Speech perception in noise, working memory, and attention in children: a scoping review

Lyan Porto (corresponding author) [1], Jan Wouters[1], Astrid van Wieringen[1,2]

[1] University of Leuven, Department of Neurosciences, Research group Experimental Oto-Rino-Laryngologie. O&N II, Herestraat 49, 3000 Leuven, Belgium.

[2] Department of Special Needs Education, University of Oslo, Norway.

Author email: lyan.porto@kuleuven.be

Abstract

Purpose: Speech perception in noise is an everyday occurrence for adults and children alike. The factors that influence how well individuals cope with noise during spoken communication are not well understood, particularly in the case of children. This article aims to review the available evidence on how working memory and attention play a role in children's speech perception in noise, how characteristics of measures affect results, and how this relationship differs in non-typical populations.

Method: This article is a scoping review of the literature available on PubMed. Forty articles were included for meeting the inclusion criteria of including children as participants, some measure of speech perception in noise, some measure of attention and/or working memory, and some attempt to establish relationships between the measures. Findings were charted and presented keeping in mind how they relate to the research questions.

Results: The majority of studies report that attention and especially working memory are involved in speech perception in noise by children. We provide an overview of the impact of certain task characteristics on findings across the literature, as well as how these affect non-typical populations.

Conclusion: While most of the work reviewed here provides evidence suggesting that working memory and attention are important abilities employed by children in overcoming the difficulties imposed by noise during spoken communication, methodological variability still prevents a clearer picture from emerging.

Keywords: speech in noise, speech perception, working memory, attention, children, scoping review

Introduction

Understanding speech in the presence of background noise or other speakers is a common occurrence in daily life. Commonly dubbed the “cocktail-party effect” (Cherry, 1953; Meister et al., 2013), the difficulties posed by having to focus on one speaker while filtering out other sound sources is ubiquitous, and the ability to overcome it is paramount to effective everyday communication. Similarly, children occupy – and acquire spoken language in – noisy or otherwise acoustically suboptimal environments such as classrooms and playgrounds. Thus, the ability to segregate sound streams is exercised from a young age.

A substantial amount of research has been done on the ability of children with typical hearing to understand speech in noise, but there is surprisingly little consensus on what factors influence speech perception measures. Many authors have pointed specifically to working memory (WM) and attention as key abilities involved in this process (Dillon & Cameron, 2021; Lewis et al., 2014; Magimairaj & Nagaraj, 2018; Söderlund & Jobs, 2016; Thompson et al., 2019), but evidence supporting these claims is mixed.

Understanding the way cognition is involved in this everyday ability as performed by children is fundamental to understanding how it develops in typical populations. Additionally, it allows us an insight into the difficulties faced by children in schools, especially children who are deaf or hard of hearing. It is a well-documented fact that noisy environments pose a particular challenge to individuals with hearing loss, including children, even when assistive hearing technology is employed (Busch et al., 2017; Crandell & Smaldino, 2000). Similarly, the way cognition and hearing interact in these situations is likely to be of interest to understanding the learning trajectories and difficulties facing children with attention and learning disorders, such as attention deficit/hyperactivity disorder and specific language impairment, among others. Due to the complex nature of the relationships between working memory, attention, and speech understanding in noise, studies shedding light on these interactions have come from a variety of different perspectives. Thus, a wide range of tasks, speech materials, maskers, and WM and attention measures are found throughout the literature, as well as different signal-to-noise ratios and modes of presentation, making direct comparisons and specific predictions challenging. Therefore, the need has arisen for a broad review of the published literature on the relationship between working memory, attention, and speech perception in noise abilities in children. In the following paragraphs, we present a brief overview of speech perception in noise, working memory, and attention, as well as their potential interactions. Afterward, we present our methodology for conducting this review.

Speech Perception in Noise (SPIN)

Within the context of this article, the term SPIN refers to any instance where an individual is required to perceive speech that is suffering environmental or transmission degradation (Mattys et al., 2012). Following Mattys et al. (2012)'s classification, environmental (or transmission) degradation occurs when the target signal is degraded either by the presence of concurrent signals or by changes to the target signal itself. One example of the former includes steady-state noise such as air conditioning or computer noise, as well as variable noise, such as traffic or construction noise, which cause energetic masking; another is non-target speech, which additionally might cause informational masking, where semantic content of the distractor interferes in the decoding of the target speech signal. On the other hand, the amplitude fluctuations present in non-steady noises and speech distractors can aid speech perception by allowing listeners to occasionally "glimpse" the target signal through the noise (Cooke, 2006). Nevertheless, and relevant to the purposes of the present investigation, degradation with energetic masking requires the listener to form separate representations of target and non-target signals and to selective attend to one over the other (Darwin, 2008; Shinn-Cunningham & Best, 2008), highlighting the importance of attention to overcome this type of degradation.

The other type of environmental degradation occurs without energetic masking by a separate stream, but rather when degradation is imposed on the target signal itself. An example is telephone and modern digital communication, which frequently filter, compress, and/or introduce temporal discontinuities to the signal for efficiency of transmission, often in ways that are detrimental to the listener's ability to perceive it. Researchers have used noise-vocoded speech and sine-wave speech as forms of degraded speech signals which, unlike degradation with energetic masking, do not require stream segregation or selective attention (Grieco-Calub et al., 2017; Mattys et al., 2012; Nittrouer et al., 2015). Signal degradation caused by hearing impairment and/or assistive hearing technologies such as hearing aids or cochlear implants can also be thought of in terms of this type of degradation. Though degradation with and without energetic masking differ in attentional demands required to overcome them, the demands the two types of degradation impose on working memory can be argued to be similar (Mattys et al., 2012; Rönnberg et al., 2013). In the context of this review, we will be discussing the role of WM in SPIN as it was proposed by the Ease of Language Understanding (ELU) model (Rönnberg et al., 2008, 2013), which posits that, in adverse conditions, working memory is key to resolving ambiguities and mismatches that arise as a result of signal degradation.

Working Memory (WM)

Different models of working memory have been proposed in the literature, but relevant to our purposes, working memory is conceptualized as a limited capacity system which allows individuals to store and manipulate information for a short period of time. The manipulation component is key to differentiating WM from short-term memory (Baddeley & Hitch, 1974; Baddeley, 2012). Most models incorporate attention as one of its central components (Baddeley & Hitch, 1974; Cowan, 1999; Engle, 2002), underscoring how these two constructs are interconnected. According to the ELU model mentioned above, working memory would be involved in speech understanding in noise or otherwise adverse conditions when rapid, automatic matches to lexical or phonological representations stored in long-term memory fail. In these circumstances, individuals would need to store mismatched phonological and semantic material and integrate it further input until such mismatches are resolved, a process which taxes working memory capacity. Thus, individuals with higher working memory capacity are predicted to have higher performance in this process, ultimately attaining better understanding of speech in adverse conditions. This model has strong empirical support from studies with adults, but little work has been done with children (Holmer et al., 2016; McCreery et al., 2019).

Attention

In the context of this review, attention is defined as the ability to select one perceptual item among others for further processing. Many of the theoretical models proposed agree that attention is as an integral part of working memory, as attention allows the individual to select which stimuli are allowed entry into working memory, to prevent it from becoming overloaded (Broadbent, 1958; Klemen et al., 2009; Lavie & Tsal, 1994; Pashler, 1998; Sörqvist et al., 2012). The ELU model above also predicts that WM capacity modulates early attentional mechanisms, with higher WM capacity leading to faster access to long-term memory representations.

Selectively attending to one source while inhibiting others requires that the individual form a "perceptual object," distinguishing it from others in the same scene based on some characteristic (Shinn-Cunningham & Best, 2008). In the context of auditory attention, features such as frequency, amplitude, timbre, and direction of the source can aid in object formation. For speech stimuli, factors such as phonology, vocabulary, and grammar can further improve this process, thus involving language abilities in the process, which might also be impaired or still in development in the case of children with and without different impairments. The presence of noise thus interferes with the ability to perceive the characteristics that might be helpful in segregating one object from the others. Additionally, when target and distractors share similar characteristics, such as being produced by speakers of similar voices, finding unique features to form and distinguish each object might be more difficult. As before, this problem is exacerbated by the inherent degradation of signals as perceived by individuals with hearing impairment.

In sum, working memory and attention – as well as connections between the two – have long and reasonably been argued to play rather central roles in the process of speech understanding in noise (Caplan & Waters, 1999; Klemen et al., 2009), alongside audiological factors such as noise and hearing, and linguistic knowledge (Nittrouer & Boothroyd, 1990). How these effects develop and manifest in children has been the subject of much research, an overview of which we aim to present below, in an effort to summarize the main findings and trends of this body of work, as well as identify limitations and future directions.

Methods

The goal of the present article is to review the existing literature on the relationship between working memory, attention, and performance in speech in noise tasks by children. Specifically, we approached this literature with the following questions in mind: (1) How do working memory and/or attention abilities affect speech-in-noise understanding by children? (2) How are the findings in the literature related to the types of measures and characteristics of each study? (3) How do these relationships differ between typically and non-typically developing populations?

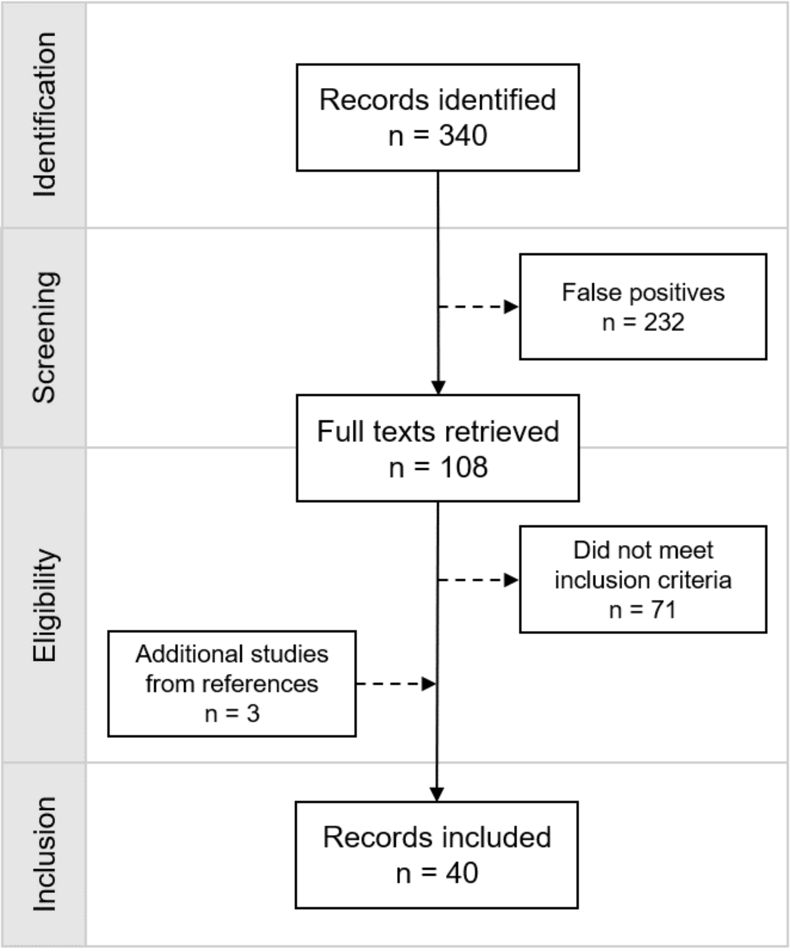
To provide an overview of this heterogenous body of work, a scoping review was chosen as a method for this investigation, as it allows us to summarize the main findings of the research on the topic at hand without the need for standardization of methods of the studies reviewed (Tricco et al., 2018). To that end, we followed the protocol outlined in the PRISMA extension for Scoping Reviews (Tricco et al., 2018), whose checklist guides the reporting below.

First, a search was conducted on the PubMed, APA PsycNet, ERIC and JSTOR databases using the string “(child*) AND speech AND (understanding OR comprehension OR recognition OR perception) AND noise AND (attention OR working memory OR executive function OR cognitive control)” and results were limited to studies published before February 2022. The search strategy was a result of discussion between all authors.

The inclusion criteria used to determine whether a study would be included in the review were as follows: (1) the study had to be performed on participants that included children (i.e., under 18 years of age); (2) the study had to report on at least one measure of speech perception in noise (as defined in the previous section); (3) the study had to report on at least one measure of either working memory or attention; (4) the study had to report some attempt to establish a relationship between WM and/or attention and SPIN measures, e.g., correlations between measures or differences between single- and dual-task conditions. Of the 340 results obtained from the search described above, 232 were excluded after title and abstract screening for failing to meet these criteria (8 being duplicates, 15 being conference proceedings, theoretical papers, or reviews; 32 not being about hearing; 24 failing criterion 1, that is, not including children; 13 not including SPIN measures; 132 not including cognitive measures of interest; and 8 not attempting to relate cognition and SPIN). Full-text analyses of the remaining 108 articles excluded another 71 articles (18 for not including a relevant SPIN measure; 38 for not including a relevant cognitive measure; and 15 for not attempting to relate cognition and SPIN). An additional three articles were found within references of the original group of studies, leading to a total of 40 studies being included in this review. This screening process was conducted by the first author only. This process is illustrated in Figure 1 below.

Then, the selected studies were categorized according to variables pertaining to each research question. For question number 1, the category was which cognitive construct was being investigated (working memory, attention, or both); for question number 2, the measures used for working memory and/or attention, as well as important details in the SPIN procedure, such as type of procedure, materials, type of noise, signal-to-noise ratios (SNRs), and modes of presentation; and for question number 3, the classification of its population (i.e., typically developing, hearing loss, auditory processing disorder, etc.) This summary, along with extra information, is presented in Table 1. In the final selection, studies between 2000 and 2021 were included.

Figure 1. Search strategy.



Of the 40 papers included in this review, 12 reported on working memory alone, 14 reported on attention alone, and 14 studies reported on both, highlighting the interrelationship between the two. Nineteen reported only studies conducted on typically developing populations, seven reported only on non-typical populations, and 13 on a mix of the two. Below are summaries of the main findings of this literature, grouped into sections guided by our research questions.

Results

This section is organized as follows: first, we will present findings from studies that investigated the relationship between WM and SPIN in children, and then those that investigated attention and SPIN. Within each section, we examine the patterns of results from common measures for the respective cognitive ability on both typically developing and then on non-typically developing participants, followed by the same examination for different characteristics of SPIN tasks. A summary of standardized measures cited is presented in Table 2, while non-standardized tasks are described when necessary.

Working Memory

WM Measures

Backwards Digit Span

Though a number of different tasks were used to assess WM capacity in children across the 26 studies reviewed in this section, the backward digit span (BDS), appeared as the most commonly used WM measure, being present in 16 (61.5%) studies. Despite its many implementations, the basic form of this task requires participants to repeat verbally or enter on a keypad the numbers presented in reverse order of presentation. The length of the list of numbers to be repeated increases until the participant fails to correctly repeat the numbers in reverse order, with the last successful length being taken as the span. As mentioned previously, this requires participants to hold the numbers in memory and manipulate them to present them in reverse order, thus being a measure of working memory. In many cases, it is combined into a single measure with forward digit span (FDS) or other WM measures, while in others, separate correlations are reported. One study, by MacCutcheon et al. (2019), used only the BDS (from the CELF-4; Semel et al., 2003) to assess WM in 39 typically hearing children aged 4-11 and found that it correlated with performance on an adapted version of the Children's Coordinate Response Measure (CCRM; Vickers et al., 2016). Participants with higher WM scores were found to have lower (better) speech reception thresholds (SRTs) than those with lower WM. Additionally, those with lower WM benefitted less from spatial separation between target and masker. Similarly, Sullivan et al. (2015) used the backwards digit recall subtest from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001), in addition to the listening recall subtest to assess WM. SPIN performance was measured using a speech comprehension task where participants were asked questions about the content of the sentences heard (adapted from the Listening Comprehension Test 2; Bowers et al., 2006). Sullivan and colleagues found a strong positive correlation between an averaged total WM score, taking both measures into account, and SPIN when speech was presented in noise.

Hsu et al. (2020) used both forward and backward digit span measures and found that the backwards, but not forward scores (from the CELF-4 Dutch version; Kort et al., 2010) correlated with response times on a noun categorization task in speech-weighted noise, as performed by children and teenagers aged 6-18 with typical hearing, consistent with the distinction between STM and WM cited previously, as well as with the prediction by the ELU model that WM is involved in SPIN.

Some studies using this measure present less clear patterns. Mealings et al. (2020) found that a similar task, Number Memory Reversed, and not Number Memory Forward from the Test of Auditory Processing Skills – Third Edition (TAPS-3; Martin & Brownell, 2005), correlated with performance on some conditions of the Listening in Spatialized Noise – Sentences (LiSN-S; Cameron & Dillon, 2008) SPIN test. However, these correlations disappeared after correction for multiple correlations. Nevertheless, the authors argue that it is unlikely that the initial correlation found was due to chance, and so maintain that such a relationship exists between SPIN and WM scores. In the same vein, Lewis et al. (2014) used the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1991) implementation of the forward and backward digit span tasks, and found that an aggregate score did not correlate with performance on a complex SPIN test wherein children had to follow instructions presented by video in the presence of noise or noise and reverberation. The authors argue, however, that it is unlikely that WM was not required to complete the task, due to its complexity, and suggested that the measure of WM used may have not been sensitive to differences in the sample studied.

Other WM measures

A number of other WM measures were common throughout the literature. The Odd One Out (OOO) subtest from the Automated Working Memory Assessment (AWMA; Alloway, 2007), considered an index of visual working memory capacity (McCreery et al., 2017; Nadler & Archibald, 2014), also appeared in a few studies included in this review. In a series of studies, McCreery and colleagues (McCreery et al. 2017; 2019; 2020) found that the OOO correlated positively with scores on different SPIN tests.

Other common measures present a more mixed pattern of findings. One such test, Listening Recall, appears in four studies in slightly different implementations, three of which find correlations with SPIN measures (McCreery et al., 2020; Sullivan et al., 2015; von Lochow et al, 2018) while one does not (Walker et al., 2019). Another one is the Auditory Working Memory subtest from the Woodcock-Johnson III Test of Cognitive Abilities (WJIII; Woodcock et al., 2007), which appears in five studies, three of which (Nagaraj et al., 2020; Strait et al., 2012; Thompson et al., 2019) find a positive correlation with SPIN scores. Strait et al. (2012) also employed a visual working memory measure from the Colorado Assessment Test (Davis & Keller, 2002). In this task, participants monitored a number of squares presented visually on screen while they changed color and were then requested to click the squares in the same or reverse order in which they changed color. Both forwards and backwards spans were measured, but only the backwards span was taken as a measure of visual working memory due to the manipulation requirement. This measure also failed to reveal any correlation with either of the two SPIN measures used in this study.

WM Measures and Non-typically developing groups

Of the six WM studies with children diagnosed with or suspected of having auditory processing disorder (APD), four used some form of the BDS as a measure, with a consistent pattern of findings. For example, Tomlin et al. (2015) used an aggregate score of both FDS and BDS (from the CELF-4), and found a moderate correlation with performance on the low-cue (i.e., more difficult) conditions of the LiSN-S SPIN task. Interestingly, when reported separately, FDS revealed a slightly stronger correlation with these scores than BDS.

These results support earlier findings by Moore et al. (2010), who similarly used an aggregate score of both forward and backward digit span tasks (from the WISC) in a large-scale study of 1469 children from the general population aged 6-12. The authors used a large sample of children without screening for APD based on the high likelihood that such a large sample would include some children who might qualify for a diagnosis. Taken as a whole, this population revealed a weak but significant correlation between WM scores and performance on a series of auditory processing tests, including a task where participants were required to repeat a vowel-consonant-vowel (VCV) nonword in speech-weighted noise. Since the authors' goal was to find evidence for a sensory processing deficit account of APD, we include these results in this section. Similarly, Kumar et al. (2021) tested 20 children with APD (i.e., scoring 2 standard deviations (SD) below the age mean in two tests or 3 SD in any one test in a battery of APD tests) aged 9-10 on a SPIN test as well as on forward, backward, ascending, and descending digit spans. Though uncommon, the latter two measures are also likely to reflect WM capacity, since they, like the BDS, require processing of information stored in memory. Half the participants then underwent SPIN training for 2-3 weeks. This training consisted of performing a closed-set one- or three-syllable word identification task presented in speech-shaped or 4-talker babble noise in varying SNRs ranging from 20 to -4 dB SNR. Training sessions lasted 30 minutes and were given 3-4 times a week. Though WM and SPIN did not correlate before training, post-training tests revealed significant SPIN improvements only for the group that underwent training, as well as significant improvements in all three working memory tasks (backward, ascending, and descending digit spans, but not forward digit spans). A correlation between WM and SPIN after training was not reported for either group.

Studies with children who are deaf or hard of hearing have also used the BDS as an index of WM capacity, but findings have not been as consistent as in studies with children with APD. Of the four such studies included in the review, two found a positive correlation, while two did not. For example, Javanbakht et al. (2021) tested children who wore bilateral hearing aids using the FDS, BDS as well as a nonword repetition task as measures of working memory, and found that all three individually correlated positively with scores on a Persian version of the BKB-SIN test (Bench et al., 1979; Etymotic Research, 2005; Moossavi et al., 2017). Likewise, Percy-Smith et al. (2020) also investigated children who wore hearing aids and children with cochlear implants (Cis) aged 4 to 10 years old. Both groups of children showed a significant and moderately strong correlation between scores on a novel SPIN test and a WM aggregate score based on forward digit span, backward digit span, and one Familiar Sequences test of the CELF-4.

On the other hand, Walker et al. (2019) used the BDS combined with two tasks from the AWMA (Odd One Out and Listening Recall) in a longitudinal study of 199 children with mild to severe hearing loss and 99 children with typical hearing. They found no correlation between an aggregate WM measure and the SPIN task used, the BKB-SIN for the children who were hard of hearing. Correlations for the children with typical hearing were not reported. Tao et al. (2014) also found no correlation between either forward or backward digit spans and performance by CI users on a SPIN test.

Two other studies included in this review deal with children suspected of belonging to a clinical group. In a study on children with ADHD and typically developing controls aged 9 and 10 years old, Söderlund & Jobs (2016) used both FDS and BDS as measures of WM, but did not report direct correlations between these scores and scores on their SPIN task. The authors do note that the ADHD group differed significantly from the control group in WM scores but not SPIN ones. Lastly, Mealings and Cameron (2019) reported that children with reading difficulties, as reported by their teachers but having no formal diagnosis, showed no significant correlations between the TAPS implementation of the BDS (Numbers Memory Reversed) and their SPIN measure – the high-cue condition of the LiSN-S.

All in all, the BDS, even when combined with other measures, emerged as a common measure for examining WM capacity as it relates to SPIN performance in both children with typical development as well as those belonging to various non-typical groups. However, it must be noted that studies differ in a number of other ways which might explain differences in results, some of which will be discussed below.

SPIN measures

In contrast to WM measures, a greater variety of SPIN measures are found in the literature, with many authors designing their own SPIN tasks instead of using standardized solutions. Thus, we are grouping studies by types of task and materials presented as speech and distractors in SPIN measures, as we found these factors greatly impact the WM demands these tasks place on children.

The most common task is asking the participant to repeat a speech stimulus presented, appearing in 20 studies. In most cases, the stimuli used in these tasks are meaningful sentences, but in some cases, isolated words or numbers are used, as well as meaningless strings of words and even nonwords, to mostly positive results. For instance, McCreery et al. (2017) used three types of stimuli in a SPIN task: monosyllabic words, syntactically correct sentences with no semantic meaning, and sequences of words with no syntax or semantics. They found that children's (aged 5-12) performance on all three types of stimuli correlated positively with WM measures. Conversely, Magimairaj et al. (2018) did not find such a correlation when SPIN performance was assessed using the meaningful BKB sentences. The authors hypothesized that the lack of semantics or syntax in McCreery et al. (2017)'s stimuli caused the engagement of working memory, whereas the relatively more predictable BKB sentences did not. In a subsequent study, McCreery et al. (2020) used the same BKB materials (with a different procedure) with 30 children aged 5-6 and 30 aged 9-10, and again found a correlation with WM scores, but only in the more difficult noise conditions, leading the authors to argue that the relationship between SPIN and WM only becomes apparent in more challenging circumstances. Consistent with this hypothesis, Thompson et al. (2019) found that one of their WM measures (auditory working memory from the WJIII) correlated with SPIN performance on the Hearing in Noise Test (HINT, Bio-logic Systems Corp.; Nilsson et al., 1994; Soli & Wong, 2008) only in the condition in which noise was collocated with the target speech stimulus, a presumably more difficult condition. Similarly, only the HINT and not the Words in Noise (WIN) test were correlated with auditory working memory in Strait et al.'s (2012) study, consistent with the authors' prediction that the WIN reflected a more purely perceptual process and thus relied less on WM than the sentences of HINT.

A few studies used speech comprehension measures to assess SPIN, such as following instructions presented verbally (Lewis et al., 2014; MacCutcheon et al., 2019), answering questions about a passage (Sullivan et al., 2015; von Lochow et al., 2018) or categorizing nouns (Hsu et al., 2020). One might argue that this ability requires more higher level processing than simple "perception." Consistent with the hypothesis that more challenging tasks require more engagement of working memory, all but one such studies found positive correlations with different WM measures.

Noise types

Another significant factor which emerged from the analysis of the literature is the type of noise used. As discussed in the Introduction, different types of noise or degradation cause different effects in speech understanding (Francart et al., 2011; Mattys et al., 2012). The studies cited here can be grouped broadly into two categories: those using speech-weighted noise or otherwise steady-state noise (13 studies, 50%), and those using irrelevant speech as noise (18, 69%). Naturally, some compare the two in different conditions and some still use reverb in combination with another form of noise.

Despite being arguably easier than speech distractors, the vast majority of studies using steady-state noises reported positive correlations between their SPIN and WM measures (but see Söderlund & Jobs, 2016; Tao et al., 2014). Results from studies using speech distractors show more mixed results. While the aforementioned study by Sullivan et al. (2015) found that WM capacity is related to SPIN performance using a "multiclassroom" noise, Lewis et al. (2014)'s study, which used a multi-talker babble noise consisting of recordings of 20 people speaking simultaneously, did not. Another previously discussed study, by Magimairaj et al. (2018), found no correlation between SPIN and WM using the BKB-SIN test which uses a 4-talker babble noise. On the other hand, Nagaraj et al. (2020) found a positive correlation using a dichotic digit test in noise as a SPIN measure, in which noise was a conversation between two adults and three children. It can be argued that such a distractor requires a lot more cognitive effort since is almost certainly caused informational as well as energetic masking, which is highly effective in children (Wightman et al., 2006). Combined with the fact that the target stimuli were isolated digits and therefore not particularly meaningful, this might explain the positive results found here.

McCreery et al. (2020) compared speech-weighted noise and a two-talker babble noise and found that only the babble noise yielded a significant correlation with WM, a finding partially supported by MacCutcheon et al. (2019), who found that children in general had worse (higher)

SRTs when the masker was a single talker than when it was speech-shaped noise. The effect of WM on SRTs in each noise type was not reported.

SPIN and WM in Non-typically developing groups

All studies with participants belonging or suspected of belonging to non-typically developing groups in our review used repetition tasks as a measure of SPIN. For example, Petley et al. (2021) typically developing children and children with listening difficulties using the LiSN-S as a SPIN measure, whose Low and High Cue scores the authors reported correlated strongly with their WM measure (the List Sorting Working Memory Test of the NIH Toolbox Cognition Battery; Weintraub et al., 2013). While children with listening difficulties performed more poorly overall in many, but not all, measures of SPIN and WM, separate correlations between SPIN and WM scores for each group were not reported. The LiSN-S, also used by Tomlin et al. (2015) to measure SPIN performance in children with APD, and by Mealings and Cameron (2019) in children with reading difficulties, uses a two-talker noise in a variety of conditions.

Two other publications used repetition of nonwords in speech-weighted noise as a measure of SPIN in studies into the factors related to APD (Ahmmed et al., 2014; Moore et al. 2010) and both found that this measure is positively correlated with same WM measure (an aggregate score of FDS and BDS from the WISC).

In addition to the aforementioned findings by McCreery et al. (2019) on children with hearing loss, a study by Klein et al. (2017) presented participants with hearing loss as well as controls with typical hearing with a series of single words in speech-weighted noise. Words varied in age of acquisition, frequency, and, in the case of nonwords, high or low phonotactic probability. WM measures correlated significantly with some of the more challenging SPIN measures (late acquired, low frequency words, nonwords with high phonotactic probability).

In general, it is clear that many factors affect findings presented in papers, and, presumably, children’s engagement of WM when perceiving (or comprehending) speech in noise. One consistent finding is that more challenging circumstances are more likely to recruit WM. Other factors could also be argued to play a role, such as SNRs or age of participants, but which for the sake of brevity could not be included in the summary above. None of the measures used were shown to be wholly ineffective at revealing the relationships we are examining, though some appear to be more sensitive than others, particularly when taking the non-typically developing groups into account. We now turn our attention to the literature investigating the impact of attention on SPIN in children.

[Table 2]

Attention

Attention measures

In total, twenty-eight papers were included which reported on the relationship between attention and SPIN performance in children. Most studies in this section (18 of 28; just over 64%), like those cited in the WM literature, use different standardized tests and attempt to find correlations between their scores and scores on SPIN tasks. Others, however, assess attention by comparing different conditions with presumably different attentional demands.

Correlation studies

Many different tests are used to look for correlations between attention measures and SPIN scores, with no clear pattern emerging. For example, McCreery et al. (2020) used both the Flanker test as well as the Dimensional Change Card Sort, two visual attention tasks from the NIH Toolbox and found that only the Flanker correlated with SPIN scores in three different noise types. In contrast, Grieco-Calub et al. (2017) found no correlation using the same implementation of Flanker test on their task and scores on a SPIN task which used the same materials (BKB sentences) as McCreery et al. (2020)'s study. One key difference is that Grieco-Calub and colleagues used noise-vocoded speech instead of noise as speech degradation. However, two other studies used noise-vocoded speech in combination with other attention tests found positive correlations (Huyck, 2018; Roman et al., 2017). One such study, Huyck (2018), used the Rapid Visual Information Processing test from the Cambridge Neuropsychological Test Automated Battery (CANTAB) eclipse-6 (CANTAB eclipse, 2013) to assess attention capabilities and found it emerged as a major factor explaining individual variability in SPIN performance.

Similarly, a study by Thompson et al. (2017) tested children aged 3-5 on a SPIN task where children were required to choose the picture corresponding to the target stimuli and assessed attention using the Leiter-R (Roid & Miller, 1997) subtest of sustained attention. Though direct correlations were not reported, the improvement with age in both SPIN and attention scores correlated with one another. In contrast, another study by Thompson and colleagues did not find a correlation between scores obtained by children aged 4-7 on a SPIN task which used the BKB sentences and attention, assessed using the same Leiter-R subtest (Thompson et al., 2019). Strait et al. (2012) also reported no correlation between performance by children aged 7-13 with or without musical training on either SPIN measure (HINT and WIN) and either visual or auditory attention as measured by the IVA-CPT, a task in which participants are asked to click a mouse when they see *or* hear one digit but not the other.

This finding is similar to that reported by Newton & Ridgway (2015) who found that the performance of children aged 6-7 on a SPIN task using the BKB sentences in familiar and novel accents did not correlate with attention scores measured by two subtests of the Test of Everyday Attention for Children (TEA-Ch, Manly et al., 2001), leading the authors to suggest that the relationship between attention and overcoming the difficulties imposed by noise and/or novel accents is not yet fully developed at this age.

Differences in performance across conditions

Studies that have assessed attention abilities by comparing performance on two conditions have more consistently found effects of attention on SPIN in children. Lewis et al. (2014), for example, compared performance on their single-talker condition with performance on their multi-talker condition, where additional speakers distracted participants from paying attention to the target speaker whose instructions they were required to follow. The authors found that the presence of additional distractors worsened performance on the main task. These findings support earlier findings by Howard et al. (2010), who found that adding a secondary task (holding series of digits in memory for later recall) worsened performance on a word repetition task relative to baseline, especially in the less favorable SNRs. Similarly, Kane et al. (2021) reported worse performance by children in a SPIN task in a condition when stimulus direction was uncertain compared to when a pretrial cue indicated from which of the 18 speakers the target stimulus would come from. In the absence of such a cue, participants had to monitor all locations to identify the target voice and then switch their attention accordingly.

One surprising finding is reported by Choi et al. (2008), who tested children aged 7-14 on a SPIN task in two conditions. In a single-task condition, participants performed this task with two lists of 25 words. In a dual-task condition, participants performed a digit recall task where they saw digits appear on screen for a few seconds and were asked to remember them to be recalled later. Then, they performed 5 trials of the SPIN task and subsequently recalled the 5 digits in order. Half the children were instructed to give priority to one task while the other half was told to prioritize the other. Interestingly, both groups showed improved performance on the SPIN task in the dual-task condition relative to single-task. Performance on the secondary task, however, decreased relative to baseline. This led the authors to argue that attentional control is still not developed in this age range, who always prioritized SPIN over the secondary task, irrespective of instructions. Also surprisingly, Nagaraj et al. (2020) found that children aged 7-12 with higher WM capacity were more likely to make intrusion errors in a dichotic digits test which required children to repeat digits present on one ear while ignoring those in the other.

Though overall performance in this measure improved as WM improved, the proportion of intrusion errors (reported as a measure of attention) also increased.

Attention Measures and Non-typically developing groups

Studies on non-typical populations have also attempted both ways of relating attention and SPIN, test score correlations and comparing conditions. As previously mentioned, studies investigating factors related to APD have grouped working memory and attention as a major contributor to the difficulties faced by children with APD (Ahmmed et al., 2014; Moore et al., 2010). More indirect evidence of this relationship is provided by Petley et al. (2021) who investigated children with APD (here termed listening difficulties; Dillon & Cameron, 2021) and typically developing children aged 6-13. The Flanker and DCCS tests from the NIH Toolbox were used and SPIN was measured with the LiSN-S sentences and select subtests of the SCAN-3:C battery (Keith, 1994a). The two groups of children were compared on these measures, and the APD group performed consistently worse on all of these measures.

However, the hallmark symptom of APD, difficulty processing sound or speech, frequently appears in children with other diagnoses as well, such as developmental dyslexia, specific language impairment, and attention deficit-hyperactivity disorder (ADHD), making differential diagnosis difficult (Dillon & Cameron, 2021; Ferguson et al., 2011; Ziegler et al., 2009, 2011). Thus, Magimairaj and colleagues included children with a range of diagnostic labels, but whose parents all reported some form of listening difficulty (Magimairaj et al., 2020). Twenty-six children aged 7-11 and age-matched controls on a dichotic digits test and an attention switching task, wherein children had to judge if a square presented on screen was big or small and press the corresponding button on screen. They were also required to keep track by counting aloud how many squares of each size they had seen and report the totals at the end. The groups differed in SPIN scores as measured by the BKB-SIN but not attention scores as measured by these two attention tasks. Similarly, both Riccio et al. (2005) and Tillery et al. (2000) investigated this relationship by testing groups of children with diagnoses of APD and with ADHD. Riccio et al. (2005) used the Auditory Figure Ground and Filtered Words subtests from the SCAN as a measure of SPIN, and the Test of Variables of Attention (TOVA; Greenberg & Crosby, 1992) as a measure of attention. The TOVA provides a number of measures that represent different aspect of attention: commission errors, omission errors, response time, and variability across the experiment. Thirty-six children aged 4-11, some of which had APD only, some had ADHD only, and some had both, completed the tasks. None of these measures correlated significantly with either SCAN subtest score. Tillery et al. (2000) tested children 32 diagnosed with both APD and ADHD. Scores on two SPIN measures and one measure of attention – the Auditory Continuous Performance Test (ACPT; Keith, 1994b) – were compared before and after treatment with methylphenidate (Ritalin), a commonly prescribed medication for ADHD, and a placebo. Half the participants received the medication first and the placebo second, with the other half receiving them in the opposite order. Both interventions increased performance on the attention task, but neither improved performance on either SPIN task, suggesting attention and SPIN performance are unrelated processes in this group.

More consistent results were found by studies using differences in performance across conditions as a measure of attention in children with hearing loss. McFadden and Pittman (2008) presented children aged 8-12 with typical hearing and with minimal hearing loss with single words and asked them to respond by saying to which of three categories – person, food, or animal – the word they had heard belonged. In the dual-task condition, a complex visual task was added, which consisted of a dot-to-dot game, where children had to connect dots numbered in increments of 3 on paper to reveal an image formed by the connecting lines. Eleven children with typical hearing performed the task and performance in the primary task (word categorization) remained unchanged by the addition of the secondary task. Data from 19 age-matched children with hearing loss, however, revealed a significant decrease in their word categorization score in the two noisy conditions of the dual-task paradigm. These findings were later corroborated by Pittman (2011), who presented children with hearing loss wearing hearing aids and age-matched controls aged 8-12 with the same tasks and conditions. Again, the addition of the secondary visual task made performance on the word categorization task significantly worse relative to baseline for children with hearing loss, even when fitted with hearing aids that increased SNRs by means of digital noise reduction.

This asymmetry between children with typical hearing and children who are deaf or hard of hearing in the way they are negatively impacted by added attentional demands on SPIN performance is not found in a study by Hicks and Tharpe (2002). Performance on a dual-task condition, when children had to activate a button when a light was turned on while performing a word repetition SPIN task was worse than performance in the single-task condition, for both groups of children equally.

Finally, one study included tested children fitted with cochlear implants (CIs). Misurelli et al. (2020) tested a group of 10 bilateral CI users aged 10 to 17. The SPIN measure presented sentences produced by a female talker through CI stimulation. Distractor sentences were produced by a male talker played on the opposite ear (contralateral), on the same ear (ipsilateral) or in both ears simultaneously (bilateral). Performance was lower in the ipsilateral and bilateral conditions than in the contralateral condition, suggesting that selectively attending to a target which is collocated with a distractor negatively affects performance.

SPIN measures

Noise types

As with the WM studies reviewed, different types of noise were used in the SPIN tasks in the attention literature as well. In addition to steady-state noises such as speech-weighted noises and white noise, and speech maskers such as multi-talker babble noises of various configurations, three studies in this section also used noise-vocoded speech. Grieco-Calub et al. (2017) tested 27 children aged 8-12 using the BKB sentences in their original form as well as in four-, six-, and eight-band noise-vocoded versions. Attention was loaded in a dual-task condition, where participants watched a series of grayscale images appearing on screen one at a time, and participants were asked to press a button when the same image appeared twice in a row. The authors also had participants perform a Flanker test and correlated those scores to performance in the other tasks. Though the authors argue that this task is a measure of executive function, it is acknowledged that it specifically targets attention and inhibition abilities. The Flanker test scores did not explain any of the variability in the dual-task condition, and performance on the SPIN task was seemingly unaffected by the introduction of the secondary task in any of the sound conditions (unprocessed, 4-, 6-, 8-band noise-vocoding). Performance on the secondary task, however, worsened in the dual-task relative to baseline, suggesting that participants had reallocated resources to the SPIN task to the detriment of the secondary task. On the other hand, Roman et al. (2017) found some correlation between attention measures and a SPIN test in which participants were required to repeat noise-vocoded words in isolation and sentences. Attention was assessed with two subtests from the Developmental Neuropsychological Assessment – Second Edition (NEPSY-II; Korkman et al., 2001), Auditory Attention, and Response Set, as well as a Talker Discrimination task using noise-vocoded speech to assess selective auditory attention, originally developed by Cleary and Pisoni (2002). In this task, participants heard a pair of sentences and had to judge whether the talker in each sentence was the same or different. Sentences produced by the two speakers were either the same sentence or different sentences, which forced the participant to ignore the meaning of the sentence and judge solely the quality of the voice. The authors report that only the Auditory Attention task from the NEPSY-II (and not the Response Set task) correlated with performance only on one condition of their noise-vocoded SPIN task (repeating words in isolation, not sentences). Additionally, both fixed-sentence and varied-sentence measures in the Talker

Discrimination task correlated with all both SPIN measures (words in isolation and words in sentences), highlighting how different measures reveal different relationships.

Two studies included both speech-weighted noise and some form of speech distractors. McCreery et al. (2020) compared speech-weighted noise, amplitude-modulated speech-weighted noise, and a two-talker babble noise and a positive correlation between scores on all three with one of the attention measures used. Kane et al. (2021), on the other hand, found that a pretrial cue indicating where the target stimulus would come from only benefitted participants aged 5-13 when the masker was a three-talker babble, and not when it was speech-weighted noise or a single talker.

Speech distractors were used in studies such as the one by Howard et al. (2010), though the exact composition of the multi-talker babble used is not reported. The authors reported a small but significant decrease in performance on the listening task when a secondary task was added. Magimairaj et al. (2018), on the other hand, reported that the multi-talker babble noise was composed of 4 speakers, and also that attention, as measured by scores on a dichotic digits test, did not explain SPIN performance. Similarly, Nagaraj et al. (2020) who, as reported earlier, used a recording of a conversation between two adults and three children as a distractor, found no relationship between their SPIN scores and their measure of attention (proportion of intrusion errors in a dichotic digits test).

SPIN and Attention in Non-typically developing groups

The literature on children who are deaf or hard of hearing has revealed a clear effect of attention on SPIN performance using both steady-state noises as well as speech distractors. Using white noise and speech-weighted noise respectively, McFadden and Pittman (2008) and Pittman (2011) found a negative effect of adding a secondary task on their primary task of noun categorization.

On the other hand, McCreery et al. (2019) found that attention scores correlated significantly only for children with hearing loss and only when reverb was also used. Hicks and Tharpe (2002) also found children with hearing loss were affected negatively by a second task when trying to repeat words presented in 20-talker babble noise.

In the case of children with or suspected of having APD, studies like Moore et al. (2010) and Ahmmed et al. (2014) used speech-weighted noise to claim a strong effect of attention, which was expanded by Petley et al. (2021)'s use of a two-talker masker in the LiSN-S. However, other studies using speech distractors did not find such relationships. Magimairaj et al. (2020) found that children APD and controls differed in SPIN scores using a four-talker distractor but not in two attention scores, while Tomlin et al. (2015) found no correlation between a two-talker masker and their attention measure (IVA-CPT). These findings support Tillery et al. (2000) and Riccio et al. (2005) results on children with APD and/or ADHD who also found no such relationship using SWN.

The results from the attention literature shows a less clear pattern of results than that of WM. Though the two constructs have long been connected (Baddeley, 2003; Dillon & Cameron, 2021), different attempts to index attention has led to contradicting findings.

Discussion

Working Memory

One of the goals of this article was to review the current knowledge on the relationship between WM and speech perception in noise. Models of speech understanding posit that working memory is crucial to resolving difficulties imposed by noise or degradation, a finding largely supported by empirical evidence from adults (Dillon & Cameron, 2021; Rönnberg et al., 2008; 2013). Thus, we expected studies to find a relationship between WM and SPIN performance also in children. The absence of such a finding might indicate that the use of WM for speech perception in noise is not yet mature in children. However, the majority of papers including

WM measures did indeed find that higher working memory leads to better speech perception in noise in children. Only seven out of 26 studies found no relationship between the two. That said, studies with the largest sample sizes often but weak to moderate or no correlation between WM and SPIN (Moore et al., 2010; Tomlin et al., 2015; Walker et al., 2019), but task difficulty constitutes a major factor in outcomes. In fact, several studies reported such correlations were stronger or only present in the more challenging conditions (e.g., less favorable SNRs, more challenging noise, less predictable language), supporting predictions that WM serves to overcome difficulties imposed by degradation (Klein et al., 2017; McCreery et al., 2020; Strait et al., 2012). Difficulties posed by "easier" conditions might be overcome with minimal engagement of WM, such that correlation between SPIN and WM measures might not reach significance (Magimairaj et al., 2018). The ELU model in particular suggests that WM would be involved when the signal perceived is degraded to the point where it cannot quickly be matched to a stored representation in long-term memory, thus impeding comprehension. Though outside the scope of this review, it is then likely that language skills also play a role in "closing the gap" between the perceived speech and stored representations when degradation is minimal, which might explain why easier conditions often do not appear to involve WM recruitment.

Thus, to uncover relationships, it is important that SPIN measures chosen are challenging enough. One factor highlighted here is the type of noise used. We expected comprehensible speech distractors to be "harder," and studies directly comparing different noise types did find that speech distractors made up of fewer speakers were more challenging than speech-weighted noise (MacCutcheon et al., 2019; McCreery et al., 2020). However, there was no overall pattern, perhaps partly due to the fact that many studies do not report important factors about how their multi-talker noises are constructed, such as how many talkers were overlaid, the gender proportion of these talkers, and whether there were temporal gaps in the noise, among others

(Francart et al., 2011). That said, it is difficult to make assertive inferences about the impact of noise on outcomes reported, since many factors interact to ultimately determine the difficulty of a SPIN task.

Attention

The literature on attention was varied both in terms of methods used as well as in terms of outcomes. While still a minority, a much greater proportion of studies reviewed, 11 out of the 28, found no evidence of a relationship between attention and SPIN performance. Perhaps significantly, four of these 11 also included WM measures and found no evidence of WM involvement, suggesting that the SPIN measures used might not have been challenging enough to elicit signs of a relationship. Furthermore, six of these included non-typically developing participants, which might indicate differences in how these factors are weighed across populations.

As explored in the Results section, studies into attention and SPIN use more varied measures of attention including differences in performance between single- and dual-task conditions, different spatial configurations for target and masker, and proportion of intrusion errors or performance variability in the SPIN task. These more "direct" measures have largely been successful in finding an effect of attentional capacity on SPIN performance (but see Nagaraj et al., 2020), which perhaps can be argued to more ecologically represent attentional demands in everyday situations where SPIN skills are required. However, due to the interconnectedness of WM and attention, it may be argued that adding a secondary task to a primary SPIN task also represents a load on WM, making it hard to tease apart the effects of loads on each construct.

Nevertheless, standardized and commercially available attention measures have also been used in the literature, not always revealing the same effects. The NEPSY, for example, revealed significant relationships between attention and SPIN in McCreery et al. (2019) and Roman et al. (2017)'s reports, while it did not in Mealings and Cameron (2019)'s. Once again, the reason might lie in differences in SPIN tasks, or their relationship to attention tasks used. Though again outside the scope of our review, different types of attention have been suggested in the literature, which may be assessed by different tests and relevant for different SPIN tasks. For example, sustained attention scores might be related to participants' ability to engage in a longer SPIN task but not reveal anything about performance in short ones, which might explain negative results reported by studies using longer attention tasks such as the ACPT or the IVA-CPT. Likewise, it has been suggested that visual and auditory attention might differ in key ways (Murphy et al., 2017), and since some attention tests are based on visual stimuli, it is possible that these do not capture the processes used to complete SPIN tasks. An underutilized solution to that problem was present in a single study (Moore et al., 2010), which reported a measure of variability in auditory processing tasks as an "intrinsic" measure of attention, arguably revealing precise attention mechanisms in real-time without the need for dedicated tasks for measuring attention.

Working Memory × Attention

Though the goal of this review was to look at how working memory and attention individually affected speech perception in noise in children, both the theoretical literature as well as the studies reviewed here suggest the two are closely related. Fourteen of the 40 studies reviewed looked at measures of both WM and attention, five of which found evidence for both. Some studies reported that WM and attention scores correlated to one another (McCreery et al., 2020; Thompson et al., 2019; Tomlin et al., 2015), even when neither score correlated with SPIN (Magimairaj et al., 2018; Nagaraj et al., 2020; but see Ahmmed et al., 2014; Mealings & Cameron, 2019). Furthermore, as mentioned previously, some tasks ostensibly measuring attention inevitably cause added load on working memory as well, such as additional distractors, unpredictable directions, or remembering sets of instructions across trials, making it difficult, if not outright impossible, to separate working memory effects from attention effects. This lays bare the need for clear theoretical frameworks that incorporate the two so that experiments can be designed to test and validate their assumptions.

Non-typically developing groups

One of the questions this article aims to investigate is how the effects of WM and attention relate to SPIN also in non-typically developing populations. This question appeared relevant also for the authors whose work we reviewed, as half (20 of 40) of studies included participants belonging to non-typical groups, with children who are deaf or hard of hearing and children with (suspected) auditory processing disorder being the most common. Many studies include only children in non-typical populations, making direct comparisons to typical populations difficult, as they often differ in other meaningful ways, too, as described previously. That said, studies investigating children who are deaf or hard of hearing generally find positive correlations in their SPIN scores to WM and attention scores (Hicks & Tharpe, 2002; Javanbakht et al., 2021; Kumar et al., 2021; McCreery et al., 2019; McFadden & Pittman, 2008; but see Tao et al., 2014; Walker et al., 2019). Some, however, do include both typically as well as non-typically developing children to allow direct comparisons, but unfortunately many do not report individual correlations for each group. Still, the general trend is that, indeed, children who are deaf or hard of hearing are affected differently, especially by attentional demands. McCreery et al. (2019) reported that a correlation between SPIN and attention was only found for their hearing loss group, while McFadden and Pittman (2008) and Pittman (2011) both report that children with hearing loss were more affected by the introduction of a secondary task. In contrast, Hicks and Tharpe (2002) found that the addition of the second task affected both groups equally.

In the case of children with APD, the matter of how to differentiate children with this diagnosis from others remains cause for debate. Thus, Moore et al. (2010)'s claim that attention was largely the driver behind clinical presentations of APD has met resistance from studies failing to find such a connection when using different measures or more general definitions of listening difficulties (Magimairaj et al., 2020; Tomlin et al., 2015). Further research is needed to resolve these conflicting results. Finally, Tillery et al. (2000)'s finding that methylphenidate did not improve children's SPIN performance is contrasted by previous findings that stimulant medication for ADHD, including methylphenidate, causes a sharp improvement in performance on auditory processing tasks by children with ADHD under the effect of methylphenidate, suggesting SPIN involves additional processes not captured by auditory processing tasks alone (Sutcliffe et al., 2006).

Clinical implications

The evidence reviewed in this article reveals general trends as well as directions for further research which both have important clinical and research implications. First, it suggests that WM recruitment is linked to task difficulty, and present task objectives and noise type as major contributors to task difficult. The effect of attention, on the other hand, appears to be best measured by tasks that directly load the attentional system, or by careful examination of which requirements are likely to be placed on it and selection of standardized tests that involve similar requirements. At the same time that we make these distinctions, we also report on evidence that suggests attention and working memory are intricately linked, and the need it presents for models and measurements that take this fact into account.

With regards to non-typical groups, the evidence reviewed provides some support for models that would predict strong recruitment of WM in children who are deaf or hard of hearing. Conversely, it exposes a still unresolved question in the APD literature with regards to whether attention is a major contributor to clinical presentations of APD. If anything, it makes clear that different populations are likely to have different requirements for WM and attention in their SPIN performance, which should be taken into account when designing studies with these populations.

Limitations

This review naturally is limited in a few key ways. Since this is a scoping review and our body of literature is very heterogenous, it was not possible to normalize studies across all variables and find the most relevant ones. Thus, we relied on our own observation of the literature, experience and interests to discuss certain aspects of tasks and participant groups, which inevitably meant others would not receive as much attention. For example, other aspects of SPIN tasks are certainly relevant for determining difficulty of a task, such as signal-to-noise ratios, modes of presentation or procedures for calculating speech perception measures. Likewise, age is a crucial factor in determining performance in all tasks discussed here, especially considering how the age ranges in this review (mostly 7-12) are important periods of neural maturation in typically developing children (Cowan, 2022). Finally, in the context of the understanding of meaningful speech in noise, language abilities are without a doubt required, and are also undergoing intense development in school-age children. Overviews of literature that take these factors into account are important for the consolidation of this body of work.

Conclusion

Though the evidence is far from unanimous, this review has found that most of the research done on the relationships between working memory, attention, and speech-perception in noise indicates that these three abilities are intricately linked. These findings were reported for a wide range of ages and clinical groups, highlighting the interconnectedness of these three skills. A particularly important contribution of this review is the recommendation that task difficulty and sensitivity of measures are accounted for when designing experiments. Nevertheless, the large variability in methods and results reported here remains a substantial factor preventing solid conclusions from being drawn.

Acknowledgements

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklowdoska-Curie Grant Agreement No. 860755 (Comm4CHILD project).

Declarations of interest: none.

References

Ahmmed, A. U., Ahmmed, A. A., Bath, J. R., Ferguson, M. A., Plack, C. J., & Moore, D. R. (2014). Assessment of Children with suspected auditory processing disorder: A factor analysis study. *Ear and Hearing*, 295-305. doi:10.1097/01.aud.0000441034.02052.0a

Alloway, T. P., Gathercole, S., & Pickering, S. J. (2006). Verbal and Visuospatial Short-Term and Working Memory in Children: Are They Separable? *Child Development, 77*(6), 1698-716. doi:10.1111/j.1467-8624.2006.00968.x

Baddeley, A. (2003). Working Memory: Looking Back and Looking Forward. *Nature Reviews Neuroscience, 4*, 829-839. doi:10.1038/nrn1201

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual review of psychology, 63*, 1-29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (pp. 47-89). Academic press.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*(3), 108-112. doi:10.3109/03005367909078884

Bowers, L., Huisingh, R., & LoGiudice, C. (2006). *The Listening Comprehension Test 2.* East moline, IL: LinguiSystems.

Broadbent, D. E. (1958). *Perception and communication.* New York: Pergamon Press.

Busch, T., Vanpoucke, F., & van Wieringen, A. (2017). Auditory Environment Across the Life Span of Cochlear Implant Users: Insights From Data Logging. *Journal of Speech, Language, and Hearing Research, 60*, 1362-1377. doi:10.1044/2016_JSLHR-H-16-0162

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

CANTAB eclipse. (2013). Cambridge, UK: Cambridge Cognition.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126. doi:10.1017/s0140525x99001788

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*, 975-979. doi:https://doi.org/10.1121/1.1907229

Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language, and Hearing Research, 51*(4), 1042-1054. doi:10.1044/1092-4388(2008/076)

Cleary, M., & Pisoni, D. B. (2002). Talker discrimination by prelingually deaf children with cochlear implants: Preliminary results. *Annals of Otology, Rhinology, & Laryngology., 111*(5 Part 2), 113-118. doi:10.1177/00034894021110s523

Cooke, M. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America, 119*, 1562-1573. doi:10.1121/1.2166600

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake, & P. Shah, *Models of Working Memory* (pp. 62-101). Cambridge, UK: Cambridge University Press.

Cowan, N. (2022). Working memory development: A 50-year asssessment of research and underlying theories. *Cognition, 224*. doi:10.1016/j.cognition.2022.105075

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,* 362-370. doi:10.1044/0161-1461.3104.362

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society, 363,* 1011-1021. doi:10.1098/rstb.2007.2156

Davidson, L. S., Geers, A. E., Hale, S., Sommers, M. M., Brenner, C., & Spehar, B. (2019). Effects of Early Auditory Deprivation on Working Memory and Reasoning Abilities in Verbal and Visuo-Spatial Domains for Pediatric CI Recipients. *Ear and Hearing, 40*(3), 517-528. doi:10.1097/AUD.0000000000000629

Davis, H. P., & Keller, F. (2002). Colorado Assessment Tests (CATs), Version 1.2. Colorado Springs, Colorado.

Dillon, H., & Cameron, S. (2021). Separating the Causes of Listening Difficulties in Children. *Ear and Hearing*, *42*(5), 1097-1108. doi:10.1097/AUD.0000000000001069

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science, 11*(1), 11-19. doi:10.1111/1467-8721.00160

Etymotic Research. (2005). *BKB-SIN Speech-in-Noise Test.*

Ferguson, M. A., Hall, R. L., Riley, A., & Moore, D. R. (2011). Communication, Listening, Cognitive and Speech Perception Skills in Children with Auditory Processing

Disorder (APD) or Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research, 54*, 211-227. doi:10.1044/1092-4388(2010/09-0167)

Francart, T., van Wieringen, A., & Wouters, J. (2011). Comparison of fluctuating maskers for speech recognition tests. *International Journal of Audiology, 50*(1), 2-13. doi:10.3109/14992027.2010.505582

Greenberg, L. M., & Crosby, R. D. (1992). A summary of developmental normative data on the TOVA ages 4 to 80+. In C. R. Riccio, C. R. Reynolds, & P. Lowe, *Clinical applications of continuous performance tests: Measuring attention and impulsive responding in children and adults.* New York: John Wiley & Sons, Inc.

Grieco-Calub, T. M., Ward, K. M., & Brehm, L. (2017). Multitasking During Degraded Speech Recognition in School-Age Children. *Trends in Hearing, 21*, 1-14. doi:10.1177/2331216516686786

Hicks, C. B., & Tharpe, A. M. (2002). Listening Effort and Fatigue in School-Age Children With and Without Hearing Loss. *Journal of Speech, Language, and Hearing Research, 45*, 573-584. doi:10.1044/1092-4388(2002/046)

Holmer, E., Heimann, M., & Rudner, M. (2016). Imitation, Sign Language Skill and the Developmental Ease of Language Understanding (D-ELU) Model. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.00107

Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology, 49*, 928-932. doi:10.3109/14992027.2010.520036

Hsu, B. C., Vanpoucke, F., Langereis, M., Dierckx, A., & van Wieringen, A. (2020). Age-Related Changes in Listening Effort for Children and Teenagers With Normal Hearing and Cochlear Implants. *Ear and Hearing, 42*(3), 506-519. doi:10.1097/AUD.0000000000000953

Huyck, J. J. (2018). Comprehension of Degraded Speech Matures During Adolescence. *Journal of Speech, Language, and Hearing Research, 61*, 1012-1022. doi:10.1044/2018_JSLHR-H-17-0252

Javanbakht, M., Moosavi, M. B., & Vahedi, M. (2021). The importance of working memory capacity for improving speech in noise comprehension in children with hearing aid. *International Journal of Pediatric Otorhinolaryngology, 147*, 110774. doi:10.1016/j.ijporl.2021.110774

Kane, S. G., Dean, K. M., & Buss, E. (2021). Speech-in-Speech Recognition and Spatially Selective Attention in Children and Adults. *Journal of Speech, Language, and Hearing Research, 64*, 3617-3626. doi:10.1044/2021_JSLHR-21-00108

Keith, R. W. (1994a). *Screening Test of Auditory Processing Disorders (SCAN).* San Antonio, TX: The Psychological Corporation.

Keith, R. W. (1994b). *ACPT Auditory Continuous Performance Test (Examiner's manual).* San Antonio: Harcourt Brace and Co.

Klein, K. E., Walker, E. A., Kirby, B., & McCreery, R. W. (2017). Vocabulary Facilitates Speech Perception in Children With Hearing Aids. *Journal of Speech, Language, and Hearing Research, 60*, 2281-2296. doi:10.1044/2017_JSLHR-H-16-0086

Klemen, J., Büchel, C., Bühler, M., Menz, M. M., & Rose, M. (2009). Auditory Working Memory Load Impairs Visual Ventral Stream Processing: Toward a Unified Model of Attentional Load. *Journal of Cognitive Neuroscience,* 437–446.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: a cross-sectional study on 800 children from the United States. *Developmental Psychology, 20*, 331-354. doi:10.1207/S15326942DN2001_2

Kort, W., Schittekatte, M., & Compaan, E. (2010). *CELF-4-NL | Test voor diagnose en evaluatie van taalproblemen.* Amsterdam: Pearson.

Kumar, P., Singh, N. K., & Hussain, R. O. (2021). Effect of speech in noise training in the auditory and cognitive skills in children with auditory processing disorders. *International Journal of Pediatric Otorhinolaryngology, 146*, 110735. doi:10.1016/j.ijporl.2021.110735

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56*(2), 183-197. doi:10.3758/bf03213897

Lewis, D. E., Manninen, C. M., & Valente, D. L. (2014). Children's Understanding of Instructions Presented in Noise and Reverberation. *American Journal of Audiology, 23*(3), 326-336. doi:10.1044/2014_AJA-14-0020

MacCutcheon, D., Pausch, F., Füllgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., . . . Ljung, R. (2019). The Contribution of Individual Differences in Memory Span and Language Ability to Spatial Release From Masking in Young Children. *Journal of Speech, Language, and Hearing Research, 62*, 3741-3751. doi:10.1044/2019_JSLHR-S-19-0012

Magimairaj, B. M., & Nagaraj, N. K. (2018). Working Memory and Auditory Processing in School-Age Children. *Language, Speech, and Hearing Services in Schools, 49,* 409-423. doi:10.1044/2018_LSHSS-17-0099

Magimairaj, B. M., Nagaraj, N. K., & Benafield, N. J. (2018). Children's Speech Perception in Noise: Evidence for Dissociation from Language and Working Memory. *Journal of Speech, Language, and Hearing Research, 61*, 409-423. doi:10.1044/2018_JSLHR-H-17-0312

Magimairaj, B. M., Nagaraj, N. K., Sergeev, A. V., & Benafield, N. J. (2020). Comparison of Auditory, Language, Memory, and Attention Abilities in Children With and Without Listening Difficulties. *American Journal of Audiology,* 1-18. doi:10.1044/2020_AJA-20-00018

Manly, T., Nimmo-Smith, I., Watson, P., Anderson, V., Turner, A., & Robertson, H. (2001). The differential assessment of children's attention: the test of everyday attention for children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry, 42*, 1065. doi:10.1111/1469- 7610.00806

Martin, N., & Brownell, R. (2005). *Test of Auditory Processing Skills - Third Edition (TAPS-3).* Novato, CA: Academic Therapy Productions.

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

McCreery, R. W., Miller, M. K., Buss, E., & Leibold, L. J. (2020). Cognitive and Linguistic Contributions to Masked Speech Recognition in Children. *Journal of Speech, Language, and Hearing Research, 63*, 3525-3539. doi:10.1044/2020_JSLHR-20-00030

McCreery, R. W., Spratford, M., Kirby, B., & Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *International Journal of Audiology, 56*(5), 306-315. doi:10.1080/14992027.2016.126670

McCreery, R. W., Walker, E. A., Spratford, M., Lewis, D., & Brennan, M. (2019). Auditory, Cognitive, and Linguistic Factors Predict Speech Recognition in Adverse Listening Conditions for Children With Hearing Loss. *Frontiers in Neuroscience, 13*, 1093. doi:10.3389/fnins.2019.01093

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools, 36*(3), 342-351. doi:10.1044/0161-1461(2008/032)

Mealings, K., & Cameron, S. (2019). Investigating Auditory Spectral and Temporal Resolution Deficits in Children with Reading Difficulties. *Journal of the American Academy of Audiology, 30*, 533-543. doi:10.3766/jaaa.17142

Mealings, K., Cameron, S., & Dillon, H. (2020). Correlating performance on the Listening in Spatialised Noise - Sentences test (LiSN-S) with the Listening in Spatialised Noise - Universal test (LiSN-U). *International Journal of Audiology, 59*(7), 519-523. doi:10.1080/14992027.2020.1753119

Meister, H., Schreitmüller, S., Grugel, L., Ortmann, M., Beutner, D., Walger, M., & Meister, I. G. (2013). Cognitive resources related to speech recognition with a competing talker in young and old listeners. *Neuroscience, 232*, 74-82. doi:10.1016/j.neuroscience.2012.12.006

Misurelli, S. M., Goupell, M. J., Burg, E. A., Jocewicz, R., Kan, A., & Litovsky, R. Y. (2020). Auditory Attention and Spatial Unmasking in Children With Cochlear Implants. *Trends in Hearing, 24*, 1-11. doi:10.1177/2331216520946983

Moore, D. R., Ferguson, M. A., Edmonson-Jones, A. M., Ratib, S., & Riley, A. (2010). Nature of Auditory Processing Disorder in Children. *Pediatrics, 126*(2), e382-e390. doi:10.1542/peds.2009-2826

Moossavi, A., Mehrkian, S., Karami, F., Biglarian, A., & Mahmoodi, B. (2017). Development of Persian version of the BKB sentences and content validity assessment. *Auditory and Vestibular Research, 26*, 27-33.

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research, 352*, 40-48. doi:10.1016/j.heares.2017.02.005

Nadler, R. T., & Archibald, L. M. (2014). The Assessment of Verbal and Visuospatial Working Memory With School Age Canadian Children. *Canadian Journal of Speech-Language Pathology and Audiology, 38*(3), 262–279.

Nagaraj, N. K., Magimairaj, B. M., & Schwartz, S. (2020). Auditory distraction in school-age children relative to individual differences in working memory capacity. *Attention, Perception, & Psychophysics, 82*, 3581-3593. doi:10.3758/s13414-020-02056-5

Newton, C., & Ridgway, S. (2015). Novel accent perception in typically-developing school-aged children. *Child Language Teaching and Therapy, 32*(1), 111-123. doi:10.1177/0265659015578464

Nilsson, M., Soli, S., & Sullivan, J. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America, 95*, 1085-1099. doi:10.1121/1.408469

Nittrouer, S., Kuess, J., & Lowenstein, J. H. (2015). Speech perception of sine-wave signals by children with cochlear implants. *Journal of the Acoustical Society of America, 137*(5), 2811. doi:10.1121/1.4919316

Pashler, H. E. (1998). *The Psychology of Attention.* Cambridge, MA: MIT Press.

Percy-Smith, L., Wischmann, S., Josvassen, J. L., Hallstrøm, M., Laplante-Lévesque, A., Sorgenfrei, M. G., & Caye-Thomasen, P. (2020). Evaluation of a sentence test in noise in children with hearing impairment. *Danish Medical Journal, 67*(1), A06190358.

Petley, L., Hunter, L. L., Zadeh, L. M., Stewart, H. J., Sloat, N. T., Perdew, A., . . . Moore, D. R. (2021). Listening Difficulties in Children with Normal Audiograms: Relation to Hearing and Cognition. *Ear and Hearing, 42*(6), 1640-1655. doi:10.1097/AUD.0000000000001076

Pickering, S. J., & Gathercole, S. E. (2001). *Working Memory Test Battery for Children.* London, England: Psychological Corporation Europe.

Pittman, A. (2011). Children's Performance in Complex Listening Conditions: Effects of Hearing Loss and Digital Noise Reduction. *Journal of Speech, Language, and Hearing Research, 54*, 1224-1239. doi:10.1044/1092-4388(2010/10-0225)

Riccio, C. A., Cohen, M. J., Garrison, T., & Smith, B. (2005). Auditory Processing Measures: Correlation with Neuropsychological Measures of Attention, Memory, and Behavior. *Child Neuropsychology, 11*(4), 363-372. doi:10.1080/09297040490916956

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale - Revised.* Wood Dale, IL: Stoelting Co.

Roman, A. S., Pisoni, D. B., Kronenberger, W. G., & Faulkner, K. F. (2017). Some Neurocognitive Correlates of Noise-Vocoded Speech Perception in Children with

Normal Hearing: A Replication and Extension of Eisenberg et al., 2002. *Ear and Hearing, 38*(3), 344-356. doi:10.1097/AUD.0000000000000393

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advance. *Frontiers in Systems Neuroscience, 7*. doi:10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47,* S99-S105. doi:10.1080/14992020802301167

Semel, E., Wiig, E., & Secord, W. (2003). *Clinical Evaluation of Language Fundamentals. Fourth Edition (CELF-4).* San Antonio, TX: The Psychological Corporation.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective Attention in Normal and Impaired Hearing. *Trends in Amplification, 12*(4), 283-299.

Söderlund, G. B., & Jobs, E. N. (2016). Differences in Speech Recognition Between children with Attention Deficits and Typically Developing Children Disappear When Exposed to 65 dB of Auditory noise. *Frontiers in Psychology, Sec. Auditory Cognitive Neuroscience*. doi:10.3389/fpsyg.2016.00034

Soli, S. D., & Wong, L. L. (2008). Assessment of speech intelligibility in noise with the Hearing in Noise Test. *International Journal of Audiology, 47,* 256-261. doi:10.1080/14992020801895136

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language, 123*(3), 191-201. doi:10.1016/j.bandl.2012.09.001

Sullivan, J. R., Osman, H., & Schafer, E. C. (2015). The Effect of Noise on the Relationship Between Auditory Working Memory and Comprehension in School-Age Children. *Journal of Speech, Language, and Hearing Research, 58*, 1043-1051. doi:10.1044/2015_JSLHR-H-14-0204

Sutcliffe, P., Bishop, D. V., Houghton, S., & Taylor, M. (2006). Effect of Attentional State on Frequency Discrimination: A Comparison of Children With ADHD On and Off Medication. *Journal of Speech, Language, and Hearing Research, 49*, 1072-1084. doi:10.1044/1092-4388(2006/076)

Tao, D., Deng, R., Jiang, Y., & Galvin, J. J. (2014). Contribution of Auditory Working Memory to Speech Understanding in Mandarin-Speaking Cochlear Implant Users. *PLoS ONE, 9*(6), e99096. doi:10.1371/journal.pone.0099096

Thompson, E. C., Krizman, J., White-Schwoch, T., Nicola, T., Estabrook, R., & Kraus, N. (2019). Neurophysiological, linguistics, and cognitive predictors of children's ability to perceive speech in noise. *Developmental Cognitive Neuroscience, 39*, 100672. doi:10.1016/j.dcn.2019.100672

Thompson, E. C., Woodruff Carr, K., White-Schwoch, T., Otto-Meyer, S., & Kraus, N. (2017). Individual differences in speech-in-noise perception parallel neural speech processing and attention in preschoolers. *Hearing Research, 344*, 148-159. doi:10.1016/j.heares.2016.11.007

Tillery, K. L., Katz, J., & Keller, W. D. (2000). Effects of Methylphenidate (Ritalin) on Auditory Performance in Children With Attention and Auditory Processing Disorders. *Journal of Speech, Language, and Hearing Research, 43*, 893-901. doi:10.1044/jslhr.4304.893

Tomlin, D., Dillon, H., Sharma, M., & Rance, G. (2015). The Impact of Auditory Processing and Cognitive Abilities in Children. *Ear and Hearing, 36,* 527-542. doi:10.1097/AUD.0000000000000172

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Levac, D., . . . Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checlist and Explanation. *Annals of Internal Medicine*. doi:10.7326/M18-0850

Vickers, D., Degun, A., Canas, A., Stainsby, T., & Vanpoucke, F. (2016). Deactivating cochlear implant electrodes based on pitch information for users of the ACE strategy. In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner, & C. Lanting, *Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp. 115-123). Cham, Switzerland: Springer.

von Lochow, H., Lyberg-Åhlander, V., Sahlén, B., Kastberg, T., & Brännström, J. (2018). The effect of voice quality and competing speakers in a passage comprehension task: perceived effort in relation to cognitive functioning and performance in children with normal hearing. *Logopedics Phoniatrics Vocology, 43*(1), 32-41. doi:10.1080/14015439.2017.1307446

Walker, E. A., Sapp, C., Oleson, J. J., & McCreery, R. W. (2019). Longitudinal Speech Recognition in Noise in Children: Effects of Hearing Status and Vocabulary. *Frontiers in Psychology, Sec. Auditory Cognitive Neuroscience*. doi:10.3389/fpsyg.2019.02421

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children - 3rd. Ed.* San Antonio, TX: The Psychological Corporation.

Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., . . . Gerson, R. C. (2013). I. NIH Toolbox Cognition Battery (CB): Introduction and pediatric data. *Monographs of the Society for Research in Child Development, 78*(4), 1-15.

Wightman, F. L., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: auditory-visual integration. *Journal of the Acoustical Society of America, 119*(6), 3940-9. doi:10.1121/1.2195121

Woodcock, R., McGrew, K., & Mather, N. (2007). *Woodcock Johnson III Tests of Cognitive Abilities.* Ithaca, IL: Riverside Publishing.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, *12*(5), 732-745. doi:10.1111/j.1467-7687.2009.00817.x

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2011). Noise on, voicing off: Speech perception deficits in children with specific language impairment. *Journal of Experimental Child Psychology, 110*, 362-372. doi:10.1016/j.jecp.2011.05.001

Figure 1. Search strategy.

Table 1. Summary of studies included in this review.

Study	Population(s)	Age range(s)	Number of participants	SPIN Measure(s)	Noise	WM Measure(s)	Attention Measure(s)	Results
Ahmmed et al., 2014	APD	6-11	110	Auditory Figure Ground, Filtered Words, Competing Words, Competing Sentences (SCAN-3:C) VCV SPIN test (IMAP)	SCAN-3:C: 8-talker babble noise, +8 dB SNR IMAP: speech-weighted noise, adaptive procedure	FDS + BDS (WISC)	Institute of Hearing Research – Cued Attention Test	Attention and WM grouped as a major factor explaining differences in SPIN.
Choi et al., 2008	TD	7-14	64	Repeat PBK words	Speech-weighted noise, +8 dB SNR		Difference between single- and dual-task conditions	Adding a second task improved SPIN performance relative to single-task, regardless of instruction to allocate attention to either task.
Grieco-Calub et al., 2017	TD	8-12	27	Repeat BKB sentences (noise-vocoded)			Difference between single- and dual-task conditions, Flanker (NIHTB-CB)	Adding a second task did not worsen SPIN performance but did worsen performance on secondary (visual) task. Flanker scores not correlated with either.
Hicks & Tharpe, 2002	TD, HL	5-11	20	Repeat PBK words	20-talker babble noise, +10, +15, and +25 dB SNR		Difference between single- and dual-task conditions	Adding a second task worsened SPIN performance in both groups equally. NH group had better scores in all conditions.
Howard et al., 2010	TD	9-12	31	Repeat AB words	Multi-talker babble noise, -4, 0, and +4 dB SNR		Difference between single- and dual-task conditions	Adding a second task worsened SPIN performance relative to baseline.

Hsu et al., 2020	TD	6-18	73	Categorize nouns	Speech-weighted noise, -3, 0, +3 dB SNR	FDS + BDS (CELF-4)		BDS and aggregate score, but not FDS alone, correlated with response times on the SPIN task.
Huyck, 2018	TD	11-22	48	Repeat noise-vocoded sentences			Rapid Visual Information Processing (CANTAB)	Attention correlated with SPIN scores.
Javanbakht et al., 2021	HL	8-12	31	Persian version of BKB-SIN	4-talker babble noise, SNRs start at +21 and lower with each list until -6 dB SNR	BDS		WM correlated with SPIN.
Kane et al., 2021	TD	5-13	Not reported	Repeat words from spatially separated speakers	Speech-weighted noise, single-talker, 3-talker babble noise, adaptive procedure		Benefit from spatial cues	Location cue improved SPIN performance in three-voice masker only.
Klein et al., 2017	TD, HL	5-12	25, 24	Repeat isolated words	Speech-weighted noise, adaptive procedure	Odd One Out (AWMA)		WM correlated with some conditions of SPIN task in the HL group. HL had worse scores in SPIN but not WM.
Kumar et al., 2021	APD	9-10	20	SPIN-IE (Indian English)	8-talker babble noise, 0 dB SNR	BDS, Ascending Digit Span, Descending Digit Span		No correlation between SPIN and WM at initial assessment. Digit spans improved following SPIN training.
Lewis et al., 2014	TD	8-12	50	Follow instructions from one or more targets	20-talker babble noise, with or without reverberation, +5 dB SNR	FDS + BDS (WISC)	Difference between single- and multiple-talker conditions	WM not correlated with SPIN. Higher attentional demands led to worse SPIN performance.
MacCutcheon et al., 2019	TD	4-11	39	Children’s Coordinated Response Measure	Speech-weighted noise or single-talker, co-located or not, adaptive procedure	BDS (CELF-4)		WM correlated with SPIN. Children with better WM benefitted more from spatial cues.

Magimairaj et al., 2018	TD	7-11	83	BKB-SIN	4-talker babble noise, SNRs start at +21 and lower with each list until -6 dB SNR	Auditory Working Memory (WJIII), Visual Digit WM Task	Dichotic Digits Test	WM and attention not correlated with SPIN measures.
Magimairaj et al., 2020	TD, LiD	7-11	85, 26	BKB-SIN	4-talker babble noise, SNRs start at +21 and lower with each list until -6 dB SNR	Auditory Working Memory (WJIII), Visual Digit WM Task	Dichotic Digits Test, Visual Attention Switching Task	The two groups differ on SPIN but not WM or attention.
McCreery et al., 2017	TD	5-12	96	Repeat meaningless sentences or disconnected words	Speech-weighted noise, adaptive procedure	Odd One Out, Counting Recall (AWMA)		WM correlated with SPIN.
McCreery et al., 2019	TD, HL	7-10	50, 56	Repeat meaningless sentences or sequences of disconnected words	Speech-weighted noise with or without reverberation, adaptive procedure	Odd One Out (AWMA)	Auditory Attention (NEPSY-II)	WM correlated with SPIN in both noise conditions for both groups. Attention correlated only with noise plus reverb and only for HL group.
McCreery et al., 2020	TD	5-6, 9-10	60	Repeat BKB sentences	Speech-weighted noise, amplitude-modulated speech-weighted noise, or 2-talker babble noise, adaptive procedure	Odd One Out, Listening Recall (AWMA)	Flanker, Dimensional Change Card Sort (NIHTB-CB)	WM correlated with SPIN. Flanker correlated with SPIN in some conditions, DCCS did not.
McFadden & Pittman, 2008	TD, HL	8-12	11-19	Categorize nouns	White noise, 0 and +6 dB SNR		Difference between single- and dual-task conditions	Adding a second task did not worsen performance on the SPIN task for the NH group, but did for the HL group.
Mealings & Cameron, 2019	RD	8-11	16	High-cue condition of the LiSN-S	Continuous irrelevant speech produced by a speaker of different gender than target,	Number Memory Reversed (TAPS-3)	Inhibition (NEPSY-II)	No significant correlation between WM, attention, and SPIN.

					spatially separated, adaptive procedure			
Mealings et al., 2020	TD	6-7	16	LiSN-S	Continuous irrelevant speech produced by a talker of different or same gender, co-located or not, adaptive procedure	Number Memory Reversed (TAPS-3)		Correlation disappears after correction; authors argue that a relationship is nonetheless likely.
Misurelli et al., 2020	HL	10-17	10	Repeat Matrix sentences	Non-target sentences produced by a talker of different gender, ipsilateral, contralateral, or bilateral, adaptive procedure		Difference between quiet and non-quiet conditions	All noise worsened SPIN performance, especially ipsilateral and bilateral.
Moore et al., 2010	TD, APD	6-11	1469	VCV non-words (IMAP)	Speech-weighted noise, adaptive procedure	FDS + BDS (WISC)	Institute of Hearing Research – Cued Attention Test	Correlation of WM and especially attention with clinical presentations of APD. APD thus likely an attentional problem.
Nagaraj et al., 2020	TD	7-12	125	Dichotic Digits Test	Non-target digits on opposite ear and conversation between adults and children, +8 dB SNR	Auditory Working Memory (WJIII)	Proportion of intrusion errors	WM correlated with SPIN and with attention. Higher WM led to more intrusion errors, but fewer errors overall.
Newton & Ridgway, 2015	TD	6-7	26	BKB sentences with familiar and novel accents	Speech-weighted noise, adaptive procedure		Creature Counting, Map Mission (TEA-Ch)	Novel accents yield higher (worse) SRTs; no correlation between SPIN and attention measures.
Percy-Smith et al, 2020	TD, HL	4-10	70	Dantale II	Speech-weighted noise, adaptive procedure	FDS + BDS (CELF-4)		WM and SPIN correlated for the HL group. Correlations not reported for the NH group.

Petley et al., 2021	TD, LiD	6-13	79, 67	Auditory Figure Ground, Filtered Words, Competing Words, Competing Sentences (SCAN-3:C), LiSN-S	SCAN-3:C: 8-talker babble noise, +8 dB SNR; LiSN-S: Continuous irrelevant speech produced by a talker of different or same gender, co-located or not, adaptive procedure	List Sorting Working Memory Test (NIHTB-CIB)	Flanker, Dimensional Change Card Sort (NIHTB-CB)	WM and attention strongly correlated with SPIN.
Pittman, 2011	TD, HL	8-12	50, 34	Repeat single words	Speech-weighted noise, 0 dB SNR		Difference between single- and dual-task conditions	Adding a second task worsened SPIN performance in both groups.
Riccio et al., 2005	APD, ADHD	4-11	36	Auditory Figure Ground, Filtered Words (SCAN-3:C)	8-talker babble noise, +8 dB SNR		Test of Variables of Attention	Attention did not correlate with SPIN.
Roman et al., 2017	TD	5-13	31	Repeat words in isolation and in sentences (noise-vocoded)			Auditory Attention, Response Set (NEPSY-II)	Attention correlated with SPIN.
Strait et al., 2012	TD (musically trained, non-musically trained)	7-13	15, 16	Hearing-in-Noise Test (HINT), Words in Noise (WIN)	HINT: speech-weighted noise, adaptive procedure, co-located or not; WIN: 4-talker babble, 24 to 0 dB SNR in 4 dB steps	Auditory Working Memory (WJIII); Visual Working Memory (Colorado Assessment Tests 1.2)	Integrated Visual And Auditory Continuous Performance Test	HINT left/right conditions correlate positively with auditory but not visual working memory; no correlations with attention measures or WIN measures.
Sullivan et al., 2015	TD	8-10	20	Answer questions about content of short stories	Spatially-separated classroom noise, -5 dB SNR	Backwards Digit Recall, Listening Recall (WM Test Battery for Children)		WM strongly correlated with SPIN.

Söderlund & Jobs, 2016	TD, ADHD	9-10	39, 10	Repeat Hagerman sentences	Speech-weighted noise, adaptive procedure	FDS + BDS (WISC)	Score, Score-DT (TEA-Ch)	ADHD children have worse WM and attention than controls, but not SPIN performance.
Tao et al., 2014	TD, HL	8-14, 6-26	21, 32	Mandarin Speech Perception Test	Speech-weighted noise, +5 dB SNR	BDS		WM not correlated with SPIN. CI users had worse WM scores.
Thompson et al., 2017	TD	3-5	59	Repeat single words produced by a male talker	Non-target meaningful sentences produced by a female talker, adaptive procedure		Attention Sustained (Leiter-R)	Improvements in attention correlated with improvements in SPIN scores between 3-5.
Thompson et al., 2019	TD	4-7	99	Hearing-in-Noise Test	Speech-weighted noise, adaptive procedure, co-located or not	Numbers Reversed, Auditory Working Memory (WJIII)	Attention Sustained (Leiter-R)	WM correlated with SPIN, but attention did not. WM and attention correlated with each other.
Tillery et al., 2000	APD, ADHD	Not reported	32	Speech-in-Noise Test	Speech-weighted noise, +10 dB SNR		Auditory Continuous Performance Test	Ritalin improved attention scores but not SPIN scores.
Tomlin et al., 2015	TD, APD	7-12	150	LiSN-S	Continuous irrelevant speech produced by a talker of different or same gender, co-located or not, adaptive procedure	FDS + BDS (CELF-4)	Integrated Visual and Auditory Continuous Performance Test	WM correlated with some conditions of the LiSN-S. Attention did not correlate.
von Lochow et al., 2018	TD	7-12	49	Answer questions about content of short stories	Multi-talker babble noise, +5 dB SNR	Competing Language Processing Test		Better WM associated with smaller difference in SPIN performance between quiet and noise conditions.
Walker et al., 2019	TD, HL	7-10	290	BKB-SIN	4-talker babble noise, SNRs start at +21 and lower with each list until -6 dB SNR	Odd One Out + Listening Recall + BDR (AWMA)		WM did not correlate with SPIN for children with HL. Correlations not reported for the NH group.

Abbreviations: TD = typical development; APD = auditory processing disorder; HL = hearing loss; RD = reading difficulties; LiD = listening difficulties; ADHD = attention deficit/hyperactivity disorder; SPIN = Speech perception in noise; SCAN-3:C = Tests for Auditory Processing Disorders for Children – 3rd Edition; VCV = vowel-consonant-vowel; IMAP = Medical Research Council Institute of Hearing Research Multi-center Auditory Processing Test; BKB-SIN = Bamford-Kowal-Bench Speech-in-Noise Test; LiSN-S = Listening in Spatialized Noise – Sentences Test; SNR = Signal-to-noise ratio; PBK = Phonetically Balanced Kindergarten; AB = Arthur Boothroyd; WM = Working Memory; FDS = Forward Digit Span; BDS = Backwards Digit Span; WISC = Wechsler Scale of Intelligence for Children; WJIII = Woodcock-Johnson III Test of Cognitive Abilities; AWMA = Automated Working Memory Assessment; BDR = Backwards Digit Recall; TAPS-3 = Test of Auditory Processing Skills – Third Edition; NEPSY-II = Developmental Neuropsychological Assessment – Second Edition; TEA-Ch = Test of Everyday Attention for Children; CANTAB = Cambridge Neuropsychological Test Automated Battery; NIHTB-CB = National Institute of Health Toolbox Cognition Battery.

Table 2. Summary of standardized tests used in studies included in this review.

Test	Subtest	Measure of	Modality	Description
Bamford-Kowal-Bench Speech-in-Noise (BKB-SIN)		SPIN	Auditory	Repeat short, meaningful sentences in the presence of 4-talker babble noise. Scoring is based on keywords. SNRs start at +21 and with each list decrease by 3 dB until -6 dB SNR.
Dantale II		SPIN	Auditory	Repeat sentences that all contain the same structure: name, verb, numeral, adjective, and object, but no semantic cues. Target level is 70 dB SPL. Noise is speech-weighted noise whose level varies adaptively. Target and noise are presented in free field from the same speaker at 0º azimuth.
Hearing in Noise Test (HINT)		SPIN	Auditory	Repeat BKB sentences presented from a loudspeaker at a fixed 65 dB SPL in speech-weighted noise using an adaptive SNR. In one condition, sound and masker come from the front speaker. In others, target comes from the front and masker comes from the side (± 90º speakers).
Medical Research Council Institute of Hearing Research Multi-center Auditory Processing (IMAP)		SPIN	Auditory	Repeat VCV nonwords presented in speech-weighted noise over headphones with adaptive SNR.
Listening in Spatialized Noise – Sentences (LiSN-S)		SPIN	Auditory	Repeat meaningful sentences presented at variable SNR in continuous irrelevant speech at constant 55 dB SPL in free field. Distractor speech can be co-located with target speech or from both + and –90° azimuth from it. The voice in target and distractor stimuli may be the same or different. In the low-cue condition, the same voice is used for both target and distractor, and both are co-located at 0° azimuth. In the talker-advantage condition, both are co-located, but different voices are used. In the spatial advantage condition, the same voice is used but target and distractor are spatially separated. In the high-cue condition, both different voices and spatial separation are used.
Tests for Auditory Processing Disorders for Children – 3rd Edition (SCAN-3:C)	Auditory Figure Ground	SPIN	Auditory	Repeat single words presented either to left or right ear against a background of multi-talker babble noise. Different versions use different fixed SNRs.
	Filtered Words	SPIN	Auditory	Repeat low-pass filtered, monosyllabic words in quiet.
	Competing Words: Directed Ear	SPIN	Auditory	Repeat words from a given ear when different words are presented simultaneously to both ears.

	Competing Words: Free Recall	SPIN	Auditory	Repeat words in any order when different words are presented simultaneously to both ears.
	Competing Sentences	SPIN	Auditory	Repeat sentences from a given ear when different sentences are presented simultaneously to both ears.
Words in Noise (WIN)		SPIN	Auditory	Repeat isolated words presented at 70 dB SPL in 4-talker babble noise. SNR starts at 24 dB SNR and every five words decreases in 4 dB steps until 0 dB SNR.
Automated Working Memory Assessment (AWMA)	Odd One Out	WM	Visual	Identify which of three shapes is different from the others and recall its position on an empty screen.
	Listening Recall	WM	Auditory	Judge if each sentence is true or false, then recall the last word in each sentence in order.
	Counting Recall	WM	Visual	Count aloud a number of red circles presented in an array with triangles. After a series of arrays, recall how many circles were counted per array in order.
	Backwards Digit Recall	WM	Auditory	Repeat a series of digits in reverse order.
Colorado Assessment Tests 1.2	Visual Working Memory	WM	Visual	Watch a group of blocks change color on a screen and then click them in the same or reverse order they changed color. The number of boxes changing color increases with each successful trial.
Competing Language Processing Test (CLPT)		WM	Auditory	Judge if each sentence is true or false, then recall the last word in each sentence in any order.
Clinical Evaluation of Language Fundamentals (CELF)	Backwards Digit Span	WM	Auditory	Repeat a series of digits in reverse order.
National Institute of Health Toolbox Cognition Battery (NIHTB-CB)	List Sorting Working Memory Test	WM	Visual and Auditory	Remember a series of stimuli presented both visually and auditorily and order them from smallest to biggest.
	Flanker	Attention	Visual	Report the direction of an arrow (or fish if participant is under 8) when the arrow is flanked by other arrows in the same or opposite direction.
	Dimensional Change Card Sort	Attention	Visual	Match a target stimulus based on color or shape, varying across trials.
Test of Auditory Processing Skills (TAPS)	Number Memory Reversed	WM	Auditory	Repeat a series of digits in reverse order.

Woodcock-Johnson – 3rd Edition (WJIII)	Auditory Working Memory	WM	Auditory	Listen to a series of nouns and numbers presented alternatingly and repeat the nouns followed by the numbers in the order presented.
	Numbers Reversed	WM	Auditory	Repeat a series of digits in reverse order.
Working Memory Test Battery for Children	Backwards Digit Span	WM	Auditory	Repeat a series of digits in reverse order.
	Listening Recall	WM	Auditory	Judge if each sentence is true or false, then recall the last word in each sentence in order.
Auditory Continuous Performance Test (ACPT)		Attention	Auditory	Raise a thumb whenever the target word is heard among other familiar words. Stimuli are presented at 50 dB SL above the participant's pure-tone-average.
Cambridge Neuropsychological Test Automated Battery (CANTAB)	Rapid Visual Information Processing	Attention	Visual	Watch a series of digits on screen and press a button when a target sequence is formed.
Institute of Hearing Research - Cued Attention Test (IHR-CAT)		Attention	Visual/Auditory	Press a button as quickly as possible in response to a target stimulus presented visually or auditorily. Target stimuli are sometimes preceded by a cue, thus allowing for four measures: auditory cued, auditory non-cued, visual cued, and visual non-cued.
Integrated Visual and Auditory Continuous Performance Task (IVA-CPT)		Attention	Visual/Auditory	Watch a series of single digits presented visually or auditorily and press a button when a "1" is presented but not a "2."
Leiter-R	Attention Sustained	Attention	Visual	Mark instances of a target presented on a page filled with target and distractor images.
Developmental Neuropsychological Assessment (NEPSY)	Auditory Attention	Attention	Auditory	Listen to a series of words and touch a button when a target word is heard.
	Inhibition	Attention	Visual	Watch a series of shapes and arrows presented. Name the shape or, in the case of arrows, indicate the direction, or still provide an alternate response, depending on the object's color.
	Response Set	Attention	Auditory	Listen to a series of words and respond according to changing instructions depending on the block.
Test of Everyday Attention for Children (TEA-Ch)	Score	Attention	Auditory	Count the number of tones played.

	Score-DT	Attention	Auditory	Count the number of tones played while monitoring a second stream for a target word.
	Creature Counting	Attention	Visual	Count the number of aliens in their burrow. When prompted, switch from counting upwards to counting downwards.
	Map Mission	Attention	Visual	Find as many as possible of a target symbol on a map within one minute.
Test of Variables of Attention (TOVA)		Attention	Visual	Press a button when an “X,” but not other letters, is presented on screen.