



Article

Differential Effects of Task-Irrelevant Monaural and Binaural Classroom Scenarios on Children's and Adults' Speech Perception, Listening Comprehension, and Visual–Verbal Short-Term Memory

Larissa Leist ^{1,*}, Carolin Breuer ², Manuj Yadav ², Stephan Fremerey ³, Janina Fels ², Alexander Raake ³, Thomas Lachmann ^{1,4}, Sabine J. Schlittmeier ⁵ and Maria Klatte ¹

¹ Cognitive and Developmental Psychology Unit, Center for Cognitive Science, Department of Cognitive Psychology, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany

² Institute for Hearing Technology and Acoustics, RWTH Aachen University, 52074 Aachen, Germany

³ Audiovisual Technology Group, Technische Universität Ilmenau, 98693 Ilmenau, Germany

⁴ Centro de Investigación Nebrija en Cognición, Facultad de Lenguas y Educación, Universidad Nebrija, 28015 Madrid, Spain

⁵ Teaching and Research Area of Work and Engineering Psychology, RWTH Aachen University, 52066 Aachen, Germany

* Correspondence: lleist@rhrk.uni-kl.de



Citation: Leist, L.; Breuer, C.; Yadav, M.; Fremerey, S.; Fels, J.; Raake, A.; Lachmann, T.; Schlittmeier, S.J.; Klatte, M. Differential Effects of Task-Irrelevant Monaural and Binaural Classroom Scenarios on Children's and Adults' Speech Perception, Listening Comprehension, and Visual–Verbal Short-Term Memory. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15998. <https://doi.org/10.3390/ijerph192315998>

Academic Editor: Paul B. Tchounwou

Received: 31 October 2022

Accepted: 26 November 2022

Published: 30 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Most studies investigating the effects of environmental noise on children's cognitive performance examine the impact of monaural noise (i.e., same signal to both ears), oversimplifying multiple aspects of binaural hearing (i.e., adequately reproducing interaural differences and spatial information). In the current study, the effects of a realistic classroom-noise scenario presented either monaurally or binaurally on tasks requiring processing of auditory and visually presented information were analyzed in children and adults. In Experiment 1, across age groups, word identification was more impaired by monaural than by binaural classroom noise, whereas listening comprehension (acting out oral instructions) was equally impaired in both noise conditions. In both tasks, children were more affected than adults. Disturbance ratings were unrelated to the actual performance decrements. Experiment 2 revealed detrimental effects of classroom noise on short-term memory (serial recall of words presented pictorially), which did not differ with age or presentation mode (monaural vs. binaural). The present results add to the evidence for detrimental effects of noise on speech perception and cognitive performance, and their interactions with age, using a realistic classroom-noise scenario. Binaural simulations of real-world auditory environments can improve the external validity of studies on the impact of noise on children's and adults' learning.

Keywords: auditory distraction; children; speech perception; listening comprehension; verbal short-term memory; irrelevant sound effect; binaural; monaural; classroom; learning

1. Introduction

Learning in classrooms is often impeded by unfavorable acoustic conditions, such as noise and reverberation [1]. A recent study in German preschool and school classrooms reported an average sound pressure level (SPL) of 66 dB $L_{A,eq}$ (A-weighted equivalent continuous SPL), and a range of 62–69 dB $L_{A,eq}$ during typical activities [2]. Other studies conducted across Europe and the US reported values ranging from 42 to 100 dB $L_{A,eq}$ [3,4] (see also Table 1 in [2]). Field studies revealed that children instructed in classrooms with high levels of indoor or external (aircraft) noise score lower in achievement tests and in ratings of well-being at school, and exhibit higher levels of annoyance due to noise [5–9]. Numerous experimental studies have analyzed effects of acute noise on children's performance in a range of auditory and non-auditory tasks. Concerning the former, it has consistently been shown that children's

language comprehension is more impaired than adults' by noise and reverberation [10–13]. Concerning non-auditory tasks, findings are less consistent for complex academic tasks such as reading and numeracy [14–16] (for review, see [17], but reliable noise-induced performance decrements have been reported for children's visual–verbal short-term memory [18–21].

However, most of these studies do not represent real-world auditory environments with respect to the type of noise and its perception by the person affected. For example, the noise maskers used in many psychoacoustic studies on speech-in-noise perception have nothing in common with a noisy classroom. In addition, in the vast majority of studies on noise effects on cognitive performance, the noise is presented monaurally (the same signal presented to both ears) over headphones, and in cases of loudspeaker presentation the same signal is sent to each speaker. These presentation formats oversimplify the multiple features of binaural hearing in complex acoustic environments, where sounds are spatially spread across the room, and sound sources change often and unpredictably. In order to represent such complex scenes in laboratory settings, one approach is binaural reproduction, wherein the interaural differences in sound reaching the ears are authentically represented, including spatial cues [22,23]. In the current study, we analyzed the effects of a realistic classroom-noise scenario on tasks requiring processing of auditory (Exp. 1) and visual information (Exp. 2) in children and adults. We aimed to assess the detrimental noise effects across age groups, and to explore whether and to what extent these effects, and the developmental change associated with them, are moderated when a realistic, binaural presentation mode is used instead of the usual, monaural presentation.

2. Experiment 1: Effects of Classroom Noise on Speech Perception and Listening Comprehension

Learning in classrooms relies heavily on oral instruction and listening in the presence of irrelevant sounds. Thus, school children are regularly faced with the requirement of focusing on a specific sound source while ignoring others [24]. Experimental studies on the effects of environmental noise on children's ability to understand speech have focused on simple speech perception tasks requiring identification of isolated speech targets in noise and/or reverberation. However, listening requirements faced by children in classrooms go far beyond pure identification. Effective listening in these situations requires storage and processing of complex oral information in working memory, while constructing a coherent mental model of the information presented [25]. There is evidence that noise may affect the storage and processing of spoken items even when the signal-to-noise ratio (SNR; signal SPL minus noise SPL) is high enough to allow perfect or near-perfect identification [26–28]. Thus, effects of noise on word identification tasks do not allow predictions of decrements in complex listening tasks.

Studying the impacts of noise and reverberation on children's speech perception in a classroom-like setting, Klatte and colleagues [29] found differential effects of foreign, single-talker speech and classroom noise without speech on word identification (word-to-picture matching) and listening comprehension (acting-out complex oral instructions). In both tasks, children were affected more than adults. In the comprehension task, both speech and classroom noise significantly reduced children's performance, with 6- to 7-year-old first-graders suffering the most, whereas adults were unaffected. Speech was more disruptive than classroom noise. In contrast, word identification was much more impaired by classroom noise when compared to speech. The authors proposed that, with the SNRs of -3 dB to 3 dB used in their study, the effects of background speech and classroom noise resulted from different mechanisms. Classroom noise masked the speech targets. Background speech was a less potential masker, but interfered with short-term memory processes that children (but not adults) rely on when listening to complex sentences. The study further revealed that the children's ratings of the sound-induced disruption were unrelated to their objective performance decline. This finding underlines that, while it is undeniable that noise has a negative impact on performance, this does not mean that a person affected is subjectively aware of these effects or feels annoyed; cf. [30].

In Klatte et al. [29], the background sounds were presented via loudspeakers located at the sides of the laboratory room. The same recording was sent to each of the eight loudspeakers, and the target signals were presented through a separate speaker located in front of the room. In the current study, we further increased the realism of the design by Klatte and colleagues [29], by including a classroom-noise scenario that is reproduced binaurally (i.e., authentically representing interaural and spatial cues). Prior studies confirmed that participants' performance in listening tasks is affected when the realism of the auditory scene is increased [31–34].

2.1. Materials and Methods

Participants: The sample consisted of 36 student volunteers (19 female) from the University of Kaiserslautern, aged between 19 and 31 years ($M = 24.9$, $SD = 3.9$ years), and 56 children recruited from a primary school in Kaiserslautern. The child sample comprised 37 second-graders (9 female), aged between 6 years, 3 months and 8 years, 2 months ($M = 7$ years, 5 months, $SD = 3$ months); and 19 third-graders (12 female) aged between 8 years, 3 months and 9 years, 7 months ($M = 8$ years, 9 months, $SD = 4$ months).

All participants were native German speakers and had normal or corrected-to normal vision and normal hearing according to either self-reports (adults) or parental reports (children). The study was approved by the Rhineland-Palatinate school authority and by the Ethics Committee of the University of Kaiserslautern. Informed written consent was provided by the adult participants and by the children's parents. Adults received either course credit or payment for participation (10 €).

Apparatus: The tasks were created in Python 3.7/Psychopy 3.1.5 [35] and executed using a 15.6-inch laptop (HP ProBook 450 G6). The screen's resolution was 1920×1080 pixels, and its refresh rate was 60 Hz. The sounds were presented via Sennheiser HD650 headphones and a Focusrite Scarlett 2i2 2nd generation audio interface.

Speech signals: The words and the instructions were read by a professional female speaker in a sound-attenuated booth and recorded with a Sennheiser MD 421-II-4 Dynamic Studio Microphone at a sampling rate of 44.1 kHz and 16-bit resolution. The recordings were loudness-normalized according to EBU R-128 [36] using Version 3.0.0 of Audacity® recording and editing software [37].

Classroom-noise scenario: The background noise represented a classroom-like auditory environment with sounds from everyday classroom activities, e.g., furniture use; desk noise, including writing and use of other stationary items; footsteps; door opening and closing; use of zippers on bags; undoing a plastic wrapper; and turning the pages of a book. To prevent any learning effects, the different noises were presented at irregular intervals as in real classroom scenarios. Some noises (e.g., writing) were played more often than others (e.g., door) to mimic their typical frequency of occurrence in reality. The selection of nonspeech sounds and the frequency with which they occurred were based on listening to recordings of lessons in medium-sized classrooms. Pink noise (-5 dB/octave decay slope) presented at $L_{Aeq,1m}$ of 41.5 dB provided a steady-state noise that simulated air-conditioning noise throughout the auditory scene.

As children frequently speak in class, multi-talker speech, consisting of four child voices talking in Hindi, which was foreign to all participants, was added to the scene. We recorded (32-bit, 44.1 kHz sampling rate) a child (8 years old) having a natural, unscripted conversation with an adult (the parent) on a range of topics for around two hours in a hemi-anechoic room. The talkers sat on chairs facing each other at a 2 m apart. Each talker wore a DPA 4066 omnidirectional headset microphone positioned 7 cm from the center of lips, as was done in previous research [38]. This recording was post-processed to remove all the adult speech parts, silences, and other artefacts. The remaining speech was segmented into smaller sentences, and the fundamental frequencies of 3/4 sentences were changed. Hence, speech from 4 child voices was created for use in the auditory scene. Individual components of the speech segments were only presented once. In the final auditory scene,

two child voices were active at any time. Following a previous study [38], the order of active talkers changed randomly.

To auralize the binaural background sounds, a classroom model was generated in SketchUp [39]. The classroom has a rectangular floor plate and flat ceiling ($9\text{ m} \times 11\text{ m} \times 2.7\text{ m}$, $W \times L \times H$), and all the room surfaces were assigned an absorption coefficient of 1 (i.e., anechoic conditions). Within the room, 12 desks and chairs were modeled in four rows. Each chair seated a child. Typical sound absorption values were assigned to each child. The listener's position referred to a child sitting in the middle of the last row with an ear height of 1 m. A total of 16 different sound sources were then placed in the room. Twelve sound sources for non-speech sounds were placed at varying distances around the listener; there were equal numbers of sound sources on the left and right sides. The directionality of the last four sound sources represented two child talkers at 3 m away at ± 30 degrees, and two talkers at 5 m away at ± 20 degrees, with an ear height of 1 m. The classroom model was auralized using RAVEN [40]. A generic head-related transfer function (HRTF) from the FABIAN dummy head [41] with a resolution of $1^\circ \times 1^\circ$ was used. Although it is well known that HRTFs differ significantly between adults and children [42], and that this difference can influence cognitive tasks [31], a generic solution of HRTFs was used in the current study, as this was the first attempt to spatially separate the sounds. Thus, in the binaural condition, the sounds were spatially spread across the room, and the order of active talkers and source locations of non-speech sounds changed randomly, as is typical in real classrooms. In the monaural condition, the sounds were presented without any spatial separation, and appeared to come from straight ahead (or inside the head) to the listener. For both sound conditions, the auralized files were mixed down to a 2-channel audio file. The headphone (Sennheiser HD 650, Wedemark, Germany) output per channel was calibrated using Brüel and Kjær Artificial Ear Type 4153 with a Brüel and Kjær Type 4190 omnidirectional microphone capsule.

For both sound conditions, the $L_{A,eq}$ of the target speech signals and the classroom noise were 60 and 63 dB, yielding a SNR of -3 dB [29].

Tasks: We used modified versions of the tasks from Klatt and colleagues [29]. As we included three sound conditions (silence, monaural noise, binaural noise), three equivalent, parallel versions of each task were constructed.

Speech Perception: A word-to-picture matching task requiring discrimination between phonologically similar words was used to measure speech perception. A total of 84 lists of four phonologically similar German nouns (for example, Kopf (head), Topf (pot), Knopf (button), and Zopf (braid)) were constructed. Each word was represented by a simple and easy-to-name colored drawing. Each trial began with a 1.5 s visual cue, followed by a spoken word. Then, the screen displayed four images in a fixed array, one representing the target word and three representing similar-sounding distractor words (see Figure 1). The position of the picture representing the target words was counterbalanced. The participant's task was to mouse-click on the picture that corresponded to the target word. There were 28 trials in each sound condition. The task was the same for both adults and children.

Listening comprehension: A paper and pencil test requiring the execution of complex oral instructions was used to assess listening comprehension. In each of the sound conditions, participants heard 8 oral instructions, such as "Male ein Kreuz unter das Buch, das neben einem Stuhl liegt" ("Draw a cross under the book that is next to the chair"). The task was to carry out the instructions on pre-prepared response sheets. On the response sheets, each instruction was represented by a row of small black-and-white drawings of the target objects (e.g., a book next to a chair) and distractor stimuli (e.g., a book next to a ball) (see Figure 2). The response sheet was also visible on the computer screen in front of the participant. A red arrow indicated the row reflecting the current instruction on the response sheet. Each instruction began with an auditory cue (bell ringing). Participants had 18 s after the end of an instruction to complete the entries on the response sheet. They were instructed to begin carrying out the instructions as soon as feasible. Scoring was based on the number of elements correctly executed according to the respective instructions. In order

to equalize task difficulty across age groups, the adults received longer and syntactically more complex instructions.

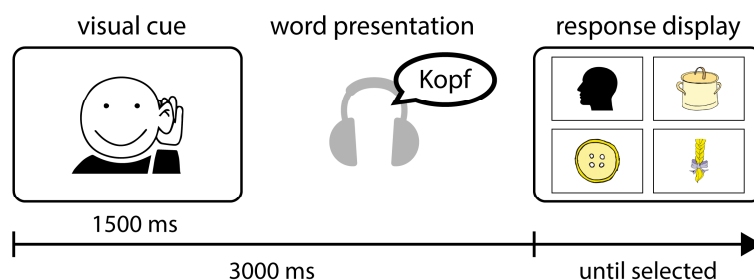


Figure 1. The experimental procedure for measuring speech perception with the word-to-picture matching task. The visual cue was displayed for 3000 ms, indicating the onset of the spoken target word presented over headphones. The target word was played over headphones 1500 ms after the onset of the visual cue. Thereafter, the response display was shown, which comprised four pictures, one representing the target word (here: Kopf (head)) and three representing phonologically similar distractor words (Topf (pot), Knopf (button), and Zopf (braid)).

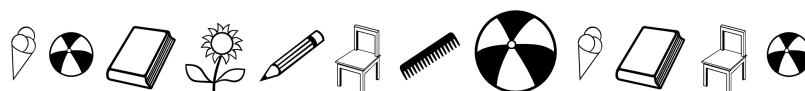


Figure 2. Exemplary trial from the listening comprehension task for the children. “Male ein Kreuz unter das Buch, das neben einem Stuhl liegt” (“Draw a cross under the book that is next to the chair”).

Subjective disturbance was assessed using a smiley scale with 4 points representing the ratings “not at all disturbed” (0), “a little disturbed” (1), “strongly disturbed” (2), or “extremely disturbed” (3).

Procedure: Both children and adults were tested in separate groups of 2 to 4 in a sound-attenuated booth at the University of Kaiserslautern-Landau. Four computer workplaces were arranged in the room, with about 4 m and partition walls between them. The walls around each workstation were equipped with posters of a primary school classroom to create a more classroom-like atmosphere. Adults received written instructions. Children were instructed orally by a researcher. Participants were informed that they should ignore the sounds and focus solely on the execution of the respective task. The experimenter stayed in the back of the laboratory room during the whole session.

Each participant performed both tasks in each of the three sound conditions (silent control, monaural, and binaural auditory classroom scene). Sound conditions were varied in blocks of trials. There were 28 and 8 trials per block in the word identification and listening comprehension tasks, respectively. The order of sound conditions and the allocation of test versions to sound conditions were counterbalanced between participants.

Each session started with a general instruction provided by the experimenter, followed by the monaural presentation of the classroom scene for 4 s to familiarize the participants with the background sound. Then, all the pictures used in the speech perception task were presented and named. Subsequently, participants performed the speech perception task. Thereafter, the listening comprehension task was instructed and performed. Both tasks started with four practice trials. In the sound conditions, the classroom-noise scenario was continuously played during the respective block of trials. Finally, the monaural and binaural auditory scenes were played for 10 s in order to complete the disturbance ratings. The session took about 40 min in total.

2.2. Results

Mean proportion correct scores and standard deviation as a function of task, sound condition, and age group are depicted in Table 1. Difference scores were calculated for each participant by subtracting proportion correct scores in noise from performance in silence.

These scores were used as dependent variables. Figure 3 depicts the mean difference scores with respect to task, sound condition, and age group.

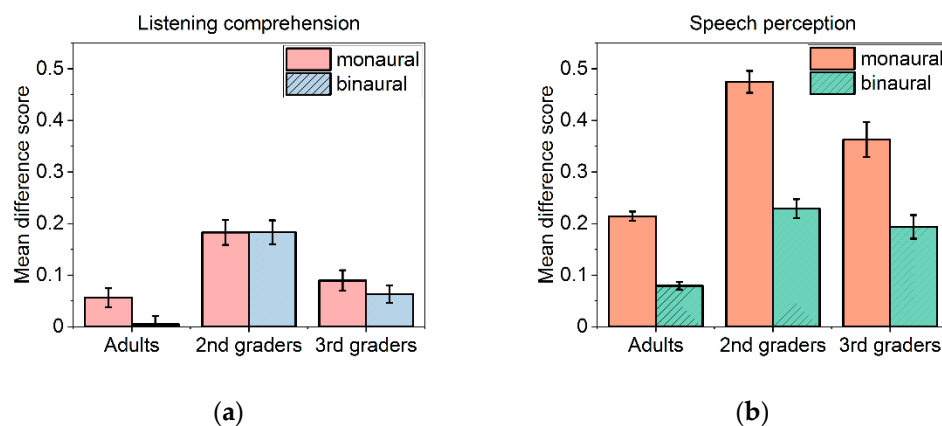


Figure 3. Mean difference scores for speech perception (a) and listening comprehension (b) with respect to age group and sound condition. Error bars denote standard errors of the mean.

Table 1. Mean proportion correct scores for speech perception and listening comprehension, and mean disturbance ratings for Experiment 1 as a function of sound condition and age group (standard deviation in parenthesis).

Task/Ratings	Sound Condition	Adults	2nd Graders	3rd Graders
		M (SD)	M (SD)	M (SD)
Speech Perception	Silence	0.99 (0.00)	0.97 (0.05)	0.96 (0.06)
	Monaural	0.79 (0.05)	0.50 (0.13)	0.61 (0.13)
	Binaural	0.92 (0.04)	0.74 (0.11)	0.77 (0.11)
Listening comprehension	Silence	0.85 (0.10)	0.94 (0.01)	0.94 (0.08)
	Monaural	0.79 (0.12)	0.76 (0.15)	0.85 (0.12)
	Binaural	0.84 (0.11)	0.76 (0.14)	0.87 (0.07)
Disturbance rating	Monaural	1.25 (0.55)	0.97 (0.93)	1.11 (0.66)
	Binaural	2.00 (0.72)	1.35 (0.98)	1.53 (0.84)

For speech perception, a 3 × 2 mixed ANOVA of the difference scores with age group (adults, second-graders, third-graders) as a between-subjects factor and sound condition (monaural, binaural) as a within-subject factor confirmed a significant main effect of the sound condition— $F(1, 89) = 222.31, p < 0.001, \text{partial } \eta^2 = 0.714$ —reflecting stronger impairment in the monaural when compared to the binaural condition; a significant main effect of age group— $F(2, 89) = 56.29, p < 0.001, \text{partial } \eta^2 = 0.558$ —reflecting stronger impairments in the children when compared to adults; and a significant interaction— $F(2, 89) = 9.19, p < 0.001, \text{partial } \eta^2 = 0.171$. Post hoc *t*-tests revealed that, in each of the age groups, the ability to recognize isolated words was less impaired in the binaural condition when compared to the monaural condition (all *p*'s < 0.001). The interaction reflects a more pronounced difference between age groups in the monaural when compared to the binaural noise condition (see Figure 1). Separate analyses per sound condition revealed significant differences between all age groups for monaural noise, and significant differences between adults and both groups of children for binaural noise (all *p*'s < 0.01), whereas second- and third-graders did not differ (*p* = 0.190).

For listening comprehension, the 3 × 2 mixed ANOVA revealed a significant main effect of age group— $F(2, 89) = 24.87, p < 0.001, \text{partial } \eta^2 = 0.359$. The effect of sound condition and the interaction were not significant ($F(1, 89) = 2.67, p = 0.106, \text{partial } \eta^2 = 0.029$ and $F(2, 89) = 1.17, p = 0.314, \text{partial } \eta^2 = 0.026$). Concerning the main effect of age, post hoc tests revealed that second-graders were more impaired than adults (*p* < 0.001) and third-graders (*p* < 0.001). No significant differences were found between third-graders

and adults ($p = 0.27$). One-sample t -tests revealed that the performance decrement due to binaural noise in adults did not differ significantly from 0; $t(35) < 1$.

In a further step, we analyzed whether speech perception in noise predicts listening comprehension in noise. In view of the small sample size for the third-graders and the fact that the adults' listening comprehension was largely unaffected by noise, the respective correlation analyses was confined to the second-graders. Correlations between proportion correct scores for speech perception and listening comprehension in noise were calculated. In the binaural condition, speech perception was significantly related to listening comprehension, $r(35) = 0.375$, $p < 0.05$, whereas in the monaural condition, speech perception and listening comprehension were unrelated ($p = 0.19$).

For the disturbance ratings, the 3×2 mixed ANOVA yielded a significant main effect of noise condition— $F(2, 89) = 33.10$, $p < 0.001$, *partial* $\eta^2 = 0.271$ —reflecting higher disturbance in the binaural when compared to the monaural condition; a significant main effect of age group— $F(2, 89) = 4.20$, $p < 0.05$, *partial* $\eta^2 = 0.09$ —but no interaction: $F(2, 89) = 2.09$, $p = 0.13$. Post hoc tests confirmed that the disturbance ratings were higher in adults when compared to second-graders ($p < 0.05$). Ratings of the third-graders did not differ from those of the second-graders and adults ($p = 1$ and $p = 0.35$, respectively). Further analyses in the second-graders and adults confirmed that, for both noise conditions, the disturbance ratings were unrelated to word identification and listening comprehension in noise (proportion correct scores) and unrelated to the actual performance decrements (difference scores) in adults (all p 's > 0.10) and children (all p 's > 0.40).

2.3. Discussion

Experiment 1 replicated the often-reported finding that children are less able than adults to understand speech in the presence of background noise [10–13,29]. In the monaural condition, word identification was more impaired in second-graders when compared to third-graders, and more impaired in third-graders than in adults. However, in each of the age groups, word identification performance was substantially less affected with binaural noise when compared to monaural presentation of the classroom-noise scenario. This indicates that children and adults may use the spatial cues inherent in the binaural scenario to support separation of the target words from the background noise. The age effect observed for monaural noise was significantly reduced (although still significant) in the binaural condition. These results suggest that the effects of classroom noise on speech perception and the developmental change associated with these effects are strongly moderated by the method of sound presentation. Especially, with a simple, monaural presentation that lacks cues to spatially separate the speech signal from the noise, impairments of speech perception due to real-life environmental noise and their increase with decreasing age might be overestimated. The dominant role of spatial cues is further confirmed by the fact that, in the study of Klatte et al. [29], the impairment of speech perception due to classroom noise was considerably lower than the effects in the monaural condition, but comparable to the binaural condition of the current study (in [29], difference scores were 0.23 in children and 0.12 in adults). This was presumably because in [29], the target words were presented through a separate loudspeaker, thereby allowing spatial separation of signal and background noise.

Concerning listening comprehension, adults showed only minor disruption in the monaural condition (5%) and no significant disruption in the binaural condition. This result replicates the findings of Klatte et al. [29] and can be attributed to the adult listeners' ability to reconstruct noise-masked elements of the speech signals using contextual cues. The age effect observed in the speech perception task was partially replicated; i.e., the second-graders were more affected by background noise when compared to third-graders and adults. However, by contrasting the results in the speech perception task, we can see that the impairment of listening comprehension performance did not differ between sound conditions (monaural vs. binaural). Furthermore, for binaural noise, children's speech perception significantly predicted listening comprehension, whereas for monaural

noise, speech perception and listening comprehension were unrelated. We may therefore conclude that, with binaural presentation of the noise, the effects on word identification allow a more valid prediction of the effects on complex listening tasks when compared to a simple monaural presentation.

Concerning the disturbance ratings, the current results replicate the findings of Klatte et al. [29] that, in view of the noise-induced performance decrements, average ratings of the children were surprisingly low. This was especially true for the monaural noise condition. Despite severe decrements in word identification, children judged this sound condition on average as “a bit disturbing.” The lowest ratings were provided by the second-graders, who showed the strongest performance impairments. Still more surprisingly, across age groups, the binaural sound scenario was judged as more disturbing, even though the speech perception decrements were much stronger in the monaural condition, and the ratings were unrelated to the actual performance decrements. Even though, with age-adequate measurement scales, children are able to provide valid ratings of general annoyance due to noise in their home and school environment [6,9,43,44], the current findings indicate that both children and adults are not aware of the detrimental effects of background noise on their listening performance. In view of this result, it is evident that teachers and researchers cannot rely on students’ ratings when evaluating the acoustic environments of classrooms.

Taken together, Exp. 1 adds to the evidence that children are more impaired than adults by noise in speech perception tasks [10–13,29], and extends this finding to listening comprehension in a realistic classroom-noise scenario. Furthermore, the results revealed that the noise effects and their interaction with age differ considerably between conditions of binaural vs. monaural presentation of the noise. In Exp. 2, we explored whether these differences hold also for cross-modal noise effects, i.e., the effects of noise on the processing of visually presented information.

3. Experiment 2: Effects of Classroom Noise on Visual–Verbal Short-Term Memory

In Experiment 2, the effects of a classroom-noise scenario on short-term memory for visual–verbal items were analyzed in children and adults. Verbal short-term memory is the ability to hold verbal information in an active state for ongoing cognitive processes. A standard task to assess verbal short-term memory requires immediate serial recall of sequences of 5 to 9 verbal items, such as words, digits, or easy-to-name pictures. In such tasks, participants usually employ a strategy called articulatory rehearsal, i.e., repetitive subvocal pronunciations of the list items in a sequential manner. This strategy is evident in children from about age 7, but there is considerable interindividual variation around this age [45,46].

In the current study, we used verbal short-term memory as a cognitive task for two reasons. First, the capacity of short-term memory is a predictor of children’s oral and written language acquisition, and short-term memory processes play a role in many school-based tasks, such as reading and spelling in the early grades, learning new vocabulary, mental arithmetic [47], and following the teachers’ instructions [48]. Second, short-term memory is especially susceptible to the adverse effects of noise. Many studies confirmed that performance in the serial recall task is reliably impaired by task-irrelevant background sounds (for recent reviews, see [21,49]). This “irrelevant sound effect” (ISE) is most pronounced with speech noise, but is also evoked by nonspeech sounds, such as tone sequences or instrumental music (for an overview cp. [50]). The ISE is reliable for coherent sound streams that consist of changing auditory elements emanating from a single source, e.g., fluent speech, sequences of different syllables or tones, and instrumental music. Sounds lacking these “changing-state” characteristics, such as continuous broadband noise, babble speech, or spectrally degraded speech, evoke minor or no disruption. Different theoretical explanations have been provided for the ISE. Some authors [18,51,52] attribute the effect to a diversion of attention away from the task and towards the sound (attention capture-account). Others oppose a role of attention, assuming that the effect results

from specific interference between processes involved in automatic, obligatory processing of the background sound and deliberate processes involved in memorizing the verbal items [53]. Within the interference account, some authors propose serial-order retention as the mechanism of disruption [54], whereas others assume that noise—especially noise with speech—interferes with storage and processing of phonological representations [55,56].

Aiming to disentangle attention-capture from interference-by-process, a number of studies on the ISE included children. In view of children's underdeveloped attention control [57], noise effects resulting from attention capturing should be more pronounced in children when compared to adults. However, this argument is not without problems, since, at least with noise-containing speech, stronger impairments in children may also indicate stronger speech-based interference due to less robust, immature phonological representations and/or maintenance strategies, i.e., articulatory rehearsal of the item sequence [19,21].

While the detrimental effects of irrelevant sounds on children's verbal short-term memory were consistently reported, the findings concerning developmental change are inconsistent. Some studies reported equivalent impairments due to background speech or mixtures of nonspeech sounds with speech in 7- to 10-year-olds and adults [19,20,58–61], whereas others found stronger impairments in the children [18,19,21,62]. Two of these studies included classroom noise [20,59]. In Klatt et al. [20], children's and adults' serial recall performance was equally affected by background speech, but only the youngest children (first-graders) were also impaired by a mixture of classroom sounds without speech. The authors attributed the age-independent effect of background speech to specific interference with the maintenance of phonological representations and the age-dependent classroom noise effect on attentional capture. Meinhardt-Injac et al. [59] used a classroom-noise scenario with speech (bits of conversation between children and adults) and found significant and equivalent impairments of serial recall performance in 8–10-year-olds, 11–12-year-olds, and adults. Across age groups, the noise effects were unrelated to participants' attention control, i.e., their ability to inhibit task-irrelevant information. This finding indicates that specific interference rather than attention capture is the mechanism underlying the noise-induced disruptions.

The vast majority of ISE studies have used simple, monaural presentation of irrelevant sounds. Only a few studies have included spatially spread sound sources. These studies provided evidence that the variation of the source location moderated the sounds' disruptive effects. Buchner et al. [63] showed that the ISE evoked by nonspeech sounds (e.g., footsteps, cries of pain, and squeaking sounds) and speech (sequences of unrelated words) played through loudspeakers from different locations was most pronounced when the sound was played from the front, i.e., from a location near the visual target display to which the participants' attention was directed. However, the source location had only a small impact on the sound-induced disruption, and the disruption evoked by speech was substantially stronger than that evoked by nonspeech sounds. These findings indicate a significant but minor role of attention capturing in the ISE in adults. Jones and Macken [64] analyzed the effects of background speech produced by six voices simultaneously. The speech was presented through loudspeakers located in a circle around the participant. The impairment of short-term memory performance was more pronounced when each voice was assigned to a separate loudspeaker (yielding six single-talker streams), when compared to assigning a mix of the six voices to each of the six loudspeakers (yielding identical streams of babble speech). Comparable results were reported in studies using dichotic vs. monaural headphone presentation of irrelevant syllables [65,66] and interpreted as evidence for specific interference through changing state speech.

In view of these findings, we might expect a stronger effect of the binaural when compared to the monaural classroom-noise scenario, through increased speech-based interference due to clearer separation of the speech streams, and/or increased attention capture due to spatially spread sound sources and changing source locations. If attention

capture is the dominant source of disruption, children should be more impaired than adults, and more impaired by the binaural than the monaural noise scenario.

3.1. Materials and Methods

Participants: The sample consisted of 40 student volunteers (24 female), aged between 19 and 32 years ($M = 24.0$, $SD = 2.5$ years), from the University of Kaiserslautern-Landau; and 69 third- and fourth-grade children recruited from a primary school in Kaiserslautern. Due to technical issues, the data of two children had to be excluded from the analysis. The final child sample consisted of 67 children (36 female), aged between 8 years, 2 months and 10 years, 3 months ($M = 9$ years, 4 months, $SD = 6$ months). Of the children, 19 had taken part in Experiment 1. All participants were native German speakers and had normal or corrected-to-normal vision and normal hearing according to either self-reports (adults) or parental reports (children). The study was approved by the Rhineland-Palatinate school authority and by the Ethics Committee of the University of Kaiserslautern. Informed written consent was provided by the adult participants and by the children's parents. Adults received either course credit or payment for participation (10 €).

Apparatus: Identical to Experiment 1.

Background noise: Identical to Experiment 1.

Task: The task required serial recall of sequences of monosyllabic German nouns presented pictorially. Pictures were used instead of written words in order to avoid confounding by the children's reading abilities. Prior studies confirmed that children and adults use verbal strategies when memorizing words presented pictorially [46,67], and that participants' strategies do not differ between pictorial and written presentation [68]. Each trial consisted of a presentation phase, a retention interval, and a recall phase. Pictures were presented one after another in a 102×73 mm rectangular black frame in the center of a white screen, with a presentation duration of 1500 ms and an interstimulus interval of 500 ms. A random interval between 1200 to 1800 ms passed by before the visual presentation of the first list item. The final list item was followed by a 5000 ms retention interval. The onset of the recall phase was signaled by the simultaneous re-presentation of all stimuli. The pictures were arranged at random in a fixed array of five (children) and eight (adults) black frames (see Figure 4). Participants had to reconstruct the serial order by using the mouse to click on the items in the presentation order. Clicking an item changed its shading, indicating that it had been selected. There was no time limit for responding and no possibility of error correction. After selection of the final item, participants were presented with a visual cue to start the next trial by pressing the space bar.

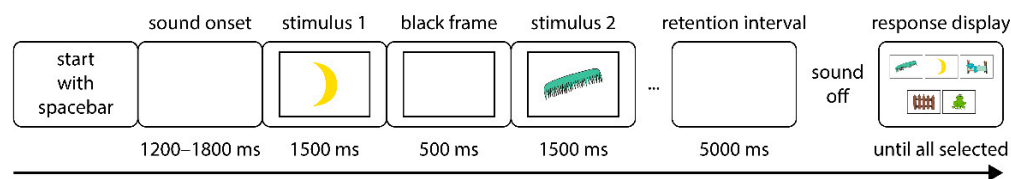


Figure 4. The serial recall task's experimental procedure. Five pictures per trial were shown to the children, whereas eight pictures per trial were presented to the adults. All pictures seen in the respective trial were randomly arranged in an array of 5 (children) and 8 (adults) frames in the response display.

Both children and adults saw colored drawings representing the monosyllabic German words *Bett*, *Bus*, *Eis*, *Frosch*, *Kamm*, *Mond*, *Pilz*, *Schal*, *Schiff*, and *Zaun* (*bed*, *bus*, *ice*, *frog*, *comb*, *moon*, *mushroom*, *scarf*, *ship*, and *fence*). The set for the adults additionally included the items *Brief*, *Haus*, *Herz*, *Hut*, *Nuss*, and *Schwein* (*letter*, *house*, *heart*, *hat*, *nut*, and *pig*). Four lists of five items (drawn out of 10) were created for the children, and six lists of eight items (drawn out of 16) for the adults. Two additional versions of each list were created using random permutations of the list items.

Procedure: Both children and adults were tested in groups of 2 to 4 in a sound-attenuated booth at the University of Kaiserslautern-Landau (see Exp. 1). Adults received written instruction. Children were instructed orally by a researcher. Participants were informed that they should ignore the sounds and focus solely on the serial recall task. At the beginning of each session, the classroom scenarios were played for 4 s, followed by the presentation of all pictures used in the task. Each picture was named by a female speaker. Following the instruction, three practice trials (one per sound condition) were performed. Thereafter, children and adults completed 24 and 48 experimental trials, respectively (8 and 16 trials per sound condition). Sound conditions (silence, classroom noise—monaural, classroom noise—binaural) were varied in blocks of trials. The order of sound conditions was balanced across participants. In the sound blocks, the sound started when the participant initiated the first trial and terminated after finishing the recall phase of the final trial of the respective block. Sounds were presented via headphones at an average level of 60 dB(A). The testing session lasted about 20 min for children and 35 min for adults.

3.2. Results

The dependent variable was the proportion of correct scores based on the number of items recalled at the correct serial position. Proportions of correct scores with respect to age group and sound condition are depicted in Figure 5a. A two-way mixed ANOVA with sound condition (silent control, monaural, binaural) as the within-subject factor and age group (adults, children) as the between-subjects factor revealed significant main effects of sound condition ($F(2, 210) = 8.69, p < 0.001, \text{partial } \eta^2 = 0.08$) and age group ($F(1, 105) = 6.73, p < 0.05, \text{partial } \eta^2 = 0.06$). Bonferroni-corrected post-hoc tests revealed that performance in both noise conditions was significantly lower when compared to the silent control condition, $p < 0.01$, whereas the noise conditions did not differ ($p = 0.99$). The main effect of age group reflects better overall performance of the adults. The sound condition \times age group interaction was not significant ($F < 1$), confirming comparable noise-induced disruption in adults and children. The analyses thus confirmed significant impairments of serial recall performance due to classroom noise. The effects did not differ between age groups, nor between monaural vs. binaural noise.

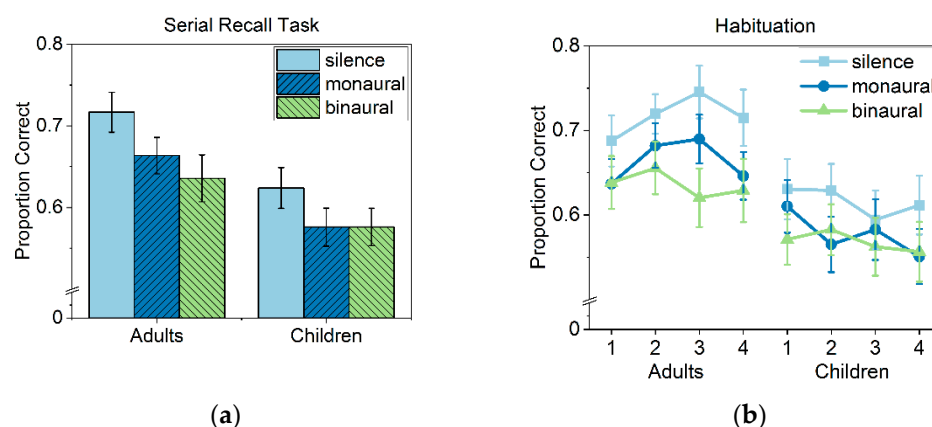


Figure 5. (a) Mean proportion correct scores as a function of age group and sound condition. (b) Mean proportion of correct scores as a function of age group and as a function of sound condition in four consecutive blocks of four (adults) and two (children) trials. Error bars denote standard errors of the mean.

Aiming to assess the role of attention in the noise-induced disruption, further analyses were performed to explore whether or not participants habituated to the noise across trials. If the noise effects result from attention capture, one might expect habituation and thereby a stronger disruption in the first when compared to the final trials performed with noise [69]. For this aim, the proportion of correct scores was calculated for four consecutive blocks of two trials (children) and four trials (adults) for each sound condition. The resulting

proportion correct scores are depicted in Figure 5b. A $3 \times 4 \times 2$ mixed ANOVA with sound condition (silent control, monaural, binaural) and trial block (block 1–block 4) as within-subject factors and age group (adults, children) as the between-subjects factor was conducted. Except for sound condition and age group (reported above), neither trial block nor any interaction reached significance (all $F < 1$). The analysis thus yielded no evidence for habituation to classroom noise in children or adults.

3.3. Discussion

In Experiment 2, the impacts of monaural and binaural classroom-noise scenarios on verbal short-term memory were examined in children and adults. The task required serial recall of words presented pictorially. Children's and adults' performances were significantly and equally impaired in both noise conditions. The magnitude of the noise effects remained stable over the course of experimental trials.

The non-significant effect of age replicates the findings of Meinhardt-Injac et al. [59], who reported significant and equivalent impairments due to classroom noise in children and adults. However, in Klatter et al. [20], only the youngest children were affected by classroom noise, whereas older children and adults were unaffected. The apparent contradiction might be attributed to the kinds of classroom noise and age groups included. In both the current and the Meinhardt-Injac et al. [59] study, the classroom noise contained speech, and the youngest children were age 9 on average, whereas Klatter et al. [20] used nonspeech classroom noise and included 6- to 7-year-old first-graders. In line with the arguments provided in these studies, we propose that the detrimental effects of the classroom-noise scenario result from separate mechanisms. The spoken parts in the noise scenarios evoke specific interference with the maintenance of the list items. This mechanism may evoke stronger disruption in children whose phonological maintenance strategies are not yet fully developed and thus more prone to speech-based interference. In addition, both nonspeech sounds and speech may impair performance through attention capture. The impact of attention capture depends on the sound's potential to grab attention (i.e., personal relevance, emotional valence, and predictability) and on the individuals' attentional abilities. Young children are more vulnerable than older children and adults due to less developed attention control. As schooling contributes considerably to children's development of attention control [70], preschool children and first-graders are especially vulnerable to noise-induced attention capture. Taken together, the findings of the current study add to the evidence provided by Klatter et al. [20] and Meinhardt-Injac et al. [59] that children aged around 9 years or older show adult-like impairments of short-term memory in the presence of classroom noise.

As outlined above, from both the attention capture and the interference account, one might expect stronger disruptive effects with binaural when compared to monaural presentation of the classroom-noise scenario. Following the attention capture account, especially in children, the binaural scenario should evoke stronger disruption because of its attention-grabbing quality (spatially distributed and changing sound source locations). However, contrary to expectation, the disruptive effects of the classroom-noise scenario did not differ with presentation mode, neither in children nor in adults. Furthermore, in both age groups, the detrimental effects of the noise scenarios remained stable across experimental trials—i.e., there was no evidence for habituation. These findings add to the evidence provided by Meinhardt-Injac et al. [59] and Klatter et al. [20] that diversions of attention play a minor role in the ISE, at least in children older than 8 years. This view was further confirmed in a recent study [21], demonstrating that, while 9-year-old third-graders were more impaired than adults by background speech in a verbal serial recall task, serial recall of nonverbal, visuo-spatial items was unaffected in both groups. As storage and processing of visuo-spatial items rely heavily on domain-general attentional resources [71,72], these findings strongly suggest that the impairment in the verbal task, and its interaction with age, reflect speech-based interference rather than a capture of attention.

Following the interference account, binaural sound scenarios comprising speech should evoke stronger disruption because the spatial separation of the sound sources fosters the segregation of the spoken parts into separate, single-talker streams [64]. Evidently, this mechanism was not at play in the current study. This might be due to the fact that, in the classroom-noise scenario used here, all spoken parts consisted of only two talkers simultaneously. Jones and Macken [64] showed that monaural presentations of single-talker speech and a mixture of two speakers evoked similar amounts of disruption in serial recall performance, whereas with six simultaneous speakers, the disruption was significantly reduced. Thus, concerning the current classroom-noise scenario, a clearer separation of the two speech streams through spatial cues might not further increase the interference, because the latter is already at its maximum (resembling that evoked by a single talker). On this assumption, stronger interference effects with binaural presentation should occur with noise scenarios containing more than two simultaneous talkers.

4. Conclusions

In the current study, the effects of a realistic classroom-noise scenario presented either monaurally or binaurally on speech perception, listening comprehension, and verbal short-term memory were investigated in primary school children and adults. In Exp. 1, across age groups, speech perception (identification of spoken words) was more impaired by monaural than by binaural classroom noise, whereas listening comprehension (acting-out complex oral instructions) was equally impaired in both noise conditions. In both tasks, children were more affected than adults. The age effect found here is in line with a number of psychoacoustic studies documenting increasing noise-induced speech perception impairments with decreasing age [10–13] and extends these findings to a realistic noise scenario and a complex listening task that more closely reflects the requirements of children faced during classroom instruction. Disturbance ratings were unrelated to the actual performance decrements, indicating that listeners' subjective reports do not allow a valid evaluation of adverse listening conditions. In Exp. 2, using a paradigm from the domain of the ISE, we found significant detrimental effects of the classroom-noise scenario on verbal short-term memory (serial order reconstruction of words presented pictorially), which did not differ with age or presentation format (monaural vs. binaural). Concerning the age-equivalence of the noise effect, the current finding replicates a recent study demonstrating comparable effects of classroom noise-containing speech in children aged 8 to 10 years and adults [59]. The lack of an effect of presentation format was contrary to our expectation. Especially for the children, we anticipated stronger impairments with the binaural condition due to increased attention capture through spatially separated and changing sound sources. The equivalence of the noise effects across age groups and presentation formats adds to the evidence that the ISE evoked by noise-containing speech results from specific interference between obligatory processing of the speech parts and deliberate processes involved in serial recall performance (i.e., maintenance of phonological representations), whereas attention capture plays a minor role.

Concerning age effects, children were more impaired by noise than adults in a complex listening task requiring processing of oral instructions, but equally impaired as adults in a task requiring processing of visually presented information. This indicates that, in verbal working memory tasks, children are more prone to distraction than adults when the targets and the irrelevant stimuli stem from the same sensory modality (unimodal interference) but are equally affected (or equally unaffected) as adults when the irrelevant stimuli and the targets originate from different sensory modalities (crossmodal interference, i.e., interference between visual and auditory information). It has been shown that selective attention (i.e., the ability to focus on the target stimuli and inhibit irrelevant distractors) is much more easily achieved in crossmodal paradigms when compared to unimodal paradigms. This holds especially for paradigms with visual targets and auditory distractors [73,74]. These findings have been attributed to top-down suppression of auditory distractors at a very early processing stage (at the level of the cochlea). As a consequence, age differences

in attention control should be largely unrelated to auditory distraction when the relevant information is visual. This has been shown in studies including older adults [73–75], but as proposed by Röer et al. [60], may also hold for children. Following this view, age differences in the ISE between children and adults should emerge when the memory items are presented auditorily instead of visually. This prediction might be tested in future studies.

Concerning the impact of the noise presentation format (monaural vs. binaural), differential effects were found only in the speech perception task. The detrimental noise effect was much stronger with monaural when compared to binaural noise, and the difference between noise conditions was especially strong in the youngest (second-grade) children. These findings indicate that, with standard, monaural presentation of the maskers, the effects of noise on speech perception in real-life listening situations and the developmental change associated with speech-in-noise perception might be overestimated. Furthermore, we found that speech perception in binaural noise significantly predicted listening comprehension in binaural noise, whereas with monaural noise, speech perception and listening comprehension were unrelated. These results suggest that studying speech-in-noise perception using binaural scenarios may have the potential to allow valid predictions of detrimental noise effects on language comprehension in everyday situations. Even though the current study yielded no evidence for an effect of the noise presentation format (monaural vs. binaural) on listening comprehension and visual–verbal short-term memory, we cannot rule out that such effects exist. In the classroom-noise scenario used here, no more than two voices are presented simultaneously, and the nonspeech sounds are confined to indoor noise, disregarding noise sources from outside such as road traffic or aircraft noise. With classroom-noise scenarios containing more unexpected and novel sounds and/or more simultaneous voices, or in samples of younger children, a significantly stronger impairment with binaural when compared to monaural noise presentation might emerge.

Despite these limitations, the current study adds to the evidence concerning developmental change in the effects of noise on speech perception, listening comprehension, and short-term memory, and extends it to a realistic classroom-noise scenario. Further research is needed to find out whether and how much the use of binaural sound scenarios in experimental studies allows more valid predictions of the effects of noise in everyday, real-life learning situations.

Author Contributions: Conceptualization, M.K. and S.J.S.; methodology, M.K.; software, L.L.; validation, L.L. and M.K.; formal analysis, L.L.; investigation, L.L.; resources, C.B., M.Y. and T.L.; data curation, L.L.; writing—original draft preparation, L.L. and M.K.; writing—review and editing, all; visualization, L.L.; supervision, M.K.; project administration, L.L., C.B. and S.F.; funding acquisition, S.F., J.F., A.R., M.K., S.J.S. and T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the project ID 444697733 and ID 401278266.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Deutsche Gesellschaft für Psychologie (DGPs, German Psychological Society) and by the Rhineland-Palatine school authority (Aufsichts- und Dienstleistungsbehörde, ADD).

Informed Consent Statement: Informed written consent was provided by the parents of the children and by the adult participants.

Data Availability Statement: All data can be found at: <https://osf.io/b9rua> (accessed on 31 October 2022).

Acknowledgments: The authors would like to thank all of the students, children, teachers, and parents for cooperating to accomplish this study. In addition, we would like to thank Michelle Turner, Fabio Sobotzki, and Markus Kurtz for assisting in data collection and technical support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Mogas-Recalde, J.; Palau, R.; Márquez, M. How Classroom Acoustics Influence Students and Teachers: A Systematic Literature Review. *J. Technol. Sci. Educ.* **2021**, *11*, 245–259. [\[CrossRef\]](#)
2. Loh, K.; Yadav, M.; Persson Waye, K.; Klätte, M.; Fels, J. Toward Child-Appropriate Acoustic Measurement Methods in Primary Schools and Daycare Centers. *Front. Built Environ.* **2022**, *8*, 688847. [\[CrossRef\]](#)
3. Shield, B.; Dockrell, J.E. External and Internal Noise Surveys of London Primary Schools. *J. Acoust. Soc. Am.* **2004**, *115*, 730–738. [\[CrossRef\]](#) [\[PubMed\]](#)
4. McAllister, A.M.; Granqvist, S.; Sjölander, P.; Sundberg, J. 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. *J. Voice* **2009**, *23*, 587–593. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Clark, C.; Head, J.; Stansfeld, S.A. Longitudinal Effects of Aircraft Noise Exposure on Children’s Health and Cognition: A Six-Year Follow-up of the UK RANCH Cohort. *J. Environ. Psychol.* **2013**, *35*, 1–9. [\[CrossRef\]](#)
6. Klätte, M.; Spilski, J.; Mayerl, J.; Möhler, U.; Lachmann, T.; Bergström, K. Effects of Aircraft Noise on Reading and Quality of Life in Primary School Children in Germany: Results from the NORAH Study. *Environ. Behav.* **2017**, *49*, 390–424. [\[CrossRef\]](#)
7. Shield, B.M.; Dockrell, J.E. The Effects of Environmental and Classroom Noise on the Academic Attainments of Primary School Children. *J. Acoust. Soc. Am.* **2008**, *123*, 133–144. [\[CrossRef\]](#)
8. Stansfeld, S.A.; Berglund, B.; Clark, C.; Lopez-Barrio, I.; Fischer, P.; Öhrström, E.; Haines, M.; Head, J.; Hygge, S.; van Kamp, I.; et al. Aircraft and Road Traffic Noise and Children’s Cognition and Health: A Cross-National Study. *Lancet* **2005**, *365*, 1942–1949. [\[CrossRef\]](#)
9. Klätte, M.; Hellbrück, J.; Seidel, J.; Leistner, P. Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: A Field Study. *Environ. Behav.* **2010**, *42*, 659–692. [\[CrossRef\]](#)
10. Buss, E.; Leibold, L.J.; Porter, H.L.; Grose, J.H. Speech Recognition in One- and Two-Talker Maskers in School-Age Children and Adults: Development of Perceptual Masking and Glimpsing. *J. Acoust. Soc. Am.* **2017**, *141*, 2650–2660. [\[CrossRef\]](#)
11. Talarico, M.; Abdilla, G.; Aliferis, M.; Balazic, I.; Giaprakis, I.; Stefanakis, T.; Foenander, K.; Grayden, D.B.; Paolini, A.G. Effect of Age and Cognition on Childhood Speech in Noise Perception Abilities. *Audiol. Neurotol.* **2006**, *12*, 13–19. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Stuart, A. Development of Auditory Temporal Resolution in School-Age Children Revealed by Word Recognition in Continuous and Interrupted Noise. *Ear Hear.* **2005**, *26*, 78–88. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Wróblewski, M.; Lewis, D.E.; Valente, D.L.; Stelmachowicz, P.G. Effects of Reverberation on Speech Recognition in Stationary and Modulated Noise by School-Aged Children and Young Adults. *Ear Hear.* **2012**, *33*, 731–744. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Boyle, R.; Coltheart, V. Effects of Irrelevant Sounds on Phonological Coding in Reading Comprehension and Short-Term Memory. *Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol.* **1996**, *49*, 398–416. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Halin, N.; Marsh, J.E.; Haga, A.; Holmgren, M.; Sörqvist, P. Effects of Speech on Proofreading: Can Task-Engagement Manipulations Shield against Distraction? *J. Exp. Psychol. Appl.* **2014**, *20*, 69–80. [\[CrossRef\]](#)
16. Kattner, F.; Hanl, S.; Paul, L.; Ellermeier, W. Task-Specific Auditory Distraction in Serial Recall and Mental Arithmetic. *Mem. Cognit.* **2022**. [\[CrossRef\]](#)
17. Klätte, M.; Bergström, K.; Lachmann, T. Does Noise Affect Learning? A Short Review on Noise Effects on Cognitive Performance in Children. *Front. Psychol.* **2013**, *4*, 578. [\[CrossRef\]](#)
18. Elliott, E.M. The Irrelevant-Speech Effect and Children: Theoretical Implications of Developmental Change. *Mem. Cognit.* **2002**, *30*, 478–487. [\[CrossRef\]](#)
19. Elliott, E.M.; Hughes, R.W.; Briganti, A.; Joseph, T.N.; Marsh, J.E.; Macken, B. Distraction in Verbal Short-Term Memory: Insights from Developmental Differences. *J. Mem. Lang.* **2016**, *88*, 39–50. [\[CrossRef\]](#)
20. Klätte, M.; Lachmann, T.; Schlittmeier, S.; Hellbrück, J. The Irrelevant Sound Effect in Short-Term Memory: Is There Developmental Change? *Eur. J. Cogn. Psychol.* **2010**, *22*, 1168–1191. [\[CrossRef\]](#)
21. Leist, L.; Lachmann, T.; Schlittmeier, S.J.; Georgi, M.; Klätte, M. Irrelevant Speech Impairs Serial Recall of Verbal but Not Spatial Items in Children and Adults. *Mem. Cognit.* **2022**. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Hammershøi, D.; Møller, H. Binaural Technique—Basic Methods for Recording, Synthesis, and Reproduction. In *Communication Acoustics*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 223–254.
23. Vorländer, M. *Auralization*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020.
24. Gheller, F.; Lovo, E.; Arsie, A.; Bovo, R. Classroom Acoustics: Listening Problems in Children. *Build. Acoust.* **2020**, *27*, 47–59. [\[CrossRef\]](#)
25. Kintsch, W. The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychol. Rev.* **1988**, *95*, 163–182. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Kjellberg, A.; Ljung, R.; Hallman, D. Recall of Words Heard in Noise. *Appl. Cogn. Psychol.* **2008**, *22*, 1088–1098. [\[CrossRef\]](#)
27. Hurtig, A.; van de Poll, M.K.; Pekola, E.P.; Hygge, S.; Ljung, R.; Sörqvist, P. Children’s Recall of Words Spoken in Their First and Second Language: Effects of Signal-to-Noise Ratio and Reverberation Time. *Front. Psychol.* **2016**, *6*, 2029. [\[CrossRef\]](#) [\[PubMed\]](#)

28. Ljung, R.; Kjellberg, A. Recall of Spoken Words Presented with a Prolonged Reverberation Time. In Proceedings of the Performance: 9th International Congress on Noise as a Public Health Problem (ICBEN), Mashantucket, CT, USA, 21–25 July 2008.
29. Klatte, M.; Lachmann, T.; Meis, M. Effects of Noise and Reverberation on Speech Perception and Listening Comprehension of Children and Adults in a Classroom-like Setting. *Noise Health* **2010**, *12*, 270–282. [CrossRef]
30. Shield, B.M.; Dockrell, J.E. The Effects of Noise on Children at School: A Review. *J. Build. Acoust.* **2003**, *10*, 97–106. [CrossRef]
31. Oberem, J.; Lawo, V.; Koch, I.; Fels, J. Intentional Switching in Auditory Selective Attention: Exploring Different Binaural Reproduction Methods in an Anechoic Chamber. *Acta Acust. United Acust.* **2014**, *100*, 1139–1148. [CrossRef]
32. Drullman, R.; Bronkhorst, A.W. Multichannel Speech Intelligibility and Talker Recognition Using Monaural, Binaural, and Three-Dimensional Auditory Presentation. *J. Acoust. Soc. Am.* **2000**, *107*, 2224–2235. [CrossRef]
33. Yost, W.A.; Sheft, S.; Dye, R. (Toby) Divided Auditory Attention with up to Three Sound Sources: A Cocktail Party. *J. Acoust. Soc. Am.* **1994**, *95*, 2916. [CrossRef]
34. Fintor, E.; Aspöck, L.; Fels, J.; Schlittmeier, S.J. The Role of Spatial Separation of Two Talkers' Auditory Stimuli in the Listener's Memory of Running Speech: Listening Effort in a Non-Noisy Conversational Setting. *Int. J. Audiol.* **2022**, *61*, 371–379. [CrossRef] [PubMed]
35. Peirce, J.W. PsychoPy-Psychophysics Software in Python. *J. Neurosci. Methods* **2007**, *162*, 8–13. [CrossRef] [PubMed]
36. European Broadcasting Union. R128-2020: Loudness Normalisation and Permitted Maximum Level of Audio Signals. Technical 734 Report. 2020. Available online: <https://tech.ebu.ch/docs/r/r128.pdf> (accessed on 12 July 2022).
37. Audacity Team Audacity(R): Free Audio Editor and Recorder [Computer Application]. Version 3.0.0. Available online: <https://audacityteam.org> (accessed on 2 December 2021).
38. Yadav, M.; Cabrera, D. Two Simultaneous Talkers Distract More than One in Simulated Multi-Talker Environments, Regardless of Overall Sound Levels Typical of Open-Plan Offices. *Appl. Acoust.* **2019**, *148*, 46–54. [CrossRef]
39. Trimble Inc. SketchUp 3D Design Software. Available online: <https://www.sketchup.com> (accessed on 18 July 2022).
40. Schröder, D.; Vorländer, M. RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments; Forum Acusticum: Aalborg, Denmark, 2011.
41. Brinkmann, F. The FABIAN Head-Related Transfer Function Data Base. 2017. Available online: <https://depositonce.tu-berlin.de/items/bff6568a-5735-4ebc-b3fa-ac10707b7beb> (accessed on 1 November 2022).
42. Fels, J.; Vorländer, M. Anthropometric Parameters Influencing Head-Related Transfer Functions. *Acta Acust. United Acust.* **2009**, *95*, 331–342. [CrossRef]
43. Waye, K.P.; van Kamp, I.; Dellve, L. Validation of a Questionnaire Measuring Preschool Children's Reactions to and Coping with Noise in a Repeated Measurement Design. *BMJ Open* **2013**, *3*, e002408. [CrossRef]
44. Quehl, J.; Bartels, S.; Fimmers, R.; Aeschbach, D. Effects of Nocturnal Aircraft Noise and Non-Acoustical Factors on Short-Term Annoyance in Primary School Children. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6959. [CrossRef]
45. Elliott, E.M.; Morey, C.C.; AuBuchon, A.M.; Cowan, N.; Jarrold, C.; Adams, E.J.; Attwood, M.; Bayram, B.; Beeler-Duden, S.; Blakstvedt, T.Y.; et al. Multilab Direct Replication of Flavell, Beach, and Chinsky (1966): Spontaneous Verbal Rehearsal in a Memory Task as a Function of Age. *Adv. Methods Pract. Psychol. Sci.* **2021**, *4*, 25152459211018187. [CrossRef]
46. Poloczek, S.; Henry, L.A.; Messer, D.J.; Büttner, G. Do Children Use Different Forms of Verbal Rehearsal in Serial Picture Recall Tasks? A Multi-Method Study. *Memory* **2019**, *27*, 758–771. [CrossRef]
47. Alloway, T.P. How Does Working Memory Work in the Classroom? *Educational Research and reviews* **2006**, *1*, 134–139.
48. Gathercole, S.E.; Alloway, T.P. *Working Memory and Learning: A Practical Guide for Teachers*; SAGE Publications: Thousand Oaks, CA, USA, 2008; ISBN 9781412936132.
49. Schlittmeier, S.J.; Marsh, J.E. Review of Research on the Effects of Noise on Cognitive Performance 2017–2021. In Proceedings of the 13th ICBEN Congress on Noise as a Public Health Problem, Stockholm, Sweden, 14–17 June 2021.
50. Schlittmeier, S.J.; Weißgerber, T.; Kerber, S.; Fastl, H.; Hellbrück, J. Algorithmic Modeling of the Irrelevant Sound Effect (ISE) by the Hearing Sensation Fluctuation Strength. *Atten. Percept. Psychophys.* **2012**, *74*, 194–203. [CrossRef]
51. Bell, R.; Röer, J.P.; Lang, A.-G.; Buchner, A. Reassessing the Token Set Size Effect on Serial Recall: Implications for Theories of Auditory Distraction. *J. Exp. Psychol. Learn. Mem. Cogn.* **2019**, *45*, 1432–1440. [CrossRef] [PubMed]
52. Cowan, N. *Attention and Memory: An Integrated Framework*; Oxford University Press: New York, NY, USA, 1995.
53. Hughes, R.W. Auditory Distraction: A Duplex-Mechanism Account. *Psych J.* **2014**, *3*, 30–41. [CrossRef] [PubMed]
54. Jones, D.; Madden, C.; Miles, C. Privileged Access by Irrelevant Speech to Short-Term Memory: The Role of Changing State. *Q. J. Exp. Psychol. Sect. A* **1992**, *44*, 645–669. [CrossRef] [PubMed]
55. Salamé, P.; Baddeley, A. Disruption of Short-Term Memory by Unattended Speech: Implications for the Structure of Working Memory. *J. Verbal Learn. Verbal Behav.* **1982**, *21*, 150–164. [CrossRef]
56. Neath, I.; Nairne, J.S.; Stevenson, A.K.; Surprenant, A.M.; Baddeley, A.; Jones, D.M. *Modeling the Effects of Irrelevant Speech on Memory*; Psychonomic Society: Chicago, IL, USA, 2000; Volume 7.
57. Wetzel, N.; Scharf, F.; Widmann, A. Can't Ignore—Distraction by Task-Irrelevant Sounds in Early and Middle Childhood. *Child Dev.* **2019**, *90*, e819–e830. [CrossRef] [PubMed]
58. Joseph, T.N.; Hughes, R.W.; Sörqvist, P.; Marsh, J.E. Differences in Auditory Distraction between Adults and Children: A Duplex-Mechanism Approach. *J. Cogn.* **2018**, *1*, 13. [CrossRef]

59. Meinhardt-Injac, B.; Imhof, M.; Wetzel, N.; Klatter, M.; Schlittmeier, S.J. The Irrelevant Sound Effect on Serial Recall Is Independent of Age and Inhibitory Control. *Audit. Percept. Cogn.* **2022**, *5*, 25–45. [[CrossRef](#)]
60. Röer, J.P.; Bell, R.; Körner, U.; Buchner, A. Equivalent Auditory Distraction in Children and Adults. *J. Exp. Child Psychol.* **2018**, *172*, 41–58. [[CrossRef](#)]
61. Schwarz, H.; Schlittmeier, S.; Otto, A.; Persike, M.; Klatter, M.; Imhof, M.; Meinhardt-Injac, B. Age Differences in the Irrelevant Sound Effect: A Serial Recognition Paradigm. *Psihologija* **2015**, *48*, 35–43. [[CrossRef](#)]
62. Elliott, E.M.; Briganti, A.M. Investigating the Role of Attentional Resources in the Irrelevant Speech Effect. *Acta Psychol.* **2012**, *140*, 64–74. [[CrossRef](#)]
63. Buchner, A.; Bell, R.; Rothermund, K.; Wentura, D. Sound Source Location Modulates the Irrelevant-Sound Effect. *Mem. Cognit.* **2008**, *36*, 617–628. [[CrossRef](#)] [[PubMed](#)]
64. Jones, D.M.; Macken, W.J. Auditory Babble and Cognitive Efficiency: Role of Number of Voices and Their Location. *J. Exp. Psychol. Appl.* **1995**, *1*, 216. [[CrossRef](#)]
65. Jones, D.M.; Macken, W.J. Organizational Factors in the Effect of Irrelevant Speech: The Role of Spatial Location and Timing. *Mem. Cogn.* **1995**, *23*, 192–200. [[CrossRef](#)]
66. Jones, D.M.; Saint-Aubin, J.; Tremblay, S. Modulation of the Irrelevant Sound Effect by Organizational Factors: Further Evidence from Streaming by Location. *Q. J. Exp. Psychol.* **1999**, *52*, 545–554. [[CrossRef](#)]
67. Steinbrink, C.; Klatter, M. Phonological Working Memory in German Children with Poor Reading and Spelling Abilities. *Dyslexia* **2008**, *14*, 271–290. [[CrossRef](#)]
68. Schiano, D.J.; Watkins, M.J. Speech-like Coding of Pictures in Short-Term Memory. *Mem. Cognit.* **1981**, *9*, 110–114. [[CrossRef](#)] [[PubMed](#)]
69. Wetzel, N.; Widmann, A.; Scharf, F. Distraction of Attention by Novel Sounds in Children Declines Fast. *Sci. Rep.* **2021**, *11*, 5308. [[CrossRef](#)]
70. McCrea, S.M.; Mueller, J.H.; Parrila, R.K. Quantitative Analyses of Schooling Effects on Executive Function in Young Children. *Child Neuropsychol.* **1999**, *5*, 242–250. [[CrossRef](#)]
71. Alloway, T.P.; Gathercole, S.E.; Pickering, S.J. Verbal and Visuospatial Short-Term and Working Memory in Children: Are They Separable? *Child Dev.* **2006**, *77*, 1698–1716. [[CrossRef](#)]
72. Morey, C.C.; Miron, M.D. Spatial Sequences, but Not Verbal Sequences, Are Vulnerable to General Interference during Retention in Working Memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **2016**, *42*, 1907–1918. [[CrossRef](#)]
73. Guerreiro, M.J.S.; Murphy, D.R.; van Gerven, P.W.M. The Role of Sensory Modality in Age-Related Distraction: A Critical Review and a Renewed View. *Psychol. Bull.* **2010**, *136*, 975–1022. [[CrossRef](#)] [[PubMed](#)]
74. Guerreiro, M.J.S.; Anguera, J.A.; Mishra, J.; van Gerven, P.W.M.; Gazzaley, A. Age-Equivalent Top-down Modulation during Cross-Modal Selective Attention. *J. Cogn. Neurosci.* **2014**, *26*, 2827–2839. [[CrossRef](#)] [[PubMed](#)]
75. Rienäcker, F.; van Gerven, P.W.M.; Jacobs, H.I.L.; Eck, J.; van Heugten, C.M.; Guerreiro, M.J.S. The Neural Correlates of Visual and Auditory Cross-Modal Selective Attention in Aging. *Front. Aging Neurosci.* **2020**, *12*, 498978. [[CrossRef](#)] [[PubMed](#)]