



Published in final edited form as:

J Occup Environ Hyg. 2016 December ; 13(12): 960–968. doi:10.1080/15459624.2016.1200723.

SPEECH INTELLIGIBILITY ASSESSMENT OF PROTECTIVE FACEMASKS AND AIR-PURIFYING RESPIRATORS

Andrew J. Palmiero¹, Daniel Symons², Judge W. Morgan III², and Ronald E. Shaffer¹

¹Centers for Disease Control and Prevention (CDC), National Institute for Occupational Safety and Health (NIOSH), National Personal Protective Technology Laboratory (NPPTL), Pittsburgh, PA, USA

²Centers for Disease Control and Prevention (CDC), National Institute for Occupational Safety and Health (NIOSH), Scott Safety Inc., Monroe, NC, USA

Abstract

Speech Intelligibility (SI) is the perceived quality of sound transmission. In healthcare settings, the ability to communicate clearly with coworkers, patients, etc. is crucial to quality patient care and safety. The objectives of this study were to 1) assess the suitability of the Speech Transmission Index (STI) methods for testing reusable and disposable facial and respiratory personal protective equipment (protective facemasks [PF], N95 filtering facepiece respirators [N95 FFR], and elastomeric half-mask air-purifying respirators [EAPR]) commonly worn by healthcare workers, 2) quantify STI levels of these devices, and 3) contribute to the scientific body of knowledge in the area of SI. SI was assessed using the STI under two experimental conditions: 1) a modified version of the National Fire Protection Association 1981 Supplementary Voice Communications System Performance Test at a Signal to Noise Ratio (SNR) of -15 (66 dBA) and 2) STI measurements utilizing a range of modified pink noise levels (52.5 dBA (-2 SNR) – 72.5 dBA ($+7$ SNR)) in 5.0 dBA increments. The PF models (Kimberly Clark 49214 and 3M 1818) had the least effect on SI interference, typically deviating from the STI baseline (no-mask condition) by 3% and 4% STI, respectively. The N95FFR (3M 1870, 3M 1860) had more effect on SI interference, typically differing from baseline by 13% and 17%, respectively for models tested. The EAPR models (Scott Xcel and North 5500) had the most significant impact on SI, differing from baseline by 42% for models tested. This data offers insight into the performance of these apparatus with respect to STI and may serve as a reference point for future respirator design considerations, standards development, testing and certification activities.

Keywords

healthcare; facemask; respiratory protection; communication; speech intelligibility

INTRODUCTION

Speech Intelligibility (SI) and clear communication among healthcare workers (HCWs) are vital components of healthcare delivery systems and occupational safety. This is especially true in times of public health emergencies and disaster situations. With approximately 13 million HCWs in the United States, clear communication in healthcare has a broad scope with significant impacts. SI is the quality or condition of speech being intelligible and clearly understood. It is well-documented that hospitals are often loud environments, which makes the issue of SI challenging. Hospital noise levels have been rising consistently since the 1960s. The background noise levels in hospitals rose from 57 dBA in 1960 to 72 dBA today during daytime hours, and from 42 dBA in 1960 to 60 dBA today during nighttime hours ⁽¹⁾. The World Health Organization (WHO) guidelines for hospital noise levels are 35 dBA during the day and 30 dBA at night in patient rooms, with recommended nighttime peaks of 40 dBA ⁽²⁾. Many studies indicate that peak hospital noise levels often exceed 85 dBA to 90 dBA ⁽³⁻⁸⁾. Noise from alarms and certain equipment that exceeds 90 dBA (for example, portable X-ray machines) are comparable to walking next to a busy highway when a motorcycle or large truck passes. These increased noise levels can pose significant challenges to communication among hospital staff and with patients. Busch examined hospital noise levels reported in 35 published research studies over the last 45 years ⁽¹⁾. It was found that not one published study reported noise levels that complied with the WHO guidelines for noise levels in hospitals ⁽¹⁾.

One challenge to SI in the healthcare setting is the increased use of facial and respiratory personal protective equipment (PPE) such as N95 filtering facepiece respirators (N95 FFR), protective facemasks (PF) and elastomeric half-mask air-purifying respirators (EAPR) necessary for worker and/or patient protection. Facial and respiratory PPE covering the mouth and nose have been reported to diminish SI ^(9,10) and impair the wearer's verbal communication ability by attenuating sound transmission, and reducing intelligibility because of muffled speech and impingement on the nasal alae ^(11,12). Compromised speech related to respirator use can occur at ambient noise levels as low as 40 dB(A) ⁽¹³⁾. Anecdotal reports of communication problems among HCWs wearing commonly used facial and respiratory PPE have led to suggestions that respirators may interfere with occupational duties ⁽¹⁴⁾. Some types of respirators have also been found to restrict articulation of the mandible ⁽¹⁵⁾. In a survey of hospital staff from Toronto, Ontario, Canada after the Severe Acute Respiratory Syndrome (SARS) outbreak in that city, 47 percent of the 2,001 respondents related that wearing of facial and respiratory PPE was associated with communication difficulty ⁽¹⁶⁾.

As part of Project BREATHE (Better Respiratory Equipment Using Advanced Technologies for Healthcare Employees), SI is designated as one of the 28 desirable performance characteristics to improve respiratory protective equipment so that "respirators should not impede, and preferably improve, the ability of others to hear the wearer's spoken words" ⁽¹⁷⁾. Project BREATHE seeks to develop respirator test methods, respirator standards for a "B95"¹ respirator class, and invent novel design features, which together will lead to commercial products that are more comfortable and useable for HCWs ⁽¹⁸⁾. Although a priority within Project BREATHE, the state of test methods for occupational interference

(which includes SI) was recently deemed not practical for inclusion in a draft B95 standard⁽¹⁹⁾. The Institute of Medicine (IOM) has also identified communication interference as a key issue to be studied and rectified in the next generation of facial and respiratory PPE developed for HCWs⁽¹⁴⁾.

Previous studies involving SI have primarily employed the use of the Modified Rhyme Test (MRT) in clinical settings or controlled listener/speaker orientations⁽²⁰⁾. The MRT is standardized under the American National Standards Institute (ANSI) S3.2 1960 and involves a listener attempting to properly identify certain key words preceded by a carrier sentence. For example, Radonovich et al. found that the odds of correctly hearing a word spoken by a HCW wearing an EAPR in an Intensive Care Unit (ICU) setting was approximately 0.46, compared with other commonly used respirators in the same location, on average⁽²⁰⁾. However, the MRT involves human test subjects, causing results to be variable and validity to be diminished based on each individual's unique variations in speech and the way listeners interpret those variations in speech. Sample size has also been small in previous studies, which may have decreased reliability⁽²⁰⁾. Human subject testing is also time-consuming, expensive, and often leads to incomplete data sets due to test subject attrition. Furthermore, the spoken words in the MRT are not industry specific, which plays a vital role in a field such as healthcare where highly technical words and sentence structures are used, as well as numerous acronyms.

Recently, an alternative method for assessing speech transmission, called the Speech Transmission Index (STI), has been applied to respirators and included in a voluntary consensus standard for Self-Contained Breathing Apparatus (SCBA) by the National Fire Protection Association (NFPA) in their 1981 Standard on Open-Circuit SCBA for Emergency Services (2013 edition). STI testing is a quantitative and objective method of speech transmission quality assessment in the presence of ambient (pink) noise. The STI was developed in the 1970s and is standardized under the International Electrotechnical Commission (IEC) Standard 60268-16. In contrast to subjective methods used to measure SI, the STI offers the advantages of being objective, highly repeatable, reproducible and time-efficient. A single value is calculated by the Modulation Transfer Function (MTF) using the signal-to-noise ratios per octave band with weightings that reflect the SI. The STI uses seven octave bands in the range of 125Hz to 8000Hz.

The objectives of this collaborative study between the National Institute for Occupational Safety and Health (NIOSH) (Pittsburgh, PA) and respirator manufacturer Scott Safety Inc. (Monroe, NC) were to 1) assess the suitability of the STI method in the NFPA 1981 standard for testing facial PPE commonly worn by HCW, 2) quantify STI levels of these devices, and 3) contribute to the scientific body of knowledge for Project BREATHE and other research in the area of SI.

¹“B95” or “biological N95” connotes protection against biological particulates as described in <http://www.publichealth.va.gov/docs/cohic/project-breathe-report-2009.pdf>

MATERIALS AND METHODS

Facial and Respiratory PPE Selection

The facial and respiratory PPE models and sizes selected and used in this study are summarized in Table I. Three types of facial PPE were tested: 1) N95 FFR, 2) PF and 3) EAPR. Each model was selected based on U.S. market share, ubiquity in healthcare workplace settings, use in previous studies⁽²⁰⁾, and tolerability among healthcare workers based on previous occupational health studies^(21,22). SI interference of these PPE have not been previously investigated to this degree under such controlled conditions as utilized in this study. Furthermore, SI interference of these types of PPE has not been extensively investigated utilizing a quantitative methodology such as the STI. It must be noted that the EAPR models tested did not contain voicemitters (also known as speaking diaphragms).

Modifications to NFPA 1981 standard

The STI component of the NFPA 1981 standard method was developed for tight-fitting, full facepiece SCBA respirators and thus is in need of validation for use on other respirator types. To assess the method for respirator and facial PPE use it was desirable to examine both repeatability measurements within each sample donning, and reproducibility measurements between each sample donning. The NFPA method requires three facepiece samples, each of which is donned on the manikin five times, with three measurements taken during each donning, for a total of 45 measurements. To assess repeatability within each donning, it is desirable to have a sample size larger than three repeats. Ideally, 28 measurements per sample and donning would need to have been performed in order to achieve the desired statistical power of 0.90. This would have led to a total of over 1,500 measurements for the six models tested. However, due to time constraints, this power level was not within the scope of this study. In an effort to increase the number of repeat measurements, the current study used three respirator samples, which were donned three times with five measurements taken during each donning. Utilizing this type of measurement protocol allowed for a better assessment of repeat measurements in a consolidated timeframe, while still maintaining a sufficient quantity of replicates.

Test System Setup

The test system setup conforming to the NFPA 1981 standard for SI testing was used for all experimental conditions. Testing took place at Technicon Acoustics (Concord, NC) inside a hemi-anechoic chamber. The test setup schematic is shown in Figure 1 per the NFPA 1981 Standard. Outside the hemi-anechoic chamber, a computer (Dell, Round Rock, TX), two graphic equalizers (Behringer-Willich, GE), pink noise generator and acoustic analyzer (NTI Audio AG, Schaan, LS) and STI test output signal were used. Inside the hemi-anechoic chamber, a K.E.M.A.R HATS Manikin (GRAS, Holte, DE), pink noise speaker and microphone (NTI Audio AG) were used. Prior to initiating each test, each model containing a nasal bar (3M 1860, 3M 1870, KC Technol 49214 and 3M 1810F) was molded to the K.E.M.A.R manikin in the same manner that a human would in order to achieve the best fit possible, not a protective fit. Figure 2 further depicts the STI testing setup inside the hemi-anechoic chamber and specifies equipment distances. All equipment was set up and

calibrated conforming to the NFPA 1981 standard each day prior to testing by staff familiar with the testing setup, system and calibration of equipment.

Speech Intelligibility Assessment

SI was measured for each PPE sample utilizing the STI under two different experimental conditions: 1) STI measurements using a modified version (3 samples \times 5 replicate measurements) of the NFPA 1981 standard and a signal to noise ratio (SNR) of -15 (66 dBA) and 2) STI measurements utilizing modified pink noise levels of 52.5 dBA (-2 SNR) to 72.5 dBA ($+7$ SNR) in 5.0 dBA increments to further characterize the SI of each facepiece. Three donnings of each facial PPE sample were performed under each experimental condition with five replicate measurements per donning. To ensure consistency, the same investigator performed each donning. Each STI value was measured directly in front of the K.E.M.A.R HATS Manikin at 1.5m per NFPA 1981. This distance was re-measured and readjusted inside the hemi-anechoic chamber when necessary to ensure the utmost accuracy (Figure 2).

Five STI measurements were taken for each sample ($n=3$) of each facepiece model (facepiece model $n=6$) per background noise decibel range (52.5, 57.5, 62.5, 67.5 and 72.5 dBA). This resulted in 25 STI measurements per sample (5 measurements \times 5 background ranges) for a total of 75 STI measurements (25 measurements \times 3 samples) per model.

DATA ANALYSIS

Data analysis was completed using Statistical Analysis Software (SAS) and Microsoft Excel. For this study it was desirable to examine both repeatability (defined as variability in measurements within each sample donning) and reproducibility (defined as variability in measurements between sample donnings).. The NFPA method requires three face piece samples, each of which is donned on the manikin five times with three measurements taken during each donning for a total of 45 measurements. To assess repeatability within each donning it is desirable to have a sample size larger than three repeats, ideally 28 measurements per sample and donning would have been performed in order to achieve the desired power of 0.90 (the implemented modification had a negligible impact on the power). This would have lead to a total of over 1500 measurements for the six models under test. The modified version of the NFPA method was utilized in order to yield more STI measurements per facepiece sample to assess the variability and repeatability. An analysis of variance (ANOVA) was performed for each respirator model comparing the means by sample (3 samples per model) and donning (3 donnings per sample) as well.

RESULTS

Repeatability & Reproducibility

To assess the repeatability of measurements, standard deviations for this testing were compared against two sets of data; one from MRT testing performed by Coyne and Barker and another STI study performed by Symons^(23,24). For the STI method to be considered a suitable SI measurement for facial and respiratory PPE worn by HCW, the standard deviation should be less than those found by Coyne and Barker using the MRT to evaluate

full facepiece air-purifying respirators⁽²⁴⁾. The standard deviations found during the tests should also be similar to those found by Symons⁽²³⁾ when testing SI of full facepiece respirators using the NFPA 1981:2013 method. At the time of this study, no STI testing had been done on the respirator models utilized in this study, so the data from full facepiece respirators was used as a comparison of the means and standard deviations. This is illustrated in the data shown in Table II.

To assess reproducibility of the measurements, the means of each sample were compared between donnings of the same sample. An analysis of variance (ANOVA) was performed for each full facepiece respirator tested comparing the means by sample and donning. The differences in means between each donning were again compared to data from full facepieces tested using the NFPA method from Symons⁽²³⁾. The differences in means should be similar to data from Symons⁽²³⁾. Upon closer examination of the standard deviations in Table III of each sample donning, it is noted that the largest standard deviation is 0.026 (3M 1860), with the average deviation across all models, samples, and donnings being 0.012. A chi-square test for SD less than 0.03 shows that the standard deviations of 41 of the 54 sample donnings were less than 0.03 STI or 3% of full scale with a 90% confidence interval.

From Coyne and Barker, the average standard deviation among listeners across all 12 full facepiece samples using the MRT method was 9.9%⁽²⁴⁾. From Symons, the average standard deviation across all three facepieces and 7 test runs was 0.014 STI when using the NFPA STI method⁽²³⁾. The standard deviations found in this study (0.012 STI) were less than those found by Coyne and Barker using MRT (9.9%) and similar to those found by Symons (0.014 STI)^(23,24).

Day 1 Testing—For the first experimental condition, the difference in means were assessed using a one-way ANOVA for each model sample. For both N95 FFR models (3M 1860 and 1870), the data from each sample donning is not statistically different ($p > 0.05$). Similarly, for the PF (KC 49214 and 3M 1818), the data from each sample donning can be considered statistically equivalent with $p > 0.05$. For the EAPR's (North 5500 and Scott Xcel), only sample 3 in each case was found to be statistically equivalent. For the North 5500, the maximum difference in means between donnings was 0.052 for sample 1 and 0.042 for sample 2. For the Scott Xcel, the maximum difference in means between donnings was 0.024 for sample 1 and 0.052 for sample 2.

Day 2 Testing—The graph in Figure 3 shows data from the second experimental condition (STI measurements utilizing modified pink noise levels of 52.5 dBA (-2 SNR) to 72.5 dBA (+7 SNR) in 5.0 dBA increments. The graph in Figure 3 illustrates that significant differences in SI exist between the different classes of PPE tested and the effect of background noise. This is shown in the boxplot illustrating STI scores by model in Figure 4 at a background noise level of 57.5 dBA (-15 SNR). Overall, the PF (Kimberly Clark 49214 and 3M 1818) had the least effect on SI interference, typically deviating from the STI baseline (no-mask condition) by 3% and 4% STI for the Kimberly Clark 49214 and 3M 1818, respectively. The N95FFR had more effect on SI interference, typically differing from baseline by 13% for the 3M 1870 and 17% for the 3M 1860. The EAPR's had the most

significant impact on SI, differing from the baseline by 42% for the Scott Xcel and 45% for the North 5500.

DISCUSSION

The STI score ranges from 0 to 1.0 as shown in Figure 5 and predicts the likelihood of words and sentence comprehension. Table IV illustrates the relationship between the STI, subjective intelligibility measures, and intelligibility ratings. A value of 0 is equivalent to a very poor SI, while 1.0 is an ideal channel with excellent speech transmission. The products tested in this study fall into the poor/fair (EAPR), good (N95 FFR), and excellent (PF) ranges.

From the results of both experimental conditions, it is seen that PF have the least impact on SI under these test methodologies followed by N95 FFR and EAPR. PF yielded the highest STI scores during both experimental conditions, most likely due to the thinness and type of materials used in their design and that they do not adhere as firmly as N95 and EAPR to the face, allowing sound to dissipate through gaps in the PF-to-face contact area. However, unlike the N95 and EAPR, the PF is not intended for respiratory protection. From a design perspective, the potential gains that could be made to both the N95 FFR class and PF class in terms of SI would be incremental.

Some of the overall trends observed here (EAPR < N95 FFR < PF) agree with those from human subject studies⁽²⁰⁾. In one set of experiments, Radonovich et al.⁽²⁰⁾ reported that the North 5500 EAPR showed a speech intelligibility rating of 72% compared to an 85% average from six disposable N95 FFRs (including the 3M 1860 used in this study) and one powered air purifying respirator. However, unlike this study, word intelligibility between PFs and N95 FFRs were not found to be statistically different. According to Table IV, the STI value of 0.45 found for the North 5500, falls in the upper range of the “poor” quality category or lower range of “fair”, corresponding roughly to an intelligibility between 67% and 87%, that is consistent with a previous study which reported an average three foot MRT score of 73% and 88% (Tables II, III Radonovich et al.⁽²⁰⁾), depending upon the experimental conditions. Likewise, the one N95 FFR (3M 1860) common to this study (average STI = 0.71) and the Radonovich⁽²⁰⁾ study resulted in an average three foot MRT score of 91%, which falls in the middle of the “good” quality category.

Based on the data from this investigation, a minimum pass/fail criterion of ≥ 0.70 (0.70 STI) with respect to the NFPA STI standard is suggested as a baseline for current facial PPE, in the development of future B95 respirator standards⁽¹⁹⁾. Facial and respiratory PPE with an STI of ≥ 0.70 fall within the “good” quality range according to IEC 60268-16 (Table IV). Further support for setting this level as a baseline can be found when comparing the results of this study with the Radonovich et al. study⁽²⁰⁾. The 3M 1860 N95 FFR with an STI of 0.71 was not found to be statistically different from control (no facial PPE) using the MRT (see Table II, Radonovich et al.⁽²¹⁾). Thus, facial and respiratory PPE with STI values ≥ 0.70 likely have little negative impact on SI.

Beyond facial and respiratory PPE design factors, controlling background noise in the healthcare environment will be paramount in improving SI in future healthcare environments. This can be accomplished through architectural design improvements, engineering controls and quieter equipment. As shown in Figure 3, with increasing background noise SI decreases and the differences between the PPE types quickly shrinks. Furthermore, a HCW using either an N95 FFR or a PF loses little with respect to SI. Thus, reducing background noise will be one of the most effective ways to improve SI in healthcare environments.

STUDY LIMITATIONS

Limitations of this study include time, non-human testing, uni-directionality and difficulty to generalize findings based on the limited number of samples used for each PPE type. A non-human modality was undertaken for this study for the purposes of objectivity and time constraints. The uni-directionality of testing was performed in order for the results to be valid, consistent and reproducible. Testing in multiple angles or distances would considerably increase the amount of time and resources needed. The surgical masks used were a flat-fold design whereas the N95 FFRs were cup-shaped (3M 1860) and tri-fold (3M 1870), rendering them further away from the face when donned. No models tested contained exhalation valves or voicemitters, which may have an effect on SI.

FUTURE WORK

If SI becomes a performance requirement for future “B95” respirators, the STI methodology used in this study could be used for testing and certification requirements and standards development. Also, independent replication of this experimental work would provide validation of the STI method as a reliable methodology for testing different types of PPE. Data could also be used for benchmark testing and design improvements for manufacturers of facial and respiratory PPE. Further investigation into a more diverse range of models and sizes under different conditions (different directions, reverberation, etc.) would be valuable contributions as well. Future studies should identify the design features that could lead to EAPRs and N95 FFRs with STI scores in the excellent range, which would allow a higher B95 standard STI pass/fail criterion to be considered (e.g., 0.75) in the future.

CONCLUSIONS

The PF models studied here (Kimberly Clark 49214 and 3M 1818) had the least effect on SI interference, typically deviating from the STI baseline measurement (no-mask condition) by 3% and 4% STI. The two N95FFR models (3M 1870 and 1860) had a greater effect on SI interference compare to the PF models, typically differing from the STI baseline measurement by 13% for the 3M 1870 and 17% for the 3M 1860. The two EAPR models (Scott Xcel 742 and North 5500) had the most significant impact on SI, differing from the baseline STI measurement by 42% for the Scott Xcel and 45% for the North 5500 shows the sounds pressure levels per octave band for each respirator model during the testing with the background noise at 52.5 dbA (-15 SNR).

Figure 6 illustrates why the EAPRs performed worse than the N95's and surgical masks. Examining the sound pressure level in the 1kHz to 4 kHz range, it is noted that the sound pressure level for both EAPR models is lower than baseline by approximately 10 dBA at 1kHz, 20 dBA at 2kHz and 15 dBA at 4kHz. This indicates that the STI signal sound pressure level has been significantly attenuated by the EAPR facepiece in these frequency ranges. Furthermore, at the 2kHz and 4 kHz frequencies, the sound pressure level of both EAPR models is nearly equivalent to the background noise level. Thus, the STI signal is masked by the background noise and intelligibility in these key frequency ranges is significantly degraded. To a lesser degree the same observation can be made about the N95 respirators (3M 1860 and 3M 1870).

The thinness and thickness of the materials used in the construction of the respirators and the materials themselves (e.g. silicone vs. polypropylene) may also be contributing factors to the variation in STI scores between respirator models and samples. The rigidity, type and number of materials used in the speech path are all contributing factors to the STI score. These factors should be taken into consideration and investigated further as the next generation of facial PPE is developed.

ACKNOWLEDGEMENTS/DISCLAIMER

The authors would like to acknowledge Dr. Raymond Roberge, Jay Parker, and Michael Parham for their helpful comments and suggestions. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health. Mention of any product name does not constitute endorsement by the National Institute for Occupational Safety and Health.

REFERENCES

1. Busch I, West J, Barnhill C, Hunter T, Orellana D, Chivukula R. Noise levels in Johns Hopkins Hospital. *J. Acoust. Soc. Am.* 2005; 118(6):3629–3645. [PubMed: 16419808]
2. Berglund, B.; Lindvall, T.; Schwelaand, DH.; Goh, TK. Protection of the human environment. World Health Organization; Geneva, Switzerland: 1999. Guidelines for community noise.
3. Aaron JN, Carlisle CC, Carskadon MA, Meyer TJ, Hill NS, Millman RP. Environmental noise as a cause of sleep disruption in an intermediate respiratory care unit. *Sleep.* 1996; 19(9):707–10. [PubMed: 9122557]
4. Balogh D, Kittinger E, Benzer A, Hackl JM. Noise in the ICU. *Intensive Care Medicine.* 1993; 19(6):343–46. [PubMed: 8227725]
5. Blomkvist V, Eriksen CA, Theorell T, Ulrich RS, Rasmanis G. Acoustics and psychosocial environment in coronary intensive care. *Occupational and Environmental Medicine.* 2005; 62:1–8.
6. Cureton-Lane RA, Fontaine DK. Sleep in the pediatric ICU: An empirical investigation. *Am. J. Crit. Care.* 1997; 6(1):56–63. [PubMed: 9116788]
7. Guimaraes H, Oliveira AM, Spratley J, Mateus M, d'Orey C, Coelho JL, Souto A, Santos NT. The noise in neonatal intensive care units. *Archives de pédiatrie.* 1996; 3(11):1065–68. [PubMed: 8952768]
8. Holmberg SK, Coon S. Ambient sound levels in a state psychiatric hospital. *Archives of Psychiatric Nursing.* 1999; 13(3):117–26. [PubMed: 10389340]
9. Abeysekera JD, Shahnava H. Ergonomics assessment of selected dust respirators: Their use in the tropics. *Appl. Ergon.* 1987; 18(4):266–272. [PubMed: 15676630]
10. Salazar MK, Connon C, Takaro TK, Beaudet N, Barnhart S. An evaluation of factors affecting hazardous waste workers' use of respiratory protective equipment. *AIHA J.* 2001; 62(2):236–245.
11. Harber P, Beck J, Luo J. Study of respirator effect on nasal-oral flow partition. *Am. J. Ind. Med.* 1997; 32:408–412. [PubMed: 9258395]

12. Lubker, JF.; Moll, KL. Simultaneous oral-nasal air flow measurements and cinefluorographic observations during speech production; Paper presented at American Speech and Hearing Association Convention; Chicago, IL. Nov 3-6. 1963
13. Caretti DM, Strickler LC. Speech intelligibility during respirator wear: influences of respirator speech diaphragm size and background noise. *Am. Ind. Hyg. Assoc. J.* 2003; 64:846–850.
14. IOM (Institute of Medicine). Preventing Transmission of Pandemic Influenza and Other Viral Respiratory Diseases: Personal Protective Equipment for Healthcare Workers. Larson, EL.; Liverman, T., editors. National Academies Press; Washington, DC: 2010.
15. Defence Scientific and Technical Laboratory. The UK General Service Respirator. Wetherell, A., editor. Porton Down, Salisbury SP4 0JQ, United Kingdom: 2003.
16. Nickell LA, Crighton EJ, Tracy CS, et al. Psychosocial effects of SARS on hospital staff: survey of a large tertiary care institution. *Can. Med. Assoc. J.* 2004; 170:793–798. [PubMed: 14993174]
17. Radonovich, L.J., Jr.; Roberge, R.; Baig, A.; Levinson, A.; Shaffer, RE.; Doerr, DF.; Davey, V. Better respirator equipment using advanced technologies for healthcare employees (Project B.R.E.A.T.H.E.): An interagency working group of the U.S. federal government. Office of Public Health and Environmental Hazards, Veterans Health Administration, U.S. Department of Veterans Affairs; Washington, DC: 2009.
18. Gosch M, Shaffer R, Egan A, Roberge R, Radonovich L. B95: A new respirator for health care personnel. *Am J Infect Control.* 2013; 41:1224–30. [PubMed: 23726655]
19. Shaffer RE, Zhuang Z, Bergman M, Fisher E, Palmiero A, Roberge R, Williams WJ, Kim JH, Sinkule E, Radonovich L, Gosch M, Egan A. Recommended requirements, test methods, and pass/fail criteria for a “B95” respirator for healthcare workers. *J. Int. Soc. Respir. Prot.* 2014; 31(1):23–42.
20. Radonovich LJ, Yanke R, Cheng J, Bender B. Diminished speech intelligibility associated with certain types of respirators worn by healthcare workers. *J. Occup. Environ. Hyg.* 2010; (7):63–70. [PubMed: 19904661]
21. Radonovich L, Cheng J, Shenal B, Hodgson M, Bender B. Respirator tolerance in health care workers. *JAMA.* 2009; 301(1):36–38. [PubMed: 19126810]
22. Roberge R, Coca A, Powell J, Palmiero A. Physiological impact of N95 filtering facepiece respirators on healthcare workers. *Resp. Care.* 2010; 55(5):569–577.
23. Symons, D. Use of Speech Transmission Index (STI) for Voice Intelligibility Testing of Respiratory Protection Devices and Proposed Measurement Method; Presented at ISRP conference; Boston, MA. 2012.
24. Coyne, K.; Barker, D. Automated Speech Intelligibility System for Head-Borne Personal Protective Equipment: Proof of Concept. Battelle Memorial Institute. Research and Technology Directorate; 2010. 20080424091

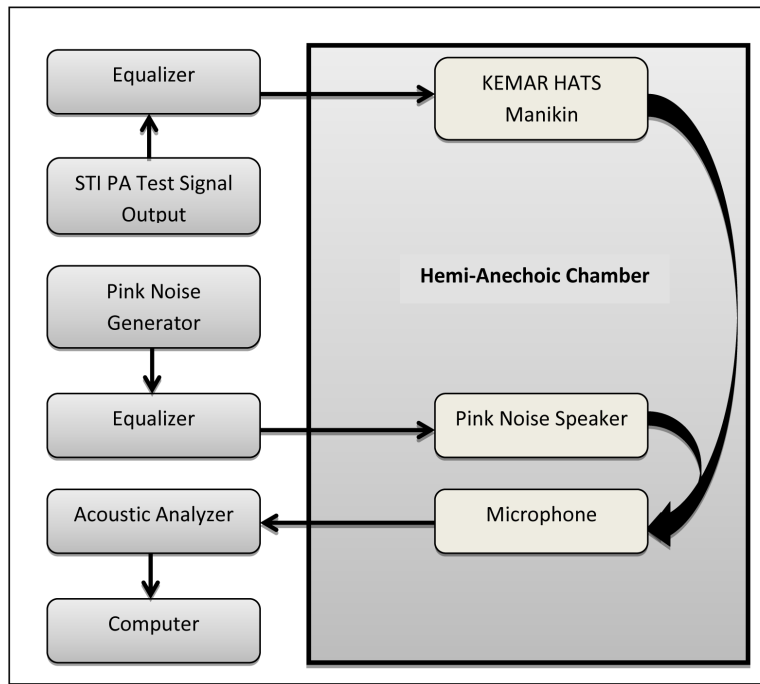


FIGURE 1. Test System Schematic

*Adapted from NFPA 1981: Standard on Open-Circuit Self-Contained Breathing Apparatus (SCBA) for Emergency Services, 2013 Edition.

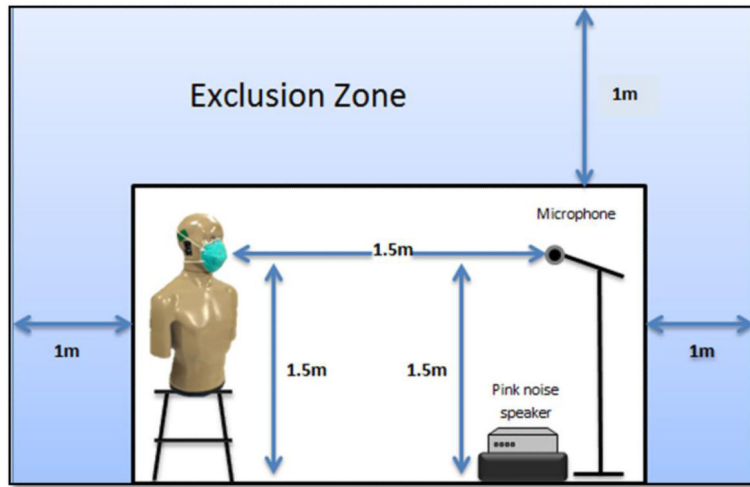


FIGURE 2.
STI Testing Setup inside Hemi-Anechoic Chamber.

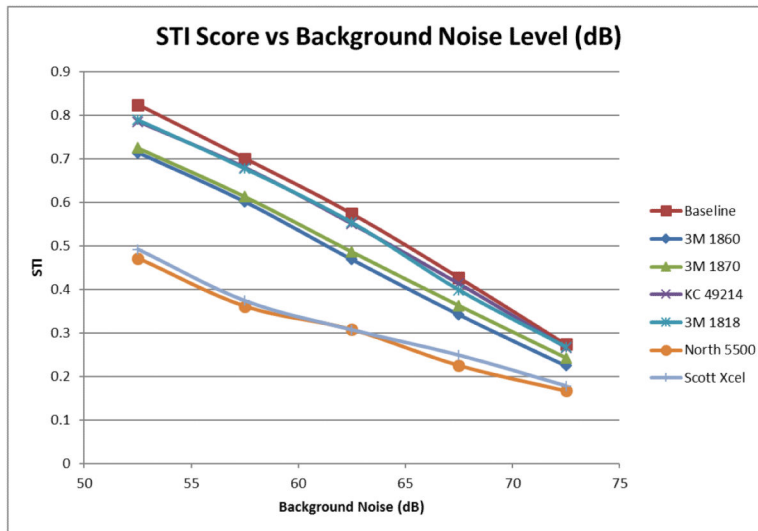


FIGURE 3.
2nd Experimental Condition.

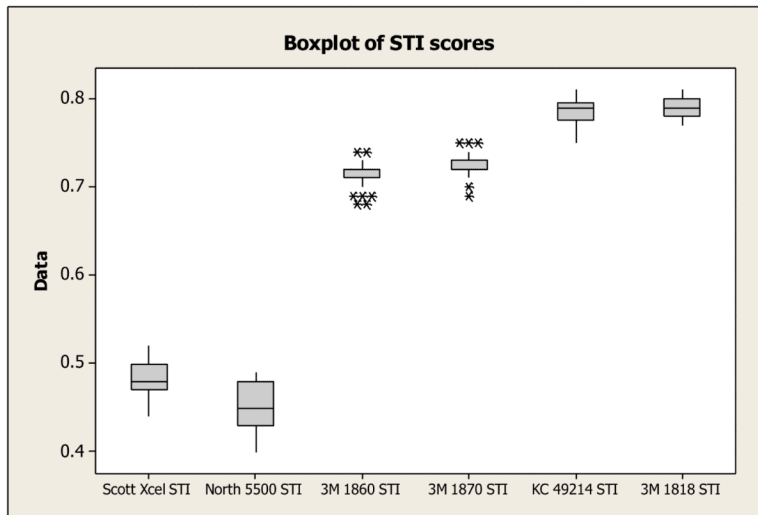


FIGURE 4.
Boxplot of STI Score by Model at -15 SNR.

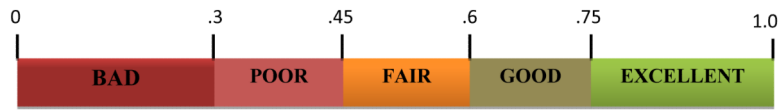


FIGURE 5.
Speech Transmission Index.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

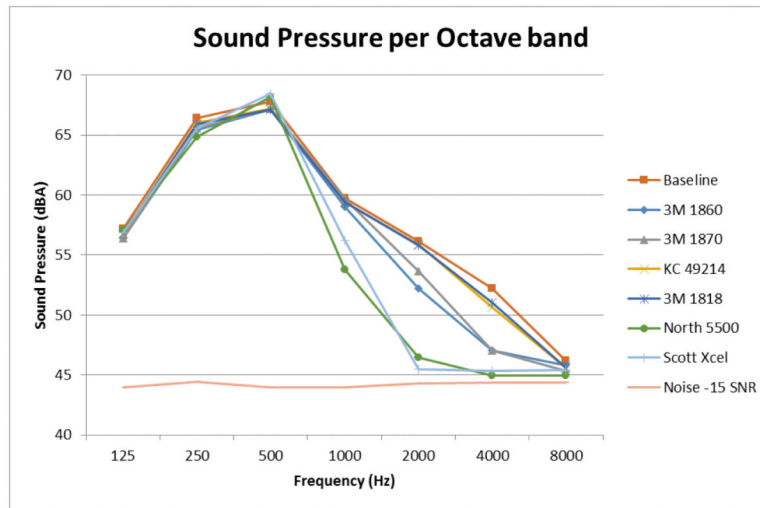


FIGURE 6.
Sound Press per Octave Band

TABLE I

Facial and Respiratory PPE Models.

Mask Type	Model	Size
N95 (FFR)	3M 1860 (70070612364)	Medium-Large
N95 (FFR)	3M 1870 (70071564523)	One Size
Surgical Mask	KC Tecnol 49214	One Size
Surgical Mask	3M 1810F (EN14683:Type II)	One Size
EAPR	North 5500	Medium (550030M)
EAPR	Scott Xcel 742 Half Mask	Medium/Large (7421-213)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE II

Comparison Data (MRT and STI).

Facepiece	Listener 1 Std Dev	Listener 2 Std Dev	Listener 3 Std Dev	Pooled Std Dev Across Listeners
3M FR-M40	8.0	11.6	6.0	8.5
Avon C50	10.2	10.9	3.5	8.2
Peltor M-TAC	7.4	5.8	9.8	7.7
Drager CDR 4500	6.8	5.7	6.5	6.3
MSA Millenium	8.5	2.6	9.4	6.8
Survivair Optifit	9.6	13.6	5.3	9.5
MSA Ultra Elite	5.0	1.9	10.0	5.6
Scott M120	5.4	5.5	7.0	6.0
North 5400	6.5	9.7	8.2	8.1
3M FR 7800B	7.4	7.9	11.6	9.0
Scott M110	8.3	4.4	9.4	7.4
Avon FM12	3.1	5.9	7.4	5.5
Pooled Std Dev				7.4

*Data aggregated from Coyne and Barker⁽²⁴⁾ Tables: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22,

Facepiece	Test Run	Mean	St Dev
1A	1	0.538	0.014
1A	2	0.555	0.012
1A	3	0.535	0.012
1A	4	0.514	0.018
1A	5	0.553	0.017
1A	6	0.540	0.012
1A	7	0.531	0.014
2A	1	0.535	0.024
2A	2	0.567	0.014
2A	3	0.536	0.012
2A	4	0.519	0.013
2A	5	0.553	0.012
2A	6	0.544	0.016
2A	7	0.539	0.015
3A	1	0.547	0.018
3A	2	0.550	0.013
3A	3	0.537	0.010
3A	4	0.533	0.010
3A	5	0.549	0.012

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Facepiece	Test Run	Mean	St Dev
3A	6	0.543	0.011
3A	7	0.525	0.011
pooled std dev			0.014
Pooled difference between means			0.016

* Data from Symons [2012] Table 2

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE III

Facial and Respiratory PPE Results.

Sample	Don	N	3M 1860		3M 1870		KC 49214		3M 1818		North 5500		Scott Xcel	
			Mean	St dev	Mean	St dev	Mean	St dev	Mean	St dev	Mean	St dev	Mean	St dev
1	1	5	0.72	0.015	0.72	0.007	0.78	0.011	0.80	0.005	0.44	0.018	0.48	0.007
1	2	5	0.71	0.005	0.72	0.016	0.78	0.020	0.79	0.011	0.47	0.009	0.49	0.012
1	3	5	0.72	0.016	0.71	0.014	0.79	0.007	0.78	0.012	0.42	0.013	0.47	0.005
2	1	5	0.72	0.008	0.72	0.005	0.79	0.017	0.80	0.011	0.43	0.016	0.47	0.018
2	2	5	0.71	0.015	0.74	0.008	0.77	0.010	0.78	0.009	0.42	0.004	0.50	0.011
2	3	5	0.71	0.026	0.72	0.011	0.79	0.015	0.79	0.015	0.46	0.012	0.45	0.018
3	1	5	0.70	0.018	0.73	0.012	0.79	0.015	0.79	0.015	0.48	0.004	0.48	0.015
3	2	5	0.72	0.005	0.72	0.011	0.78	0.013	0.79	0.008	0.46	0.015	0.48	0.021
3	3	5	0.71	0.007	0.73	0.015	0.79	0.008	0.79	0.011	0.48	0.017	0.51	0.017
Pooled Mean and Standard Deviation			0.71	0.013	0.72	0.011	0.78	0.013	0.79	0.011	0.45	0.012	0.48	0.014

TABLE IV

Relationship Between STI, Subjective Intelligibility Measures and Intelligibility Ratings.

STI Value	Quality according to IEC 60268-16	Intelligibility of Syllables in %	Intelligibility of Words in %	Intelligibility of Sentences in %
0 - 0.3	bad	0 - 34	0 - 67	0 - 89
0.3 - 0.45	poor	34 - 48	67 - 78	89 - 92
0.45 - 0.6	fair	48 - 67	78 - 87	92 - 95
0.6 - 0.75	good	67 - 90	87 - 94	95 - 96
0.75 - 1	excellent	90 - 96	94 - 96	96 - 100

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript