JASA ARTICLE

Speech-in-noise testing: Innovative applications for pediatric patients, underrepresented populations, fitness for duty, clinical trials, and remote services[a)]

Victoria A. Sanchez,[1,b)] Michelle L. Arnold,[2] David R. Moore,[3] Odile Clavier,[4] and Harvey B. Abrams[2]

[1]*Department of Otolaryngology–Head and Neck Surgery, University of South Florida, 12901 Bruce B. Downs Boulevard, MDC 73, Tampa, Florida 33612, USA*

[2]*Department of Communication Sciences and Disorders, University of South Florida, Tampa, Florida 33612, USA*

[3]*Communication Sciences Research Center, Cincinnati Children's Hospital, Cincinnati, Ohio 45229, USA*

[4]*Creare LLC, Hanover, New Hampshire 03755, USA*

ABSTRACT:

Speech perception testing, defined as providing standardized speech stimuli and requiring a listener to provide a behavioral and scored response, has been an integral part of the audiologic test battery since the beginning of the audiology profession. Over the past several decades, limitations in the diagnostic and prognostic validity of standard speech perception testing as routinely administered in the clinic have been noted, and the promotion of speech-in-noise testing has been highlighted. This review will summarize emerging and innovative approaches to speech-in-noise testing with a focus on five applications: (1) pediatric considerations promoting the measurement of sensory and cognitive components separately; (2) appropriately serving underrepresented populations with special attention to racial, ethnic, and linguistic minorities, as well as considering biological sex and/or gender differences as variables of interest; (3) binaural fitness for duty assessments of functional hearing for occupational settings that demand the ability to detect, recognize, and localize sounds; (4) utilization of speech-in-noise tests in pharmacotherapeutic clinical trials with considerations to the drug mechanistic action, the patient populations, and the study design; and (5) online and mobile applications of hearing assessment that increase accessibility and the direct-to-consumer market. https://doi.org/10.1121/10.0014418

(Received 6 December 2021; revised 24 August 2022; accepted 15 September 2022; published online 25 October 2022)

[Editor: James F. Lynch] Pages: 2336–2356

I. INTRODUCTION

The audiologic test battery has traditionally been associated with the assessment of pure-tone air- and bone-conduction thresholds and recognition measurements of words in quiet. However, the tests comprising this battery do not adequately reflect self-perceived hearing handicap, nor do they substantially inform rehabilitative success (Walden and Walden, 2004; Davidson *et al.*, 2021). In contrast to the traditional audiometric test battery, best audiologic practices call for measures of speech understanding in noise (American Academy of Audiology, 2006; British Society of Audiology, 2016), particularly for those individuals who report difficulties communicating in environments with multiple speakers or other types of background noise. Further, speech-in-noise (SiN) testing has gained recognition as an assessment that may be better equipped to describe functional hearing difficulties within the World Health Organization framework (Humes, 2019; Vermiglio and Fang, 2022). The traditional audiometric battery does not adequately predict speech recognition performance in noise, as threshold and performance measurements in quiet only provide insight into the audibility domain of hearing (Wilson, 2011). Beyond decreased audibility, the distortion aspect of sensorineural hearing loss cannot be estimated without measures of higher-level auditory processing, such as a SiN assessment (Killion, 2002). Furthermore, as we explore the underlying pathophysiology related to hearing loss, such as the more recent discovery of synaptopathy and the field's increased understanding of damage related to noise-induced hearing loss (NIHL) (e.g., Kujawa and Liberman, 2015), we need innovative tools that better measure auditory function. Therefore, as the field of audiology expands, clinical and investigational tools that are sensitive for detection of hearing function are needed—including those that involve SiN testing.

A variety of SiN tests are available to clinicians and researchers (Table I). Some of the most popular, clinically available SiN tests include the Hearing-in-Noise Test (HINT; Nilsson *et al.*, 1994), the Words-in-Noise Test (WIN; Wilson, 2003; Wilson *et al.*, 2003), and the Quick Speech-in-Noise Test (QuickSIN; Killion *et al.*, 2004). As evident in Table I, in addition to the most commonly used assessments, there are numerous other SiN tests available that vary in characteristics. The extensive repertoire of

[a)]This paper is part of a special issue on Noise-induced Hearing Disorders: Clinical and Investigational Tools.

[b)]Electronic mail: vasanchez@usf.edu

https://doi.org/10.1121/10.0014418

TABLE I. List of available SiN tests, including the name of the test, characteristics of the target signal and the competing noise, and some additional relevant considerations to paradigm, potential usage, question being asked, and the resources available.

Test	Target signal	Noise	Considerations
AzBio Sentence Test	Sentences	Four-talker babble	Manual administration at a single SNR level; 250 sentences in 33 lists spoken by four different speakers
Bamford–Kowal–Bench Speech-in-Noise Test			
BKB-SIN	High context sentences	MTB[a]	Descending paradigm; 50% threshold calculated with full performance function; designed for pediatric and CI patients
Digits-in-noise tests			Closed-set of familiar and frequent numbers; available in many languages; suitable for children
DIN[b]	Digits (1–10)	SSN[c]	Fully automated and adaptive; can be administered via mobile/remote devices; diotic or dichotic administration possible
DIN[d]	Digits (1–6, 8–10)	MTB	Descending paradigm, 50% threshold calculated with full performance function
Hearing-in-Noise Test			
HINT HINT-C	High context sentences	SSN	Adaptive manual paradigm; high contextual cues; entire sentence scored; computerized option available; normed for children 6–12; available in multiple languages
S-HINT	Spanish sentences	SSN	
Listening in Spatialized Noise-Sentences test			
LiSN-S	Low and high context sentences	Distracting sentences	Automatic adaptive paradigm; noise is distracting sentences with same speaker voice or different voice; allows for difference scores to evaluate talker, context, and spatial advantages; developed for children
Modified Rhyme Test			
MRT80	Rhyming monosyllable words	SSN	Automatic adaptive paradigm; two 80-word lists; factorial design with two speech levels, two signal to ratios, and diotic and dichotic conditions
Oldenburg Sentence Tests			
American English Matrix test	Five-word sentence	SSN	Automatic adaptive paradigm; available in >15 languages; fixed syntactical structure of a sentence and limited contextual cues
Portable automated rapid testing			
PART	CRM[e] sentences	MTB varying spatial cues	Automatic adaptive paradigm; testing done in co-located and spatially separated conditions to evaluate spatial release from masking
Quick Speech-in-Noise test			
QuickSIN	IEEE sentences	Four-talker babble	Descending paradigm; 50% threshold calculated with full performance function; constant noise level; limited context cues
Speech Recognition in Noise Test			
SPRINT	Monosyllabic words	Six-talker babble	Fixed SNR; scored as percent correct; 200 NU6 monosyllabic words (male speaker)
Words-in-Noise test			
WIN	Monosyllabic words	MTB	Descending paradigm, 50% threshold calculated with full performance function; monosyllabic NU6 words (female speaker); low context cues; normed for children 6 years and older
S-WIN	Bisyllabic Spanish words	MTB	Spanish version of the WIN test paradigm

[a]Multi-talker babble (MTB).
[b]DIN from Motlagh Zadeh *et al.* (2019).
[c]Speech-shaped noise (SSN).
[d]DIN from Wilson and Weakley (2004).
[e]Coordinate response measure (CRM).

evaluative options can make the selection of a SiN test, as well as appropriate interpretation of related outcomes, challenging. Thus, careful consideration of factors that influence SiN performance is needed by many audiences, including clinicians, researchers, and industry partners. As described in previous works (e.g., Miller, 1951; Bocca and Calearo, 1963; Wilson and McArdle, 2005), factors that influence SiN performance can be simplified into two categories: (1) intrinsic factors, or factors related to the listener, and (2) extrinsic factors, or factors related to the SiN test and test administration. Intrinsic factors include the listener's hearing sensitivity, developmental age, sex, race and ethnicity, language, cognitive abilities, motivation, and listening effort. Extrinsic factors include the testing materials with respect to the recorded speaker, level of context for the target stimuli, stimuli frequency spectrum, type of background noise, degree of informational masking (IM), physical sound-level paradigms, test environment, and method of

administration. The effects of these intrinsic and extrinsic factors on recognition performance are greatly debated.

Although a multitude of factors can influence SiN performance, many measures have been extensively validated, are suitable for both clinical and research protocols, and can be used for multiple populations and aims. To fully appreciate the flexibility and innovations in available SiN measures, this review discusses facets of the utility of SiN testing. Our approach to the review included engaging subgenre experts who synthesized relevant and recent papers in the various applications. We selected applications that are not commonly reviewed and that could provide unique opportunities for overlap of these not very common applications. The covered application suggestions and innovations include use with pediatric patients, considerations to sex differences, use with linguistically diverse populations, evaluations of fitness for duty, usefulness in clinical trials, and flexibility of use via mobile and remote services. Different audiences, such as clinicians, researchers, or industry partners, may initially read for a specific section of interest but could benefit from the other sections unexpectedly. For example, if an industry partner wanted to review the SiN applications for clinical trial utility, they may become aware of important considerations for underserved patient populations, considerations for pediatrics or service members, and how to potentially incorporate remote options in clinical trials. This wide-scope review is intended to uniquely allow for interaction among the different sections and for many audiences, including clinicians, researchers, and industry partners.

II. SPEECH TESTING IN CHILDREN

One of the most fascinating aspects of child development is language acquisition. This appears to be so robust in most children that it is easy to forget the small but significant portion, perhaps 5%–10% (Bishop and Leonard, 2000), who are left behind. The extent to which hearing difficulties contribute to this percentage is unclear, but children with mild or greater hearing loss are clearly at risk if untreated (Tomblin *et al.*, 2020). There is growing evidence that subclinical hearing loss and other types of hearing and listening difficulties are also associated with developmental language disability (Dillon and Cameron, 2021; Hunter *et al.*, 2020; Petley *et al.*, 2021). Of the difficulties that go undetected or untreated in current clinical practices, speech hearing in challenging environments is of primary concern because of its key role in human communication. We do not know the full extent to which the difficulties are of environmental or genetic origin or whether they are related to noise exposure. Nevertheless, it is critical that policymakers and caregivers be aware of the need for early detection, characterization, and intervention for difficulties in the perception of speech. Here, we present a range of findings, and their implications, that represent important innovations in pediatric clinical speech testing and interventions. We also present evidence on effects of noise exposure in childhood.

A. Neonatal testing and intervention

Measuring and remediating hearing in young children is extremely challenging but crucial for improving speech, language, and academic outcomes, as perhaps best appreciated by the benefits of expeditious fitting of cochlear implants to infants with profound hearing loss (Kral *et al.*, 2019). Newborn hearing screening has had a tremendous impact on the detection of moderate to profound hearing loss [pure-tone average (PTA) >40 dB sound pressure level (SPL)] and delivery of early intervention in the form of hearing aids and cochlear implants (Yoshinaga-Itano, 2003). Recent major innovations in this field include improving the sensitivity and specificity of neonatal test procedures, the use of speech sounds to predict outcomes, and the development of behavioral speech interventions to improve language.

With the correct ear and device calibration, it is now possible to measure otoacoustic emissions (OAEs) and auditory steady-state responses (ASSRs) down to 10–15 dB hearing level (HL) in newborns (Sininger *et al.*, 2018). It is around this HL range that longer-term academic, language, and cognitive problems in untreated older children can become significant (Moore *et al.*, 2020). Using the ASSR and multiple-frequency, narrowband, modulated chirp stimuli up to 14 dB, lower thresholds were found in less time than required for current standard-of-care automated auditory brainstem responses (ABRs; Sininger *et al.*, 2018). A notorious challenge of neonatal screening is the high false positive rate due to fluid in the middle ear, so improved response specificity for detecting sensorineural hearing loss is needed. Possible solutions to this problem, using complex acoustic stimuli, such as wideband reflectance (Hunter *et al.*, 2008; Sininger *et al.*, 2018) and multifrequency OAEs (Blankenship *et al.*, 2018), rather than tones, show promise for optimizing specificity sensitivity and potential relevance for speech processing.

Another innovative approach to newborn hearing screening has been envisioned by Kraus and White-Schwoch (2016). Speech sounds [consonant-vowel (CV) syllables: e.g., "da"] in quiet or in noise generate frequency following responses (FFRs) in infants (Anderson *et al.*, 2015) that resemble equivalent responses in older children, though more sluggish, variable, and small (Thompson *et al.*, 2021). Nevertheless, group responses revealed the same richness of information in time and amplitude that grew to become more secure and increased in power between 3 and 10 months of age (Fig. 1). Interestingly, the chirps used by Sininger *et al.* (2018) to measure ASSRs were spectrotemporally modulated and could thus, like Kraus's 40 ms "da," also be considered a speech-like stimulus that is near-ready for full clinical trials. Together, these approaches should lead to the next generation of newborn hearing tests to predict language outcomes.

Turning to intervention, recent evidence indicates that behavioral methods are promising. Infants nurtured from birth on a rich diet of clear speech and language, delivered in an instinctively caring, affective, and vibrant manner,

https://doi.org/10.1121/10.0014418

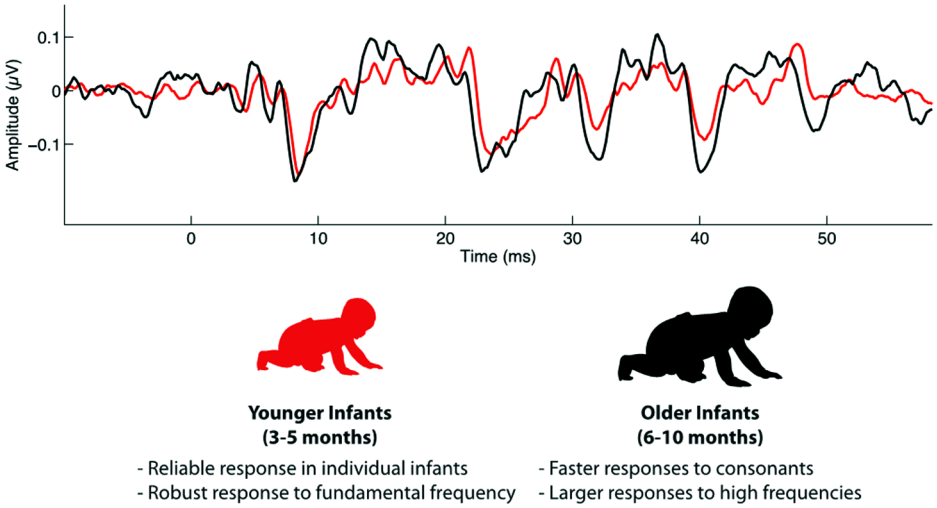

FIG. 1. (Color online) FFRs from younger and older infants show that ear and brainstem mechanisms of envelope and formant speech perception are present and near maturity at or soon after birth. Further development occurs over the first few years. Reprinted with permission from Kraus and White-Schwoch, Hear. J. **69**(11), 44–46 (2016). Copyright 2016 Lippincott Williams and Wilkins Ltd. (Kraus and White-Schwoch, 2016).

have enhanced ability to acquire social, intellectual, and communication skills that will serve them throughout life. One study involved coaching caregivers of typically developing infants to use "parentese," an instinctive, enhanced communication style, during the early months of life. Relative to a control group, the infants had enhanced conversational turn-taking and language skills at 18 months (Ferjan Ramírez *et al.*, 2020). Unfortunately, it follows that those lacking adequate communicative opportunities, whether through hearing loss, childhood learning disorders, or social/economic deprivation, may not flourish without intervention. However, similar strategies may be used as therapy for communication difficulties. For example, video and coaching techniques have been used to promote positive parenting of 9–14-month-old infants with early signs of autism spectrum disorder (Whitehouse *et al.*, 2021). Compared to usual care practices, such coaching resulted in a lower rate of autism diagnosis at age 3 years. To assure access to speech during the extremely critical period of infancy, regular testing of children's speech perception and production may be a vital contribution to preventing neurodevelopmental disorders. Speech-based FFR, as above, could be an objective surrogate for behavioral measures of speech perception throughout infancy.

B. Interaction between speech perception, language, and cognition

While a broad-spectrum audiogram is useful for determining the sensitivity of hearing, measuring speech intelligibility against competing stimuli should be an essential part of a pediatric audiological assessment. However, increasingly complex hearing tests simultaneously interrogate elements of cognition, notably attention, memory, and, in the case of speech, language. These two aspects of hearing and cognition are particularly critical in testing children, since their cognitive skills are developing alongside their auditory skills. In fact, physiological measures in animals and humans suggest that the coding of tones in the ear and central auditory system matures early, perhaps by age 2–3 years in humans (Ponton *et al.*, 1992). However, the development of speech coding extends into later childhood (∼8 years; Thompson *et al.*, 2021), while that of non-speech and SiN perception extends into adolescence (∼11 years; Corbin *et al.*, 2016; Moore *et al.*, 2011), and speech-in-speech perception matures even later (∼14 years; Corbin *et al.*, 2016; Moore *et al.*, 2019). These differences between objectively measured neural coding and behavioral performance appear to be due to the limited ability of younger children to perform increasingly complex behavioral tasks consistently. Using conditioned tasks (e.g., play audiometry; Suzuki and Ogiba, 1961) or sensitized behavioral observations (e.g., observer-based psychoacoustics; Olsho *et al.*, 1987) under stringent control, at least some children appear able to perform non-speech discriminations in early infancy. Around 6 years of age, a small proportion of children were found to perform pure-tone frequency discrimination to adult levels, while most children did not achieve that level of performance until much later (Halliday *et al.*, 2008). Most, if not all, of the later development appears to be cognitive (including language), and this factor also plays a major role in early development.

A recent physiological study suggests that most of the coding, or "sensory" aspect of speech perception seems to develop linearly between 3 and 8 years of age, as determined by the midbrain-dominated FFR (Thompson *et al.*, 2021). A multiplicity of changes in response spectrum, amplitude, phase, stability, and resting activity were observed. Interestingly, individual differences across age were apparent only in spectral coding, suggesting that this aspect of sensation may be a key differentiator of performance. However, the relatively small individual differences stand in stark contrast to the behavioral variability presented in the previous paragraph. It also appears that factors not measured by the audiogram, such as spectrotemporal dynamics, are critical for the development of speech perception. The detection of spectrotemporal modulation of tones is also a good predictor of SiN intelligibility in adults (Bernstein *et al.*, 2013).

Auditory cognitive maturation consists of two components: (1) the ability to decode the target stimulus based on

sensory input and (2) the ability to perform the given task. Decoding the target against a speech masking stimulus is sometimes referred to as "informational masking" (IM). IM is higher (more problematic) in children than in adults, as evidenced by the greater difficulty children have hearing a target speech stimulus against a speech masker (e.g., two-talker babble) than against a speech-shaped masker (see Table I; Corbin *et al.*, 2016). Inability to perform a perceptual task is commonplace and natural for younger children, but this has often been confused in the auditory development literature with an inability to hear, since it may be difficult to distinguish between hearing and other components of cognition, including attention, memory, and emotional status. In addition, background noise directly interferes with attention and learning in children. These aspects of listening should not be downplayed, since they are part of everyday perception, but it is confusing, at best, to refer to them as an "auditory processing disorder" (Moore, 2018).

C. Behavioral tests of speech perception

Because the sensory and cognitive components of hearing are highly interwoven but may require different intervention approaches, especially in children, it is useful to be able to measure them separately. This may be achieved for all the basic elements of hearing (e.g., pitch, duration, level) by designing subtractive or derived measures, where the results of two tests that differ only in a single dimension are numerically compared. Examples include temporal and spectral resolution (Moore, 2012; Moore *et al.*, 2010). In these tasks, cognitive factors of attention, memory, and linguistics are held constant for the two tests. Only the sensory elements differ, allowing a pure measure of sensory performance. Remarkably, detection threshold on these derived measures of resolution hardly changes across the age range 6–12 years, whereas the constituent tests, incorporating cognitive elements, show clear maturation across the same age range.

When using speech stimuli, additional linguistic factors are introduced that may reduce the effectiveness of the approach (Petley *et al.*, 2021). For example, in the Listening in Spatialized Noise-Sentences (LiSN-S) test (Cameron and Dillon, 2007; Table I), there are two "advantage" measures that involve the comparison with a "low cue" condition, where target and distractor sentences are presented from the same frontal, spatial direction (0°) and use the same talker. For spatial advantage, the comparison task is target at 0°, with same talker distractors at ±90°. For talker advantage, the comparison task uses different talkers for target and distractors, but all stimuli at 0°. The spatial advantage measure appears to be free of cognitive influences but, in contrast to the resolution measures, improves with maturation up to about 14 years old (Cameron *et al.*, 2011) and differs between children with and without a history of otitis media (Graydon *et al.*, 2017). A spatial "pattern score," incorporating the spatial advantage measure, was developed (Cameron and Dillon, 2011) as a quantitative clinical measure of the benefit of adding virtual spatial cues to the two conditions of the LiSN-S lacking those cues. The addition of a cutoff criterion for the pattern score defined a "spatial processing disorder" (SPD). In addition to otitis media, SPD has been associated with reported difficulties hearing in noisy environments (Cameron *et al.*, 2012). SPD is one of a small number of reports of experience-dependent, auditory sensory deficits in humans who have clinically normal PTAs. The talker advantage contrasts two different talkers as the distracting speech. In this instance, both the acoustic and linguistic signatures of the distractors differ, so unlike the spatial advantage with which it shares a similar maturation pattern (Cameron *et al.*, 2011), the talker advantage may be dependent on both acoustic and non-acoustic differences between the distractors.

The simpler digits-in-noise (DIN) test (Smits *et al.*, 2013; Van den Borre *et al.*, 2021; see Table I) has become very widely used, as it is suitable for automated self-testing, as reviewed elsewhere in this article. It may be completed with adult assistance from as young as 4 years old and independently from about 6 years old (Koopmans *et al.*, 2018; Moore *et al.*, 2019; Denys *et al.*, 2021). Automated self-testing via internet-connected, commonly available devices, such as smartphones, lack of other equipment (e.g., sound booth, audiometer, or other specialist electronics), and general ease of use make this test suitable for several pediatric applications for which audiologist-led tests will not work or are unavailable. For example, in school settings, the DIN may be used as a rapid screen for hearing loss. In addition, the DIN can become essential in some rural or low- and middle-income countries where no other hearing healthcare is available. In these settings, innovative modifications of the standard DIN may be used to provide testing sensitized to high frequency hearing loss (Motlagh Zadeh *et al.*, 2019) or testing that enables segregation of routine hearing loss (i.e., mild/moderate bilateral sensorineural) from hearing loss requiring referral to a specialist (unilateral or conductive loss; De Sousa *et al.*, 2021). Although these behavioral tests and others (e.g., Table I), have the potential to provide much needed diagnosis from a young age, age-specific ranges of normal recognition performance are sometimes recommended (Wilson *et al.*, 2010), and these behavioral tests cannot provide a diagnosis in infancy when interventions are most needed.

D. Noise exposure and hearing loss in childhood

Incubators in neonatal intensive care units (NICUs), noise-generating sleep devices, and personal music players are three sources of potential NIHL in childhood that have received significant research attention. While numerous studies have pointed toward a link between noise levels inside infant incubators in the NICU and NIHL (Falk and Farmer, 1973; Monson *et al.*, 2020), no direct evidence for NIHL has been established (Wachman and Lahav, 2011). Recently, it has been suggested that modern, quieter

incubators may, in fact, reduce exposure to desirable language stimulation in the NICU (Monson *et al.*, 2020).

The output sound level of infant "sleep machines" was measured by Hugh *et al.* (2014), following concerns that sustained exposure may induce NIHL. Three of 14 machines tested produced noise >85 dBA at 30 cm, a level that if played continuously exceeds recommended adult occupational noise limits. As the authors suggest, such a level may overestimate that needed to produce NIHL in infants. However, there does not appear to have been any direct evidence that use of these machines does cause hearing loss in children.

Although widely publicized, the effect of music experience on childhood hearing remains unclear. The World Report on Hearing (WHO, 2021) recently suggested that more than 1×10^9 people are at risk for NIHL from listening to loud music. In contrast, Couth *et al.* (2020) found that early career musicians (18–27 years old), with an average 13 years of music experience and 15 h per week of practice, had sensitive mean tone thresholds of 0–5 dB HL across 0.25–8 kHz. They also had normal SiN SNR thresholds. Another study of college students found no statistical association between recreational noise exposure (predominantly music) and tone thresholds, or words-in-noise speech-to-babble ratios (Le Prell *et al.*, 2018).

E. Summary of pediatric speech testing and effects of noise

SiN testing is starting to take its rightful place as a gold standard in pediatric audiology. It has been found capable of providing additional information beyond the audiogram in neonatal electrophysiological testing and in behavioral testing in older children. As in other applications in adults, it provides a more holistic and applied measure of hearing. In children, it can provide information on history of hearing loss and evidence of processing problems in the brain, including those occurring beyond the conventional auditory system and in multimodal areas of the cerebral cortex. Finally, it can help delineate problems that are primarily sensory, and may thus be managed by some form of amplification, from those that are primarily cognitive, and may thus benefit from another approach such as pharmacological interventions. Despite reasonable concerns that neonatal incubators, sleep machines, and loud music can impair hearing in children, we currently lack convincing evidence that they do. One thing that is agreed upon, however, is that a rich environment of speech is needed from the earliest possible age to promote optimal language development in children.

III. ACCOUNTING FOR DIVERSITY IN SPEECH PERCEPTION IN NOISE TESTING FOR CLINICAL TRIALS

Understanding of sex- and gender-specific effects and of how the social constructs of race and ethnicity influence the outcomes of clinical trials have been recognized in the U.S. and other high-resource nations as research priorities (Alvidrez *et al.*, 2019; Clayton, 2018; Havinsky *et al.*, 2018; National Institute for Health Research, 2020; Shansky and Murphy, 2021). While frequently included as control variables in clinical trials investigating hearing loss, few studies have disaggregated results from appropriately sized samples of racial, ethnic, and linguistic minorities or included biological sex and/or gender as variables of interest.

A. Sex as a biological variable in hearing performance

In 2015, the National Institutes of Health (NIH) passed a regulation that consideration of sex (a biological classification, coded in DNA, which we refer to as "male" and "female"), including a compelling, scientific rationale for single-sex studies, must be included in proposals for funding consideration (National Institutes of Health, 2015). Further, the NIH also endorses the consideration of gender (socially constructed roles, expressions, and identities, which we refer to as "men" and "women") in health research, stating that studies that include data on both sex and gender provide the most complete picture on the influence of respective biological and social influences (NIH Office of Research on Women's Health, 2021).

Studies have demonstrated evidence of sex- and gender-specific differences in hearing ability, providing support for the argument of including sex and gender as variables of interest in hearing-related clinical trials. Sex differences in humans have been demonstrated at the biological level as revealed by electrophysiologic studies. For example, in a comparison of click-evoked ABRs, females displayed earlier peak waveform responses compared to males (Krizman *et al.*, 2012). Thornton *et al.* (2019) demonstrated sex-specific differences in sensorimotor activity, as measured by electroencephalography (EEG), in response to a speech discrimination task, suggesting that males and females utilize different cognitive strategies for speech perception. In animal studies of biomarkers, such as estrogen expression, estrogen has been found to be a protective factor against age-related hearing loss, suggesting that hormonal changes associated with female aging contribute to decreases in hearing sensitivity in older age (e.g., Simonoska *et al.*, 2009; Shuster *et al.*, 2019).

Less understood are how the social construct of gender and the intersection of sex and gender influence speech perception across gender identities. While sex and gender are different constructs, disentangling their individual effects is complex and has not been adequately explored in hearing-based clinical trials. Further confusion stems from a lack of consistency and conflation of the definitions of sex and gender in the extant literature, which is a significant limitation with regard to interpretation of findings. However, differences have been demonstrated in behavioral studies that ask participants to self-report "sex" or "gender," such as longitudinal panel surveys, likely reflecting this intersection. A longitudinal study using data from the Baltimore Study of Aging revealed a significantly faster rate of decline in pure-

tone hearing sensitivity in men compared to women, earlier onset of hearing loss in men compared to women, and frequency-specific differences in hearing sensitivity between men and women (Pearson *et al.*, 1995). In a cross-sectional analysis of Medicare beneficiaries participating in the Health Aging and Body Composition (Health ABC) study, Helzner *et al.* (2005) found that hearing loss was most common in White men, likely a result of history of occupational noise exposure. Likewise, nationally representative cross-sectional data of adults aged 20–69 years included in the 1999–2004 waves of the National Health and Nutrition Examination Survey (NHANES) demonstrated that men had significantly higher odds of bilateral, unilateral, and high-frequency pure-tone hearing loss compared to women {odds ratios [ORs] = 2.4 [95% confidence interval (CI): 1.7–3.5]; 2.0 [95% CI: 1.4–2.9], and 5.5 [95% CI: 4.0–7.5], respectively} (Agrawal *et al.*, 2008). Despite these known differences in hearing sensitivity, sex and gender biases are still evident in published hearing research, with data skewed toward studies focused on only or primarily male participants (Lauer and Schrode, 2017; Pittman *et al.*, 2021).

Few studies have investigated how SiN ability differs based on sex or gender. While frequently used as a covariate or confounder to determine whether interactions impact SiN ability, most previous works in this area do not stratify results based on sex or gender. However, some evidence exists that demonstrates differences in SiN perception and ability. In a study examining gender differences in acceptable noise levels among 25 men and 25 women, Rogers *et al.* (2003) found that men had a mean 6 dB higher comfortable listening level and tolerated a mean of 7 dB more background noise than women. A more recent study (Yumba, 2022) explored gender differences in SiN ability in new hearing aid users. Using the Hagerman matrix sentences test (Hagerman and Kinnefors, 1995) presented in a sound field, aided participants listened to a series of sentences under different hearing aid noise reduction settings. Findings revealed that women listeners had lower SNRs compared to men when noise reduction was activated at the 80% correct performance level but that men had marginally lower SNRs when noise reduction was activated at the 50% correct performance level. Taken together, results suggest a complex relationship between gender, background noise level, SNR, and noise reduction algorithms, which should be considered in future trials.

To foster consideration of sex and gender as variables of interest in clinical trials, several recommendations have been put forth for all aspects of study design, conduct, and interpretation and dissemination of results. The NIH provides guidance, including development of sex-specific hypotheses during study conception; consideration of effect size and power when determining the number of males and females to include in a study; determination of whether hormonal changes in the female estrous cycle are relevant for study conduct and interpretation of results; and disaggregation of results on the basis of sex, which it poses to contribute to future meta-analyses (National Institutes of Health, 2021). In terms of dissemination of results, the Sex and Gender Equity in Research (SAGER) guidelines provide both general and section-specific recommendations for article writing (summarized in Table II) (Heidari *et al.*, 2016). These recommendations should be adopted by hearing researchers who conduct clinical trials when designing studies and interpreting and disseminating their results.

TABLE II. SAGER guidelines. Table content was adapted from Heidari *et al.* (2016).

Article section	SAGER recommendation
Title and Abstract	Title and abstract should specify if only one sex or gender was under investigation.
Introduction	Hypothesized sex and/or gender differences should be stated if applicable.
Method	Detailed information regarding how sex and/or gender were considered in the study design should be included. If sex and/or gender were not considered, a scientific or theoretical justification of why should be included.
Results	Results data should be stratified by sex and/or gender if applicable.
Discussion/ Conclusion	The potential implications of sex and/or gender (or the lack of consideration thereof) should be included in the Discussion. Lack of consideration of sex and/or gender should be identified as a limitation, if not included in the study design.

B. Underrepresented and linguistically diverse populations

In addition to understanding of sex- and gender-specific differences, investigation of race and ethnicity and how these factors relate to outcomes is needed. As language(s) spoken is inherently tied to ethnicity (Fought, 2011), we include in this group linguistic minorities, such as those who are monolingual non-English speakers (as the majority of work in this area focuses on English speakers and stimuli), as well as multilingual speakers. There is an overall lack of diversity in terms of representation of racial, ethnic, and linguistic minorities in clinical trials of hearing loss. One aim of a recent systematic review by Pittman *et al.* (2021) was to describe the racial/ethnic representation in U.S.-based clinical trials of hearing loss management. Results revealed that of 125 studies meeting inclusion criteria, only 16 studies reported race/ethnicity of participants, and of those that did, only five studies included 30% or greater participants from racial or ethnic minority backgrounds. This is particularly unsettling considering the relatively large contribution language(s) spoken and understood has on speech understanding for bilingual listeners (e.g., Cutler *et al.*, 1989; von Hapsburg and Peña, 2002; Rogers *et al.*, 2006), who are likely not represented in U.S. studies that focus on primarily White, non-Hispanic samples. For example, in the U.S., the two largest linguistic minorities are also from racial and ethnic minorities: Spanish speakers from Hispanic/LatinX backgrounds and Chinese speakers (Mandarin and Cantonese) from Asian backgrounds (U.S. Census Bureau, 2015), of whom approximately 2.9×10^6 are bilingual (U.S. Census Bureau, 2015).

There are several measurement factors unique to bilingual listeners to consider when evaluating speech perception, and unlike sex and gender influences, the effects of bilingualism are well represented in the extant literature. While a comprehensive review of these factors is outside the scope of this paper, briefly, they include age of acquisition of first and subsequent language(s), length of language immersion, language use (how much and in what capacity each language is used), listening proficiency in each language (self-rated), language dominance (relative competency of the two languages), and language of the speech perception test stimuli (Carlo, 2009; Kilman *et al.*, 2014; Shi and Sánchez, 2010). For bilingual speakers who are fluent in multiple languages, performance is similar to monolingual listeners in quiet, but the impact of the measurement factors above introduces the greatest amount of variability on speech perception in noise tasks.

Generally, bilingual listeners demonstrate worse performance on speech perception tasks compared to monolingual speakers, even when bilingual speakers appear to perform well in both languages (termed the "monolingual advantage"; Bialystok *et al.*, 2003; Shi and Sánchez, 2010). This advantage is magnified when assessing speech perception in noise and increases with worsening SNRs (i.e., the difference in performance between monolingual and bilingual speakers is greater, with monolingual speakers performing better at poorer SNRs) (Rogers *et al.*, 2006; Tabri *et al.*, 2011). It should be noted that listeners who report early second and subsequent language acquisition have better performance in their non-native language than late listeners both in quiet and in noise (Mayo *et al.*, 1997; Shi and Sánchez 2010; Tabri *et al.*, 2011). Thus, these factors should be considered when including bilingual participants in any hearing-related clinical trial where speech perception ability is evaluated.

There are multiple options for selecting appropriate speech perception materials for use in clinical trials that include bilingual participants. Cross-culturally adapted word and sentence lists exist for testing in quiet and in noise in Spanish, Chinese, and other languages. Psychometrically validated speech audiometry materials [speech recognition threshold (SRT); word recognition in quiet] are commercially available in Latin American Spanish (Carlo *et al.*, 2020) and Cantonese (Nissen *et al.*, 2011). Further, the use of English-language digits to obtain SRTs in non-English speakers familiar with the numbers has been shown to produce SRTs equivalent to native English speakers in those with normal hearing (Ramkissoon *et al.*, 2002). Several SiN materials are also available in multiple languages, including those that have undergone a cross-cultural adaptation process to ensure that they are linguistically and culturally relevant, such as the HINT, available in Latin American Spanish, Cantonese, and Mandarin (Soli *et al.*, 2002; Wong and Soli, 2005; Wong *et al.*, 2007), and the WIN test, available in Latin American Spanish (Carlo, 2009). DIN tests are an option for materials that do not rely heavily on linguistic content; however, poorer performance for non-native bilingual participants when tested in English has been demonstrated, so caution should be used when interpreting results (Marinova-Todd *et al.*, 2011).

IV. FUNCTIONAL SIN TESTING

A. Speech recognition in the work environment

Noise-induced hearing loss is the most prevalent, irreversible worldwide occupational hazard (Smith, 1998). In the United States, approximately 16% of all workers experience hearing difficulty [see Themann and Masterson (2019), Chen *et al.* (2020), or Lie *et al.* (2016) for recent reviews of occupational noise exposure and its effects on hearing]. In the U.S. military, hearing loss and damage was the most prevalent disability for soldiers returning from Iraq and Afghanistan (Le Prell and Clavier, 2017; Institute of Medicine of the National Academies, 2006; Yankaskas, 2013; Grant *et al.*, 2021). Hearing loss affects a person's ability to hear and recognize speech, which in turn leads to communication handicaps (Le Prell and Clavier, 2017). Difficulty hearing speech will affect a worker's ability to communicate and understand commands and tasks. In some industries and for some job functions, the ability to hear sounds and speech may be essential to performance. For example, communication over the telephone or a radio is hearing-critical (HC) because all information is obtained through the hearing modality. In some cases, other modalities (e.g., vision, tactile feedback) are effective in supplementing hearing, and these tasks no longer qualify as HC (e.g., a flashing light that always accompanies a warning tone). In other instances, these modalities only serve to support the hearing component, and the task remains HC (e.g., lip-reading). Furthermore, HC tasks may take place in noisy and distracting environments. These environments make the hearing tasks much more difficult for hearing-impaired listeners who may have difficulty not only with hearing complex signals, such as speech, but also processing the speech against a background of masking noise. In addition, SRTs may vary depending on whether the speech and noise are delivered monaurally or binaurally (Soli and Vermiglio, 1999).

In an early investigation on the impacts of reduced speech intelligibility on worker performance, researchers at the Department of Defense (Garinther and Peters, 1990) showed that well-defined gunnery task performance degraded significantly with reduced speech intelligibility, despite a limited vocabulary set. In a more recent but similar experiment (Keller *et al.*, 2017; Sheffield *et al.*, 2017), sailors were asked to perform a number of specific tasks in a command and control center environment where speech intelligibility was systematically degraded by reducing the SNR. The results showed degraded performance on both objective and subjective measures as the SNR decreased. These experiments, conducted in environments that could be controlled and manipulated, demonstrate that speech intelligibility, which is affected by both hearing loss and the noise environment, can impact workers' performance.

B. Auditory fitness for duty (AFFD)

AFFD is defined as the possession of hearing abilities sufficient for safe and effective job performance (Tufts *et al.*, 2009). AFFD is also referred to as functional hearing screening or hearing fitness for duty. AFFD tests are designed to measure one's ability to detect, recognize, localize, and understand complex, real-world signals. Certain jobs, particularly those involving arduous duties or a high standard of human reliability, such as military personnel, law-enforcement officers, and first responders, require sufficient functional hearing ability to perform HC tasks that cannot be accomplished without the use of the hearing modality. In some settings, managers and occupational health providers need to know whether and how a hearing impairment might affect the ability of workers to perform HC duties in their environment. However, there is currently no standard accepted system in place for measuring or predicting functional hearing performance in the workplace. The current metric used in hearing conservation programs, which capture a large portion of individuals working in noisy environments, is the audiogram. However, pure-tone audiometry does not provide an accurate measure of functional hearing ability. From the audiogram, the speech intelligibility index (SII; ANSI, 2017) can be calculated. Developed as a predictor of speech recognition, the SII, however, is limited to stationary noise environments and does not provide specific information on the impact the individual's specific hearing loss may have on job performance. In the last 20 years, several studies have attempted to address this issue, and some tests have emerged as potential measures of functional hearing ability.

Giguère *et al.* (2008) developed functional hearing assessment tools and protocols to screen personnel in HC tasks as part of their job with the Canadian Coast Guard and Conservation and Protection sections of the Department of Fisheries and Oceans Canada (DFO). Their approach followed seven steps: (1) they identified hearing requirements and measured the noise environments for jobs with HC tasks; (2) they selected potential screening measures that could assess functional hearing abilities; (3) they studied how performance on these screening measures related to performance in real noise environments in normal hearing individuals and (4) in hearing impaired individuals; (5) they validated their model of prediction of functional performance with subject matter experts (SME); (6) they applied the model to establish functionally based criteria for HC jobs; and (7) they validated the functionally based criteria with DFO incumbents. Theirs was a very detailed modeling approach, combining speech production and perception parameters to evaluate workplace communication requirements (Laroche *et al.*, 2005). The main downside to their approach is that it was extremely labor intensive in order to validate the tools and establish acceptable requirements for the DFO. Their findings were confined to workers from this particular occupation.

A similar approach was used by Soli *et al.* (2018a) to assess functional hearing requirements in California prison guards (first study), in United States Army occupations (second study), and with police officers in Ontario (third study). One objective of these studies was to develop a way to predict performance in relevant, real-world noise environments using the SRT and the characteristics of the real-world noise. For this purpose, they used the extended speech intelligibility index (ESII; Soli *et al.*, 2018b; Rhebergen *et al.*, 2006; Rhebergen *et al.*, 2014), which was developed for the analysis of non-stationary noise. The ESII is calculated with the noise spectrum level analyzed in "snapshots" instead of the long-term average noise spectrum with which the SII is calculated. Forward-masking threshold was also included in the ESII to account for speech performance in gated noise (Rhebergen *et al.*, 2006). In all three studies reported in Soli *et al.* (2018a), the researchers aimed to predict performance in real-world noise using the SRTs from the HINT and the ESII calculated with the real-world noise samples, and they attempted to demonstrate the feasibility by then measuring the speech intelligibility of the subjects in real-world noise. First, they used the HINT to obtain a speech intelligibility performance curve for each subject by measuring word intelligibility at three different SNRs, including their SRTs (50% intelligibility). Next, they computed the ESII of their real-world noise samples. The performance curve yielded a performance slope that is usually relatively linear at intelligibility levels between 30% and 80%. Using this performance curve, they mapped the performance intensity function (intelligibility versus SNR) to intelligibility versus ESII for each real-world noise with representative equivalent continuous level (LAeq) values sampled from calibrated recordings. Finally, they measured the actual ESII-intelligibility performance intensity function for each subject and compared the prediction to the measured function. All three studies showed good agreement for normal and impaired hearing subjects, though sample sizes were small for all three studies.

Another study by van der Hoek-Snieders *et al.* (2021) investigated the validation of a signal detection test to assess AFFD in locomotive engineers in the Netherlands. The authors developed a task- and job-specific test to assess the engineers' ability to detect two acoustic warning signals in Dutch train cabins. Their study assessed the correlation between audiometric thresholds, SiN test performance, and the SNR at which subjects detected the target signal 50% of the time. In general, they found moderate agreement with the conventional tests, demonstrating that the signal detection test, a job-specific test with realistic targets and noise levels, provides additional insight into an individual's ability to perform their job function. Once again, this was a test developed for a very specific occupation.

Most recently, researchers at the Walter Reed National Military Medical Center and at the Hearing Center of Excellence attempted to develop a new, closed-set SiN test that could be used to predict performance in relevant occupational noise environments, using an approach slightly different from that presented in the studies above (Brungart *et al.*, 2021). They developed a clinical version of the

https://doi.org/10.1121/10.0014418

Modified Rhyme Test (MRT) with two 80-word lists with a design that incorporates two speech levels, two SNRs, and two binaural conditions. This new Military Occupational Hearing Test is now being validated in a study comparing performance on the test to performance in a speech recognition task involving audio recordings made during training exercises with real-world speech communication and environmental noise accompanied by a photo depicting the scene. Subjects are asked multiple choice questions related to the auditory scene as a way to assess their ability to recognize speech in a typical training environment.

The coordinate response measure (CRM; Brungart, 2001; Moore, 1981) is another proposed AFFD assessment. The test requires listeners to identify a number and color from sentences that start with a specific call sign. Eight talkers speak at the same time, each with a different call sign, color, and number. This test has the potential for good face validity for some job functions, such as pilots, since it is representative of the types of target and maskers that may be present in real scenarios. Semeraro *et al.* (2017) have developed a version of the CRM in British English, with speech-spectrum noise as the masker, as a potential AFFD test for British military personnel. In their study, they found good test-retest reliability and sensitivity to hearing impairment.

The development of suitable AFFD tests is complex because it needs to be both relevant to a particular job function and also standardized to ensure reliability and consistency across listeners with varying degrees of hearing loss. Furthermore, these tests need to show some correlation with performance on the job to differentiate workers whose impairment directly affects their ability to complete job-critical auditory tasks. This is a very active area of research for both military and civilian applications. Several tests have been proposed or evaluated for AFFD, including the HINT (Soli and Wong, 2008), the Speech Recognition in Noise Test (SPRINT) used in the U.S. Military for AFFD (Brungart *et al.*, 2017), CRM (Semeraro *et al.*, 2017), and the MRT. The first two require administration by an audiologist or trained administrator who scores the words or sentences recognized by the listener, while the last two can be easily self-administered through an automated test paradigm. Self-administered tests are advantageous for large scale deployment through computer- or tablet-based delivery, which also allows for better randomization. Additional research is still needed to validate test performance against job performance for both normal hearing and impaired listeners.

V. SIN TESTING IN CLINICAL TRIALS

Novel approaches for treating hearing loss are being developed given the high prevalence of sensorineural hearing loss and the negative impact that untreated hearing loss has on personal, socioeconomic, and public health. Outside of rehabilitation options, such as hearing assistive technologies, there are no available medicinal treatments for permanent hearing loss. Recent traction in drug discovery and delivery has significantly contributed to the state of hearing and pharmaceutical science. Significant resources are being allocated to the pursuit of more treatment options for those with hearing loss. In a 2019 review paper, there were more than 40 pharmaceutical and biotechnology companies developing inner ear therapies and potential drug delivery systems [for review and meta-analysis, see Schilder *et al.* (2019b)]. There are 23 therapeutics in clinical trials and over 56 potential treatments in the preclinical pipeline (Isherwood *et al.*, 2021). Recent advancements in biotech discovery have made it possible to conduct first-in-kind clinical trials to evaluate potential pharmacotherapies to address hearing loss (Sanchez, 2018; Schilder *et al.*, 2019a). Thus, only recently has there been enough advancement in drug discovery and delivery to have clinical trials where speech perception measures would be appropriate as study outcomes. As potential pharmacotherapeutics advance through the discovery pipeline, speech perception measurements can be used, in part, to evaluate the safety and efficacy of these investigational treatments.

While it is essential to assess both safety and efficacy as potential pharmacotherapeutic advances through the pipeline toward the market, the evaluation process is considerably different when seeking safety as compared to efficacy outcomes. Determining the safety and efficacy of a pharmacotherapeutic can be influenced by the measurement selected and the performance on that measurement. When evaluating speech recognition performance, a full view of the psychometric function and a detailed look at the most linear portion of that function should be attempted. As stated by Egan (1948), speech perception performance is not "uniformly sensitive" over the range of possible performances. Figure 2 provides an example of psychometric functions for a SiN task where performance is plotted as a function of the SNR. The most linear portion of the psychometric function is likely where variable manipulations will have the most dramatic effects, allowing for valuable insights into the data set. Near the 50% point of the performance function is the most linear and where performances are most variable, contrasted to the extremes of the function, where

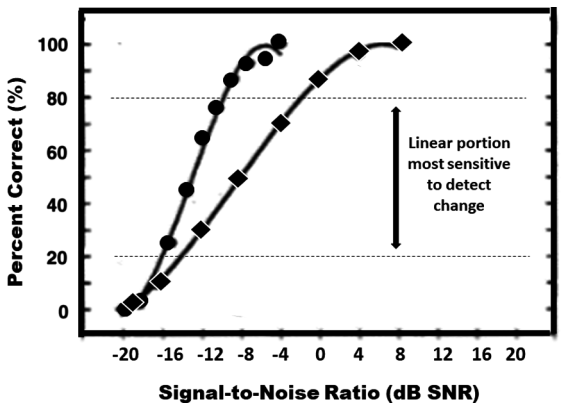

FIG. 2. Example psychometric functions highlighting the most linear portion of the function between the floor (bottom dashed line) and ceiling (top dashed line).

performances are least variable because the measurement can be too easy (ceiling effect) or too difficult (floor effect) (Egan, 1948; Wilson and Margolis, 1983). Selecting a test that provides a full psychometric function, like a SiN test at multiple SNRs, can play a valuable role in both safety and efficacy evaluations. Use of SiN testing is advocated for in clinical trials because it is considered the "stress test" of the auditory system (Wilson, 2011). Such a stress test allows for performance challenges and a full psychometric function to provide more insight into changes caused by a therapeutic agent.

A. SiN tests for clinical trial safety evaluation

In early phase clinical trials, when safety is the primary outcome, there may be a need for SiN testing. SiN measures can be used to identify and/or monitor any reduction in performance that may be considered a drug-related adverse event. Some drugs are known to be ototoxic; therefore, assessing for any changes to hearing, balance, and/or tinnitus is required as part of an ototoxicity monitoring protocol. Current monitoring guidelines, such as the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE; National Cancer Institute, 2009), American Speech-Language-Hearing Association (1994), and others (Brock *et al.*, 1991; Chang, 2011), promote pure-tone audiometry for monitoring. However, although pure-tone audiometry is often considered the standard for ototoxicity monitoring, SiN testing may prove to be more sensitive at identifying functional changes in hearing as compared to a traditional audiometric battery (Konrad-Martin *et al.*, 2018; Humes, 2019). It is not the purpose of this review to discuss the pitfalls of ototoxicity monitoring, but rather to highlight that SiN tests may allow more sensitive measures to monitor possible adverse events in clinical trials. Furthermore, with advancements in remote testing, it may be possible to obtain SiN performance outside of the traditional audiology booth testing (see Sec. VI), allowing for additional flexibility on the part of the study design, facility resources, and patient compliance (Konrad-Martin *et al.*, 2021). The use of valid and automated measures testing, such as DIN tests [for a review, see Van den Borre *et al.* (2021)] or the Portable Automated Rapid Testing (PART; Gallun *et al.*, 2018), can be efficient tools and very informative regarding functional changes. The added flexibility of automated SiN testing could prove to be a boon in clinical trials outside of those solely concerned with hearing. For example, there are many oncology clinical trials that fail to measure hearing function due to the inability to obtain an audiology assessment, especially at baseline before the ototoxic agent is delivered. Automated SiN tests would allow for assessment without having to add on qualified study personnel to complete audiological examinations and therefore contribute to a more global understanding of pharmacological safety.

B. SiN tests for clinical trial efficacy evaluation

Just as SiN testing can be valuable in clinical trial safety evaluations, SiN testing can also contribute significant value to efficacy evaluations of pharmacotherapeutics. In the last ten years, multiple pharmacotherapeutic agents designed to address hearing loss or associated difficulties have progressed to Phase II clinical trials (Isherwood *et al.*, 2021). Despite the growing number of trials and the significance of this research, there is no gold standard SiN assessment recommended by the Food and Drug Administration (FDA) as a primary or secondary outcome. Evidence of improved SiN performance in a clinical trial would indicate functional improvements that can be argued to be more ecologically valid beyond basic assessments such as pure-tone audiometry and speech-in-quiet performance. Although caution should be taken if using an unvalidated novel measurement in a clinical trial, novel and innovative uses of already available and validated instruments can be considered. Pharmacotherapeutic approaches for hearing are developing rapidly as the biotech field advances, requiring creativity and flexibility when designing evaluative protocols for new pharmacotherapeutics using existing testing measures. Thus, the novelty here lies in the use of SiN tests to appropriately measure desired outcomes instead of innovations of the measures themselves. By providing an overview of SiN testing in clinical trials, including trials that are anticipated to begin soon, are in progress, or have been completed, additional perspectives are provided on the current state of science. Selecting the best SiN assessment to utilize in a clinical trial can be difficult and requires careful consideration of the primary and secondary outcomes of the trial. The test selected should be sensitive to the drug mechanism of action, the trial design, and the patient population evaluated.

C. Selecting a SiN test with respect to the drug mechanism of action

The SiN assessment selected for a clinical trial should be sensitive to the drug mechanism of action. Through advancements in multiple areas, including microscopy, molecular biology, and biochemistry, scientists and researchers have been able to add to the body of knowledge surrounding the cellular and molecular mechanisms of hearing loss, as well as the associated signaling pathways. These scientific advances allow for the identification of targets that could be influenced by pharmacotherapeutics. These targets are the focus of drug development, thus leading to target indications or a drug's mechanism of action. The drug's mechanism of action intends to fix, repair, reproduce, or enhance a specific biomedical element, such as proteins, neurotransmitters, or structure functionality. Understanding the target indications provides some of the necessary framework for selecting a SiN assessment that will be needed to measure the desired efficacy outcome appropriately.

To exemplify, therapeutic mechanisms of action that influence the peripheral auditory system would likely benefit from assessments that are less influenced by higher-up processing or cognitive influence. A potential choice would be a word-level test that would task the peripheral fine-coding instead of a sentence-level test that could be influenced by contextual cues. The WIN test (Wilson, 2003) has been shown to be sensitive at quantifying hearing performance

with fewer cognitive and linguistic influences (Wilson *et al.*, 2007). The WIN test is perhaps one of the most extensively studied SiN tests with high intra- and inter-session test-retest reliability (Wilson, 2007). The WIN test is composed of monosyllabic words from the NU-No.6 corpus (Tillman and Carhart, 1966) presented in multi-talker babble and provides an open-set word-recognition task without linguistic context. The 50% point, which is calculated with the Spearman–Kärber equation (Finney, 1952), is defined in terms of SNR required for 50% recognition. There are several clinical trials completed or presently enrolling participants where the WIN test is an outcome measure (see examples from ClinicalTrial.gov identifiers NCT041120116, NCT04129775, NCT04462198).

Although the WIN test is advocated as an appropriate test in some clinical trials, the WIN stimuli have a limited frequency spectrum. This may be a concern if a regenerative therapeutic was likely to influence the ultra-high frequency regions in the cochlea due to drug delivery and absorption in the most basal cochlear regions. If the drug mechanism of action were to influence this area of the cochlea, speech testing materials with extended high-frequency (EHF) stimuli should be used as they would better reveal treatment-related effects. EHF speech cues provide advantageous cues, and listeners can, on average, detect the loss of speech spectral energy beyond 13 kHz (Hunter *et al.*, 2020). EHF energy serves as a salient cue to determine the direction a talker is facing, an ability proposed to be important for determining whether one should attend to a speech message, allowing for detection and segregation of the speech to be heard (Monson *et al.*, 2020). The presence of EHF spectral energy has been shown to enhance speech perception in noise, in part due to its detection and segregation contributions (Motlagh Zadeh *et al.*, 2019). Therefore, it follows that we should expect that listeners who are provided a treatment that improves EHF hearing will benefit from the spectral energy provided in this range. There are a variety of tests that include EHF energy, such as some of the DIN test options (Motlagh Zadeh *et al.*, 2019). There remains a challenge to determine how and in which conditions perception of EHF energy supports speech perception and, further, whether a pharmacotherapeutic agent can improve EHF hearing and ultimately speech perception. However, if the drug mechanism of action suggests improvements in EHF, then the SiN test selected should consider these spectral cues.

Other SiN tests could be considered if the drug mechanism of action does not affect the auditory periphery but instead has an influence on auditory processing beyond the cochlea. Pharmacotherapeutics that may influence auditory processing or cortical processing may be more adequately assessed by using materials that assess more complex processing, such as spatially separated SiN, auditory working memory tests, or a dual-competing task paradigm. An example of such a test that allows for evaluating the differential contributions of peripheral and central hearing is the aforementioned LiSN-S (Cameron and Dillon, 2007; Dillon *et al.*, 2014). The LiSN-S uses binaural listening in a variety of situations with shifting speaker location, voice cues, and varying contextual cues and determines how these cues influence a listener's ability to understand SiN. A total of five different scores are generated using the performance obtained with the varying test conditions. For information regarding norms and more in-depth discussion of protocol and scoring, see Cameron *et al.* (2011). The LiSN-S test was one of many assessments selected for a clinical trial conducted for patients with age-related hearing loss (ClinicalTrials.gov identifier NCT02345031), where the drug mechanism of action was intended to enhance neural synchrony and auditory processing. In addition to the LiSN-S, the primary outcome measure for the above-mentioned trial was QuickSIN (Killion *et al.*, 2004). The QuickSIN utilizes sentences presented against increasing four-speaker babble to calculate a SNR loss, or dB SNR loss. The SNR loss score demonstrates the amount of difference in dB between the speech signal and competing noise for the patient to be able to understand the target stimulus compared to normal listeners. Our current thinking is the use of multiple SiN tests can be informative for cross-checking and confirmation of treatment effects.

D. Selecting a SiN test with respect to trial design

Many times, clinical trials would benefit from research or lab-based measurements that require sophisticated equipment and/or protocols. In the instance of multi-site studies, these measurements are often not feasible. For example, studies that involve a potential investigational medicinal product that may influence spatial hearing or speech understanding in multidirectional noise would benefit from a calibrated array of speakers 360° around the participant. However, array speakers are not feasible or available in most clinics or centers and thus become impractical as an outcome measure for larger, multisite trials. Specialized, and often expensive, equipment is only part of the issue posed by the benefit of sophisticated auditory assessment. Access to sufficient audiological support and lack of facilities that can act as appropriate measurement sites, such as a sound booth, can also create issues for site staff and clinical researchers.

Conveniently, recent developments have produced measurement tools that mimic the free-field environment, allowing multi-site studies access to assessments they may otherwise be unable to complete. Innovative testing software and technology that allows for refined testing protocols without additional organizational and practical challenges have broadened the availability of SiN testing to many multisite or small studies without access to the funds or space to host the required technology. Examples of such tests include the LiSN-S and PART tests already introduced. These computer-based tests, administered via calibrated headphones, allow for administration without a sound booth, speaker array, or even the immediate presence of an audiology clinician. Of course, there are limitations to such tests; for example, they are best used in instances of systemic drug administration, where a binaural evaluation is appropriate. Conversely,

studies that use unilateral intratympanic injections likely would need to utilize monaural measures.

A clinical trial designed to evaluate restoration after NIHL or prevention of NIHL may need SiN testing that is available for administration in the field or when the noise exposure recently occurred. Examples of testing in the field or remotely are described elsewhere in this review. A recent report indicated three clinical trials are currently being conducted for NIHL prevention (Isherwood *et al.*, 2021), and a few of these NIHL clinical trials report the use of SiN testing as an outcome measure (Le Prell, 2021). Other reviews in the literature have also advocated for SiN tests and suprathreshold auditory evoked potentials as metrics monitoring efficacy in NIHL clinical trials [see Le Prell and Brungart (2016) and Le Prell and Lobarinas (2015)].

E. Selecting a SiN test with respect to the patient population

Patient population is another factor that can influence SiN performance as well as dictate the type of SiN evaluation that is appropriate for the given target indications of the trial. Mentioned above, a recent systematic review by Pittman *et al.* (2021) sought to describe the racial/ethnic representation in U.S.-based hearing-related clinical trials. Results revealed there is a lack of diversity in clinical trials, and the inclusion of racial/ethnic and sex diversity will contribute to the advancement of effective treatments, informative to advancement of pharmacogenetics, and improve hearing health equity.

While it may require additional resources and complicate study design, the importance of including varied sex and race participants in a clinical trial cannot be overstated. There are multiple SiN tests available in several languages to allow for more diversity in clinical trials. Testing materials available in multiple languages are necessary when studying a diverse population, as is awareness of cultural and socioeconomic factors. Furthermore, when clinical trials reach Phase III or the market advancement stage, the therapy may need to be evaluated across the world in multiple countries; thus, the availability of assessments in multiple languages is a necessity. One such SiN test that is available in 14 different languages is the Matrix test, although test performance differs by language (Kollmeier *et al.*, 2015). The American English version of the Matrix test (AEMT) is a sentence-based SiN test that uses five-word sentences presented in masking background noise. Each sentence is comprised of the same structure, i.e., name, verb, number, adjective, and noun. The sentences are grammatically correct but semantically unpredictable, therefore making them less likely to be correctly guessed if not heard properly. The Matrix is administered adaptively, and the SRT is determined by averaging the SNR at 50% correct performance for 20 presented sentences.

Other patient factors to be taken into consideration are age, language fluency, cognitive abilities, and type, degree, and configuration of hearing loss. Age of participants affects both the duration and difficulty of tests. When evaluating a pharmacotherapeutic designed for children, age and developmental stage must be considered to avoid both floor and ceiling effects. Cognition and the ability to provide an appropriate response require similar considerations. Finally, the type, degree, and configuration of hearing loss can have a significant effect on testing measures selected and expected outcomes. For example, when assessing patients with cochlear implants, the Minimum Speech Test Battery (MSTB; Nilsson *et al.*, 1996) has been vetted for that population and has been established as the gold standard for evaluating the auditory status of these patients. The MSTB consists of recorded consonant-nucleus-consonant (CNC) words (Peterson and Lehiste, 1962) and subsets of the AzBio (Spahr and Dorman, 2004) and Bamford–Kowal–Bench Sentence-In-Noise (BKB-SIN; Etymotic Research, 2005; Bench *et al.*, 1979) tests. The MSTB recording format allows the speech and noise to be presented from separate loudspeakers in the sound field at different (fixed) presentation levels. When used in a clinical trial to assess the hearing status of cochlear implant patients (ClinicalTrials.gov identifier NCT02832128), the MSTB allowed for SiN testing that could be referenced back to other validated work from patients with cochlear implants (Fabry *et al.*, 2009; Gifford *et al.*, 2008).

F. Final thoughts on SiN testing in clinical trials

SiN testing has a growing presence in clinical trials. Test measures are often listed on ClinicalTrials.gov, allowing for a review of previous selected trial measurements that can be useful during the selection process for future trials. Not only is the selection of a test important, but critical properties for that selection include how the test was used, how the test was presented (e.g., monaural versus binaural, presentation level, number of stimuli influencing statistical power), and the expected patient response. These are some of the factors that should be considered when identifying an appropriate SiN test for a clinical trial. For a further review of critical speech perception test properties and administration considerations, see Theunissen *et al.* (2009). Even with careful selection of testing protocols and parameters, there are limitations to SiN tests in clinical trials. Primarily, these measures require a behavioral response, and little is known about the amount of effort required to generate the observed performance even when proper administration is used with considerations for the therapeutic mechanistic action, trial design, and patient populations. Therefore, pairing correlated measures of speech perception performance with more objective measures, such as autonomic indices of listening effort evaluations (e.g., pupillometry) and/or physiologic functioning of the cochlea and auditory processing tracts (e.g., auditory evoked potentials), could help contextualize patients' behavioral performances. SiN testing likely best reflects the hearing difficulties that prompt patients to present to clinics or participate in clinical trials; however, it does not help identify the underlying pathology. Combining behavioral performance and objective physiological

https://doi.org/10.1121/10.0014418

measurements will enable an understanding of both the degree of communication difficulty experienced by the listener and the possible connected pathology, thus, allowing clinicians and researchers to better interpret results for functionality and pathophysiology.

The above review of considerations for the use of SiN tests in clinical trials provides a valuable resource for scientists working in drug development, industry partners, and clinicians, such as ENT physicians and audiologists. Clinicians have an important role to play in educating patients about the current state of novel therapies, understanding the measurements used in the clinical trial that determined safety and efficacy, and partnering with other scientists or industry partners in the development and evaluation of promising novel interventions. Several hurdles must be overcome before pharmacotherapeutics will be available to those with auditory disorders, including a better understanding of the pathophysiology of hearing disorders, diagnosing and monitoring patients, successful delivery of the therapeutics, and evaluating these potential therapies with the appropriate outcome measures as discussed here.

VI. ONLINE AND MOBILE APPLICATIONS

While SiN testing, despite its compelling rationale as detailed in this paper, does not appear to enjoy widespread utilization in traditional clinical settings, SiN testing is being increasingly utilized in online and mobile hearing testing spaces, which have proliferated due to the combination of a growing mobile health industry and the ubiquitous nature of personal computers (PCs), tablets, and smartphones. The following review of online and mobile applications of SiN testing should be of particular interest to clinicians as they attempt to navigate the growing numbers and varieties of remote testing options. Online and mobile hearing testing serve several purposes: to provide ready access to hearing health facilities for underserved populations (Visagie *et al.*, 2015; Sandström *et al.*, 2016; Swanepoel, 2020); as a method for hearing aid manufacturers to identify potential consumers; for direct-to-consumer sales of hearing aids (e.g., Eargo, Lively, Bose); and, recently, as a safe alternative to in-person testing during the COVID-19 pandemic. Self-administered SiN testing has also been developed for occupational noise identification purposes (Leensen *et al.*, 2011a; Leensen *et al.*, 2011b; Leensen and Dreschler, 2013a; Leensen and Dreschler, 2013b; Sheikh *et al.*, 2017). Shown in Fig. 3, online SiN testing applications vary as a function of the delivery platform (e.g., PC, tablet, mobile device, mail and return systems), the stimulus (e.g., phonemes, words, sentences, environmental sounds, numbers), the test procedure (e.g., fixed levels, adaptive algorithms), the output display (e.g., audiogram, narrative, score, scale), and recommended action (e.g., contact provider, purchase device).

The nature of SiN tests makes them particularly applicable for mobile applications as the score, depending upon the specific scoring algorithm, is based on the SNR, or intensity at which 50% of the target stimuli, delivered at the user's comfortable listening level, are correctly identified, eliminating the need for calibrated stimuli or a sound-treated enclosure. Indeed, recent innovations in "boothless audiometry" [e.g., KUDUwave, Wireless Automated Hearing Test System (WAHTS)] have enabled remote hearing testing to be conducted in less than optimal environments or when preferred clinical resources are unavailable or impractical (Gates *et al.*, 2021).

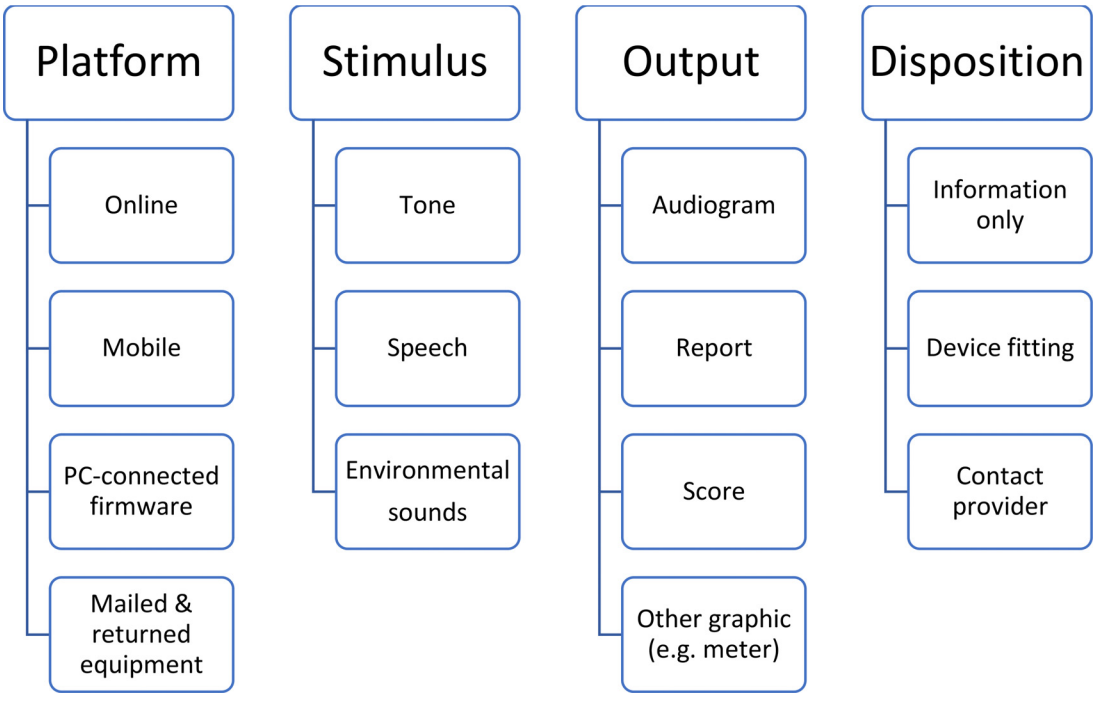

FIG. 3. (Color online) A suggested taxonomy for online/mobile hearing testing options.

A. DIN test

Among the available options for mobile or PC-based SiN testing, the Digit Triplet Test (DTT) and previously mentioned DIN algorithms have gained considerable attention as digits provide a closed-set choice of targets and minimize the semantic contributions to understanding target stimuli. While there are several variations to the DIN and DTT, the test design usually consists of randomly presented numbers, often three at a time, presented in a background of continuous or speech babble noise. The user's task is to identify the digits on a virtual keypad presented on the PC or smartphone screen. In the adaptive version of the test, if the numbers are correctly identified, another three digits are presented at a lower speech or higher noise level (i.e., lower SNR); if incorrectly identified, the level of the speech is increased or level of the noise decreased (i.e., higher SNR). The sequence continues until a stopping criterion is achieved (e.g., the SNR at which 50% of the triple-digit combinations are correctly identified after some predetermined number of presentations).

DIN testing traces its origin to a national telephone hearing screening program in the Netherlands developed by Smits and colleagues (Smits *et al.*, 2004; Smits *et al.*, 2006) and later adopted by a number of other countries. The American English version was developed and described by Watson *et al.* (2012). At a fail criterion of >20 dB HL PTA at 500, 1000, and 2000 Hz, the specificity and sensitivity of the American English version were 0.80 and 0.83, respectively. The correlation between the PTA and the DIN score (SRT threshold) was $r = 0.74$. A follow-up validation study of the American English version of the DIN telephone screening test (adopted as the U.S. National Hearing Test; NHT) was conducted on a large (1379 ears) veteran population (Williams-Sanchez *et al.*, 2014). At a fail criterion of >25 dB HL PTA at 500, 1000, and 2000 Hz, the specificity and sensitivity of the NHT were 0.87 and 0.54, respectively. When 4000 Hz was added to the PTA, the sensitivity and specificity values were 0.81 and 0.65, respectively. A PC version has been validated by Folmer *et al.* (2017), who assessed DIN scores, pure-tone thresholds, and self-perceived handicap as measured by the Hearing Handicap for the Elderly (HHIE; Newman *et al.*, 1990) among 40 community dwelling adults. Their sensitivity and specificity analysis revealed the area under the curve (AUC) as 0.96, 0.97, and 0.97 for both ears, the right ear, and the left ear, respectively. The data also indicated a positive correlation between the PTA average at 0.5, 1, 2, and 4 kHz and the DIN test scores (Pearson $r = 0.86$ for the right ears and 0.83 for the left ears). There was also a positive correlation between the DIN and the Hearing Handicap for Adults (HHIA) scores (Pearson $r = 0.77$ for the right ear and 0.73 for the left ear). The development and validation of a South African English smartphone version of the DTT is described by Potgieter *et al.* (2015). The SRT and speech recognition properties of the smartphone version compared favorably with telephone-based versions of the test. Furthermore, a comparison among five different headphone types failed to reveal any significant differences among them, making the test suitable for the home environment.

A comprehensive scoping review of the DTT was recently conducted by Van den Borre *et al.* (2021), who evaluated 39 articles that met their review criteria. The aim of the review was to explore the effects of the variations associated with the administration of the DTT. The variables under study included the speech material (number of syllables that constituted the digits as determined by the test language), noise type (speech-weighted, fluctuating, multitalker babble), platform (telephone, headphones, earbuds), number of trials, scoring method for correct response (complete triplet, single digit), starting SNR, presentation method (monaural; dichotic, diotic), test procedures (adaptive algorithm), and target population (children, cochlear implant, and hearing aid users). The authors concluded that, in general, and despite the variations in the test administration and structure, the literature supports the DTT as a test that yields results with steep psychometric functions; high measurement precision, sensitivity, and specificity; and strong SRT-PTA correlation. While variations in stimuli, test environment, and testing procedures can have small effects on test quality metrics, the DTT has been shown to represent "a highly reliable and efficient means for measuring the loss of functional hearing ability, as well as a basis for estimating the hearing loss that is obtained through the traditional audiogram" [Van den Borre *et al.* (2021), p. 16], which is encouraging for its use in mobile applications.

One potential limitation of remote SiN testing might be its lack of diagnostic precision. Two recent papers, however, have evaluated the effect of modifying the procedure or adding the results of pure-tone threshold testing to the DTT to improve its diagnostic specificity. De Sousa *et al.* (2020b) compared diotic with antiphasic presentations of the South African English version of the DIN test among individuals with normal hearing, unilateral or asymmetric sensorineural hearing loss (SNHL), symmetric SNHL, unilateral or asymmetric conductive hearing loss (CHL), and symmetric CHL. The findings of their investigation revealed that the antiphasic presentation was more sensitive (i.e., poorer SRTs as a function of poorer ear PTAs) to each of the three types of hearing loss (asymmetric SNHL, symmetric SNHL, CHL) than the diotic presentation. In a follow-up study, De Sousa *et al.* (2020a), evaluated the sensitivity of a diotic presentation of the DIN combined with poorer ear PTAs to distinguish between hearing loss type among 158 adults with confirmed CHL ($n = 36$) or SNHL ($n = 122$). Results revealed better SRTs among those with CHL than SNHL for each of the audiometric frequencies tested as well as for the PTA. The combination of low frequency PTA, DIN SRT, and age achieved the highest sensitivity for distinguishing between SNHL and CHL. The findings of these two studies are clinically meaningful as the equipment required for bone conduction or middle ear diagnostic evaluation purposes is not often available in remote settings, particularly in the home.

https://doi.org/10.1121/10.0014418

In addition to identifying the specific site of lesion, questions arise as to the sensitivity of a DIN test to distinguish among degrees of auditory function in the absence of pure-tone testing. Armstrong *et al*. (2020) examined the correlation between degrees of hearing loss, as determined by PTA and SRTs, as measured by the DIN test. The authors also attempted to establish "optimum" SRT cutoff points associated with three categories of hearing status levels—"normal," "insufficient performance," and "poor performance." Analyses revealed the correlation between the PTA and SRT to be 0.65 (95% CI = 0.63–0.67) and the optimum cutoff points to be <–5.55 dB (normal), >–5.55 dB and <3.80 dB (insufficient performance), and >3.80 dB (poor performance).

Fundamental changes in healthcare delivery to include the increasing availability of remotely delivered services, coupled with widespread availability of PCs and smartphones, has resulted in ample opportunities for individuals to self-test their hearing. SiN testing in general and DIN algorithms in particular appear well-suited for online and mobile applications and address some of the shortcomings associated with conventional hearing testing in the clinic as predictors of rehabilitation success. The results of DIN testing appear to be generally immune to the effects of variations in test administration, and recent studies have demonstrated that DIN can distinguish between types and degrees of hearing loss. Research investigating the psychometric properties of DIN testing reveals relatively high sensitivity and specificity values and positive correlations between DIN scores, PTAs, and self-perceived hearing handicap, making the DIN tests particularly effective as hearing screening tools that can be administered over different platforms.

VII. OVERALL DISCUSSION AND CONCLUSION

Although multiple tests are available, utilization of SiN testing among clinicians appears to be low, despite good justification for their use (Taylor, 2003; Wilson, 2004; Wilson and McArdle, 2008; Mueller, 2016). Various sections of this review indicate the need for SiN testing, including innovative applications to maximize the test's utility. Intrinsic and extrinsic factors can have a significant effect on recognition performance, which must be acknowledged and understood by clinicians and researchers looking to utilize SiN assessments. Analyses comparing different SiN tests on the same listeners, such as Wilson *et al*. (2007), McArdle and Wilson (2008), and McArdle *et al*. (2005), allow for direct review of the strengths and weaknesses in the assessments and provide insight into proper interpretation of the varying performances. Some of these comparison analyses have evaluated the QuickSIN, WIN, BKB-SIN, DIN, and HINT tests and shown the sensitivity of the tests among listeners with normal hearing and listeners with hearing loss and/or difficulty hearing in noise. Such comparisons can support the proper selection of a SiN test.

Outside of selection consideration, available SiN assessments allow additional flexibility that offers ease of access for not only clinicians and researchers but also listeners. Testing materials offered in multiple languages and with considerations for differing levels of auditory impairment make these tests accessible to a wider range of the population, benefiting both proctors and participants in assessing auditory ability appropriately and with consideration for individual differences. In addition, remote access to testing materials and assessments allows for versatility in staffing and facility requirements, which increases the available population for testing as well as improving access to auditory evaluation for patients who may otherwise have geographic or socioeconomic barriers to equivalent healthcare, including traditional audiometric assessment.

When considering flexibility and variety of applications, assessment needs, population considerations, and access to quality evaluations, the limitations of standard audiometry become clear. Standard, traditional audiometry measures of pure-tone thresholds and word recognition in quiet do not provide accurate measures of functional hearing ability and have limited use in special patient populations. In addition, they inaccurately measure changes, especially in real-world, out of the clinic testing situations where listeners may be confronted by challenging listening situations that are not represented in traditional audiometry booth measures. In the last 20 years, several studies have highlighted the need for assessments that are more ecologically valid, inclusive, and applicable. SiN tests have emerged as potential measures that fit these criteria, and there are demonstrated innovative uses for these assessments in both research and clinical environments.

While there are many validated, and several emerging, SiN assessments available to clinicians and researchers, there is no singular encompassing evaluation that covers all needs in every environment. Even though the tests all include a speech and noise component, there are vast differences not only in parameters but also in stimuli and outcome measures between assessments. Therefore, it follows that a single test cannot and, at this point, should not be recommended for all assessment needs and populations. Considerations for each of the above-discussed specialty purposes and populations should be an integral part of the selection process for SiN tests. Understanding both intrinsic and extrinsic factors is paramount when choosing an appropriate test, and it is our intention for the above review to act as a guide for clinicians and researchers interested in including innovative SiN tests or innovative uses of SiN tests that meet their needs and intentions.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the help of Mr. Keehan White (Creare LLC) in the final production of this manuscript. D.M. received funding from National Institute on Deafness and Other Communication Disorders (NIDCD), NIH, Grant Nos. DC014078 and DC016241 and the Cincinnati

Children's Research Foundation. He also received support from the NIHR Manchester Biomedical Research Centre. D.M. is a paid scientific advisor to the hearX Group and to Otonomy Inc. He received personal support from NIH Grant Nos. R01DC014078 and R01DC018734 and from the NIHR Manchester Biomedical Research Center during preparation of this review. O.C. received funding from the Defense Health Department under SBIR Phase III contract W81XWH-21-C-0082. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Army Medical Research and Development Command (USAMRDC). O.C. is a member of the board of directors and holds partial ownership of Edare LLC, manufacturer of WAHTS. M.A. received funding from the NIDCD, NIH [Loan Repayment (LRP) Award Nos. R21DC018655 and R01DC019408-01], and the National Institute on Aging (NIA), NIH (Grant No. R01AG055426-01), as well as salary from the University of South Florida. H.A. is a consultant to Lively Hearing Corporation, for which he receives compensation and restricted shares. V.S. received funding from NIH Grant Nos. R01DC019408-01, R01AG055426-01, and R21DC018655 and receives a salary at the University of South Florida. V.S. reports research support, funding, and/or paid consultancy from Frequency Therapeutics Inc., Otonomy Inc., Autifony Therapeutics Ltd., Pipeline Therapeutics, Oticon Medical A/S, and Sonova Holding AG.

Agrawal, Y., Platz, E. A., and Niparko, J. K. (**2008**). "Prevalence of hearing loss and differences by demographic characteristics among US adults: Data from the National Health and Nutrition Examination Survey, 1999–2004," Arch. Intern. Med. **168**(14), 1522–1530.

Alvidrez, J., Castille, D., Laude-Sharp, M., Rosario, A., and Tabor, D. (**2019**). "The National Institute on Minority Health and Health Disparities research framework," Am. J. Public Health **109**, S16–S20.

American Academy of Audiology (**2006**). "Guidelines for the audiologic management of adult hearing impairment," https://mk0audiologyotvgk5ps.kinstacdn.com/wp-content/uploads/2021/05/haguidelines.pdf_53994876e92e42.70908344.pdf (Last viewed September 24, 2021).

American Speech-Language-Hearing Association (**1994**). "Audiologic management of individuals receiving cochleotoxic drug therapy," https://www.asha.org/policy/gl1994-00003/ (Last viewed October 10, 2022).

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (**2015**). "Development of subcortical speech representation in human infants," J. Acoust. Soc. Am. **137**(6), 3346–3355.

ANSI (**2017**). ANSI/ASA S3.5-1997 (R2017). *Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).

Armstrong, N. M., Oosterloo, B. C., Croll, P. H., Ikram, M. A., and Goedegebure, A. (**2020**). "Discrimination of degrees of auditory performance from the digits-in-noise test based on hearing status," Int. J. Audiol. **59**(12), 897–904.

Bench, J., Kowal, A., and Bamford, J. (**1979**). "The BKB (Bench–Kowal–Bamford) sentence lists for partially-hearing children," Br. J. Audiol. **13**, 108–112.

Bernstein, J. G., Mehraei, G., Shamma, S., Gallun, F. J., Theodoroff, S. M., and Leek, M. R. (**2013**). "Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners," J. Am. Acad. Audiol. **24**(4), 293–306.

Bialystok, E., Majumder, S., and Martin, M. M. (**2003**). "Developing phonological awareness: Is there a bilingual advantage?," Appl. Psycholinguist. **24**, 27–44.

Bishop, D. V. M., and Leonard, L. B. (**2000**). *Speech and Language Impairments in Children* (Psychology, Hove, UK).

Blankenship, C. M., Hunter, L. L., Keefe, D. H., Feeney, M. P., Brown, D. K., McCune, A., Fitzpatrick, D. F., and Lin, L. (**2018**). "Optimizing clinical interpretation of distortion product otoacoustic emissions in infants," Ear Hear. **39**(6), 1075–1090.

Bocca, E., and Calearo, C. (**1963**). "Central hearing processes," in *Modern Developments in Audiology*, edited by J. Jerger (Academic, New York), pp. 337–370.

British Society of Audiology (**2016**). "Common principles of rehabilitation for adults in audiology practices," https://www.thebsa.org.uk/resources/common-principles-rehabilitation-adults-audiology-services/ (Last viewed September 24, 2021).

Brock, P. R., Bellman, S. C., Yeomans, E. C., Pinkerton, C. R., and Pritchard, J. (**1991**). "Cisplatin ototoxicity in children: A practical grading system," Med. Pediatr. Oncol. **19**(4), 295–300.

Brungart, D. S. (**2001**). "Evaluation of speech intelligibility with the coordinate response measure," J. Acoust. Soc. Am. **109**, 2276–2279.

Brungart, D. S., Makashay, M. J., and Sheffield, B. M. (**2021**). "Development of an 80-word clinical version of the modified rhyme test (MRT_{80})," J. Acoust. Soc. Am. **149**, 3311–3327.

Brungart, D. S., Walden, B., Cord, M., Phatak, S., Theodoroff, S. M., Griest, S., and Grant, K. W. (**2017**). "Development and validation of the Speech Reception in Noise (SPRINT) Test," Hear. Res. **349**, 90–97.

Cameron, S., and Dillon, H. (**2007**). "Development of the Listening in Spatialized Noise-Sentences Test (LISN-S)," Ear Hear. **28**(2), 196–211.

Cameron, S., and Dillon, H. (**2011**). "Development and evaluation of the LiSN and learn auditory training software for deficit-specific remediation of binaural processing deficits in children: Preliminary findings," J. Am. Acad. Audiol. **22**(10), 678–696.

Cameron, S., Glyde, H., and Dillon, H. (**2011**). "Listening in Spatialized Noise-Sentences Test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age," J. Am. Acad. Audiol. **22**(10), 697–709.

Cameron, S., Glyde, H., and Dillon, H. (**2012**). "Efficacy of the LiSN and Learn auditory training software: Randomized blinded controlled study," Audiol. Res. **2**(1), e15.

Carlo, M. A. (**2009**). "A review of the effects of bilingualism on speech recognition performance," Perspect. Hear. Hear. Dis. Res. Diag. **13**, 14–20.

Carlo, M. A., Wilson, R. H., and Villanueva-Reyes, A. (**2020**). "Psychometric characteristics of Spanish monosyllabic, bisyllabic, and trisyllabic words for use in word-recognition protocols," J. Am. Acad. Audiol. **31**, 531–546.

Chang, K. W. (**2011**). "Clinically accurate assessment grading ototoxicity," Laryngoscope **121**(12), 2649–2657.

Chen, K.-H., Su, S.-B., and Chen, K.-T. (**2020**). "An overview of occupational noise-induced hearing loss among workers: Epidemiology, pathogenesis, and preventive measures," Environ. Health Prev. Med. **25**(1), 65.

Clayton, J. A. (**2018**). "Applying the new SABV (sex as a biological variable) policy to research and clinical care," Physiol. Behav. **187**, 2–5.

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (**2016**). "Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers," Ear Hear. **37**(1), 55–63.

Couth, S., Prendergast, G., Guest, H., Munro, K. J., Moore, D. R., Plack, C. J., Ginsborg, J., and Dawes, P. (**2020**). "Investigating the effects of noise exposure on self-report, behavioral and electrophysiological indices of hearing damage in musicians with normal audiometric thresholds," Hear. Res. **395**, 108021.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (**1989**). "Limits on bilingualism," Nature **340**, 229–230.

Davidson, A., Marrone, N., Wong, B., and Musiek, F. (**2021**). "Predicting hearing aid satisfaction in adults: A systematic review of speech-in-noise tests and other behavioral measures," Ear Hear. **42**(6), 1485–1498.

Denys, S., Wouters, J., and van Wieringen, A. (**2021**). "The digit triplet test as a self-test for hearing screening at the age of school-entry," Int. J. Audiol. **61**(5), 408–415.

De Sousa, K. C., Smits, C., Moore, D. R., Myburgh, H. C., and Swanepoel, D. W. (**2020a**). "Pure-tone audiometry without bone-conduction thresholds: Using the digits-in-noise test to detect conductive hearing loss," Int. J. Audiol. **59**(10), 801–808.

De Sousa, K. C., Smits, C., Moore, D. R., Myburgh, H. C., and Swanepoel, D. W. (**2021**). "Diotic and antiphasic digits-in-noise testing as a hearing screening and triage tool to classify type of hearing loss," Ear Hear. **43**(3), 1037–1048.

De Sousa, K. C., Swanepoel, W., Moore, D. R., Myburgh, H. C., and Smits, C. (**2020b**). "Improving sensitivity of the digits-in-noise test using antiphasic stimuli," Ear Hear. **41**(2), 442–450.
Dillon, H., Cameron, S., Tomlin, D., and Glyde, H. (**2014**). "Comments on 'Factors influencing tests of auditory processing: A perspective on current issues and relevant concerns' by Tony Cacace and Dennis McFarland," J. Am. Acad. Audiol. **25**, 699–703.
Dillon, H., and Cameron, S. (**2021**). "Separating the causes of listening difficulties in children," Ear Hear. **42**(5), 1097–1108.
Egan, J. P. (**1948**). "Articulation testing methods," Laryngoscope **58**(9), 955–991.
Etymotic Research (**2005**). *BKB-SIN™ Speech-in-Noise Test Version 1.03* (Etymotic Research, Inc., Elk Grove Village, IL).
Fabry, D., Firszt, J. B., Gifford, R. H., Holden, L. K., and Koch, D. B. (**2009**). "Evaluating speech perception benefit in adult cochlear implant recipients," Audiol. Today **21**(3), 36–43.
Falk, S. A., and Farmer, J. C. (**1973**). "Incubator noise and possible deafness," Arch. Otolaryngol. **97**(5), 385–387.
Ferjan Ramírez, N., Lytle, S. R., and Kuhl, P. K. (**2020**). "Parent coaching increases conversational turns and advances infant language development," Proc. Natl. Acad. Sci. U.S.A. **117**(7), 3484–3491.
Finney, D. J. (**1952**). *Statistical Method in Biological Assay* (Charles Griffen, London).
Folmer, R. L., Vachhani, J., McMillan, G. P., Watson, C., Kidd, G. R., and Feeney, M. P. (**2017**). "Validation of a computer-administered version of the digits-in-noise test for hearing screening in the United States," J. Am. Acad. Audiol. **28**(2), 161–169.
Fought, C. (**2011**). "Language and ethnicity," in *The Cambridge Handbook of Sociolinguistics* (Cambridge University, Cambridge, UK), pp. 238–258.
Gallun, F. J., Seitz, A., Eddins, D. A., Molis, M. R., Stavropoulos, T., Jakien, K. M., Kampel, S. D., Diedesch, A. C., Hoover, E. C., Bell, K., Souza, P. E., Sherman, M., Calandruccio, L., Xue, G., Taleb, N., Sebena, R., and Srinivasan, N. (**2018**). "Development and validation of Portable Automated Rapid Testing (PART) measures for auditory research," Proc. Mtgs. Acoust. **33**(1), 050002.
Garinther, G., and Peters, L. J. (**1990**). "Impact of communications on armor crew performance," Army Res. Dev. Acquis. Bull. **70**, 1–5.
Gates, K., Hecht, Q. A., Grantham, M. A. M., Fallon, A. J., and Martukovich, M. (**2021**). "Hearing health care delivery outside the booth," Perspect. ASHA Spec. Interest Groups **6**(5), 1123–1136.
Gifford, R. H., Shallop, J. K., and Peterson, A. (**2008**). "Speech recognition materials and ceiling effects: Considerations for cochlear implant programs," Audiol. Neurotol. **13**(3), 193–205.
Giguère, C., Laroche, C., Soli, S. D., and Vaillancourt, V. (**2008**). "Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test," Int. J. Audiol. **47**, 319–328.
Grant, K. W., Kubli, L. R., Phatak, S. A., Galloza, H., and Brungart, D. S. (**2021**). "Estimated prevalence of functional hearing difficulties in blast-exposed service members with normal to near-normal-hearing thresholds," Ear Hear. **42**(6), 1615–1626.
Graydon, K., Rance, G., Dowell, R., and Van Dun, B. (**2017**). "Consequences of early conductive hearing loss on long-term binaural processing," Ear Hear. **38**(5), 621–627.
Hagerman, B., and Kinnefors, C. (**1995**). "Efficient adaptive methods for measuring speech reception threshold in quiet and in noise," Scand. Audiol. **24**(1), 71–77.
Halliday, L. F., Taylor, J. L., Edmondson-Jones, A. M., and Moore, D. R. (**2008**). "Frequency discrimination learning in children," J. Acoust. Soc. Am. **123**(6), 4393–4402.
Havinsky, O., Springer, K. W., and Hunting, G. (**2018**). "Beyond sex and gender difference in funding and reporting of health research," Res. Integr. Peer Rev. **3**(1), 6.
Heidari, S., Babor, T. F., De Castro, P., Tort, S., and Curno, M. (**2016**). "Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use," Res. Integr. Peer Rev. **1**(1), 2.
Helzner, E. P., Cauley, J. A., Pratt, S. R., Wisniewski, S. R., Zmuda, J., Talbott, E. O., de Rekeneire, N., Harris, T. B., Rubin, S. M., Simonsick, E. M., Tylavsky, F. A., and Newman, A. B. (**2005**). "Race and sex differences in age-related hearing loss: The Health, Aging and Body Composition study," J. Am. Geriatr. Soc. **53**(12), 2119–2127.
Hugh, S. C., Wolter, N. E., Propst, E. J., Gordon, K. A., Cushing, S. L., and Papsin, B. C. (**2014**). "Infant sleep machines and hazardous sound pressure levels," Pediatrics **133**(4), 677–681.
Humes, L. E. (**2019**). "The World Health Organization's hearing-impairment grading system: An evaluation for unaided communication in age-related hearing loss," Int. J. Audiol. **58**(1), 12–20.
Hunter, L. L., Monson, B. B., Moore, D. R., Dhar, S., Wright, B. A., Munro, K. J., Zadeh, L. M., Blankenship, C. M., Stiepan, S. M., and Siegel, J. H. (**2020**). "Extended high frequency hearing and speech perception implications in adults and children," Hear. Res. **397**, 107922.
Hunter, L. L., Tubaugh, L., Jackson, A., and Propes, S. (**2008**). "Wideband middle ear power measurement in infants and children," J. Am. Acad. Audiol. **19**(4), 309–324.
Institute of Medicine of the National Academies (**2006**). *Noise and Military Service: Implications for Hearing Loss and Tinnitus* (National Academies Press, Washington, DC).
Isherwood, B., Goncalves, A. C., Cousins, R., and Holme, R. (**2021**). "The global hearing therapeutic pipeline: 2021," Drug Discov. Today **27**(3), 912–922.
Keller, M. D., Ziriax, J. M., Barns, W., Sheffield, B., Brungart, D., Thomas, T., Jaeger, B., and Yankaskas, K. (**2017**). "Performance in noise: Impact of reduced speech intelligibility on Sailor performance in a Navy command and control environment," Hear. Res. **349**, 55–66.
Killion, M. C. (**2002**). "New thinking on hearing in noise: A generalized articulation index," Semin. Hear. **23**(1), 57–76.
Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (**2004**). "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **116**(4), 2395–2405.
Kilman, L., Zekveld, A., Hällgren, M., and Rönnberg, J. (**2014**). "The influence of non-native language proficiency on speech perception performance," Front. Psychol. **5**, 651.
Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., and Wagener, K. C. (**2015**). "The multilingual matrix test: Principles, applications, and comparison across languages: A review," Int. J. Audiol. **54**(Suppl 2), 3–16.
Konrad-Martin, D., Bennett, K. O. C., Garinis, A., and McMillan, G. P. (**2021**). "A randomized controlled trial using automated technology for improving ototoxicity monitoring in VA oncology patients," Am. J. Audiol. **30**(3S), 870–886.
Konrad-Martin, D., Poling, G. L., Garinis, A. C., Ortiz, C. E., Hopper, J., O'Connell Bennett, K., and Dille, M. F. (**2018**). "Applying U.S. national guidelines for ototoxicity monitoring in adult patients: Perspectives on patient populations, service gaps, barriers and solutions," Int. J. Audiol. **57**(Suppl. 4), S3–S18.
Koopmans, W. J. A., Goverts, S. T., and Smits, C. (**2018**). "Speech recognition abilities in normal-hearing children 4 to 12 years of age in stationary and interrupted noise," Ear Hear. **39**(6), 1091–1103.
Kral, A., Dorman, M. F., and Wilson, B. S. (**2019**). "Neuronal development of hearing and language: Cochlear implants and critical periods," Annu. Rev. Neurosci. **42**, 47–65.
Kraus, N., and White-Schwoch, T. (**2016**). "Newborn hearing screening 2.0," Hear. J. **69**(11), 44–46.
Krizman, J., Skoe, E., and Kraus, N. (**2012**). "Sex differences in auditory subcortical function," Clin. Neurophysiol. **123**, 591–597.
Kujawa, S. G., and Liberman, M. C. (**2015**). "Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss," Hear. Res. **330**(Pt. B), 191–199.
Laroche, C., Giguère, C., Vaillancourt, V., and Soli, S. (**2005**). *Development and Validation of Hearing Standards for Canadian Coast Guard Seagoing Personnel and Conservation & Protection Seagoing and Land-Based Personnel Phase II* (Fisheries and Oceans Canada, Ottawa, ON, Canada).
Lauer, A. M., and Schrode, K. M. (**2017**). "Sex bias basic preclinical noise-induced hearing loss research," Noise Health **19**, 207–212.
Leensen, M. C., de Laat, J. A. P. M., and Dreschler, W. A (**2011a**). "Speech-in-noise screening tests by internet, part 1: Test evaluation for noise-induced hearing loss identification," Int. J. Audiol. **50**(11), 823–834.
Leensen, M. C., de Laat, J. A., Snik, A. F., and Dreschler, W. A. (**2011b**). "Speech-in-noise screening tests by internet, part 2: Improving test sensitivity for noise-induced hearing loss," Int. J. Audiol. **50**(11), 835–848.

Leensen, M. C., and Dreschler, W. A. (**2013a**). "The applicability of a speech-in-noise screening test in occupational hearing conservation," Int. J. Audiol. **52**(7), 455–465.
Leensen, M. C., and Dreschler, W. A. (**2013b**). "Speech-in-noise screening tests by internet, Part 3: Test sensitivity for uncontrolled parameters in domestic usage," Int. J. Audiol. **52**(10), 658–669.
Le Prell, C. G. (**2021**). "Investigational medicinal products for the inner ear: Review of clinical trial characteristics in ClinicalTrials.gov," J. Am. Acad. Audiol. **32**, 670–694.
Le Prell, C. G., and Brungart, D. S. (**2016**). "Speech-in-noise tests and supra-threshold auditory evoked potentials as metrics for noise damage and clinical trial outcome measures," Otol. Neurotol. **37**(8), e295–e302.
Le Prell, C. G., and Clavier, O. H. (**2017**). "Effects of noise on speech recognition: Challenges for communication by service members," Hear. Res. **349**, 76–89.
Le Prell, C. G., and Lobarinas, E. (**2015**). "Strategies for evaluating antioxidant efficacy in clinical trials assessing prevention of noise-induced hearing loss," in *Free Radicals in ENT Pathology* (Springer, New York), pp. 163–192.
Le Prell, C. G., Siburt, H. W., Lobarinas, E., Griffiths, S. K., and Spankovich, C. (**2018**). "No reliable association between recreational noise exposure and threshold sensitivity, distortion product otoacoustic emission amplitude, or word-in-noise performance in a college student population," Ear Hear. **39**(6), 1057–1074.
Lie, A., Skogstad, M., Johannessen, H. A., Tynes, T., Mehlum, I. S., Nordby, K.-C., Engdahl, B., and Tambs, K. (**2016**). "Occupational noise exposure and hearing: A systematic review," Int. Arch. Occup. Environ. Health **89**(3), 351–372.
Marinova-Todd, S. H., Siu, C. K., and Jenstad, L. M. (**2011**). "Speech audiometry with non-native English speakers: The use of digits and Cantonese words as stimuli," Can. J. Speech Lang. Pathol. Audiol. **35**, 220–227.
Mayo, L. H., Florentine, M., and Buus, S. (**1997**). "Age of second-language acquisition and perception of speech in noise," J. Speech. Lang. Hear. Res. **40**, 686–693.
McArdle, R., and Wilson, R. H. (**2008**). "Selecting speech tests to measure auditory function," ASHA Leader **13**(12), 5–6.
McArdle, R. A., Wilson, R. H., and Burks, C. A. (**2005**). "Speech recognition in multi-talker babble using digits, words, and sentences," J. Am. Acad. Audiol. **16**(9), 726–739.
Miller, G. A., Heise, G. A., and Lichten, W. (**1951**). "The intelligibility of speech as a function of the context of the test materials," J. Exp. Psychol. **41**(5), 329–335.
Monson, B. B., Rock, J., Cull, M., and Soloveychik, V. (**2020**). "Neonatal intensive care unit incubators reduce language and noise levels more than the womb," J. Perinatol. **40**(4), 600–606.
Moore, T. J. (**1981**). "Voice communication jamming research," in *AGARD Conference Proceedings 331: Aural Communication in Aviation*, Neuilly-Sur-Seine, France, pp. 2:1–2:6.
Moore, D. R. (**2012**). "Listening difficulties in children: Bottom-up and top-down contributions," J. Commun. Disord. **45**(6), 411–418.
Moore, D. R. (**2018**). "Editorial: Auditory processing disorder," Ear Hear. **39**(4), 617–620.
Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., and Ferguson, M. A. (**2011**). "Development of auditory processing in 6- to 11-yr-old children," Ear Hear. **32**(3), 269–285.
Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., and Riley, A. (**2010**). "Nature of auditory processing disorder in children," Pediatrics **126**(2), e382–e390.
Moore, D. R., Whiston, H., Lough, M., Marsden, A., Dillon, H., Munro, K. J., and Stone, M. A. (**2019**). "FreeHear: a new sound-field speech-in-babble hearing assessment tool," Trends Hear. **23**, 2331216519872378.
Moore, D. R., Zobay, O., and Ferguson, M. A. (**2020**). "Minimal and mild hearing loss in children: Association with auditory perception, cognition, and communication problems," Ear Hear. **41**(4), 720–732.
Motlagh Zadeh, L., Silbert, N. H., Sternasty, K., Swanepoel, W., Hunter, L. L., and Moore, D. R. (**2019**). "Extended high-frequency hearing enhances speech perception in noise," Proc. Natl. Acad. Sci. U. S. A. **116**(47), 23753–23759.
Mueller, H. G. (**2016**). "Speech-in-noise testing for selection and fitting of hearing aids: worth the effort?," https://www.audiologyonline.com/articles/signia-expert-series-speech-in-18336 (Last viewed September 24, 2021).
National Cancer Institute (**2009**). *Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0* (National Institutes of Health, Bethesda, MD).
National Institute for Health Research (**2020**). "Improving inclusion of under-served groups in clinical research: Guidance from the NIHR INCLUDE project," https://www.nihr.ac.uk/documents/improving-inclusion-of-under-served-groups-in-clinical-research-guidance-from-include-project/25435 (Last viewed September 29, 2021).
National Institutes of Health (**2015**). "Consideration of sex as a biological variable in NIH-funded research," https://grants.nih.gov/grants/guide/notice-files/not-od-15-102.html (Last viewed September 28, 2021).
National Institutes of Health (**2021**). "Consideration of sex as a biological variable in NIH-funded research: Guidance document," https://orwh.od.nih.gov/sites/orwh/files/docs/NOT-OD-15-102_Guidance.pdf (Last viewed September 28, 2021).
Newman, C. W., Weinstein, B. E., Jacobson, G. P., and Hug, G. A. (**1990**). "The Hearing Handicap Inventory for Adults: Psychometric adequacy and audiometric correlates," Ear Hear. **11**(6), 430–433.
NIH Office of Research on Women's Health (**2021**). "Sex/gender influence in health and disease," https://orwh.od.nih.gov/sex-gender/sexgender-influences-health-and-disease (Last viewed September 29, 2021).
Nilsson, M. J., McCaw, V. M., and Soli, S. (**1996**). *Minimum Speech Test Battery for Adult Cochlear Implant Users* (House Ear Institute, Los Angeles).
Nilsson, M., Soli, S. D., and Sullivan, J. A. (**1994**). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. **95**, 1085–1099.
Nissen, S. L., Harris, R. W., Channell, R. W., Conklin, B., Kim, M., and Wong, L. (**2011**). "The development of psychometrically equivalent Cantonese speech audiometry materials," Int. J. Audiol. **50**(3), 191–201.
Olsho, L. W., Koch, E. G., Halpin, C. F., and Carter, E. A. (**1987**). "An observer-based psychoacoustic procedure for use with young infants," Dev. Psychol. **23**(5), 627–640.
Pearson, J. D., Morrell, C. H., and Gordon-Salant, S. (**1995**). "Gender differences in a longitudinal study of age-associated hearing loss," J. Acoust. Soc. Am. **97**(2), 1196–1205.
Peterson, G. E., and Lehiste, I. (**1962**). "Revised CNC lists for auditory tests," J. Speech Hear. Disord. **27**, 62–70.
Petley, L., Hunter, L. L., Motlagh Zadeh, L., Stewart, H. J., Sloat, N. T., Perdew, A., Lin, L., and Moore, D. R. (**2021**). "Listening difficulties in children with normal audiograms: Relation to hearing and cognition," Ear Hear. **42**(6), 1640–1655.
Pittman, C. A., Roura, R., Price, C., Lin, F. R., and Nieman, C. L. (**2021**). "Racial/ethnic and sex representation in US-based clinical trials of hearing loss management in adults: A systematic review," JAMA Otolaryngol. Head Neck Surg. **147**, 656–662.
Ponton, C. W., Eggermont, J. J., Coupland, S. G., and Winkelaar, R. (**1992**). "Frequency-specific maturation of the eighth nerve and brainstem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs)," J. Acoust. Soc. Am. **91**(3), 1576–1586.
Potgieter, J. M., Swanepoel, D. W., Myburgh, H. C., Hopper, T. C., and Smits, C. (**2015**). "Development and validation of a smartphone-based digits-in-noise hearing test in South African English," Int. J. Audiol. **55**(7), 405–411.
Ramkissoon, I., Proctor, A., Lansing, C. R., and Bilger, R. C. (**2002**). "Digit speech recognition thresholds (SRT) for non-native speakers of English," Am. J. Audiol. **11**, 23–28.
Rhebergen, K. S., Pool, R. E., and Dreschler, W. A. (**2014**). "Characterizing the speech reception threshold in hearing-impaired listeners in relation to masker type and masker level," J. Acoust. Soc. Am. **135**, 1491–1505.
Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (**2006**). "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise," J. Acoust. Soc. Am. **120**(6), 3988–3997.
Rogers, C. L., Lister, J. J., and Febo, D. M. (**2006**). "Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing," Appl. Psycholinguist. **27**, 465–485.
Rogers, D. S., Harkrider, A. W., Burchfield, S. B., and Nabelek, A. K. (**2003**). "The influence of listener's gender on the acceptance of background noise," J. Am. Acad. Audiol. **14**(7), 372–382.
Sanchez, V. A. (**2018**). "Emerging interventions for age-related hearing loss: Review of the CLARITY-1 trial through the description of the pharmaceutical development phases," Perspect. ASHA Spec. Interest Groups **3**(6), 51–60.

Sandström, J., Swanepoel, d., Carel Myburgh, H., and Laurent, C. (**2016**). "Smartphone threshold audiometry in underserved primary health-care contexts," Int. J. Audiol. **55**(4), 232–238.
Schilder, A. G. M., Su, M. P., Blackshaw, H., Lustig, L., Staecker, H., Lenarz, T., Safieddine, S., Gomes-Santos, C. S., Holme, R., and Warnecke, A. (**2019a**). "Hearing protection, restoration, and regeneration: An overview of emerging therapeutics for inner ear and central hearing disorders," Otol. Neurotol. **40**(5), 559–570.
Schilder, A. G. M., Su, M. P., Mandavia, R., Anderson, C. R., Landry, E., Ferdous, T., and Blackshaw, H. (**2019b**). "Early phase trials of novel hearing therapeutics: Avenues and opportunities," Hear. Res. **380**, 175–186.
Semeraro, H. D., Rowan, D., Van Besouw, R. M., and Allsopp, A. A. (**2017**). "Development and evaluation of the British English coordinate response measure speech-in-noise test as an occupational hearing assessment tool," Int. J. Audiol. **56**, 749–758.
Shansky, R. M., and Murphy, A. Z. (**2021**). "Considering sex as a biological variable will require a global shift in science culture," Nat. Neurosci. **24**, 457–464.
Sheffield, B., Ziriax, J., Keller, M. D., Barns, W., and Brungart, D. (**2017**). "The impact of reduced speech intelligibility on reaction time in a naval combat environment," Proc. Hum. Factors Ergonom. Soc. Annu. Meet. **61**, 1570–1574.
Sheikh, R. M., Leensen, M. C. J., de Laat, J., and Dreschler, W. A. (**2017**). "Laboratory evaluation of an optimized internet-based speech-in-noise test for occupational high-frequency hearing loss screening: Occupational earcheck," Int. J. Audiol. **56**(11), 844–853.
Shi, L., and Sánchez, D. (**2010**). "Spanish/English bilingual listeners on clinical word recognition tests: What to expect and how to predict," J. Speech. Lang. Hear. Res. **53**, 1096–1110.
Shuster, B. Z., Depireux, D. A., Mong, J. A., and Herzano, R. (**2019**). "Sex differences in hearing: Probing the role of estrogen signaling," J. Acoust. Soc. Am. **145**(6), 3656–3663.
Simonoska, R., Stenberg, A. E., Duan, M., Yakimchuk, K., Fridberger, A., Sahlin, L., Gustafsson, J., and Hultcrantz, M. (**2009**). "Inner ear pathology and loss of hearing in estrogen receptor-beta deficient mice," J. Endocrinol. **201**, 397–406.
Sininger, Y. S., Hunter, L. L., Hayes, D., Roush, P. A., and Uhler, K. M. (**2018**). "Evaluation of speed and accuracy of next-generation auditory steady state response and auditory brainstem response audiometry in children with normal hearing and hearing loss," Ear Hear. **39**(6), 1207–1223.
Smith, A. W. (**1998**). "The World Health Organization and the prevention of deafness and hearing impairment caused by noise," Noise Health **1**(1), 6–12.
Smits, C., Goverts, S. T., and Festen, J. M. (**2013**). "The digits-in-noise test: Assessing auditory speech recognition abilities in noise," J. Acoust. Soc. Am. **133**(3), 1693–1706.
Smits, C., Kapteyn, T. S., and Houtgast, T. (**2004**). "Development and validation of an automatic speech-in-noise screening test by telephone," Int. J. Audiol. **43**(1), 15–28.
Smits, C., Merkus, P., and Houtgast, T. (**2006**). "How we do it: The Dutch functional hearing-screening tests by telephone and internet," Clin. Otolaryngol. **31**(5), 436–440.
Soli, S. D., Amano-Kusumoto, A., Clavier, O., Wilbur, J., Casto, K., Freed, D., Laroche, C., Vaillancourt, V., Giguère, C., Dreschler, W. A., and Rhebergen, K. S. (**2018a**). "Evidence-based occupational hearing screening II: Validation of a screening methodology using measures of functional hearing ability," Int. J. Audiol. **57**, 323–334.
Soli, S. D., Giguère, C., Laroche, C., Vaillancourt, V., Dreschler, W. A., Rhebergen, K. S., Harkins, K., Ruckstuhl, M., Ramulu, P., and Meyers, L. S. (**2018b**). "Evidence-based occupational hearing screening I: Modeling the effects of real-world noise environments on the likelihood of effective speech communication," Ear Hear. **39**(3), 436–448.
Soli, S., and Vermiglio, A. (**1999**). "Assessment of functional hearing ability for hearing critical jobs in law enforcement: A report prepared for the California Peace Officers Standards and Training Commission. Human Communication Sciences and Devices Department" (House Ear Institute, Los Angeles).
Soli, S. D., Vermiglio, A., Wen, K., and Filesari, C. A. (**2002**). "Development of the Hearing in Noise Test (HINT) in Spanish," J. Acoust. Soc. Am. **112**, 2384.
Soli, S. D., and Wong, L. L. (**2008**). "Assessment of speech intelligibility in noise with the Hearing in Noise Test," Int. J. Audiol. **47**, 356–361.
Spahr, A., and Dorman, M. F. (**2004**). "Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices," Arch. Otolaryngol. Head Neck Surg. **130**, 624–628.
Suzuki, T., and Ogiba, Y. (**1961**). "Conditioned orientation reflex audiometry," Arch. Otolanryngol. **74**, 192–198.
Swanepoel, W. (**2020**). "eHealth technologies enable more accessible hearing care," Semin. Hear. **41**(2), 133–140.
Tabri, D., Chacra, K. M. S. A., and Pring, T. (**2011**). "Speech perception in noise by monolingual, bilingual and trilingual listeners," Int. J. Lang. Commun. Disord. **46**, 411–422.
Taylor, B. (**2003**). "Speech-in-noise tests," Hear. J. **56**(1), 40–42.
Themann, C. L., and Masterson, E. A. (**2019**). "Occupational noise exposure: A review of its effects, epidemiology, and impact with recommendations for reducing its burden," J. Acoust. Soc. Am. **146**(5), 3879–3905.
Theunissen, M., Swanepoel, D. W., and Hanekom, J. (**2009**). "Sentence recognition in noise: Variables in compilation and interpretation of tests," Int. J. Audiol. **48**(11), 743–757.
Thompson, E. C., Estabrook, R., Krizman, J., Smith, S., Huang, S., White-Schwoch, T., Nicol, T., and Kraus, N. (**2021**). "Auditory neurophysiological development in early childhood: A growth curve modeling approach," Clin. Neurophysiol. **132**(9), 2110–2122.
Thornton, D., Harkrider, A. W., Jenson, D. E., and Saltuklaroglu, T. (**2019**). "Sex differences in early sensorimotor processing for speech discrimination," Sci. Rep. **9**(1), 392.
Tillman, T. W., and Carhart, R. (**1966**). "An expanded test for speech discrimination utilizing CNC monosyllabic words. Northwestern University Auditory Test No. 6," Technical Report SAM-TR-66-55 (USAF School of Aerospace Medicine, Brooks Air Force Base, TX).
Tomblin, J. B., Oleson, J., Ambrose, S. E., Walker, E. A., McCreery, R. W., and Moeller, M. P. (**2020**). "Aided hearing moderates the academic outcomes of children with mild to severe hearing loss," Ear Hear. **41**(4), 775–789.
Tufts, J. B., Vasil, K. A., and Briggs, S. (**2009**). "Auditory fitness for duty: A review," J. Am. Acad. Audiol. **20**(09), 539–557.
U.S. Census Bureau (**2015**). "Detailed language spoken at home and ability to speak English for the population 5 years and over: 2009-2013," https://www.census.gov/topics/population/language-use.html (Last viewed June 15, 2021).
Van den Borre, E., Denys, S., van Wieringen, A., and Wouters, J. (**2021**). "The digit triplet test: A scoping review," Int. J. Audiol. **60**(12), 946–963.
van der Hoek-Snieders, H. E., Houben, R., and Dreschler, W. A. (**2021**). "Measuring auditory fitness in locomotive engineers: Development and validation of a signal detection test," Ear Hear. **42**, 1313–1320.
Vermiglio, A. J., and Fang, X. (**2022**). "The World Health Organization (WHO) hearing impairment guidelines and a speech recognition in noise (SRN) disorder," Int. J. Audiol. **61**, 818–825.
Visagie, A., Swanepoel, D. W., and Eikelboom, R. H. (**2015**). "Accuracy of remote hearing assessment in a rural community," Telemed. J. E. Health **21**(11), 930–937.
von Hapsburg, D., and Peña, E. D. (**2002**). "Understanding bilingualism and its impact on speech audiometry," J. Speech Lang. Hear. Res. **45**, 202–213.
Wachman, E. M., and Lahav, A. (**2011**). "The effects of noise on preterm infants in the NICU," Arch. Dis. Child. Fetal Neonatal Ed. **96**(4), F305–F309.
Walden, T., and Walden, B. (**2004**). "Predicting success with hearing aids in everyday living," J. Am. Acad. Audiol. **15**, 342–352.
Watson, C. S., Kidd, G. R., Miller, J. D., Smits, C., and Humes, L. E. (**2012**). "Telephone screening tests for functionally impaired hearing: Current use in seven countries and development of a U.S. version," J. Am. Acad. Audiol. **23**(10), 757–767.
Whitehouse, A. J. O., Varcin, K. J., Pillar, S., Billingham, W., Alvares, G. A., Barbaro, J., Bent, C. A., Blenkley, D., Boutrus, M., Chee, A., Chetcuti, L., Clark, A., Davidson, E., Dimov, S., Dissanayake, C., Doyle, J., Grant, M., Green, C. C., Harrap, M., Iacono, T., Matys, L., Maybery, M., Pope, D. F., Renton, M., Rowbottam, C., Sadka, N., Segal, L., Slonims, V., Smith, J., Taylor, C., Wakeling, S., Wan, M. W., Wray, J., Cooper, M. N., Green, J., and Hudry, K. (**2021**). "Effect of preemptive intervention on developmental outcomes among infants showing early signs of autism: A randomized clinical trial of outcomes to diagnosis," JAMA Pediatr. **175**(11), e213298.
WHO (**2021**). *World Report on Hearing* (World Health Organization, Geneva, Switzerland).

Williams-Sanchez, V., McArdle, R. A., Wilson, R. H., Kidd, G. R., Watson, C. S., and Bourne, A. L. (**2014**). "Validation of a screening test of auditory function using the telephone," J. Am. Acad. Audiol. **25**(10), 937–951.
Wilson, R. H. (**2003**). "Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance," J. Am. Acad. Audiol. **14**, 453–470.
Wilson, R. H. (**2004**). "Adding speech-in-noise testing to your clinical protocol," Hear. J. **57**(2), 10–19.
Wilson, R. H. (**2011**). "Clinical experience with the words-in-noise test on 3430 veterans: Comparisons with pure-tone thresholds and word recognition in quiet," J. Am. Acad. Audiol. **22**(7), 405–423.
Wilson, R. H., Abrams, H. B., and Pillion, A. L. (**2003**). "A word-recognition task in multitalker babble using a descending presentation mode from 24 dB to 0 dB signal to babble," J. Rehabil. Res. Dev. **40**(4), 321–328.
Wilson, R. H., Carnell, C. S., and Cleghorn, A. L. (**2007**). "The Words-in-Noise Test with multitalker babble and speech-spectrum noise maskers," J. Am. Acad. Audiol. **18**, 522–529.
Wilson, R. H., Farmer, N. M., Gandhi, A., Shelburne, E., and Weaver, J. (**2010**). "Normative data for the Words-in-Noise Test for 6-to 12-year-old children," J. Speech Lang. Hear. Res. **53**(5), 1111–1121.
Wilson, R. H., and Margolis, R. H. (**1983**). "Measurements of auditory thresholds for speech stimuli," in *Principles of Speech Audiometry*, edited by D. F. Konkle and W. F. Rintelmann (University Park, Baltimore, MD), pp. 79–126.
Wilson, R. H., and McArdle, R. (**2005**). "Speech signals used to evaluate functional status of the auditory system," J. Rehabil. Res. Dev. **42**(4 Suppl 2), 79–94.
Wilson, R. H., and McArdle, R. (**2008**). "Change is in the air—in more ways than one," Hear. J. **61**, 10–15.
Wilson, R. H., and Weakley, D. G. (**2004**). "The use of digit triplets to evaluate word-recognition abilities in multitalker babble," Semin. Hear. **25**(01), 93–111.
Wong, L. L. N., and Soli, S. D. (**2005**). Ear Hear. **26**, 276–289.
Wong, L. L. N., Soli, S. D., Liu, S., Han, N., and Huang, M. (**2007**). "Development of the Mandarin Hearing in Noise Test (MHINT)," Ear Hear. **28**, 70S–74S.
Yankaskas, K. (**2013**). "Prelude: Noise-induced tinnitus and hearing loss in the military," Hear. Res. **295**, 3–8.
Yoshinaga-Itano, C. (**2003**). "Early intervention after universal neonatal hearing screening: Impact on outcomes," Ment. Retard. Dev. Disabil. Res. Rev. **9**(4), 252–266.
Yumba, W. K. (**2022**). "Influences of listener gender and working memory capacity on speech recognition in noise for hearing aid users," Speech Lang. Hear. **25**(2), 112–124.