Development of Hearing Standards for the Selection of Police Constables and Correctional Officers

C. Laroche[1], V. Vaillancourt[1], C. Giguère[1], S. Soli[2]

[1] Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, CANADA, claroche@uottawa.ca, vvaillan@uottawa.ca, cgiguere@uottawa.ca

[2] House Research Institute, Los Angeles, California, USA, ssoli@hei.org

ABSTRACT

Police constables perform hearing-critical (HC) jobs in noisy environments. Hearing fitness standards are necessary to ensure that constables possess the necessary auditory skills to perform their work safely and efficiently. The Ontario Ministry of Community Safety and Correctional Services is reviewing its standards for applicants to better take into account measures of functional hearing. Stage I of this project involved better defining HC tasks performed by constables. In Stage 2, the Hearing in Noise Test together with analyses based on the ESII were used to predict the ability to communicate effectively with speech in real-world noise environments, using quality audio recordings made in various work environments. The developed model, validated using a sample of normal-hearing and hearing-impaired individuals, shows better predictive ability for purely energetic maskers than for informational maskers. Finally, in Stage 3, the predictive model was used to set scientifically-based minimum hearing standards for the screening of new applicants. The proposed standards and screening protocol must be validated by the Ministry before implementation.

INTRODUCTION

A paper at the last ICBEN meeting provided an update on fitness standards for hearing-critical (HC) jobs (Laroche et al., 2011). Efforts in developing scientifically-based hearing standards and protocols for screening have been ongoing, and an example pertaining to Ontario's police constables is presented in this paper.

Police constables perform HC jobs. The Ontario Ministry of Community Safety and Correctional Services is reviewing the hearing standards of the Constable Selection System (CSS, 2004) to better define critical components of medical fitness essential to safe and effective police operations.

Policing tasks are detailed in a 1992 report by Hay Management Consultants, in which their importance and frequency is also rated. A 1995 report by Shaw Consulting describes the physical demands of policing duties and provides a set of recommendations for tests and standards. It should be noted that while auditory tasks are less comprehensively defined than visual tasks, they share several features: they are often performed in noisy environments and involve a number of auditory skills and abilities (speech communication, sound localization, and sound detection). If an individual lacks these skills in a sufficient amount, it may constitute a safety risk for this individual, as well as for fellow workers and the general public.

Scientific models have been developed in different workplace environments (Bhérer et al., 2005) that allow one to assess auditory skills of employees based on diagnostic measures of hearing such as pure-tone audiograms. The Constable Selection System currently uses such an approach (CSS, 2004). Three criteria (Criteria I, IA and II) are based on different pure-tone thresholds meeting a certain

cut-off value. Criteria II also includes a word discrimination test in quiet and in noise, but the cut-off value is somewhat arbitrary, and without experimental validation. Indeed, such tests often have little or undocumented accuracy or are unable to make sufficiently adequate predictions of real life auditory performance to be used for individual personnel actions regarding HC jobs.

An alternative approach is used in the current project. The HINT (Nilsson et al., 1994), a measure of functional hearing ability with adequate reliability and sensitivity, and with established norms, was selected for use in screening applicants for HC jobs. The HINT has previously been used for screening/evaluation purposes in various occupational settings (e.g., Goldberg, 2001; Laroche et al., 2005; Giguère et al., 2008; Vaillancourt et al., 2011; Montgomery et al., 2011). Empirically-based validation studies were then performed to establish the relationship between scores on the screening test and performance in laboratory simulations using audio recordings made in real environments in which HC tasks occur. Minimum performance requirements were then established based on proportional likelihoods of effective speech communication.

The goal of this project was to develop a valid, reliable and legally defensible hearing assessment standard and protocol for Ontario's constable applicants. The project was conducted in three steps: Stage 1 – Identification of HC tasks and environments; Stage 2 – Laboratory validation with real noise environments; and Stage 3 – Analyses of real-world noise environments and establishment of normative data.

It should be noted that a similar approach has been used to set standards for the state of California Corrections Standards Authority (Montgomery et al., 2011) and during a large-scale contract with Fisheries and Oceans Canada to set new hearing standards for Canadian Coast Guard (CCG) and Conservation & Protection (C&P) personnel (Laroche et al., 2005).

STAGE 1 – IDENTIFICATION OF HC TASKS AND ENVIRONMENTS

Identification and validation of HC tasks

In the 1992 report by Hay Management Consultants, the importance and frequency of 167 policing tasks are rated on a scale of 1 to 5. In a preliminary analysis, 128 tasks deemed HC were identified by the research group. As a result of a validation by police constables from four different police services across Ontario asked to indicate whether or not hearing was judged essential to the successful completion of each task, 11 additional tasks were identified, for a total of 139 HC tasks performed by police constables. To yield a manageable set of approximately 50 HC tasks for further analysis within a focus group, only tasks with a mean importance rating greater than 3.85 were selected.

Identification of environments, required hearing abilities and factors which can affect performance

To establish hearing standards, the underlying communication situations from which screening criteria are derived must correspond to the actual HC tasks in the workplace. Five Subject Matter Experts (SMEs), defined as experienced police constables with a minimum of 5 years working in the field or other individuals highly knowledgeable about the field operations and work environments (such as supervisory personnel), took part in a focus group meeting during which they were asked, for each of the previously identified HC tasks, to: 1) identify the functional hearing abilities used to perform the task (sound detection, sound recognition, sound

localization and speech communication), 2) rate the criticality of each hearing ability in successfully completing the task, 3) identify the environments in which the task is performed, 4) specify the communication/listening parameters associated with the task at hand (for example the maximum distance from the sound/speech source, the minimal vocal effort of the talker, the ability or not to repeat a command), and 5) determine if some complicating factors (time-criticality, multitasking, equipment interference, etc.) make the task more difficult. In doing so, they were asked to think about typical situations in which the task is performed.

STAGE 2 – LABORATORY VALIDATION WITH REAL NOISE ENVIRONMENTS

Noise recordings

A total of 221 noise recordings were carried out in various environments and during simulation exercises at the Ontario Police College in Aylmer. In an attempt to classify the sound files into a manageable and restricted set of noise environments for use in the laboratory validation studies, all files were grouped into 8 categories (and 25 subclasses): 1) Freeway, 2) Inside Public Spaces, 3) Outdoor Public Spaces, 4) Public Transport, 5) Urban Street, 6) Police Incident: Domestic/Indoor, 7) Police Incident: In Vehicle, and 8) Police Incident: Outside.

A second focus group meeting was held with 4 SMEs to validate the noise recordings and the above classification, and to identify in which of these environments the HC tasks identified and validated during Stage 1 are performed. The SMEs were also asked to rate the proportion of time spent in each subclass within a given environment and also across the various environments as a whole. This exercise was also repeated using ratings of importance (criticality) of the hearing component in the accomplishment of tasks. It was noted that while some policing activities seldom occur in a given environment subclass yet require high levels of hearing for safe and accurate performance, others are frequent but require lower levels of hearing abilities.

Laboratory validation

The goal of the laboratory validation was to verify that HINT screening is predictive of the likelihood of effective communication with speech in real-world noise environments, and to demonstrate that it is equally predictive for normally hearing and hearing-impaired subjects with mild to moderate sensorineural hearing loss. The method chosen for validation is based on the Speech Intelligibility Index (SII) (ANSI S3.5-1997 R2012), more specifically the Extended SII (ESII: Rhebergen & Versfeld, 2005) since very few of the real-world noise environments where constables must perform HC job tasks are stationary. While ESII calculations have yet to be standardized, they have been used successfully to predict speech intelligibility and speech reception thresholds in non-stationary real-world noise environments (Rhebergen et al., 2006; Rhebergen et al., 2008; Soli, 2012).

The likelihood of effective speech communication in a particular real-world noise environment is defined as the proportion of time during which the value of the ESII equals or exceeds the ESII needed by an individual to achieve a specified level of intelligibility. For example, an individual with normal hearing may require an ESII of 0.34 to achieve 80% intelligibility, which is defined as effective speech communication. If short time intervals in a real-world noise environment have ESII values that exceed 0.34 only 50% of the time, the likelihood of effective speech communication for that individual in that noise environment is 50%.

Twenty-four individuals with normal hearing and five individuals with sensorineural hearing loss ranging in degree from mild to moderate took part in the validation study. In the standard 65-dBA HINT noise, a 20-sentence list was used to determine the speech reception threshold (SRT) [signal-to-noise ratio (SNR) for 50% correct sentence recognition] and to estimate percent correct word identification at threshold (usually about 70%). This score, along with air-conduction pure-tone thresholds, were entered in a Matlab routine developed to predict speech intelligibility in the stationary standard HINT noise and provide corresponding ESII values. Three other sentence lists were then presented at various fixed SNR: 1) at threshold, 2) at a SNR predicted to provide 50% word intelligibility, and 3) at a SNR predicted to provide 25 % word intelligibility. After each list, the actual percent correct intelligibility score was calculated and entered into the software, which provided the target SNR to be used next, and its corresponding ESII value. Finally, an additional list was used to measure the SRT and estimate word identification at threshold a second time. Generally, the first SRT value was used in plotting the SII-Intelligibility function, unless it was poorer than the second SRT value, in which case the latter was used. From the SII-Intelligibility function, ESII values corresponding to 25, 50 and 75% correct word intelligibility could be extrapolated. These estimated ESII targets were then used to validate the predictive model in the real-world noise recordings.

The average SRT for the 24 subjects with normal hearing was -2.4 dB SNR with a standard deviation of 0.8 dB, as compared with the norm of -2.6 dB SNR with a standard deviation of 1.0 for the American English HINT when speech and noise are presented frontally (Vermiglio, 2008). The close agreement of the SRTs in the current study with the published norms in the same experimental condition confirms the consistency of the HINT test materials and protocol.

For the standard HINT noise, average word intelligibility at the SRT was 73.3%, as estimated during the adaptive threshold measurement. The average word intelligibility, as measured at the SRT, was 78.8%, which is within the measurement error of the HINT (Vermiglio, 2008). Average intelligibility at the 50% target was 50.7%, and 21.6% at the 25% intelligibility target. When these measures were used to define the individual SII-Intelligibility functions for subjects with normal hearing, the range of SII values predicting the same target intelligibility was approximately 0.10, again indicating the consistency of the test materials and assessment protocol.

A total of 75 recordings across 12 environment subclasses were retained for the validation study. Four-sec noise samples were presented at their respective levels as recorded in the field, to make the listening tests as realistic as possible. However, all samples exceeding 85 dBA were previously excised to avoid any potential hazards of exposing subjects to loud noises and to limit startle reactions. For each target ESII value (25, 50 and 75%) a different 20-sentence list was presented in each of 2 distinct real-world noises, for a total of 6 noises per subject, and measured word intelligibility was recorded. During testing, the software tool uses an iterative process for adjusting the speech presentation level in each noise segment to obtain predicted intelligibility, taking into account: 1) the best ear pure-tone threshold at each frequency, 2) the SII-Intelligibility function, and 3) the characteristic ESII value of each noise segment.

To take into account the classification of noise recordings based on speech content or masker type in the analyses, all data pairs of predicted vs measured word intelligibility scores (144 data points) were grouped into three different categories (interference caused primarily by noise, interference caused primarily by babble, and

interference caused by clear speech). Within each of these categories, the distributions of predicted errors were computed for each target intelligibility (25%, 50% and 75%), after removal of the most extreme outliers (10%), and a linear regression function relating predicted and measured intelligibility was fit using the remaining data points. The slope and goodness of fit for the regression functions determine the accuracy with which intelligibility, and thus the likelihood of effective speech communication, can be predicted in real-world noise environments.

Figure 1 displays predicted and measured word intelligibility scores obtained in the three types of noise environments. Results show that SII-Intelligibility functions are most accurate in predicting intelligibility in real-world noise environments when the noise contains little or no speech. Panel a displays the results for environments where the interference was primarily produced by noise, with little babble or clear speech. In these environments variance accounted for (r^2) was high, 0.91, and the slope of the regression function relating predicted and measured intelligibility was close to unity, 0.99. If the SII-Intelligibility functions were perfectly predictive, the r^2 value and the slope would both be 1.00. The SII-Intelligibility functions also slightly over-predict intelligibility by 2.53%.

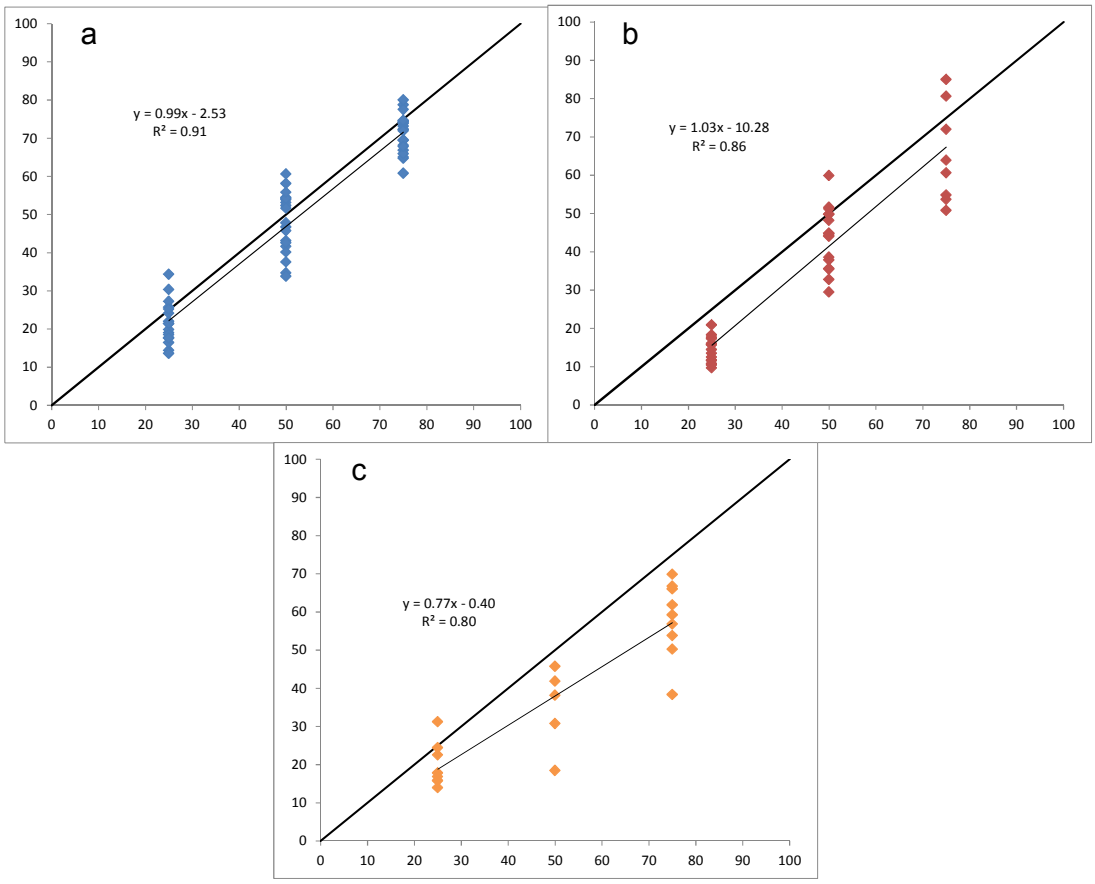

Figure 1: Predicted word intelligibility (abscissa) and measured word intelligibility (ordinate) for each type of noise environment. The most extreme 10% of prediction errors were eliminated from these analyses. Panel a displays the results for environments where the interference was primarily produced by noise, with no babble or clear speech. Panel b displays the results for environments where the interference was primarily babble, and Panel c displays the results for environments where the interference consisted of clear speech. Regression equations and r^2 values showing the relationship between predicted and measured intelligibility are displayed in each panel.

Panel b displays the results for environments where the interference was primarily produced by babble, with little noise or clear speech. In these environments intelligibility predictions based on the SII-Intelligibility functions were less accurate (r^2 = 0.86), especially for high levels of intelligibility, and the amount of over-prediction (10.28%) was consistently greater; although the slope of the regression function remained close to unity. This pattern of results is even more evident in Panel c, where interference was primarily produced by clear speech. Again, prediction errors are largest for moderate and high levels of intelligibility, and the slope of the regression function is significantly less than unity.

Five hearing impaired subjects were tested (3 mild and 2 moderate hearing losses) using the same protocol as used for subjects with normal hearing. Although the hearing impaired subjects were tested in all three types of interference, the primary results of interest are for the environments where the primary source of interference was noise without clear speech content. These data, although limited, show that the SII-Intelligibility functions for hearing impaired individuals, like those for individuals with normal hearing, can be used to accurately predict intelligibility in real-world noises relevant to law enforcement activities when the noise is not comprised of clear speech content. As with the results for normally hearing individuals, the r^2 value and the slope of the regression function are close to unity, and the amount of under-prediction is less than 1.0%.

Intelligibility in real-world noise environments relevant to law enforcement activities can thus be accurately predicted using a small set of intelligibility measures obtained in a stationary reference noise (HINT noise). These findings provide support for the use of SRTs and measures of speech intelligibility in noise for screening individuals who perform, or will perform, law enforcement activities in such noise environments. Note, however, that these statements apply primarily to real-world environments where the source of interference is noise that does not contain substantial linguistic content (e.g. vehicle and equipment noise, urban street noise, the unintelligible babble of large crowds at festivals or sporting events, the ambient noise in a busy bar or restaurant, etc).

When the source of interference is primarily clear speech with intelligible linguistic content, these screening measures are unreliable and over-predict intelligibility. Such findings are not unexpected, in that there is a substantial body of scientific literature that distinguishes the masking effects of interference that does not contain linguistic content (energetic masking) from the masking effects of interference with intelligible linguistic content (informational masking). Informational maskers cause more interference with intelligibility than energetic maskers that are equally loud. The SII-Intelligibility function is based on principles of energetic masking, which may explain why intelligibility is over-predicted by large amounts when clear speech is the primary source of interference. These findings suggest that screening methods based on measures of speech intelligibility in a stationary reference noise may not be appropriate for predicting intelligibility in most noise environments where clear speech is the primary source of interference. However, data from the current study show that clear speech interference usually occurs at moderate levels (60-70 dB LAeq) that may not reduce the likelihood of effective speech communication by a large amount. On the other hand, energetic maskers typically occur at higher levels that may reduce the likelihood of effective speech communication, even for individuals with normal hearing. One of the principles associated with setting appropriate hearing screening criteria is to ensure that individuals whose hearing

meets these criteria are able to perform HC tasks in challenging environments produced by relatively high level energetic maskers. Thus, the remaining analyses in support of the proposed screening criteria focus on the most challenging noise environments, which in most cases are produced by energetic maskers.

STAGE 3 – ANALYSES OF REAL-WORLD ENVIRONMENTS AND ESTABLISHMENT OF NORMATIVE DATA

ESII analyses of the sound recordings from the most challenging environments where HC job tasks must be performed by police officers were performed to estimate the likelihood of effective speech communication by individuals with normal functional hearing ability, as determined from the HINT SRT thresholds. These estimates were calculated using appropriate communication distance (1 meter) and levels of vocal effort (normal, raised, loud, and shouted). Previous work (Montgomery et al., 2011) suggests that effective communication will occur for this prototypical individual in all segments where the ESII value ≥ 0.30. This is also consistent with the findings of other studies (e.g. Houtgast & Festen, 2008). Thus, the proportion of segments whose ESII values are ≥ 0.30 defines the likelihood of effective communication for each environment.

Calculations were repeated for a range of elevated HINT SRT thresholds to determine how much the likelihood of effective speech communication decreased with increasing HINT thresholds. Previous research has shown that every 1 dB change in SNR near the HINT SRT results in a change in the ESII value of 0.03 (Amano-Kusumoto et al., 2012). These considerations suggest that a prototypical hearing impaired individual with a HINT SRT 1 dB above normal will require an ESII value ≥ 0.33, rather than 0.30. The proportion of segments with ESII values ≥ 0.33 defines the likelihood of effective communication for this individual. Proportional likelihoods in relation to the likelihoods for individuals with normal functional hearing ability can be established. For example, if the likelihood of effective communication for a normally hearing individual in a noisy environment is 0.80, and the likelihood for an individual with elevated HINT thresholds is 0.60, the proportional likelihood for this individual is 0.75, i.e., 0.60/0.80. In other words, this individual is expected to communicate with 75% of the effectiveness of a normally hearing individual under the same conditions.

The degree of risk that is acceptable for police constables must be determined by the Ministry of Community Safety and Correctional Services. For example, a greater than 20% reduction in the proportional likelihood of effective speech communication may be an unacceptable risk, especially given that even individuals with normal hearing may not always achieve high levels of effective speech communication in noise environments where the most challenging HC tasks are performed. The recommended screening criteria will be based on these considerations, and on the precedent established by other law enforcement and public safety organizations (e.g., Goldberg, 2001; Laroche et al., 2005; Montgomery et al., 2011).

Assuming this 20% cut-off, a hearing screening criterion based on the HINT Composite score, which is a weighted combination of the Noise Front, Noise Right, and Noise Left SRTs, can be established. In fact, following discussions with the Ministry of Community Safety and Correctional Services, two screening criteria will be proposed. A screening criterion defined by the Composite SRT in noise is based on the need for effective speech communication in background noise environments where challenging HC job tasks are performed, while a screening criterion based on

the SRT in quiet addresses the additional need to understand soft and whispered speech, as well as speech originating from behind doors or through windows. A screening protocol will also be established in collaboration with the Ministry, with options to include or not audiometric pure-tone threshold testing in addition to HINT testing in quiet and in noise.

REFERENCES

Amano-Kusumoto A, Soli SD, Yoon Y-S, Pal I (2012). Predicting speech intelligibility in real-world noise environments for screening of functional hearing ability. 2012 International Hearing Aid Research Conference (IHCON).

American National Standards Institute (2012). Methods for calculation of the speech intelligibility index. ANSI S3.5-1997 (R2012). New York: American National Standards Institute.

Bhérer L, Deshaies P, Fortier P, Laroche C, Tremblay V, Turcotte F (2005). Medical Examination of Hearing Required by Quebec Regulations: A Relevance Analysis. Canadian Acoustics 33:27-35.

Giguère C, Laroche C, Soli SD, Vaillancourt V (2008). Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test. Int J Audiol 47:319-328.

Goldberg RL (2001). Hearing Guidelines for the California Commission on Peace Officer Standards and Training. Sacramento, CA: California Commission on Peace Officer Standards and Training.

Hay Management Consultants and Dennis Strong & Associates. Police Constable Selection. Community Recruiting – Selecting Constables for the Future. Final Report Phase 1 – Volume 1. June 1992.

Houtgast T, Festen JM (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. Int J Audiol 47:287-295.

Laroche C, Giguère C, Vaillancourt V, Soli SD (2005). Development and validation of hearing standards for Canadian Coast Guard Seagoing Personnel and C&P and land-based personnel. Phase II, Final report to Department of Fisheries and Oceans under Contract No. F7053-000009.

Laroche C, Giguère C, Vaillancourt V, Soli SD (2011). Update on fitness standards for hearing-critical jobs. Proceedings of the 11th International Congress on Noise as a Public Health Problem (ICBEN), London, UK (24-28 July 2011).

Ministry of Community Safety and Correctional Services Policing Services Division. Selection Systems and Appointments Unit. Ontario's Constable Selection System. Hearing Performance Standard. March 2004.

Montgomery S, Soli SD, Meyers LS, Alvarez L, Hunley KA (2011). Hearing standards for selection of entry-level Correctional Officers. Sacramento, CA: Corrections Standards Authority, Standards and Training for Corrections Division.

Nilsson M, Soli SD, Sullivan JA (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.

Rhebergen KS, Versfeld NJ (2005). A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. J Acoust Soc Am 117:2181-2192.

Rhebergen KS, Versfeld NJ, Dreschler WA (2006). Extended speech intelligibility index for the prediction of the speech reception thresholds in fluctuating noise. J Acoust Soc Am 120:3988-3997.

Rhebergen KS, Versfeld NJ, Dreschler WA (2008). Prediction of the intelligibility for speech in real-life background noises for subjects with normal hearing. Ear Hear 29:169-172.

Shaw Consulting & N. Gledhill Consulting. Constable Selection Project. Report on Job Review Phase. Medical, Physical, Skills and Abilities. Recommendations, Tests and Standards. Volume #1 – Report. March 1995 (Revised 11/95).

Soli SD (2012). Use of the Extended Speech Intelligibility Index (ESII) and the Hearing In Noise Test (HINT) to quantify functional hearing ability. 6th International Symposium on Middle-Ear Mechanics in Research and Otology, sponsored by Kyungpook National University, Daegu, Korea.

Vaillancourt V, Laroche C, Giguère C, Beaulieu MA, Legault JP (2011).Evaluation of Auditory Functions for Royal Canadian Mountain Police Officers. J Am Acad Audiol 22(6):313-31

Vermiglio AJ (2008). The American English Hearing In Noise Test. Int J Audiol 47:386-387.