-exercise. *Journal of Electrocardiology*, *63*, 57–63. https://doi.org/10.1016/j.jelectrocard.2020.10.002

Yadav, M., Cabrera, D., Kim, J., Fels, J., & de Dear, R. (2021). Sound in occupied open-plan offices: Objective metrics with a review of historical perspectives. *Applied Acoustics*, *177*, 107943. https://doi.org/10.1016/j.apacoust.2021.107943

Zwicker, E., & Fastl, H. (2013). *Psychoacoustics: Facts and Models*. Springer Science & Business Media.

Acknowledgements

Words cannot express my gratitude to all the people and support I have received while working on my doctoral project. This dissertation would not have been possible without these people, and thus, I would like to thank them explicitly for their continuous motivation and support.

This endeavor would not have been possible without Prof. Janina Fels: Thank you for supervising and guiding my work, for all your helpful advice, and for giving me the opportunity to pursue my research interests with your valuable support. I would like to sincerely thank Prof. Kerstin Persson Waye, who accepted to be my second reviewer, for the insightful discussions and feedback shaping my research.

By being part of the Institute of Hearing Technology and Acoustics (IHTA) and carrying the ITA spirit in my heart, I had the privilege to call many amazing people my colleagues, contributing professional and personal advice in many different ways. Many thanks to Prof. Michael Vorländer, Karin Charlier, Ellen Vergöls, Rolf Kaldenbach, Lukas Aspöck, Gottfried Behler, and Uwe Schlömer who have been a strong backbone within the institute. Special thanks to Marco Berzborn and Cosima Ermert, who not only shared an office with me but also all the ups and downs of a doctoral project. I will truly miss our conversations. Thanks should also go to Josefa Oberem, Hark Braren, Shaima'a Doma, Elie Abi Raad, Lukas Vollmer, Carolin Reimers, and Julia Seitz for all the helpful discussions, challenging thoughts, and technical support.

Driving cutting-edge research requires inspiring ideas and innovative perspectives. For this, I would like to acknowledge my collaborators in the various projects and publications, especially Sophie Nolden, Maria Klatte, and Manuj Yadav.

I am also grateful for all the valuable contributions made by my students: Julian Burger, Simon Klein, Moritz Bender, Leonie Schmitz, Frederike Rust, Jana Berger, Nathalie Gierden, Christina Holländer, Karina Katke, Jacob Beißel, Iris Glosauer, Roser Batlle, Beier Ding, and Xinya Xu.

I would be remiss not to mention all the pre- and primary schools in Aachen that support the studies in this doctoral project and all the parents, children, and people who participate in the listening experiments. I also had the pleasure of working with a great consortium within the Equal-Life project, which added new impulses to my doctoral project. Furthermore, I would like to acknowledge the Equal-Life project, the HEAD-Genuit-Stiftung, and the ERS of the RWTH Aachen University for funding the studies executed within my doctoral thesis.

Finally, I am extremely grateful to my mother and brother for their never-ending support every step of the way, and especially to Marcel Krüger for his love, his endless patience, and being my safety net.

Curriculum Vitae

Personal Data

	Karin Loh
02 November 1990	born in Aachen, Germany

Education

31 January 2025	Doctoral Examination for Dr.-Ing.
10/2017 – 01/2025	Doctoral Candidate Institute for Hearing Technology and Acoustics RWTH Aachen University, Aachen, Germany
10/2014 – 05/2017	M. Sc. Electrical Engineering – Biomedical Engineering RWTH Aachen University, Aachen, Germany
01/2016 – 05/2016	Erasmus Exchange Program – Bioinformation Technology Aalto University, Espoo, Finland
09/2010 – 09/2014	B. Sc. Electrical Engineering – Computer Engineering RWTH Aachen University, Aachen, Germany
06/2010	General Higher Education Entrance Qualification (Abitur) Städt. Gymnasium Eschweiler, Eschweiler, Germany

Professional Experience

02/2021 – 12/2023	Chief engineer Institute for Hearing Technology and Acoustics RWTH Aachen University, Aachen, Germany
08/2017 – 12/2023	Research assistant Institute for Hearing Technology and Acoustics RWTH Aachen University, Aachen, Germany
06/2016 – 10/2016	Biomedical engineering intern Cochlear Benelux NV Mechelen, Belgium
10/2015 – 12/2015	Biomedical engineering intern Abiomed Europe Aachen, Germany

Aachen, Germany, June 18, 2025