2009

Evaluating the effectiveness of fit-testing of hearing protection devices in northern West Virginia coal mines

Alexa Michele Quinn
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Quinn, Alexa Michele, "Evaluating the effectiveness of fit-testing of hearing protection devices in northern West Virginia coal mines" (2009). *Graduate Theses, Dissertations, and Problem Reports*. 2778.
<https://researchrepository.wvu.edu/etd/2778>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

**Evaluating the Effectiveness of Fit-Testing of Hearing Protection Devices in Northern
West Virginia Coal Mines**

Alexa Michele Quinn

**Thesis Submitted to the
College of Engineering and Mineral Resources
at West Virginia University**

**in partial fulfillment of the requirements
for the degree of**

Master of Science

in

Industrial Hygiene

**Steven Guffey, Ph.D., Chair
Warren R. Myers, Ph.D.
Mohammed Abbas Virji, M.Sc., Sc.D.**

Department of Industrial & Management Systems Engineering

**Morgantown, West Virginia
2009**

Keywords: Noise Reduction, Fit-Test, Hearing Protector, Mining Noise

ABSTRACT

Evaluating the Effectiveness of Fit-Testing of Hearing Protection Devices in Northern West Virginia Coal Mines

Alexa Michele Quinn

Fit testing is frequently described as a solution to the problem of detecting ineffective hearing protection devices and thus reducing noise exposures in workers. This study evaluates the effectiveness of a fit-test conducted at mine sites with cap mounted ear muffs actually worn by the tested miners. The dependent variable was noise reduction, and the independent variables included subject, noise source frequency distribution, orientation to noise source, instruction, and earmuff brand.

A series of eight 30-second tests were conducted for each worker to evaluate the effects of the independent variables. The tests were not randomized due to time constraints. Each of 30 workers was tested once for each test condition. The noise reduction (NR) was determined by setting up four dosimeters with one microphone on each shoulder and one at the entrance to each ear.

NR values for the right ear were modestly (3.3 dBA) but significantly higher ($p < 0.0002$) than for the left ear. The average of left and right ear NR values from fit-testing in this study showed substantial between-subjects variability (8.5 dBA) and within-subjects variability (ranging from 1.1 to 12.3 dBA). As expected, NR values were significantly ($p < 0.0001$) but modestly higher (0.3 dBA) for pink noise than for simulated underground noise. The mean NR for different brands of muffs varied from 13.4 dBA to 25.6 dBA with the overall effect of brand being highly significant ($p < 0.0004$). The lowest and highest visual fit categories of a fit rating had no false positives but many false negatives. The middle categories were uncorrelated to NR values. The orientation of the subject to the source had modest effects on NR (< 1 dBA) that were highly significant ($p < 0.008$) because of the large sample size. Brief instruction showed no significant benefits for these experienced users of earmuffs.

The findings here support multiple repetitions of individualized fit-testing for every employee.

TABLE OF CONTENTS

Abstract.....	ii
Table of Contents.....	iii
Table of Figures	iv
List of Tables.....	v
Introduction	1
Hearing Protection in the Mining Industry	1
Fit-testing Methods.....	4
Literature Review	6
MIRE, F-MIRE and REAT	6
Variability of Attenuation in HPDs	8
Apparatus.....	9
Methods and Procedures.....	13
Noise Sources	13
Using Dosimeters versus an Octave Band Analysis	13
Methodology for Test Subjects.....	14
Results Subjects 1-5 (Group A)	19
Results Subjects 6-30 (Group B)	21
NR Left versus NR Right.....	23
Within Subjects Variability for NR Values.....	27
Between Subjects Variability for NR Values.....	29
Earmuff Types	33
Visual Fit-rating	35
Orientation.....	39
Effects of Repositioning.....	41
Instruction.....	43
Conclusions	46
Caveats	48
References.....	49
Appendices	51

TABLE OF FIGURES

Figure 1 Noise spectra frequency distributions used for subject fit-test.....	9
Figure 2 Fit-test set up in mining office	11
Figure 3 Illustration of microphones used in study set up on a human subject	14
Figure 4 Illustration of adjustability of cap-mounted earmuffs.....	18
Figure 5 Plot of difference between NR Left and NR Right values	21
Figure 6 NR, average of Left and Right Ears versus Subject number	25
Figure 7 Distribution of average NR values for 5 dBA intervals for subjects 4-30	27
Figure 8 Image of visibly poor fit for Subject 27.....	35
Figure 9 Illustration of NR for the right ear versus the fit-rating for the right ear	37
Figure 10 Illustration of NR for the left ear versus the fit-rating for the left ear	38
Figure 11 NR values for before and after instruction	42
Figure 12 NR values for before and after instruction	44

LIST OF TABLES

Table 1	Description of earmuff types used in the study	12
Table 2	Explanation of visual fit evaluation rating applied to the first four tests.....	17
Table 3	Explanation of corrections made to Subject numbers.....	19
Table 4	Mean difference between NR left and NR right for Subjects 4 through 30	22
Table 5	Descriptive statistics for NR Left and NR Right.	23
Table 6	Average NR values (in dBA) for each subject for all tests	26
Table 7	Descriptive Statistics for NR values averaging all tests.	29
Table 8	Number in each range of median NR values	30
Table 9	T-test results for comparison of average NR values for two noise spectra.....	31
Table 10	Analysis of variance for noise spectra	31
Table 11	Explanations of conditions for each test	32
Table 12	Average noise reduction for each type of earmuff	33
Table 13	Analysis of variance for earmuff type for all tests with muffs on	34
Table 14	Average noise reduction for each fit-rating (1-5) for each ear	36
Table 15	Average NR results for the effects of orientation	39
Table 16	Analysis of variance for orientation and subject for all tests with muffs on	40
Table 17	Descriptive statistics for the difference between two fits.....	41
Table 18	T-test results for Test 2 versus Test 6 showing effects of re-fit.....	42
Table 19	NR values for before and after instruction (Group B, Tests 2 and 7)	43
Table 20	Comparison of Test 2 and Test 7 with respect to the effect of instruction	45

INTRODUCTION

Hearing Protection in the Mining Industry

Exposure to high levels of noise is routine for many underground mine workers due to the tasks and equipment associated with underground mining activities. The National Occupational Health Survey of Mining (NOHSM) conducted from 1984 to 1989 revealed “unacceptably” high risk of developing noise induced hearing loss (NIHL) at the current PEL of 90 dBA (as an 8-hour time weighted average, or TWA₈) and a “significant risk” at noise levels just below the TWA₈ (Joy, 2007). From 2000 to 2004 the mean dose percent of the PEL recorded by Mine Safety and Health Administration (MSHA) during inspections was 83.0% or 88.7 dBA (Joy, 2007). At this dose, there is still a risk of NIHL. As a result of the studies on NIHL and noise levels in mines, MSHA now requires workers to participate in a hearing conservation program if they are found to have exposures exceeding a 50% PEL dose or 85 dBA (Labor, 2000).

This study focuses on the effectiveness of a fit-test on cap-mounted earmuffs, which are commonly used in underground mines. The two participating mines in this study, a northern West Virginia mine and a southeastern Pennsylvania mine, had different requirements for hearing protection devices (HPDs).

The West Virginia mine required HPDs for all workers and provided ed earmuffs for employees. The mine operator provided mostly Peltor Optime 95 type cap-mounted earmuffs, and most of the volunteers from this mine used this type of muff (see Table 1 for types of muffs and Noise Reduction Rating). This muff requires drilling a hole in the miner’s helmet in order to mount the earmuff. All installation, maintenance and periodic inspections of the earmuffs were conducted on site by a designated mine employee. Earmuffs were usually replaced on an annual basis, depending on the frequency of use of the muff. The Pennsylvania mine did not have strict requirements for HPDs. In fact, most of the miners did not wear earmuffs on a regular basis. The volunteers who did not own earmuffs were provided cap-mounted muffs for the purposes of this study.

Hearing Protection Devices

There are several issues that can affect the selection of a HPD. One important decision factor in choosing the most effective HPD for a workplace is the frequency distribution of the noise exposures. The noise spectra used in this study include a pink noise frequency distribution and a frequency distribution similar to that found in underground mining operations. Most noise exposure in the mining industry is dominated by low frequencies (Durkt, 1993). This means that the noise levels at frequencies ranging from 200 to 800 Hz are greater than the noise levels at frequencies above 800 Hz. A similar spectrum was used in a study done by George Durkt when evaluating the effectiveness of HPDs in surface mining (Durkt, 1993). Generally, earmuffs provide better protection against higher frequencies (2000 to 8000 Hz) than low frequencies (125 to 2000 Hz). In frequencies from 125 to 2000 Hz, leakage of noise through the HPD occurs resulting in poorer protection. This can be attributed to the vibration of the cup that occurs at lower frequencies (Zannin, 2005).

Even the ideal HPD will not block all noise from transmitting to a worker (Berger, 1982). There are several pathways that sound can enter the body. Air conduction, or the movement of a sound wave that excites the cochlear hair cells in the ear canal, is the most common pathway for noise to come in contact with the ear. This is also the pathway in which an un-occluded ear would receive most of its noise exposure. The second pathway involves bone conduction. Bone conduction occurs when the sound waves come in contact with the body causing vibration of the skull or ear canal, which transfers to the inner ear and excites the cochlea. In order to block both methods of transferring sound to the ear, one would have to employ a device blocking sound from reaching the head entirely, not just the ear. Even then, some sound would transfer through the body to reach the ear. Berger, 1982, states that the decibel level reaching the ear due to bone conduction can be up to 60 dB (at 1 kHz) lower than when sound is transferred through air conduction. Adding a HPD to this scenario introduces two more pathways of conducting noise to the inner ear. In addition to air conduction and bone conduction, noise leakage can occur by penetrating the HPD material or by entering the enclosure of the HPD through leaks in the seal of the muff cup. The fit-ratings (see Table 2) in this study are based on the last pathway.

Another issue to consider when selecting a HPD is the Noise Reduction Rating (NRR). This rating system was developed by the Environmental Protection Agency (EPA) to help determine the effectiveness of the hearing protection device. The NRR is determined by the manufacturer and is displayed on HPD packaging to indicate the predicted attenuation from the HPD. This number is determined using the Real Ear Attenuation at Threshold (REAT) method for fit-testing of HPD (discussed below). However, many studies have shown that the NRR values do not show consistent accuracy in predicting attenuation, especially in the field (Durkt, 1993, Behar, 1985). The EPA is currently working on a new test method to increase the validity of the NRR system (Myers, 2008).

A major obstacle in finding true NRR for HPDs is the lack of field data due to the inconvenience that comes with measuring in the field. One of the methods commonly used to test the effectiveness of hearing protectors is to perform a fit-test. The two standard fit-test evaluation methods of HPDs include a REAT method and a Microphone in Real Ear (MIRE) method.

One of the challenges in evaluating hearing protector effectiveness in the field is the inconvenience of the noise measurement apparatus. Many instruments that are used to measure noise reduction (NR) are not intrinsically safe in an explosive atmosphere, will not provide accurate readings due to the physical nature of the tasks, or are too bulky to add on to the miner in a small space. For this reason, this study was conducted in onsite mining offices. This setting does not have diffuse sound fields, which is not compliant with the American National Standards Institute (ANSI) MIRE method (ANSI S12.42-1995). The ANSI 12.42-1995 requires a diffuse sound field for MIRE tests. However, the test set up is consistent with the methods used by other researchers applying the same method of fit-testing, and should not present a significant difference in the observed results (Casali, 1995).

Fit-testing Methods

Fit-testing of HPDs is not common in most workplaces today, but the idea of fit-testing to determine NR for the individual has been around since the 1980s (Berger, 2007). ANSI requires compliance with the REAT method for determining the attenuation and NRR of HPDs. This method measures the difference in hearing thresholds with and without hearing protection. In other words, the subject is given audiograms, with ears unoccluded, followed by a second audiogram with the subject wearing a HPD. The insertion loss (IL) is then determined from the difference in the hearing thresholds. This method requires an audiometric test booth or specialized headphone equipment, which can be expensive. The method also requires minimal background noise, which can be hard to obtain in many workplaces. In addition, the process of REAT is time consuming and the results rely considerably on the subject's capacity to identify his/her own hearing threshold, which can result in up to 5 dB variability (Berger, 2007). The limitations of REAT are a primary reason for the lack of fit-testing in the field. Despite the limitations, REAT remains the most widely accepted method for fit-testing hearing protection (Berger, 1986).

A second approach to fit-testing hearing protection is the MIRE method. This method involves using an in-ear microphone to determine the sound levels inside the occluded ear versus the sound levels outside the ear measured simultaneously. The difference between these two values yields the NR of the HPD. The MIRE method was developed for, and is intended to be used, in a reverberatory chamber or diffuse sound field. However, many studies have been conducted that validate the MIRE method when used in a non-reverberatory chamber. Berger, 2007, examined noise reduction results of ear plugs using the Field MIRE method (F-MIRE, which is essentially the MIRE method in a non-reverberatory chamber) and compared them to the insertion loss values of the REAT methods. He found that the NR values from the F-MIRE tests were consistent with the actual REAT values. The fit-test procedure used in this study was similar to the ANSI S12.6-1997 Method B, with some exceptions made to fit the needs of the study. For example, the study adopted an F-MIRE (Berger, 2007) design by determining the arithmetic difference between the A-weighted noise level inside and outside the ear. The study was also conducted in a directional, not diffuse, sound field, which is consistent with the F-MIRE design. The subjects in this study were assumed not to be

naïve in their knowledge of HPDs because most of the subjects wore HPDs on a daily basis. This is different from both the REAT and MIRE methods which require the subjects to have no prior knowledge of how to fit their HPD. This shifts the focus of the study to individuals that are familiar with HPDs. Despite their experience they may still get a poor fit. Each subject was asked to fit the device themselves without instruction for all tests except for the test after instruction on fitting.

LITERATURE REVIEW

MIRE, F-MIRE and REAT

The 1/3 octave band REAT method of testing the effectiveness of HPDs has been considered the highest standard among fit-test methods and was the method used to develop the NRR (Berger, 2007). However, due to the limitations in using the REAT method outside of a laboratory setting, many studies have been conducted to validate other fit-testing methods. Casali (1995) explored the results of two psychophysical (REAT method) tests versus two physical (MIRE method) tests to determine if the fit-test methods that are convenient for use in the field give attenuation results that are similar to the standard REAT tests. Psychophysical tests are subjective in that they rely heavily on the cooperation of the individual taking part in the fit-test. The physical tests require no active participation because a microphone takes the measurements inside and outside the ear to determine the sound levels. Casali's study compared MIRE insertion loss (IL), MIRE noise reduction, pure-tone REAT and 1/3 octave band REAT fit-test methods, the last being the accepted standard method. The difference between the two REAT tests ranged from 0.5 dB to 3.5 dB and the two MIRE tests were not significantly different. The overall difference between the MIRE tests and the REAT tests was significant, ranging from 0.3 dB to 6.0 dB. Casali also found that the differences in attenuation between the 1/3 REAT and the MIRE tests were consistent across frequencies, with the IL and NR showing higher attenuation from 500 to 6300 Hz. The NR was corrected for the transfer function of the open ear (TFOE) to make the IL and NR comparable. The TFOE correction is required due to an amplification or attenuation of noise by some parts of the ear, such as the pinna or ear canal. This value is varied for individuals. The authors concluded that the difference in the pure-tone REAT and the MIRE noise reduction versus the standard 1/3 octave band REAT results was low enough to allow the use of either technique in the field.

Another study comparing the standard 1/3 octave REAT method with a field type MIRE, or F-MIRE fit-test method was conducted by Berger (2007). This study evaluated a method for validating an F-MIRE system and identifying its limitations and variability. The study focused on developing an F-MIRE system that is comparable with measuring

devices currently in use. The F-MIRE design employed a dual-element microphone probe that was fitted into foam ear plugs. This design was used to measure noise reduction by taking measurements inside and outside the ear. The study employed three field tests methods: standard REAT, F-MIRE and a non-acoustic (pressure and seal) test. Berger specifically states "...on the average, the F-MIRE predictions are reliable indicators of the REAT values. However, review of the scatter plots indicates that errors for a single measurement on one individual can exceed 10 dB." With proper compensation factors, the F-MIRE system could be a reliable predictor of the REAT value with a variability of 3.3 dB with the majority of the variability coming from the ability to repeat the fit of the HPD (Berger, 2007).

There are limitations, and in some cases, more sources for error in the MIRE method. One of the benefits of the REAT method is that it accounts for other pathways of transmission of noise, such as bone conduction, whereas the microphones used in the MIRE method do not measure these sounds. Berger (2005) compared REAT to MIRE and acoustical test fixtures (ATFs). ATFs are essentially manikins that are used in place of human subjects. Generally, the effect of the physiological noise masking that occurs with the REAT method provides higher attenuation at frequencies below 500 Hz. This phenomenon occurs because physiological noise is amplified in an occluded ear at these lower frequencies and the amplification can reach magnitudes of up to 6 dB (Berger and Kerivan, 1983). The MIRE method is not able to measure this amplification because the noise pathway is essentially through the head of the individual. If a microphone were capable of measuring this vibration, it would most likely be too sensitive to take accurate measurements for the purposes of a fit-test, and error could increase dramatically (Berger, 2005). Thus, using the MIRE method gives higher attenuation values at frequencies higher than 1 kHz in comparison to the REAT measurements.

Another key issue discussed by Berger (2005) is the placement of the microphone in the ear when using the MIRE method. Many in-ear fit-test measurements employ an altered HPD, as in the probed microphone fitted into a hole in the center of an earplug. One of the challenges in using an in-ear microphone with a HPD is keeping the microphone from interfering with the fit of the HPD. Berger states that proper placement of microphones include placing the microphone at the entrance of the ear canal or

penetrating the HPD so that the microphone becomes part of the HPD (Berger, 2005). In this study, the microphone was held at the entrance to the ear canal by a metal wire hook, and the microphone wire was placed between the earmuff cup and the subject's skin. This was preferred over altering the HPD because the majority of the HPDs in this study belonged to the subjects. The effects of the microphone wire on NR were determined to be minimal in studies at West Virginia University (Lewin, 2009 – personal communication).

Berger concluded that while the REAT method still remains the 'gold standard,' subjective measurements such as the MIRE method can be useful in certain situations. For instance, the MIRE method can measure a wider range of sound levels including levels above the threshold, but it requires larger corrections in the data than the REAT method.

Variability of Attenuation in HPDs

Many studies have discussed the variability in attenuation of HPDs, which has been attributed to several influences. This variability occurs in both laboratory settings and field study settings. Durkt (1993) compared laboratory measurements of noise reduction of two different noise spectra (a pink noise spectrum and a low frequency dominated simulated underground noise spectrum). The study included evaluation of eleven different types of earmuffs and compared the values to the field measurements using the same HPDs. The field measurements were taken during the workers normal workday at several mining locations with wireless microphones using the F-MIRE method. The laboratory measurements revealed noise reduction values consistent with the NRR values for the pink noise spectrum. However, the noise reduction for the simulated underground noise spectrum was less than the NRR. The results from the field portion of the study revealed noise reduction values similar to the laboratory simulated underground noise spectrum. The findings of Durkt's study revealed noise frequency distributions as a strong factor in variation of the NR in different earmuffs. In conclusion, Durkt specifically states "... (A)ttempting to predict a hearing protector's performance with a great degree of certainty appears to be minimal" (Durkt, 1993).

The variability is not limited to field settings. Fit-tests using the ANSI S3.19-1974 standard repeatedly produce a wide variation in HPD fit results with standard deviations

greater than 4 dB (Voix, 2007). Berger (2005) discusses the variability in attenuation results in his comparison of the REAT test method versus the F-MIRE test method. The study concluded that a large part of the variability in attenuation results was due to the repeatability of the fit for the ear plug.

APPARATUS

The two sound sources for the study were recorded with a Nagra ARES-PII+ digital audio recorder in a reverberatory chamber in the industrial hygiene lab at West Virginia University. The reverberatory chamber has a diffuse sound field and is located in the Industrial Hygiene Laboratory at the West Virginia University Mineral Resources Building. The chamber was built to be MIRE compliant. The frequencies recorded in the reverberatory chamber were configured using the OROS OR38 Analyzer and Recorder, in conjunction with NVGate software. The sound system that provided the sound to the reverberatory chamber employed a Behringer Europower Model EP1500 amplifier.

Figure 1 Noise spectra frequency distributions used for subject fit-test

Figure 1 shows a plot of the frequency distributions of the two noise spectra used in this study. The pink line represents pink noise, and the blue line is a general representation of expected frequency distribution in an simulated underground mine setting (figure provided by Dr. Alton Burks, doseBusters, State College, PA).

Figure 2 Fit-test set up in mining office

The fit-test set up is illustrated in Figure 2. The fit-tests were conducted in the main mining offices at each of the mines. The previously recorded noise sources were played back using an Infiniti Primus Model #160 speaker in conjunction with an amplifier and a laptop computer. There were four Larson Davis dosimeters used for each subject. Two of the Larson Davis dosimeters were customized to be used with the doseBusters USA microphones (State College, PA) which were utilized for the in-ear measurements. The remaining two Larson Davis dosimeters were equipped with standard microphones which were used to take measurements at the shoulder. To ensure the recorded sound was being played back at the targeted sound level of 90 to 95 dBA, a Larson Davis sound level meter model 831 was used to take measurements near the ear of the subject before the start of the test. All microphones were operated at the same location at the same time for microphone tests to ensure comparisons between microphones could be made. All microphones reported the same level.

Table 1 Description of earmuff types used in the study

Muff Brand	NRR	Attachment Type	Users	Miner Owned	Investigator Owned	Notes
Peltor Optime 95	20 dB	Bolt Attachment	20	20	0	Low Profile
Bilsom Thunder T1H	23 dB	Snap In Attachment	5	1	4	
Bilsom Thunder T3H	27 dB	Snap In Attachment	1	0	1	
Peltor Optime 98	23 dB	Bolt Attachment	1	1	0	Low Profile
Peltor Optime 101	24 dB	Bolt Attachment	1	1	0	Deep Cup
Bilsom 728	25 dB	Snap In Attachment	1	0	1	
3M Premium 1450	23 dB	Snap In Attachment	25 (All)	0	25 (All)	Investigator supplied muff

One of the primary objectives in the study was to perform the fit-test using the miners' own HPD. However, six of the subjects who participated in the study did not own or use a HPD on a daily or regular basis. In these cases, the investigators supplied the earmuffs including the Bilsom Thunder T1H, Bilsom Thunder T3H, and the Bilsom 728. Information on the HPDs used in this study is listed in Table 1.

METHODS AND PROCEDURES

Noise Sources

The noise source for the fit-testing was recorded in a MIRE compliant reverberatory chamber prior to the fit-testing using an OR38 noise analyzer. Because it was desirable for this study to achieve measurements with sounds that have a wide range of frequencies, two different noise spectra were used. The first sound source was pink noise, where the sound at each frequency was recorded at the same decibel level (90 dBA). The second sound source involved a frequency distribution based on typical underground mining noise (see Figure 1). The recorded noise was played back by a single speaker at 90-97¹ dB in a non-reverberatory room at the mining site. Because the non-reverberatory room produces directional sound, subjects were seated no more than 18 inches from the noise source for every test. The same two sound sources (pink and simulated underground mining noise) were used for all test subjects.

Using Dosimeters versus an Octave Band Analysis

The NR here was measured using dosimeters and an A-weighted scale. Other fit-test studies have used dB linear, or an octave band analysis to determine NR. The dBA measurement was used here for several reasons. First, compliance with OSHA's noise standard is defined in dBA. In addition, instruments that are capable of measuring dB linear inside the ear are often expensive, bulky, and inconvenient. By using a less expensive, more compact device, employers may be more encouraged to practice fit-tests as part of their hearing conservation program.

Test runs comparing the dB linear measurements from the OROS analyzer were compared with dBA measurements from the Larson Davis dosimeters. Mingyu Wu at West Virginia University conducted this investigation (Wu, 2009 – personal communication).

¹ The simulated underground mining frequency distribution was recorded at an overall lower decibel level than the pink noise frequency spectrum. Therefore, the playback of the simulated underground noise source was consistently 2-3 decibels lower than that of the pink noise source.

Methodology for Test Subjects

The methodology used for the fit-test was as follows:

- Prepare equipment
 - o Set up speaker, subject chair, amplifier, and noise source accordingly (subject chair centered in front of speaker at 18 inches and speaker 36 inches from floor on stand).
- Ensure speaker, amp and noise source are properly connected.
- Test noise source with sound level meter at 18 inches from speaker to ensure sound will reach subject at 90-97 decibels.
- Prepare microphones and dosimeters (includes setup, calibration and test run).
 - o Set up two custom Larson Davis dosimeters with doseBusters microphones and ear hook.
 - o The in-ear microphones were secured with a hook device, illustrated in Figure 3.
 - o Set up two regular Larson Davis dosimeters with microphones for shoulder measurements.

Figure 3 Illustration of microphones used in study set up on a human subject

Prepare subject for fit-test

- Inform subject of procedure for fit-test, the general purpose of the study, and benefits or possible negative effects of the testing.
- Take facial measurements in accordance with ANSI S12.6-1997 Annex D.
- Place subject in chair with ear canal 18 inches from the sound source.

- Place regular Larson Davis dosimeter microphone in center of each shoulder.
- Hang custom Larson Davis dosimeter microphones in each ear with hook. Instruct subject to don hearing protection as s/he normally would.
- Ensure subject is properly set up for test.
- Begin fit-test series – eight tests.² The majority of the subjects' test lengths were 30 seconds and tests were done in the same order for each subject.
 - o Test No 1: Subject faces source with pink noise spectrum.
 - o Test No 2: Subject faces source with simulated underground noise spectrum.
 - o Test No 3: Subject turns to left; right ear fronts source with simulated underground noise spectrum.
 - o Test No 4: Subject turns to right; left ear fronts source with simulated underground noise spectrum.
 - o Test No 5: Subject faces source with ears unoccluded; remove hearing protection, underground noise spectrum.
 - o Test No 6: Subject faces source and replaces hearing protection; underground noise spectrum (repetition of Test No 2 after re-fitting).
 - o Test No 7: Subject faces source; instructed on fit of hearing protection (see below for explanation of instruction) with underground noise spectrum.

² Not all tests were conducted for every volunteer. The first five subjects had varying tests and test lengths. More discrepancies in the first five subjects are discussed in the results section.

- Test No 8: Subject faces source with control hearing protection ('control' earmuffs provided by WVU lab); underground noise spectrum.

The subjects were tested within an hour either before or after their shift at their convenience. All tests were conducted in the main mining office on site to minimize background noise. The subjects were asked to don their own hearing protection device (HPD) and fit it themselves with no instruction or guidance from the investigator. This allowed for a true evaluation of customary practices in donning HPDs. In those cases, the workers did not use a HPD on a regular basis, and therefore did not have hearing protection available at the time of the study. In this case, the investigators supplied a HPD from the West Virginia University Industrial Hygiene Laboratory. All HPDs involved in this study were cap-mounted earmuffs. Each subject was equipped with four microphones, one in each ear, and one on each shoulder, to measure the noise reduction associated with the HPD. The microphones placed in the ear were situated in the opening of the ear canal (See Figure 3).

Table 2 Explanation of visual fit evaluation rating applied to the first four tests for each subject

Rating	Explanation of Rating System for Apparent Quality of Fit
1	Muff cup sitting on top of ear and/or gap between muff cap and skin visible with ear visible in gap and locking mechanism of cap-mounted muff not observable; very light compression of muff against skin
2	Cup encompassing ear; ear visible through gap between earmuff and skin on front and/or back side of muff; locking mechanism of cap mounted muff not observable; light compression of muff against skin
3	Cup fully encompassing ear; ear not at all visible; visible gap on front or back (not both); locking mechanism observed but not effective; light compression of muff against skin
4	Cup fully encompassing ear; ear not at all visible; small visible gap on front or back (not both); locking mechanism of muff functioning; sufficient compression/pressure of muff against skin
5	Cup fully encompassing ear; ear not at all visible; no visible gaps anywhere around muff, locking mechanism of muff obviously functioning; deep compression of muff on skin

Factor Affecting Fit	1	2	3	4	5
Muff cup sitting on ear	X				
Gap + ear showing	X	X			
Light compression	X	X	X		
Gap + ear NOT showing			X	X	
Medium compression				X	
No gap + ear NOT showing					X
Deep compression					X

All subjects where Caucasian males between the ages of 20 and 55. The investigator recorded the model of the earmuff and a visual evaluation of the fit-rating of each ear on a scale from 1 to 5, with 1 being worst case and 5 being optimal. (See Table 2 for a description of fit-ratings).

Explanation of Instruction

For Test 7, the individuals were given an approximately 30 second instruction on how to fit the HPD properly. The duration of the instruction was short because of the limited time allotted for fit-tests. The verbal instruction included a demonstration of the mechanics of the cap-mounted earmuffs. Figure 4 demonstrates the ways an HPD can be adjusted. The distance of the muff cap from the point of installation can be adjusted vertically (illustrated by the solid green arrow). In addition, the earmuff cup can be tilted towards the front of the cap or the rear of the cap (illustrated by blue dashed arrow).

Figure 4 Illustration of adjustability of cap-mounted earmuffs

The subject was also instructed to press on the cup of the earmuff to the head until the “clicking” mechanism of the HPD was heard. The instructor then emphasized the seal fit, advising the subjects to attempt to achieve the best skin to earmuff cup seal possible. The subjects were told not to consider their own comfort in this fit.

RESULTS SUBJECTS 1-5 (GROUP A)

Table 3 Explanation of corrections made to Subject numbers

Subject No	Group	Notes
1	A	Sampling error, not included in statistical analysis
2	A	Sampling error, not included in statistical analysis
3	A	Sampling error, not included in statistical analysis
4	B	Subject 4 = Subject 19; Subject 4 omitted due to error; Subject 19 reported as Subject 4
5	B	Subject 5 = Subject 20; Subject 5 omitted due to error; Subject 20 reported as Subject 5
6	B	None
7	B	None
8	B	None
9	B	None
10	B	None
11	B	None
12	B	None
13	B	None
14	B	None
15	B	None
16	B	Subject 16 =Subject 22; Subject 16 and 22 data is reported here
17	B	None
18	B	None
19	NA	Values are reported under Subject 4
20	NA	Values are reported under Subject 5
21	B	None
22	NA	Values are reported under Subject 16
23	B	None
24	B	None
25	B	None
26	B	None
27	B	None
28	B	None
29	B	None
30	B	None

Statistical analysis was conducted using Data Desk[®] software (Data Description, Inc, Ithaca, NY).

There was a problem with the microphone under the right earmuff during the first sampling session (includes subjects 1-5 or sampling Group A). Group A subjects also employed the use of tubing as an extension of the in-ear microphone. This tubing showed an approximate amplification of 5 dBA in NR when comparing microphone data (see Appendix for more information). For these two reasons, the first five subjects were excluded from the statistical analysis for all conditions except for the fit-rating discussion. Subjects in Group A also displayed very poor NR values. The average NR for subjects in Group A was -2.0 dBA. See Appendix for more information. Several subjects in Group A had a visually poor fit, so the evaluation of fit-rating was included with a caveat. Descriptive statistics for Group A are displayed in the Appendix.

The fit-rating was recorded for each ear, so the data could be analyzed for the left ear only for the first five subjects. The first five subjects are referred to as Group A throughout this study. Subjects were numbered in the order they participated in the study. There were three subjects who participated in the study twice: Subjects 4, 5 and 16. Because the data was invalid for the first five tests, the second tests for Subjects 4 (also Subject 19) and 5 (also Subject 20) were inserted and used as Subjects 4 and 5 instead of 19 and 20. The third replication was Subject 16, who is also Subject 22. This individual has the notation of Subject 16 and both data sets are under that subject number. This correction of subject number leaves no information for Subjects 19, 20 and 22. See Table 3 for table of subject information. Because of the microphone error, the total sample size for all tests except for the fit-rating is 25.

RESULTS SUBJECTS 6-30 (GROUP B)

The remaining Subject (numbers 4-30, excluding 19, 20 and 22) are considered group B. See Table 3 for explanation of corrected Subject numbers. Group B was used to determine effects of independent variables on values of NR, including Subject, Left-or-right ear, noise source Spectra, noise source Orientation, Instruction, “Refit,” and Brand of earmuff.

Figure 5 Plot of difference between NR Left and NR Right values for individual test for each Subject in Group B (excludes effects of orientation; see Table 11 for description of conditions for each test)

Table 4 Mean difference between NR left and NR right for Subjects 4 through 30 (includes all tests with muffs on and neutral orientation (0°))

Subject	Count	Mean	StdDev	Earmuff
4	5	-3.8	3.6	Peltor Optime95
5	5	-5.5	2.8	Peltor Optime95
6	5	-3.5	5.7	Peltor Optime95
7	5	12.3	4.9	Peltor Optime95
8	5	0.5	1.6	Peltor Optime95
9	5	1.8	2.0	Peltor Optime95
10	5	2.1	3.7	Peltor Optime95
11	5	-3.8	2.6	Peltor Optime95
12	5	0.9	2.8	Bilsom T1H
13	5	-1.3	1.6	Bilsom T3H
14	5	-12.0	10.5	Bilsom T3H
15	5	-0.1	4.2	Peltor Optime95
16	10	-0.6	5.4	Peltor Optime95
17	5	3.4	2.2	Peltor Optime101
18	5	-10.0	3.0	Peltor Optime95
21	5	-8.0	4.3	Peltor Optime95
23	5	4.5	1.8	Peltor Optime95
24	5	-1.7	12.1	Peltor Optime98
25	5	-1.6	1.0	Peltor Optime95
26	5	-10.9	15.7	Bilsom T1H
27	5	3.0	3.3	Bilsom T1H
28	5	-16.5	16.2	Bilsom 728
29	5	-14.6	12.0	Bilsom T1H
30	5	-17.1	5.5	Bilsom T1H

As shown on Figure 5 and Table 4, if all independent variables other than Subject are ignored, variability is high both within and between subjects.

NR Left versus NR Right

The NR for subjects 4 through 30 varied greatly between subjects, within subjects, and between the two ears. Many individual subjects had significant differences in NR between the left and right ear for the same fit conditions. Since the right and left ears are independently adjustable, a difference between the NR for the left and right ears is plausible.

In some subjects, the difference in NR between the right and left ear was evident even after repositioning the protector. Other subjects that displayed large differences in NR between ears for the same fit had an improved NR after refitting of the muff. Figure 5 plots NR for the left ear versus NR for the right ear with respect to test number. This value should center around zero, but there are consistent outliers, usually specific to the subject number. Table 4 gives the mean difference between NR left and NR right for each subject.

Table 5 Descriptive statistics for NR Left and NR Right. Mean difference shows average for NR Right is higher.

NR Value	Mean	StdDev	p-value
NR Left	16.4	11.0	-
NR Right	19.4	11.0	-
Mean Difference	-3.3	-	p<0.0002

The t-test confidence intervals for NR Left versus NR Right showed a range from -5.1 to -1.6 dBA for a 95% confidence interval for all tests conducted at the neutral orientation (0°). This difference had a highly significant p-value of 0.0002 with a mean paired difference of -3.3 dBA, favoring higher noise reduction for the right ear (see Table 5).

Whether for repeated testing of the same subject or for testing of many subjects, one might expect an average difference between left and right ear to be zero and the deviations to be randomly distributed. However, since the right and left ear muffs are independently adjusted for cap-mount muffs (see Figure 4), it is also possible that differences in the dexterity of left and right hands would affect left and right muffs

differentially. Since the results here showed higher protection for the right ear, it is likely (but not proven since handedness was not determined in this study) that this finding is valid, because most individuals are right-handed. An additional possibility for the difference involves the placement of the muff on the hard hat. With the snap-mounted muff (all Bilsom models), the installation of the earmuff is limited to one position. However, the muffs that require installation via bolt attachment (Peltor models) have much less restriction in the placement of the earmuff on the helmet. This presents the possibility for different placement for the left and right earmuff on the same helmet. While the earmuffs are still individually adjustable, the placement of the earmuff on the helmet may limit the adjustability and therefore have an effect on the NR. This factor may have had an effect on several results in this study, including the overall average NR and variability. Future studies should investigate the variability attributable to improper mounting of the earmuff on the miner's cap.

Average NR Values

Previous discussion used the NR for the left and right ear independently. The remaining sections explore the average NR of the left and right ear.

Figure 6 NR, average of Left and Right Ears versus Subject number for all test conditions with muffs on at the orientation facing the source (neutral, or 0°)

Table 6 Average NR values (in dBA) for each subject for all tests with earmuffs on at the orientation facing the source (neutral, or 0°)

Subject	Count	Mean	Median	StdDev	Min	Max	Range	Earmuff
4	7	28.3	27.8	1.5	26.5	31.1	4.5	Peltor Optime95
5	7	12.7	10.6	7.3	6.4	28.2	21.8	Peltor Optime95
6	7	23.3	23.3	2.5	20.3	27.5	7.2	Peltor Optime95
7	7	22.8	22.5	3.2	18.8	28.6	9.8	Peltor Optime95
8	7	21.2	21.2	2.7	17.8	25.4	7.6	Peltor Optime95
9	7	25.9	26.3	2.1	22.2	27.9	5.7	Peltor Optime95
10	7	21.5	21.6	3.5	16.6	26.5	10.0	Peltor Optime95
11	7	29.1	30.4	2.5	24.6	30.9	6.4	Peltor Optime95
12	7	24.0	23.4	1.3	22.6	25.7	3.1	Bilsom T1H
13	7	24.4	24.4	3.8	19.7	31.2	11.5	Bilsom T3H
14	7	24.7	23.6	5.1	19.6	33.8	14.2	Bilsom T3H
15	7	26.0	27.0	2.1	23.0	28.2	5.2	Peltor Optime95
16	14	19.0	21.5	10.1	3.7	31.2	27.5	Peltor Optime95
17	7	26.7	28.4	10.8	4.3	36.4	32.1	Peltor Optime101
18	7	8.8	5.8	8.6	1.2	25.9	24.7	Peltor Optime95
21	7	10.4	8.4	7.5	3.0	22.1	19.1	Peltor Optime95
23	7	25.7	25.9	1.1	23.9	27.3	3.4	Peltor Optime95
24	7	13.8	10.3	7.0	8.1	25.4	17.4	Peltor Optime98
25	7	19.5	18.9	1.6	18.1	22.7	4.6	Peltor Optime95
26	7	7.5	2.8	12.3	-5.4	23.3	28.7	Bilsom T1H
27	7	11.3	8.7	9.0	-0.6	28.9	29.5	Bilsom T1H
28	7	22.3	22.7	3.7	17.3	27.7	10.4	Bilsom 728
29	7	18.6	21.2	7.5	9.5	26.2	16.7	Bilsom T1H
30	7	16.2	18.1	6.3	2.3	22.0	19.7	Bilsom T1H

Within Subjects Variability for NR Values

Standard deviations within subjects ranged from 1.1 dBA to 12.3 dBA, which is an extreme range. Average of the left and right NR average values ranged from -5.4 to 36.4 dBA. This finding, along with the wide range of average NR values, suggests that data for individuals matters, and that assigning one NR value over all subjects would not be effective.

Histogram of Average NR Values

Figure 7 Distribution of average NR values for 5 dBA intervals for subjects 4-30 and all tests with muffs on

About 2.3% (4 values out of 175³) of the average NR results within subjects were negative values. See Figure 7 for a histogram of average NR values. All of the negative

³ The total of 175 comes from all 25 subjects and all 7 tests with subjects wearing hearing protection.

average NR values were confined to subjects 26 and 27. The negative NR values for the right and left ear independently were scattered throughout subjects and were observed mainly during tests with the simulated underground spectrum as the noise source. This finding suggests that the lower frequency dominated spectrum can possibly increase noise levels when using earmuffs. Eleven values were negative for the NR in the right ear, which includes subjects 18, 21, 24, 26, 29, and 30. These numbers do not include subjects 1-5. Including subjects 1-5 increases the percentage of negative NR values to around 8.8%. The majority of average NR values (54.8%) fell within the 20-30 dBA range.

Between Subjects Variability for NR Values

Table 7 Descriptive Statistics for NR values averaging all tests with earmuffs on and all Subjects in Group B.

Average Noise Reduction, dBA (n=175)	
Mean	20.1
Standard Deviation	8.5
Minimum	7.5
Maximum	29.1

As shown on Table 7, the average NR when all 25 subjects and tests with earmuffs are combined was 20.1 dBA with a standard deviation of 8.5 dBA. The range of average NR values was 7.5 to 29.1 dBA. Assuming miners are exposed to ambient levels around 95 to 100 dBA TWA, the NR found here would reduce the average internal TWA to around 75 to 80 dBA.

Variability reported in noise measurements commonly falls in the range of 2-5 dBA (Voix (2007); Berger (2007)). The wide variation apparent here within subjects (8.5 dBA) suggests that even with one individual using one type of earmuff the repeatability of NR results is still highly variable, and that fit-tests should be repeated for each subject. A future study might attempt to determine the number of fit-tests required on one individual before a consistent NR level is determined.

Table 8 Number in each range of median NR values when considering all subjects and tests with muffs on

Median NR Range (dBA)	Number of Values
<10	4
10 – 20	4
20 – 30	15
30 – 40	1

One method of evaluating the effectiveness of a fit-test would be to look at the median NR values. If you were to define an “acceptable” fit-test to include median NR values above 10 dBA, and reject fits below 10 dBA, this study would reject 4 subject fits (see Table 8 for distribution of the medians). This could be used as a method of identifying and eliminating poor fits during fit-tests.

Frequency Distribution of the Source

Table 9 T-test results for comparison of average NR values for two noise spectra over Test 1 and Test 2

Noise Source	Mean	StdDev	p-value
Pink (n=25)	20.0	8.3	-
Underground (n=25)	19.3	8.9	-
Mean Difference	0.3	-	p<0.78

As shown on Table 9, the means for the same conditions were slightly different when pink noise was used (20.0 dBA) and simulated underground noise was used (19.3 dBA), a modest mean difference of 0.3 dBA with the pink noise spectrum showing higher NR. A Student t-test showed a p-value of 0.78 for comparison of the two spectra. The modest difference found between the two spectra is not statistically significant.

Table 10 Analysis of variance for noise spectra

Source	df	Sums of Squares	Mean Square	F-ratio	Probability
Constant	1	20945.7	20945.7	1905.1	0.0001
Noise Source	1	24.0679	24.0679	2.189	0.1515
Subject No	23	3530.85	153.51	13.963	0.0001
Error	25	274.87	10.99	-	-
Total	49	3829.79	-	-	-

ANOVA found that the noise spectra was not significant in the overall scheme of the study ($p < 0.15$) whether Subject was included or not (Table 10). Therefore, the magnitude of the mean difference between the two noise spectra was neither practically important nor statistically significant. This is different from the findings of Durkt (1993). This could be because the dosimeter was used to measure dBA instead of performing an octave band analysis. If these measurements were recorded at each frequency, one would expect a significant difference between noise sources. In any case, since frequency distribution had no effect on the overall results, it was omitted as a factor in

subsequent analyses. See the Methods section for discussion on dBA versus dB linear as it applies to this study.

Table 11 Explanations of conditions for each test

Test No.	Muffs	Noise Spectra	Orientation	Instruction	Owner
1	On	Pink	Neutral (0°)	No	Miner
2	On	Underground	Neutral (0°)	No	Miner
3	On	Underground	Right Ear (270°)	No	Miner
4	On	Underground	Left Ear (90°)	No	Miner
5	Off	Underground	Neutral (0°)	No	Miner
6	On	Underground	Neutral (0°)	No	Miner
7	On	Underground	Neutral (0°)	Yes	Miner
8	On	Underground	Neutral (0°)	Yes	Investigator

Table 12 Average noise reduction for each type of earmuff⁴

Muff Type	Users	Tests	Average NR (dBA)	StdDev
Peltor Optime95	15	15	21.3	7.8
Peltor Optime98	1	6	24.6	14.6
Peltor Optime101	1	6	13.4	8.1
Bilsom T1H	5	30	16.4	8.0
Bilsom T3H	2	12	25.4	3.6
Bilsom 728	1	6	22.2	3.1
3M Premium 1450	25	150	25.6	4.1

Earmuff Types

The types of earmuffs and the respective NR values are listed in Table 12. A total of 20 miners out of 30 used the Peltor Optime 95. However, because the first five subjects were omitted from statistical analyses, and subjects 1 through 5 all used the Peltor Optime 95, the true number of subjects considered in the statistical analyses that used this muff is 15 out of 25 total subjects. Sample sizes are extremely different for the muff type variable, which is heavily weighted with one type of muff. See Table 12 for sample sizes for muff types. The number of tests for muffs varied from 6 to 150, and the mix of test conditions were not the same for all muff brands (Table 12).

The West Virginia University industrial hygiene lab provided the three Bilsom models to miners who did not normally wear an earmuff (n=6). Because these subjects did not use earmuffs on a daily basis, they could be considered naïve in the knowledge of properly donning hearing protection. However, this variable was not considered in the scheme of

⁴ Average NR only for tests when subject is wearing the earmuff (for volunteer muffs include tests 1, 2, 3, 4, 6, 7 and for investigator supplied muff includes test 8).

this study. Two subjects in the study owned the Bilsom model earmuffs and used them on a daily basis.

Table 13 Analysis of variance for earmuff type for all tests with muffs on

Source	df	Sums of Squares	Mean Square	F-ratio	Probability
Constant	1	57855.7	57855.7	1108.8	0.0001
Muff Type	6	1393.26	232.21	4.45	0.0004
Error	118	6156.88	52.18	-	-
Total	124	7550.15	-	-	-

ANOVA (see Table 13) shows that muff type considered alone was highly significant ($p < 0.0004$). When the analysis was restricted to the two muff types used by 15 or more subjects, muff type was still highly significant.

The 3M Premium 1450 earmuff was used as a comparison to the earmuffs belonging to the miners. All subjects were asked to don this same earmuff for test number 8 in the fit-test series. Test number 8 employed the simulated underground noise spectrum and the neutral position (facing the source). Test 8 was conducted after the subject was given instruction (Test 7). See Table 11 for description of conditions for each test. Although the results for the investigator supplied muff provided one of the best NR averages compared to the miners muffs, the average NR value is not exceptionally better. Some of the differences in NR between muffs were significant, but the differences were not of practical importance. However, two muff types (Peltor Optime101 and Bilsom T1H) were substantially lower than the others.

Visual Fit-rating

Figure 8 Image of visibly poor fit for Subject 27. Ear is visible through gap in back of earmuff

The investigator examined the HPD for each ear of the subject for the appearance of detectable gaps in the fit of the earmuff cup to the face. An explanation of the fit-ratings is listed in Table 2. Figure 8 is an example of a poor fit (fit-rating of 2) for the left ear. The fit for the right ear for this subject was rated a 4.

Table 14 Average noise reduction for each fit-rating (1-5) for each ear. This average does not exclude the first five subject numbers.

Average NR for Each Fit-rating		
Rating	Right	Left
1	No Data	10.9
2	-3.0	18.6
3	16.6	17.5
4	15.8	18.6
5	26.7	27.8

The fit-rating was very different for different subjects. The fit of the earmuff on both ears was evaluated for the first four tests (Figures 9 and 10). Table 14 shows the average NR values for each fit rating. The fit-rating analyses were done using all 30 sets of subject data (Group A and Group B) for the left ear because this data was available. The data for the right ear was not available for the first five subjects, so the analysis for the right ear includes 25 subjects. The high variability in fit-rating is probably due to the fact that not all individuals had the same facial and neck dimensions. Some individuals would have deep compression simply because their neck was larger, and vice versa. A handful of subjects also had long hair, which made it difficult to see leaks or gaps.

Figure 9 Illustration of NR for the right ear versus the fit-rating for the right ear

The illustration of NR right versus fit-rating (Figures 9 and 10) shows a slight trend in increasing NR with an increased fit-rating. The results show that extreme ratings appear to be strong indicators of NR whereas ratings 3 and 4 have varied NR.

Figure 10 Illustration of NR for the left ear versus the fit-rating for the left ear

From the information presented here, one could assume that a “visually excellent fit,” with a fit-rating of 5, is likely to give a reduction value greater than 20 dBA. However, from these results, you cannot conclude that the opposite is true. In Figure 10, the fit-rating and NR for the left ear shows values as high as around 11 dBA for a fit-rating of 1. Additional studies could be conducted to determine if this visual evaluation of hearing protector fit is effective. If employees can be trained in recognizing problems with a co-workers earmuff fit, corrections could be made in the field and potentially decrease noise exposures from HPD leakage.

Table 15 Average NR results for the effects of orientation

Average NR for Each Orientation			
Orientation	n	Mean	StdDev
270° - Right Ear Facing Source	25	15.5	9.8
0° - Neutral Position: Facing Source	125	21.5	7.8
90° - Left Ear Facing Source	25	17.6	8.9

Results from Student T-test	Mean Difference	p-value
Mean Difference (90° - 270°)	-0.1	p= 0.008
Mean Difference (90° - 0°)	-0.4	p= 0.0001
Mean Difference (270° - 0°)	-0.4	p= 0.0001

Orientation

A Student t-test was conducted to determine the differences between each of the three orientations. Each non-neutral position (90° and 270°) had a sample size of 25 (one test for each subject) with a sample size of 125 for the neutral position (0°). Because the NR at both ears was measured, there would be no expected differences in average NR between the two non-neutral orientations (270° or 90°). The differences between the non-neutral positions and the neutral position should be observable because the sound is approaching the ear differently in a directional sound field. Table 15 shows the observed results for the effect of orientation.

Table 16 Analysis of variance for orientation and subject for all tests with muffs on

Source	df	Sums of Squares	Mean Square	F-ratio	Probability
Constant	1	70725.1	70725.1	2048.2	0.0001
Orientation	2	927.28	463.64	13.43	0.0001
Subject	23	6611.5	287.46	8.32	0.0001
Error	149	5145.07	34.53	-	-
Total	174	12683.9	-	-	-

The difference in average NR between two non-neutral positions was 0.1 dBA, with the average NR for the 90° position (left ear facing source) slightly higher than the 270° position (right ear facing source). When compared to the neutral position (0°), both non-neutral positions showed a difference of 0.4 with the non-neutral positions giving slightly lower NR than the neutral orientation. Both of these differences were highly significant ($p=0.0001$). Despite the fact that the differences in means between the orientations were found to be highly significant, it would be practical to disregard the modest deviations found here. Mean differences were less than 0.5 dBA for all three orientations, which is essentially a negligible difference to a practitioner trying to assess the effectiveness of a hearing protector. The effects of orientation were also shown to have significant effects on the overall results ($p<0.001$). See Table 16 for ANOVA for orientation.

Effects of Repositioning

Table 17 Descriptive statistics for the difference between two fits (Test 2 and Test 6) after repositioning the earmuff

Subject	NR Fit 1	NR Fit 2	Difference
4	23.7	23.2	0.5
5	22.5	21.5	1.0
6	20.8	21.9	-1.1
7	26.1	27.4	-1.3
8	21.9	18.6	3.3
9	30.5	30.0	0.5
10	25.7	23.1	2.5
11	24.4	23.2	1.2
12	23.1	24.1	-1.0
13	28.1	27.6	0.5
14	13.4	3.7	9.7
15	34.1	4.3	29.8
16	5.2	5.8	-0.6
17	29.4	26.8	2.6
18	10.2	10.6	-0.4
21	8.4	5.3	3.1
23	29.0	28.8	0.2
24	26.3	25.9	0.4
25	10.3	9.7	0.6
26	20.5	18.1	2.4
27	-1.9	18.7	-20.6
28	8.7	8.0	0.8
29	22.7	24.3	-1.7
30	13.5	26.2	-12.6

Figure 11 NR values averaged between left and right ear for before and after instruction (includes all Group B subjects, and Tests 2 and 6)

Previous studies have found that NR varies greatly as muffs are removed and donned again (Berger, 2007). In this study, Test 2 (simulated underground noise) and Test 6 (re-fit with simulated underground noise) are the only two tests that differ only by having a refit between the tests. Tests 7 and 8 both required re-fitting of the HPD, but there were other factors influencing that data (instruction and a different earmuff type), so we will consider the mean difference between Tests 2 and 6. The sample size for each test was 25. Table 17 and Figure 11 show the NR values for both tests for each subject, and the difference between the two fits.

Table 18 T-test results for Test 2 versus Test 6 showing effects of re-fit

Test	Mean	StdDev	p-value
Fit 1 – Test 2 (n=25)	19.8	9.0	-
Fit 2 – Test 6 (n=25)	18.9	8.5	-
Mean Difference	0.8	-	p<0.61

The t-test results (Table 18) for the comparison of Tests 2 and 6 shows that the modest difference of 0.8 dBA between fittings was due to chance ($p = 0.61$).

Instruction**Table 19** NR values for before and after instruction (Group B, Tests 2 and 7)

Subject	NR Before	NR After	Difference
4	29.0	27.0	-2.0
5	10.2	12.7	2.5
6	23.7	21.8	-1.9
7	22.5	24.5	2.0
8	20.8	21.2	0.4
9	26.1	26.3	0.2
10	21.9	21.6	-0.3
11	30.5	30.4	-0.1
12	25.7	25.0	-0.7
13	24.4	24.4	0.0
14	23.1	33.8	10.7
15	28.2	27.0	-1.2
16	13.5	5.3	-8.2
17	34.1	23.8	-10.3
18	5.2	13.7	8.5
21	8.4	19.1	10.7
23	26.3	26.0	-0.3
24	10.3	25.4	15.1
25	20.5	19.5	-1.0
26	-1.9	19.0	20.9
27	8.7	14.4	5.7
28	22.7	17.7	-5.0
29	13.5	25.8	12.3
30	17.7	18.4	0.7

Figure 12 NR values averaged between left and right ear for before and after instruction (includes all Group B subjects, and Tests 2 and 7)

Tests 2 and 7 were the same except that Test 2 was done before instruction and Test 7 was done after. Both were done using the simulated underground noise frequency facing the source (neutral position). As shown on Table 19 and Figure 12, the effects of instruction varied by subject. Thirteen out of 25 subjects had slightly improved NR after instruction and 6 out of 25 had slightly degraded NR values.

Table 20 Comparison of Test 2 and Test 7 with respect to the effect of instruction

Instruction		Mean	n	StdDev	p-value
No Instruction	(Test 2)	19.8	25	9.0	-
Instruction	(Test 7)	22.1	25	6.2	-
Mean Difference		-2.3	-	-	p=0.057

As shown in Table 20, the mean NR before instruction was 19.8 dBA and the mean after instruction was 22.1, an improvement of 2.3 dBA. A Student t-test of the paired data showed that the difference was not statistically significant ($p=0.057$). On the other hand, Subject 17 showed a very large fall in NR after instruction. If the data for this subject is removed from the analysis, the effect of instruction is highly significant ($p<0.025$). Removing the most substantial negative change (Subject 17) as well as the highest positive change (Subject 26), the difference still shows significance at $p<0.05$.

CONCLUSIONS

The results of this study suggest the following conclusions for the conditions studied:

1. NR values from fit-testing show substantial between-subjects variability (8.5 dBA). This variability can be explained by some parameters, including frequency of the noise source, instruction of the individual on how to don a HPD, and quality of the fit of the HPD. However, the most significant source for variability lies in the individual subject. This finding strongly suggests a fit-test procedure that requires testing for each individual.
2. NR values from fit-testing also show substantial within-subjects variability (1.1-12.3 dBA). The portion of the variability attributable to differences in test conditions could not be determined due to the lack of true replications.
3. NR values for the right ear were modestly (3.3 dBA) but significantly higher ($p < 0.0002$) than for the left ear. The differences here could be due to the fact that most people have right-handed dominance, which could have an effect on fitting of the HPD. This could also be attributable to the inconsistency in installation of the earmuff on the cap for the bolt mounted earmuffs.
4. NR values were significantly ($p < 0.0001$) but modestly higher (0.7 dBA) for pink noise than for simulated underground noise. While this difference is significant, it is minimal, and not of practical importance.
5. The brand of muffs was shown to have a significant effect on the NR results ($p < 0.0004$). However, different groups of subjects wore different ear muff brands, making it difficult to determine how much of the brand effect could be attributed to differences between subjects.
6. The highest and lowest categories of the visual fit-rating had no false positives but many false negatives. The middle categories were largely uncorrelated to values of NR.

7. The orientation of the subject to the source had moderate, (<1 dBA) but highly significant ($p < 0.008$), differences in mean NR values with the neutral position producing the highest values.
8. Brief instructions showed no significant benefits for these experienced users of muffs.

An overall conclusion is that the average of multiple fit-tests is likely to be a much better predictor of on the job protection than the results of a single fit test. The findings also suggest that NR values are extremely variable.

A fit-test program should also employ periodic monitoring to evaluate NR over time. A one-number system for assigning NR to one particular HPD is not effective. Instead, the one-number system should be a combination of a rating for each individual and their preferred HPD. This will allow a noise reduction rating to be assigned to the worker and the device, not the just device. In addition, a more thorough training and instruction should be employed to educate workers of the factors that can contribute to poor NR in a HPD. Future studies might explore the effects a more intensive level of instruction, number of repetitions required to determine a consistent NR for one individual and earmuff placement when using muffs that require the bolt attachment. Additional research should also be conducted on the relationship between NR in the left and right ears to determine if this is a consistent problem throughout cap-mounted earmuff wearers.

Further research might also attempt to investigate the relationship between NR and participation in hearing conservation programs. In this study, the two mines had very different attitudes and policies on hearing conservation. It would be interesting to see how effective a 'good attitude' towards hearing conservation could be on NR.

Caveats

There were limitations to the study that tend to weaken the strength of the conclusions. First, the test order was the same for each subject, allowing for the possibility of order effects. For example, subjects may do may poorly with each succeeding test due to fatigue or diminishing motivation. There was no indication of such a diminishment, but it is possible that the effect was masked due to different tests.

Another issue effecting the analysis of the data was the non-factorial study design in which each subject would have been tested for every combination of conditions. Instead, time constraints during sampling limited the number of possible tests. Rather than eliminating possibly important variables, we chose to make the study a series of two-factor studies. This made it impossible to consider many possible interactive effects of independent variables on the observed outcomes.

REFERENCES

1. ANSI. (1997). Methods for Measuring the Real-Ear Attenuation of Hearing Protectors. *American National Standards Institute, S12.6 1997* . New York, NY.
2. Behar, A. (1985, January-February). Field Evaluation of Hearing Protectors. *Noise Control Engineering Journal* , 24 , 1, 13-18.
3. Berger, E. H. (2007). Methods of Developing and Validating a Field-MIRE Approach for Measuring Hearing Protector Attenuation. *Spectrum*, 24, Suppl.1 .
4. Berger, E. (1982). Hearing Protector Performance: How They Work - and - What Goes Wrong in the Real World. *E A R Log* (5) . Aero Company.
5. Berger, E. (1986). Review and Tutorial- Methods of Measuring the Attenuation of Hearing Protection Devices. *Journal of the Acoustical Society of America* (79(6)), 1655-1687.
6. Berger, E. (2005). Preferred Methods for Measuring Hearing Protector Attenuation. *The 2005 Congress and Exposition on Noise Control Engineering*.
7. Casali, J. G., Mauney, D. W., & Burks, J. A. (1995, July). Physical vs. Psychophysical Measurement of Hearing Protector Attenuation - a.k.a. MIRE vs REAT. *Sound and Vibration* .
8. Durkt, G. (1993). Field Evaluation of HPDs at Surface Mining Environments. Department of Labor, US Mine Safety and Health Administration.
9. Joy, G. J., & Middendorf, P. J. (2007, January). Noise Exposure and hearing Conservation in U.S. Coal Mines - A Surveillance Report. *Journal of Occupational and Environmental Hygiene* , 4 , 26-35.
10. Labor, U. D. (2000, September 13). Federal Mine Safety and Health Act. 30 CFR 62.160 Hearing Protectors .
11. Myers, B. (2008, June). NRRs for Hearing Protectors: A Change is Coming. *Occupational Health & Safety* .

12. NIOSH. (1996, May). Results from the National Occupational Health Survey of Mining (NOHSM). Publication No. 96-136. Cincinnati, OH: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.
13. Voix, J., & Hager, L. (2007). Individual Fit Testing of Hearing Protection Devices. First European Forum on Efficient Solutions for Managing Occupational Noise Risks .
14. Zannin, P. H., & Gerges, S. N. (2005, September 29). Effects of Cup, Cushion, Headband Force, and Foam Lining on the Attenuation of an Earmuff. International Journal of Industrial Ergonomics .

APPENDIX

Appendix I: Microphone Tests with and without Tubing as an Extension of the In-Ear Microphone

Initial microphone tests revealed an approximate 5 dBA amplification of noise for the in ear microphone. This is illustrated in Figures A and B.

Figure A & Figure B Results from microphone test done before and after the tests on the first five subjects. This test run employed the tubing as an extension of the in-ear microphone (Muff Left Ear). This showed a ~5 dB amplification throughout all microphone tests.

Further testing was conducted to evaluate the amplification effects of three inch tubing as an extension of the doseBusters microphones. Five microphones were used for the microphone test followed by an in ear test. The microphone test included two new doseBusters microphones, one old doseBusters microphone and two standard Larson Davis microphones. The microphones were all attached to a stand 18 inches in front of the noise source. Data was collected first without tubing on the microphones for 70 seconds of pink noise and then 70 seconds of underground frequency distribution. Then the tubing was placed on the microphones and the test was repeated with the pink and underground noise. The tubes were then plugged on the open end with an inactive microphone to block the noise and the test was again repeated. The plugs were then removed from the tubing and the test was repeated.

In Ear Microphone Test

A volunteer was used in this test to illustrate differences between the tubing extension of the microphone and the metal ear clip. The first tests were done with pink and underground noise with tubing attached to the microphone under the right ear muff and the microphone on the left ear was held in place with a metal ear clip. The tubing was then removed from the microphone on the right ear muff and a metal clip was used. T

Results

Microphone Test (outside ear)

Figure C Sound level readings from five microphones measuring noise in a directional sound field with and without a tubing extension on the in-ear microphones. No subject was used here; microphones were placed next to each other using a stationary pole to measure noise

In Ear Microphone Test

Figure D Sound level readings from four microphones measuring noise in a directional sound field with and without a tubing extension on the in-ear microphones. Human subject was used here; a microphone was placed in each ear and on each shoulder

It seems from this set of data that the tubing causes the slight amplification when the underground noise is used in conjunction with an earmuff. For this reason, the data obtained using the tubing as an extension of the in-ear microphone (Group A subjects) was not used in the study.

Appendix II: NR Average Values for Group A

Figure E Average NR of the left ear for each subject in Group A for each test. Group A subjects did not participate in all tests that Group B subjects participated in

Average NR for Each Subject for Each Test (dBA)					
Subject	Test 1 (Pink)	Test 2 (UG)	Test 3 (270°)	Test 4 (90)°	Test 6 (Re-Fit)
1	-2.9	-7.6	-	-	-
2	-4.2	-9.1	-	-	-9.2
3	2.4	3.1	-	-	1.0
4	7.1	6.9	2.4	2.3	5.6
5	3.5	1.6	-9.7	-4.6	-4.7
Average	1.2	-1.0	-7.3	-1.2	-1.8
Overall NR Average for All Subjects in Group A				-2.0	