

University of Southampton Research Repository ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

University of Southampton, 2010

Faculty of Engineering, Science and Mathematics

Institute of Sound and Vibration Research

**TEST-RETEST RELIABILITY OF REAL EAR
MEASUREMENTS FOR OPEN CANAL HEARING AID
FITTINGS**

By Wala' Alaqrabawi

**A dissertation submitted in partial fulfilment of the requirements for
the degree of MSc by instructional course**

Declaration

I, Wala' Alaqrabawi, declare that this project is my own work, except where acknowledged, and the project was done according to the principles for ethical treatment of human subjects as approved for this research by the Ethics Committee at the Institute of Sound and Vibration Research, University of Southampton.

Acknowledgements

I would firstly like to thank my supervisor Emma Mackenzie for her continual support and guidance throughout my project. I would also like to thank all the 20 participants for their cooperation. Finally, I would like to thank my family and Ahmad for all their encouragement throughout my MSc Audiology course.

Abstract

When fitting hearing aids with conventional earmoulds using real ear measurements (REM) it is recommended to use the modified pressure with concurrent equalization method whereby a reference microphone is used to monitor the REM loudspeaker output level. The review of the literature indicates good test-retest reliability for REM using conventional earmoulds. However, in REM using open-fit, another sound field equalization method (the modified pressure with stored equalization method) should be used to minimize any inaccurate measurements resulting from the amplified sound leakage when the open-fit hearing aid is used. In addition the sound is delivered via a generic ear tip rather than a custom made earmould. Therefore, the test-retest reliability for conventional occluded earmould REM cannot be generalized to the open-fit measures.

Twenty otologically normal participants were tested to investigate short-term test-retest reliability for open-fit REM, and for conventional earmoulds at both 0° and 45° head-to-loudspeaker azimuths by performing repeated measurements at the same participant visit but after removing both the probe tube and the hearing aid and reinserting them. It was found that open-fit REM have good short-term test-retest reliability (mean 1.57 dB, SD 1.10 dB) at both azimuths and are in agreement with measured (mean 2.12 dB, SD 1.45 dB) and reported (SD ranged from 1 to 3.2 dB) conventional earmould REM test-retest variability values in the 0.25 to 4 kHz frequency range (Ringdahl A & Leijon A, 1984, The reliability of insertion gain measurements using probe microphones in the ear canal, *Scandinavian Audiology*, 13, 173-8). This finding is clinically significant as open fittings are widely fitted using REM to hearing impaired patients.

Table of Contents

1.0 INTRODUCTION.....	9
1.1 Real ear measurements and their use in hearing aid fittings.....	9
1.2 Different measures in the REM	9
1.3 Pre-real ear measurements considerations	10
1.3.1 Sound field equalization	10
1.3.2 Loudspeaker position	12
1.3.3 The probe tube placement	14
1.3.4 Testing environment	15
1.4 Test-retest reliability of REM.	16
1.4.1 Test-retest reliability of REM within the same tester	16
1.4.2 Test-retest reliability of REM between testers.....	20
1.5 Open-fit hearing aids.....	21
1.5.1 Background	21
1.5.2 Advantages of open-fit hearing aids	23
1.5.3 Problems with sound field calibration in open-fit REM.....	24
1.5.4 Test-retest reliability for open-fits	26
1.6 Research question.....	27
2.0 METHOD.....	28
2.1 Aim.....	28
2.2 Design.....	28
2.3 Hypotheses.....	29
2.4 Sample Selection.....	30
2.4.1 Sample size calculation	30
2.4.2 Participants.....	31
2.4.3 Inclusion and exclusion criteria	31
2.5 Stimulus.....	32
2.6 Equipment and Apparatus.....	32
2.7 Test room set up.....	33
2.8 Daily checks.....	34
2.9 Test protocol.....	35
2.9.1 Preparation	35
2.9.2 Screening.....	35
2.9.3 Real Ear Measurements	36
2.10 Pilot Study.....	38
2.11 Data Management	38
3.0 RESULTS.....	39
3.1 Introduction.....	39
3.2 Descriptive Statistics.....	39
3.3 Data Distribution.....	41
3.4 Inferential Statistics.....	42
3.4.1 Significance of difference between test and retest measures.....	42
3.4.2 Analysis of variance.....	42
3.4.3 Effect of gender on the test-retest difference	43
3.4.4 Effect of ear on the test-retest difference	43
3.4.5 Effect of loudspeaker-to-head azimuth on the test-retest difference	44
3.4.6 Effect of earmould type on the test-retest difference.....	46

3.4.7 Effect of the stimulus frequency on the test-retest difference	47
3.4.8 Interaction effects.....	47
4.0 DISCUSSION.....	49
4.1 The effect of gender on the test-retest difference for REIG.....	49
4.2 The effect of ear on the test-retest difference for REIG	49
4.3 The test-retest reliability for REIG using open-fit.....	50
4.4 The test-retest reliability for REIG using conventional earmould.....	51
4.5 The effect of loudspeaker-to-head azimuth on the test-retest difference for REI..	52
4.6 The effect of earmould type on the test-retest difference for REIG	53
4.7 The effect of stimulus frequency on the test-retest difference for REIG.....	54
4.8 Clinical implications.....	55
4.9 Limitations.....	55
4.10 Further research.....	57
5.0 CONCLUSIONS.....	58
6.0 REFERENCES.....	59
7.0 APPENDICES.....	66

List of Figures

Figure 1: Illustration of the substitution method of sound field equalization for probe-microphone measurements. [From Mueller (1992) used with permission]	11
Figure 2: Illustration of the modified-pressure method of sound field equalization for probe-microphone measurements. [From Mueller (1992) used with permission]	12
Figure 3: Difference between eardrum- and probe-measured SPL at eight frequencies as a function of distance of the probe from eardrum. [From Hawkins & Mueller (1992) used with permission]	15
Figure 4: Comparison of mean test-retest SD difference for REIG between Killion and Revit (1987) and Ringdahl and Leijon (1984) results.	18
Figure 5: An open-fit hearing aid	21
Figure 6: Measuring tool used to select correct Corda2 tube length.	33
Figure 7: Test room setup.	34
Figure 8: Probe tube placement during the REUG measure.	37
Figure 9: Probe tube and hearing aid placement during REIG measure for open-fit (a) and conventional earmould (b).	38
Figure 10: Mean REIG test-retest difference for open-fit at 0° and 45° azimuths. Error bars show +/- 1.0 SD	40
Figure 11: Mean REIG test-retest difference for conventional earmould at 0° and 45° azimuths. Error bars show +/- 1.0 SD	41
Figure 12: Open-fit mean test-retest difference at 0° and 45° across frequency. Error bars show +/- 1.0 SD	45
Figure 13: Conventional earmould mean test-retest difference at 0° and 45° across frequency. Error bars show +/- 1.0 SD.	45
Figure 14: Mean REIG test-retest difference for open-fit and conventional earmould. Error bars show +/- 1.0 SD.	46

List of Tables

Table 1: The eight experimental conditions that were examined for each participant	29
Table 2: Hearing level Thresholds across frequencies that were used in hearing aid fitting.....	32
Table 3: Mean and SD in dB for REIG of open-fit test-retest difference at 0 ⁰ and 45 ⁰ across frequency.....	39
Table 4: Mean and SD in dB for REIG of conventional earmould test-retest differences at 0 ⁰ and 45 ⁰ across frequency.....	40
Table 5: Results of 32 paired-sample t-tests examining the difference between test and retest REIG measurements for all conditions across frequency.	42
Table 6: Mean and SD for the males and females data averaged across conditions. ..	43
Table 7: Mean and SD for the right and left ears data averaged across conditions.	44
Table 8: Mean and SD for 0 ⁰ and 45 ⁰ azimuths data averaged across conditions.	44
Table 9: Mean and SD for open-fit and conventional earmould data averaged across conditions.	46
Table 10: Results of 16 paired-sample t-tests examining the test-retest difference between open-fit and conventional earmould at each frequency for both azimuths....	48

1.0 INTRODUCTION

1.1 Real ear measurements and their use in hearing aid fittings

Real ear measurements (REM) using a probe tube microphone have become the standard and preferred clinical procedure for hearing aid fittings. In REM, the device loudspeaker produces an acoustical stimulus that is transmitted to the patient's hearing aid. A thin silicon probe tube, which is placed near the tympanic membrane and attached to the probe measuring microphone, then measures the sound pressure level near the patient's tympanic membrane (Lantz et al., 2007).

These measurements take into account the performance of the whole electro-acoustical system for each patient. The electro-acoustical system includes the hearing aid, the tubes, sound filters, earmould, the patient's external ear canal, and any head and torso acoustic effects. REM, unlike measurements which are done using the 2 cm³ coupler, measure hearing aid's output or gain in the patient's external ear (Ackley et al., 2007; Lantz et al., 2007). REM are needed to adjust the patient's hearing aid in order to get the patient's prescribed and desired gain near his/her tympanic membrane (Swan & Gatehouse, 1995; Harrowven, 1998).

1.2 Different measures in the REM

REM results can be presented in two different ways; in terms of real ear response which measures the exact output sound pressure level (SPL) delivered in the ear canal, or in terms of real ear gain which is a relative measure of the difference between the input and output levels (Pumford & Sinclair, 2001; BSA & BAA, 2007; Ackley et al., 2007). However, these two different terms involve the same REM procedure. Different input levels are produced; soft, moderate, and loud levels to ensure audible and comfortable sound amplification in all sound levels (Ackley et al., 2007).

The first REM performed is the real ear unaided gain (REUG) which is " the difference in dB, between the SPL at the measurement point and the test signal level, as a function of frequency, with an un-occluded ear canal" (BSA & BAA, 2007), this is done by placing a probe tube microphone in the ear canal without placing the earmould and the hearing aid. The following REM, is the real ear aided gain (REAG) which is " the difference in dB as a function of frequency between the SPL at the measurement point and test signal level with the hearing aid in-situ and turned on" (BSA & BAA, 2007).

These two measurements are performed to incorporate the calculation of the real ear insertion gain (REIG) which is "the difference in dB as a function of frequency, between the real-ear aided gain and the real-ear unaided gain" (BSA & BAA, 2007). The REIG is then adjusted to match the selected prescription target which is suitable for the patient's specific needs. The hearing aid gain should match the prescribed target within the recommended tolerance values, which are ± 5 dB at 0.25, 0.5, 1 and 2 kHz and ± 8 dB at 3 and 4 kHz (BSA & BAA, 2007).

1.3 Pre-real ear measurements considerations

Prior to conducting the real ear measurements for hearing aid fitting, many premeasurement steps should be done to prepare both the equipment and the patient. Special care should be taken during these premeasurement steps as they might affect the validity and reliability of the subsequent measurements (Mueller, 1992).

1.3.1 Sound field equalization

Sound field equalization is the procedure in which the acoustic signal at a specific point in space is being controlled so the amplitude remains at the desired level across frequencies. Two equalization methods are commonly used; the first one is the substitution method, also called the stored equalization method, shown in Figure 1. In this method a single measurement microphone is used in measuring the SPL for each frequency at a point where the centre of the patient's head will be positioned, without the presence of the patient (Mueller, 1992; Lantz et al., 2007). This calibration

measurement is then used as a reference for the following unaided and aided measurements. In this method the patient's head should be positioned on that specific point and remain still during all the following measures, as any movement may affect the REM accuracy (Mueller, 1992; Lantz et al., 2007).

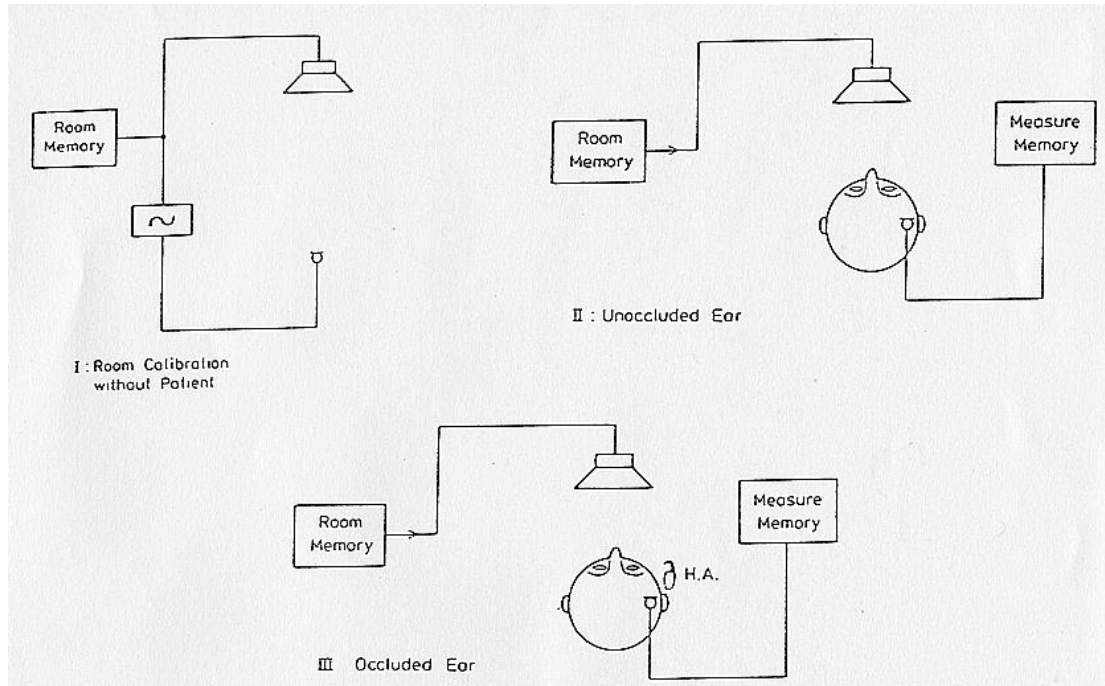


Figure 1: Illustration of the substitution method of sound field equalization for probe-microphone measurements. [From Mueller (1992) used with permission]

The second equalization method is the modified pressure method, also called modified pressure with concurrent equalization method, shown in Figure 2. This method implements the use of a reference microphone positioned at the patient's tested ear. The reference microphone measures the SPL from the loudspeaker. It then adjusts the loudspeaker output levels to maintain the desired levels throughout the measurements (Mueller, 1992; Lantz et al., 2007).

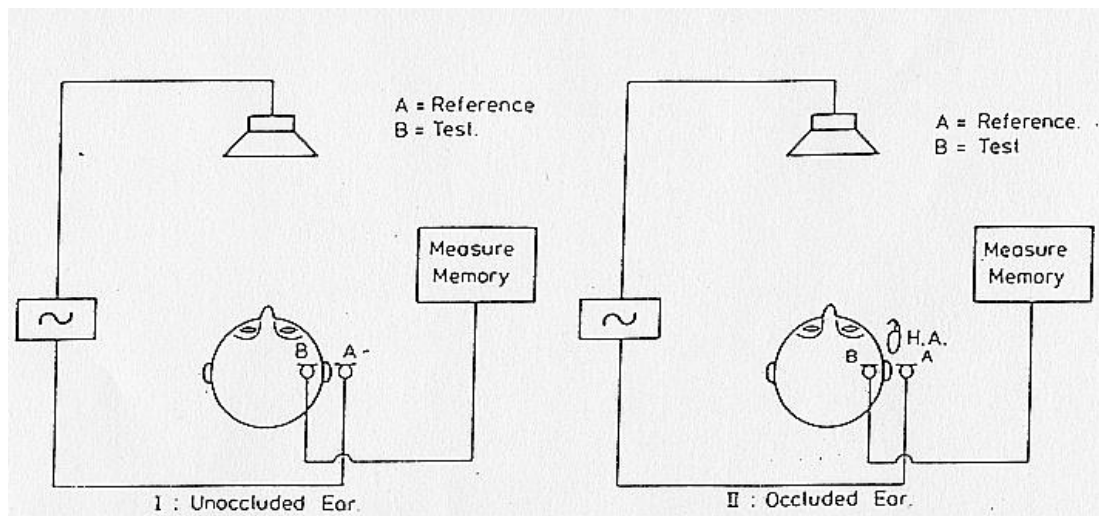


Figure 2: Illustration of the modified-pressure method of sound field equalization for probe-microphone measurements. [From Mueller (1992) used with permission]

1.3.2 Loudspeaker position

Loudspeaker position relative to the patient is another important consideration when conducting the REM (Hawkins & Mueller, 1986, 1992; Ickes et al., 1991). Both distance from and azimuth to loudspeaker can significantly affect the results. A close distance (less than 40 cm) between the patient head and the loudspeaker could bother the patients and make them tend to move back (Mueller, 1992; Stone & Moore, 2004). Moreover, in this placement any small position changes significantly change both intensity and phase of the sound reaching the patient's ear, due to the sound reflections from the pinna, head and torso (Stone & Moore, 2004). However a large distance such as, 1.5 m increases the effect of sound reverberation and background noise on the test results (Hawkins & Mueller, 1992; Stone & Moore, 2004). Moreover, as the distance increases the effect of the head movements during testing on the reliability of the test results will be more pronounced (Ringdahl & Leijon, 1984; Killion & Revit, 1987; Valente et al., 1990). According to these two considerations many studies recommended a distance of around 0.5 m to be used (Hawkins & Mueller, 1992; Killion & Revit, 1993; Stone & Moore, 2004, BSA & BAA, 2007).

Regarding the loudspeaker to patient azimuth, many studies found that 0° and 45° have the best test-retest reliability (Hawkins & Mueller, 1986; Mueller et al., 1992;

Pumford & Sinclair, 2001). However, a recent study by Stone and Moore (2004) indicates that a 0° positioning is more reliable and precise in clinical conditions than the 45° positioning. Stone and Moore (2004) findings are not in agreement with Killion and Revit (1987) who found that a 0° azimuth resulted in greater variability than a 45° azimuth. Positioning the loudspeaker at 45° decreases the effect of unintended head motion and room reflections (Killion & Revit, 1987). However, this disagreement could be due to their methodological differences; in the first study KEMAR manikin was used while repetitive measurements were performed on 10 participants using foam ear-tips in the second study. Moreover, Killion and Revit did not remove the probe tube between their repetitive measures persistently, which decreased their test-retest variability. In addition to that, the methodology in their study minimises the errors resulting from changes in the participant's head position relative to the loudspeaker. Minimizing errors was achieved by asking the participants to "sight across the tip of his or her nose, selecting a spot on the wall for each eye's line of sight" (Killion & Revit, 1993). The variations in the placement of the probe microphone between measurements, and the variations in the BTE hearing aid exact position in the ear were the main reasons for the small variability in the Killion and Revit (1987) study (Stone & Moore, 2004). In clinical practice, the loudspeaker-to-head azimuth would not be achieved as described in Killion and Revit (1987) and the measurements are not usually performed in anechoic conditions with the absence of the tester during measurements as they used.

The errors that result from the loudspeaker distance and azimuth are more significant when the substitution equalization method is used. Therefore, the patient must be placed in a specific location throughout measurements and any movement will decrease the REM accuracy. Since the effect of head movements will not be detected due to the absence of the reference microphone (Hawkins and Mueller, 1986; Mueller, 1992; Pumford & Sinclair, 2001; Shaw, 2010). However, in the modified pressure method any deviation from the desired SPL due to the small head movements is instantly adjusted by the reference microphone. Moreover, any change in the reflection and diffraction effects from the patient's torso and body movement are taken into account (Hawkins & Mueller, 1992; Shaw, 2010).

1.3.3 The probe tube placement

Probe tube placement in the patient's ear canal relative to the tympanic membrane should be taken into consideration for the REM. To ensure “no-effect” of the standing waves and accurate measuring of the high frequency components, a placement of the probe tube tip at a distance of approximately 5 mm from the tympanic membrane should be made. This placement gives accurate measures within 2 dB at the tympanic membrane up to 8000 Hz. However, a 10 mm distance could cause up to 10 dB inaccuracy in results at 8000 Hz (Mueller, 1992; Pumford & Sinclair, 2001).

Sullivan (1988) stated that the standing waves, which are generated from the interaction between acoustic waves in the ear canal and their reflections from the surface of the tympanic membrane, will decrease the measured response. This effect is more pronounced on frequencies greater than the quarter wavelength resonant frequency of patients' ear canal (more than 3000 Hz). Therefore it was suggested to reduce the distance to within 5 mm from the tympanic membrane (Hawkins & Mueller, 1986; Gilman & Dirks, 1986; Dirks & Kincaid, 1987; Sullivan, 1988). To ensure this desired distance between the tympanic membrane and the probe tube tip a clinical visually inserted method can be used. This includes a placement of the probe tube in a constant insertion depth from the patient's intertragal notch. This insertion depth is varied according to different age and gender of patients.

The average length of the adult ear canal is 25 mm and the distance between the adult intertragal notch and the opening of their external ear is usually 10 mm. According to these averages, an insertion depth of 30 mm beyond the intertragal notch should be used to achieve a desired 5mm distance from the tympanic membrane. As a general guideline for adult females and males, the insertion depth of the probe tube should be 28 and 30 mm respectively (Hawkins & Mueller, 1992; Pumford & Sinclair, 2001). Figure 3 shows different probe tube distances from the tympanic membrane and the errors that will be present for frequencies from 1 to 8 kHz. The figure also displays that as the frequency increases more error occurs even in small distances.

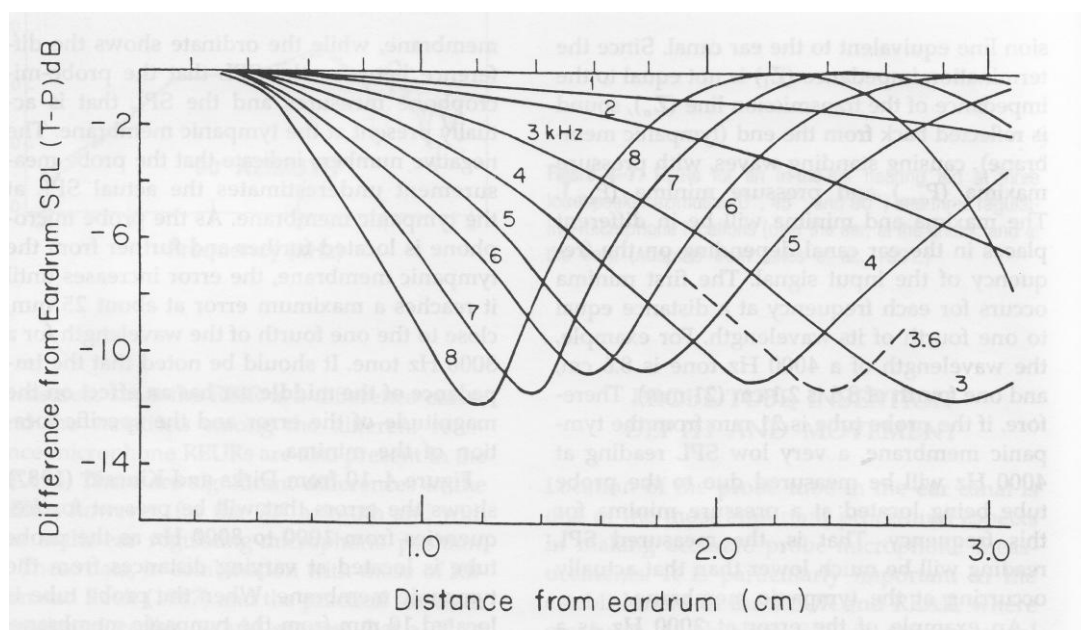


Figure 3: Difference between eardrum- and probe-measured SPL at eight frequencies as a function of distance of the probe from eardrum. [From Hawkins & Mueller (1992) used with permission]

1.3.4 Testing environment

Minimum ambient noise and sound reverberation are the two main conditions for REM environment. Thus, there is no need to do the REM in a sound-attenuating room (BSA & BAA, 2007), but it should be done in a quiet room, in which the test results are not affected by more than 1 dB at any frequency as a result of ambient noise. Moreover, the signal level for all frequencies should be at least 10 dB more than the noise floor (BSA & BAA, 2007).

According to BSA and BAA recommendations (2007), the position of the loudspeaker should not be at the back of a table. Also no large flat reflecting surfaces should be present near the patient. These surfaces should be at least 1 m away from the patient and the loudspeaker. Positioning of the loudspeaker facing the centre of the room is the most suitable position, as this position lessens the sound reverberation effect. However, a 1 m distance away from the corner is acceptable (BSA & BAA, 2007).

Hawkins and Mueller (1986) reported that the difference in non-reverberant and reverberant sound field conditions significantly affects the reliability of

measurements. However, Tecca (1990) reported no significant differences between the two conditions. Sandlin (2000) stated that this disagreement between the two previous studies resulted from the differences between their test conditions. Among other differences, the reverberant booth used in the study by Hawkins and Mueller (1986) had many reflective surfaces near the measurement location, while this was not the case in the study by Tecca (1990).

1.4 Test-retest reliability of REM

1.4.1 Test-retest reliability of REM within the same tester

Test reliability refers to the ability of the test to provide similar results for identical conditions when testing a person more than once (Valente et al., 1990). If the measurement is not reliable, then the test will not be valid and consequently considered not suitable for use. To consider a test as reliable, a high correlation between the test and retest results should be obtained (Valente et al., 1990).

According to Tecca (1990), there are three methods to investigate the test-retest reliability of real ear measurements (REM). The first method is immediate test-retest variability in which repetition of testing is performed on the same patient visit without turning off the system and without removing the probe tube or the hearing aid. Although there is no data available, Tecca stated that in immediate test-retest variability method, a change of more than 2 dB is considered to be significant (Hawkins & Mueller, 1992).

The second method is short-term test-retest variability in which the repetition of measurements is performed at the same patient visit but after removing both the probe tube and the hearing aid and reinserting them (Tecca, 1990). Previous research has shown relatively good short-term test-retest reliability for real ear measurements (REM) using occluding earmoulds (Hawkins & Mueller, 1992). All the following studies investigated the short-term REM test-retest reliability for non open earmoulds and the majority of their data were collected utilizing the modified pressure with concurrent equalization method in which the reference microphone is active during

testing. In the following review of the literature only standard deviation values were reported because the majority of the following studies only used the standard deviation values to present their results.

Ringdahl and Leijon (1984) examined the short-term test-retest difference for the REIG measurement in 20 normal participants. The measurements were performed using special laboratory equipment under anechoic conditions with participants seated at 0.6 metre distance and 0^0 azimuth relative to the loudspeaker. For all participants the same examiner performed the measures using the same behind ear (BTE) hearing aid and individually customized earmoulds. Their findings indicated acceptable test-retest reliability. The standard deviations ranged from 1 to 3.2 dB for frequencies between 0.25 and 4 kHz, and as they measure the higher frequencies (more than 4000 Hz) the standard deviation increased up to 4.7 dB. Dillon and Murray (1987) examined the test-retest difference in 8 hearing impaired participants and used test procedure and conditions that are similar to those in the study by Ringdahl and Leijon (1984). Their results were in agreement with those for Ringdahl and Leijon (1984); there was generally less than 1 dB standard deviation difference between their results across 0.25 to 6 kHz frequency range.

Another study by Killion and Revit (1987) examined the test-retest difference 5 times on each of the 10 participants of their study. In their study both 0^0 and 45^0 loudspeaker-to-head azimuth at 0.5 metre distance were used. The standard deviation of their results at 0^0 azimuth ranged from 1 to 2.3 dB in 0.25 and 4 kHz frequency range, and increased up to 3.5 dB at 6 kHz. These results revealed considerably better test-retest reliability than reported by both Ringdahl and Leijon (1984) and Dillon and Murray (1987). Figure 4 displays the mean test-retest standard deviation results for both Ringdahl and Leijon (1984) and Killion and Revit (1987) studies. Several procedural variables utilized by Killion and Revit (1987) were responsible for the lower test-retest variability seen in their results. The first variable was the instructions they conducted to the participants of their study to reduce their head movement effect on results (see section 1.3.2). The second variable was removing the hearing aid and the eartip and then replacing them without removing the probe tube constantly between each trial, which decreases the test-retest variability. The third variable was the special care that was taken in locating the probe tube in the ear canal and the use

of a well fitted foam earmould. Moreover, the tester was in a separate room throughout the measurements time to reduce reverberations and motions' effects in the test room, which was the fourth variable.

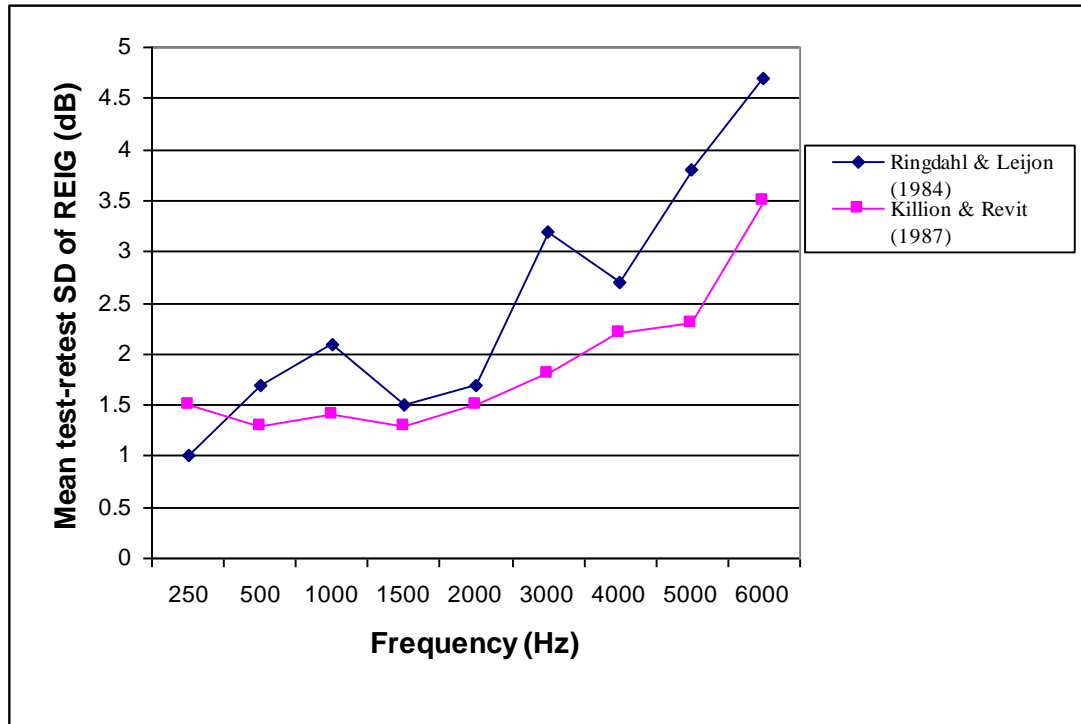


Figure 4: Comparison of mean test-retest SD difference for REIG between Killion and Revit (1987) and Ringdahl and Leijon (1984) results.

Barlow et al. (1988) investigated the test-retest variability for REM with three repeats in 15 hearing impaired participants (ranging from mild to profound) utilizing their own earmoulds and hearing aids. Their findings were considerably similar to that reported by Killion and Revit (1987). They found up to 2.2 dB mean test-retest difference standard deviation at 4 kHz, however; they did not mention a lot of their procedural considerations such as, the location of the loudspeaker relative to the participant and their testing environment. Agreeing with previous studies, Valente et al. (1990) and Hawkins et al. (1991) reported a mean test-retest standard deviation reaching up to approximately 2.2 dB in 1 to 4 kHz frequency range, for the REIG. This high agreement between these studies is due to their similar procedural considerations and their particular attention to positioning of the probe tube in the ear canal. However, Humes et al. (1988) found higher intra-tester test-retest variability than that reported by the previously mentioned studies. In Humes et al. (1988) study three different probe tube systems were compared. They reported up to 5.99 dB intra-

tester test-retest standard deviation at 0.5 to 4 kHz frequency range. Their higher test-retest variability may be due to their large loudspeaker-to-participant distance (1.5 m) compared to the other studies such as, 0.5, 0.6 and 1 m distances in the studies by Killion and Revit (1987), Ringdahl and Leijon (1984), and Dillon and Murray (1987) respectively.

Feigen et al. (1989) observed the test-retest reliability for two times repeated real ear occluded response. Measurements were performed using insert earphones and non-customized probe tips. Their results indicated a good test-retest reliability ranging from 0.5 up to 2.1 dB test-retest standard deviations in 0.25 to 8 kHz frequency range, and their results were consistent with Killion and Revit (1987).

In all these studies increasing the frequency increases the test-retest variability which might be because the high frequencies are more sensitive to subtle changes in the probe tube position in the ear canal (Khanna & Stinson, 1985; Hellstrom & Axelsson, 1994). In addition, a greater influence of standing waves on probe tube measures performed at points remote from the tympanic membrane, results when the wavelength of sound decreases with the increasing frequency (Hawkins & Mueller, 1986). However, poorer test-retest reliability was also observed in low frequencies (0.2 and 0.5 kHz) in some studies (Hawkins, 1987; Humes et al., 1988; Feigen et al., 1989; Valente et al., 1990). This variability may be accounted for by the inability to adequately seal the ear canal by using unsuitable earmoulds and earplugs (Hawkins, 1987; Feigen et al., 1989).

The third method for test-retest variability investigation is the long-term test-retest variability in which the test is repeated at another patient visit (Tecca, 1990). Hawkins and Mueller (1992) postulated that the differences are expected to be a little more than that for short term variability as a result of any slight differences in middle ear pressure and calibration values with time, even though there was no available data to prove this. Hawkins (1987) investigated long-term REIG test-retest reliability using KEMAR manikin. In his study 5 repeated careful measurements, with a specific attention paid to the placement of the probe tube in the ear canal, were done on different days (see section 1.4.2).

1.4.2 Test-retest reliability of REM between testers

All the previously mentioned studies examined the test-retest reliability done by single tester. Regarding the inter-tester reliability Hawkins (1987) examined long-term test-retest REIG reliability across 6 testers. Each tester performed measurements at five separate times on the same participant using the same earmould, hearing aid, and probe tube system. Findings indicated that across testers mean test-retest differences ranged from 2.2 dB at 0.5 kHz up to 7 dB at 6 kHz. Humes et al. (1988) on the other hand, examined short-term inter-tester test-retest reliability. Their findings indicated better across testers mean test-retest differences ranging from 1.45 to 4.06 dB for 0.5 to 4 kHz frequency range. Differences in findings between these two studies are most likely due to their methodological differences, for example, in Humes et al. (1988) two examiners performed the measurements on 12 participants each using foam earplugs and 3 different probe tube systems, while Hawkins (1987) examined the test-retest reliability across 6 testers on a single participant using customized earmould and one probe tube system.

On the contrary to the previous studies, Valente et al. (1990) found no significant difference in the short-term mean test-retest differences across their two testers. The main reason for their lower test-retest variability may be because that they used 0.3 m loudspeaker to participant distance, while Humes et al. (1988) used 1.5 m distance. This increased distance increases the possibility of reflected sound waves to interfere with REM (Valente et al., 1990; Hawkins & Mueller, 1992; Stone & Moore, 2004). Moreover, in the study by Valente et al. (1990), the procedure entailed that special care was taken in positioning the probe tube; it was taped at a point just below the tragus and the participants were instructed to focus on a specific point during the measurement to control the head movements. On the other hand Humes et al. (1988) did not mention any specific procedural consideration to control the participants head movements. In addition, in the study by Valente et al. both in the ear (ITE) and in the canal (ITC) hearing aid users were tested, while in the study by Humes et al (1988) normal hearing participants fitted with earplugs attached to BTE Hearing aid were tested.

1.5 Open-fit hearing aids

1.5.1 Background

Open-fit hearing aid (shortly open-fit), also known as open canal (OC) and open ear hearing aid, is a small non-occluding, non custom, and minimally visible ear-tip, placed in the ear canal and connected to a behind-the-ear (BTE) hearing device by a thin transparent tube (Muller, 2006). Figure (5) shows an example of an open-fit hearing aid.



Figure 5: An open-fit hearing aid.

It is mainly used to allow natural external low frequency sounds to enter the ear canal unamplified (Yanz, 2006). Open-fit hearing aids have become a fast growing and popular phenomenon in recent years in Europe and United States (Fabry, 2007; Scheweiter & Jesse, 2006; Taylor, 2006; and Staples & Aiken, 2006). Open-fit's popularity comes from its cosmetic design in combination with the presence of advanced digital properties such as feedback reduction (Taylor, 2006). Strom (2006) reported that during the year 2006 in the United States (US), 24 percent of behind the ear hearing aids dispensed were open fit. Open-fits were responsible for a growth in

the total sales of BTE hearing aids in the US from a percentage of 26% in 2004 to 66% in the first half of 2010 (Flattens, 2010).

The probable number of hearing loss patients in UK is approximately 9 million patients. The majority of them, around 8 million, have a hearing loss with a mild to moderate severity (RNID, 2008). For these patients, hearing aid fittings will benefit them by helping them restore some degree of their hearing ability (Yanz, 2006). In the UK, only 2 million patients use hearing aids and about a quarter of them do not use their hearing aids frequently (RNID, 2008).

Approximately 20 million Americans who need hearing aids and would benefit from them are not using them (Kochkin & Marketrak, 1994). Kochkin and Marketrak (1994) stated that cosmesis, comfort, cost, performance and patients' beliefs of no need for hearing aid were the main reasons for that. Fortunately, the invention of open-fit hearing aids has helped significantly to handle most of these problems (Kochkin & Marketrak, 1994; Goode & Krusemark, 1999).

The primary candidates of open-fit hearing aids are patients with normal or nearly normal hearing thresholds in low frequencies and a mild to moderate sloping sensory neural high frequency hearing loss (Kuk et al., 2005). Generally in open-fit hearing aids, frequencies less than 1000 Hz are not amplified because any amplification will escape from the un-occluded ear canal (Goode & Krusemark, 1999).

The major limitation of open-fits is the presence of feedback, particularly near the resonant frequency of the ear canal (around 3000 Hz), and this may make the patients hear the unamplified sound through the open-fit before receiving the hearing aid amplified sound (Flynn, 2003). A sensation of reverberant sound can result which bothers the hearing aid user. Because of that, an appropriate feedback cancellation system implementation is important (Flynn, 2003).

1.5.2 Advantages of open-fit hearing aids

The several advantages of open-fit hearing aids make them familiar and widely used in clinical fitting. One advantage of the open-fit hearing aid is that it significantly reduces the probability of occlusion effect which happens frequently with closed earmoulds. Occlusion effect occurs when the sounds of low frequency, produced by the hearing aid users' own voice or other signals such as chewing, are delivered to the ear by bone conduction. Conventional earmoulds, which occlude the ear canal, increase the loudness of those sounds which in turn, bothers the hearing aid user and increases discomfort (Revit, 1992).

Kiessling et al. (2005) investigated the reduction of occlusion effect in the open-fit hearing aids. The findings indicated no significant differences in occlusion effect between the open unaided ear and the open-fit aided measures. However, generalization of this study results is not possible since only one type of the open-fit hearing aid was used.

Another study done by Mackenzie (2006) investigated the occlusion effect in three different open-fit hearing aid models. The findings indicated no significant differences between the unaided and aided conditions in the three different open-fit hearing aid models at low frequencies. In addition, participants reported the same natural own voice quality during the two conditions. According to these findings, the open-fit hearing aid is an effective method in handling the occlusion effect problem in hearing aid users. Moreover, the participants in the study by Mackenzie's (2006) highlighted their comfortable feeling when the open-fit hearing aid was in its place in the ear canal. This is in agreement with Goode & Krusemark's (1999) findings which indicated that the open-fit hearing aid users were significantly satisfied in cosmetic appearance and comfort features of their open-fit hearing aids. The small, thin tube and stylish design features of the open-fit hearing aids make them cosmetically accepted (Dillon, 2001; Mueller & Ricketts, 2006; Staples & Aiken, 2006).

Better sound localization ability in open-fit hearing aids compared to non open-fit designs was reported by Noble et al. (1998). Noble et al. (1998) suggested that listeners with good low frequency thresholds may be more reliant on time and phase cues in sound localization. Closed earmoulds negatively affected horizontal localisation performance, however; the open and sleeve earmould conditions restored performance to unaided levels. This was attributed to the fact that low frequency time and phase cues were undistorted in both of the open earmoulds conditions (Noble et al., 1998).

The various advantageous features of open-fit aids make their users more satisfied with their hearing aids than non open-fit hearing aid users. Moreover, this open invisible ear-tip replaces the traditionally used earmould, thus; there is no need for making an ear impression before the fitting session (Kiessling et al., 2005). Open-fit hearing aids are easily fitted in less than 30 minutes time without a need for doing an ear impression, which save time and effort (Goode & Krusemark, 1999).

With a recent growth in the open-fit sales, it is important to compare the satisfaction and benefits of the open-fit with the non open-fit hearing aids. Taylor (2006), Gnewikow & Moss (2006), Goode & Krusemark (1999) and Christensen & Matsu (2004) used subjective evaluation for the hearing aids performance comparison. All these studies found no significant difference in total satisfaction between open-fit and non open-fit, however; in specific satisfaction aspects open-fit hearing aids were significantly superior to the non open-fit hearing aids. These aspects were the localization, comfort, reduction of occluding effect, and better own voice quality.

1.5.3 Problems with sound field calibration in open-fit REM

If the traditional modified pressure with concurrent equalization method is used with open-fit hearing aids to control the signal, significant error will result (Lantz et al, 2007). This error in measurement is caused by the unwanted adjustments from the reference microphone as the reference microphone is responsible for continually adjusting the loudspeaker output level to deliver the required signal at the patient's ear. In open-fit systems, the reference microphone detects the escaped sound from the

open ear canal in addition to the output loudspeaker sound. As a result, it will decrease the output from the loudspeaker due to the increase in the detected sound pressure level (Staples & Aiken, 2006; Shaw, 2010). Generally these unwanted adjustments resulted in reducing hearing aid gain in the high frequencies and this leads to inaccurate fitting (Lantz et al, 2007).

To solve this inaccurate measurement problem an alternative equalization method has been developed. The alternative method is the modified pressure with stored equalization which is a modification of the modified pressure method. This method requires equalization of the sound field during the unaided measurements using the reference microphone, while the subsequent aided measurements are done without turning on the reference microphone. Therefore, the loudspeaker output will be fixed and will not be adjusted in correspondence to any head movement. This method increases the potential of having unreliable results, as any small patients' head movement throughout testing may change the delivered sound pressure level to the ear. The change in sound pressure level happens due to a change in head and torso reflected sound waves (Ricketts & Mueller, 2009). According to this, accurate positioning of the patient's head is required throughout the REM (Shaw, 2010).

Shaw (2010) examined the effect of using the modified pressure with stored equalization method on the accuracy of REUG measurements. In his procedure the REUG for 20 normal middle ear function participants was measured twice, once before fitting the contra lateral non-test ear with non-functioning hearing aid and once after. Participants were instructed to keep still throughout the measurements to reduce the head movement effect on the test variability. Measurements were done for both 0^0 and 45^0 loudspeaker-to-participant azimuth. His findings indicated no significant difference between both azimuths with less than 1.5 dB standard deviation test-rest difference in 0.25 to 4000 kHz frequency range. His results were not consistent with Stone and Moore (2004) who found more accuracy in results at 0^0 azimuth. This inconsistency may result from their procedural differences. For example, Shaw (2010) recruited real participants and monitored the participants' head movement effect, while Stone and Moore did their measurements using KEMAR manikin, and did their testing in several horizontal and vertical loudspeaker positions.

1.5.4 Test-retest reliability for open-fits

Only one study done by Ricketts & Mueller (2009) investigated the REM short-term test-retest reliability for open-fit hearing aids. The study included only two participants and each was tested twice in each of three different REM systems. The stored equalization method was used as it is necessary to be used during open-fit measures as explained earlier.

The study results showed that there were no significant differences among the three different systems and that the test-retest findings were in agreement with Hawkins et al. (1991) findings. The standard deviations for the mean test-retest differences ranged from 0.5 to 2.9 dB in the frequency range from 1000 to 4000 Hz with the greatest variance at 4000 Hz. Ricketts and Mueller (2009) stated that the similarity in results between modified pressure concurrent equalization method, which was used in Hawkins et al. (1991) study, and their stored equalization may have resulted from the fact that the participants were instructed not to move their heads but to keep them in the same exact locations during the test time instead. Ricketts and Mueller (2009) test-retest reliability findings for open-fit were consistent with that found for conventional earmoulds, as there was less than 1 dB standard deviation difference between their findings and conventional earmould studies' findings (see section 1.4).

However, in Ricketts & Mueller (2009) only two participants were tested. Moreover, in their procedure they did not mention information such as, the distance and the azimuth between the participant and the loudspeaker. These pieces of information are important for the repeatability of the test. They only mentioned that they followed each systems recommended procedure and it could be anywhere between 0.5 and 1 metre as most manufacturers recommend this range of distance (Hawkins & Mueller, 1992). Different distances could affect the test-retest reliability differently (especially in open-fit REM as this distance might significantly affect the results (Hawkins & Mueller, 1992; Shaw, 2010), so it is important to know the exact distance. Their small sample size and unclear method made their results unrepresentative and thus could not be generalized. Therefore, more research is needed to investigate the test-retest reliability for REM using open-fit.

1.6 Research question

Over the past two decades, real ear measurement (REM) has become the standard procedure of hearing aids fitting, to ensure the hearing aid is adjusted according to patients' specific needs. Therefore, measuring the validity and repeatability of REM is important. Many studies reported a good test-retest reliability of REM using occluding earmould hearing aids (Ringdahl & Leijon, 1984; Killion and Revit, 1987; Hawkins et al., 1991; Hawkins & Mueller, 1992; Mackenzie et al., 2004). Recently, a great up-spread of the open-fit hearing aids was reported, as the open-fit hearing aid users are more satisfied and comfortable in using open-fits compared to conventional earmoulds. In REM using open-fit, a different sound field equalization method is used than that used for conventional earmould fittings. This change in the REM procedure could affect the reliability of open-fit REM results. Therefore, the test-retest reliability for conventional occluded earmould REM can not be generalized to the open-fit measures. After reviewing the literature and to the best of the author's knowledge, only one study by Ricketts and Mueller, (2009) has been done to examine open-fit test-retest reliability. However, in their study they used a small sample size (2 participants), and they was not clear in their methodology (see section 1.5.4). In view of that, more research to examine the test-retest reliability for REM of open-fit is needed.

This study is designed to investigate short-term test-retest reliability for open-fit REM, and to compare the test-retest variability results for conventional earmoulds. It is hypothesized that open-fit REM have a good short-term test-retest reliability and the test-retest variability values are in agreement with occluded earmoulds REM test-retest variability values.

2.0 METHOD

2.1 Aim

The aim of this study was to investigate the short-term test-retest reliability of REM using open-fit devices when compared to REM using conventional earmoulds.

2.2 Design

A repeated measures design was used to find the test retest reliability of the REIG using the open-fit and conventional ear moulds at azimuths of 0 and 45 degrees. Three independent variables: earmould type (open-fit and conventional earmould), azimuth (0 and 45 degree), and ear (left and right) are considered in this study. The dependent variable is the difference in dB SPL between the first and the second REIG measured following the replacement of the probe tube, hearing aid and the conventional earmould/open-fit.

Modification of the Latin square generated by a randomisation program was used to randomize the independent variables between the participants and to decide the test order of the conditions for each participant (Brown, 2005, see Appendix A). This was to reduce any order bias and to allow the test to be more sensitive to the experimental manipulating (Brown, 2005). Each participant underwent 4 conditions with two repeats for each earmould type within the same session (total of 8 conditions). Table (1) shows the eight conditions that were examined.

Table 1: The eight experimental conditions that were examined for each participant

Condition	Earmould type and head-loudspeaker azimuth during REM
1	REM using open-fit at 0° azimuth – test 1
2	REM using open-fit at 0° azimuth – test 2
3	REM using open-fit at 45° azimuth – test 1
4	REM using open-fit at 45° azimuth – test 2
5	REM using conventional earmould at 0° azimuth – test 1
6	REM using conventional earmould at 0° azimuth – test 2
7	REM using conventional earmould at 45° azimuth - test 1
8	REM using conventional earmould at 45° azimuth – test 2

2.3 Hypotheses

The study aimed to examine the following hypotheses. These were based on the findings of the previous studies discussed in the introduction.

Primary hypotheses

1. The difference between the test and retest REIG measurements will not be statistically significant for open-fit across all conditions (based upon Ricketts & Mueller, 2009).
2. The difference between the test and retest REIG measurements will not be statistically significant for conventional earmould across all conditions.
3. The test-retest difference values for REIG will not be significantly different between open-fit and conventional earmould (based upon Ricketts & Mueller, 2009).

Secondary hypotheses

4. The test-retest difference values for REIG will not be significantly different between 0° and 45 ° loudspeaker-to-head azimuths (based upon Shaw, 2010).
5. The test-retest difference values for REIG will be significantly different across frequency range.
6. There will not be a statistically significant difference between the test-retest difference values for females and males.
7. There will not be a statistically significant difference between the test-retest difference values for right and left ears.

2.4 Sample Selection

2.4.1 Sample size calculation

Sample Power program was used prior to data collection to determine the appropriate sample size for this study based upon the following criteria:

- A within subject repeated measure design.
- An effect size of 3.6 dB (based on the mean difference of four times measured REIG at an azimuth angle of zero degrees using the conventional earmould in the frequency range 0.25-6.0 kHz found by Ringdahl and Leijon, 1984).
- A standard deviation of the mean difference of 3.65 dB (based on Ringdahl and Leijon, 1984).
- A statistical power of 0.8.
- A significance level of 0.05.

The power analysis indicated that 15 participants are needed to achieve a 0.8 statistical power and a 0.05 significance level. However, 20 participants will be included in this study to increase the statistical power to 0.97.

2.4.2 Participants

Participants were 20 volunteer adults (5 males and 15 females) aged between 22 and 50 years recruited from the student and staff population of the University of Southampton. The mean age for participants was 26 years with a standard deviation of 6 years. One ear of each participant was randomly selected with 10 right ears and 10 left ears tested.

2.4.3 Inclusion and exclusion criteria

All participants were required to have normal middle ear function. This was identified using a screening form (see Appendix B). Otoscopic examination and tympanometry were performed for all participants to ensure normal external ear canal and normal middle ear pressure and compliance. Participants with excessive wax or any external or middle ear pathologies such as otitis media or tympanic membrane perforation were not suitable to take part in this study. Participants with tinnitus were excluded because prolonged hearing aid placement and stimuli presentation may exacerbate the tinnitus.

Prior to testing each participant was required to sign the consent form to participate in the experiment (see Appendix C) after reading the experiment information sheet (see Appendix D).

Participants were required to have their own conventional earmould. Aural impression taking session was required for the participants with no suitable earmould.

The experiment approval was obtained from the Human Experimentation Safety and Ethics Committee of the Institute of Sound and Vibration Research at the University

of Southampton on 1/7/2010 (Approval number 1113). The risk assessment approval was also obtained.

2.5 Stimulus

According to ISVR recommended procedure two different stimuli were used. For conventional earmould measures, FFT speech modification stimulus at 65 dB level was used. While for open-fit REM a swept pure tone at 65 dB level was used.

2.6 Equipment and Apparatus

A GSI calibrated tympanometer, calibrated according to BS EN 60645:2005 standards, and a Heine 2000 otoscope with disposable speculae were used for external and middle ear screening prior to the experiment.

REM were performed using a previously calibrated Aurical Plus diagnostic and fitting test box by GN Otometrics connected with the PC and REM headset. All measures were performed using NOAH Aurical REM software. Six inch flexible silicone probe microphone tubes with a 1 mm internal diameter were used to perform the REM. An Oticon Spirit Zest mini BTE hearing aid was fitted using the click and fit function on the Oticon fitting software according to a mild to moderate sloping hearing loss audiogram. The audiogram hearing level thresholds are shown in Table 2.

Table 2: Hearing level Thresholds across frequencies that were used in hearing aid fitting.

Frequency (Hz)								
	250	500	1000	2000	3000	4000	6000	8000
Hearing threshold (dB HL)	15	20	35	40	50	50	50	50

The Oticon Spirit Zest hearing aid was a digital signal processing programmable hearing aid with eight channels and it was usually used with mild to moderate hearing loss patients.

Each participant undertook the REM with a 1mm vent customized ear mould and with a suitable Oticon Corda2 ear tube and open-fit eartip. Ear tube length and eartip size were determined using the Oticon Corda measuring tool and visual inspection respectively. The Oticon Corda measuring tool is shown in Figure 6. It was placed horizontally and the tube size was chosen according to the number in line with the top of the ear canal opening.

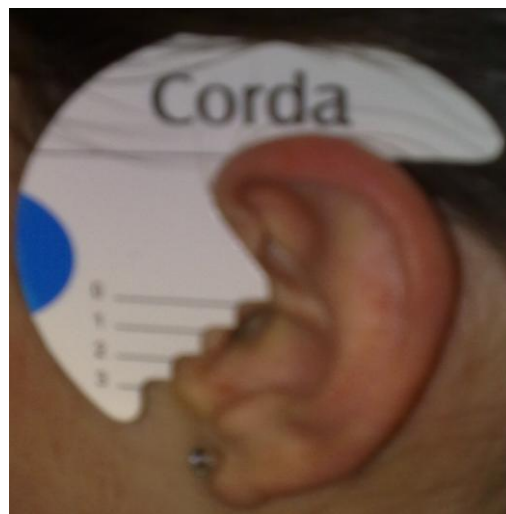


Figure 6: Measuring tool used to select correct Corda2 tube length.

In this study two azimuths were tested, 45° and 0° azimuths. To achieve a precise loudspeaker-to-head azimuth position with restricted head movement, a conventional laser pointer was used. The laser pointer was attached to the top of the participant's head with a headband. Tape markers for each azimuth were positioned on the test room wall and the laser light was pointed to the test azimuth throughout testing.

2.7 Test room set up

All experiments were carried out in a sound proof anechoic booth in the audiology clinic in ISVR. The test room had a carpeted floor, drapes covering walls, and acoustically treated ceiling by sound absorptive foam cones. Figure 7 shows the test room setup. The noise floor in the booth of the present study was less than 30 dBA. However, this environment is not necessary for carrying out REM testing hence it is not affected by a small amount of ambient noise (Tecca, 1990).



Figure 7: Test room setup.

According to the BSA and BAA guidance on the use of REM (2007), participants were seated at a 0.5 m distance between his/her nose tip and the centre of loudspeaker at 0 degree azimuth. To ensure this placement tape marks were used to determine the chair, table and Auricle placement. Moreover a 0.5 m long string connected to the loud speaker was used to check the distance between the participant's nose tip and the loudspeaker before testing. A height adjustable chair was used to ensure that the participants' ear canal and the Aurical loudspeaker were on the same level.

2.8 Daily checks

A daily Tympanometry check was done prior to the testing using the 2cc and 5cc cavity to ensure an accurate measurement. Hearing aid setup and battery were checked daily. Visual inspection of Aurical wires, headset, loudspeaker and buttons were performed daily.

2.9 Test protocol

2.9.1 Preparation

Prior to testing, daily equipments check, room arrangement, and equipment positioning were performed. Tape markers were placed on the test room floor and table to ensure the same placement of the testing chair, table and Aurical device for all participants.

Participants information were entered onto AuditBase System 3 and a mild to moderate sloping hearing loss configuration was entered and saved for all participants. An Oticon hearing aid with a new battery was connected via an Oticon connection led to the Aurical. The hearing aid had been previously programmed using Genie 2008-1 fitting in NOAH-3 software. The hearing aid was automatically fitted according to the mild to moderate sloping hearing loss configuration (see Table 2).

Prior to participant placement, probe tube calibration was performed at 0° azimuth and 0.3 m distance from the loudspeaker. This calibration was done to make the probe tube acoustically invisible and to have no effect on the subsequent measurements. For each participant, a new probe tube was used which had a sliding marker to determine the appropriate insertion depth. In this study, 28 mm and 30 mm probe tube insertion depths beyond the tragus were used for females and males respectively; this was to ensure an approximately 5 mm distance between the probe tube tip and the tympanic membrane.

2.9.2 Screening

Each participant was given an information sheet (see Appendix D), which highlighted the aim of experiment and outlined the participants' task. After that each participant was required to complete a screening form (see Appendix B), if the participants answered "yes" to any of the questions, they were excluded from the study. Participants were then required to read and sign the consent form (see Appendix C).

In accordance to BSA ear examination recommended procedure (2010) otoscopic screening was carried out to rule out any REM and tympanometry contraindications and to ensure a clear ear canal and an intact tympanic membrane. Tympanometry was then carried out to assess the middle ear function. Tympanometry results were considered to be within normal limits if it was in agreement with BSA recommended values (1992); -50 to 50 daPa middle ear pressure and within 0.3 to 1.6 cm³ middle ear compliance. Presence of excessive wax, ear infection or tympanic membrane perforation excluded the subject from the study.

2.9.3 Real Ear Measurements

For each participant an appropriate ear tube length and eartip size were chosen and checked using Oticon Corda measuring tool and visual inspection respectively (see section 2.6).

The participant was placed at 0° azimuth and at a 0.5 m distance between the participant's head and the Aurical loudspeaker. The laser pointer was placed on the participant's head using a headband. The participant was asked to try to maintain the laser dot fixed on the wall marker as much as possible to ensure that the desired azimuth was kept unchanged throughout testing. The ear hanger where the probe tube is attached was hooked over the participant's ears. Then a careful safe probe tube insertion was performed using otoscopy to check correct probe tube placement; lying flat in the ear canal floor without any obstruction. From the software settings and protocols a specific setting utilizing modified pressure with stored equalization method was selected for open-fit REM. On the other hand, for conventional earmould REM a modified pressure with concurrent equalization method was selected. Checking for the appropriate stimulus (see section 2.6) was done and the REUR was measured. Figure 8 shows the probe tube placement during the REUG measurement.

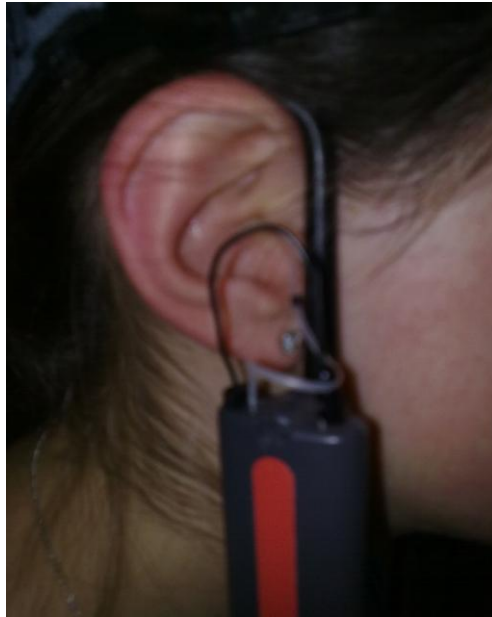
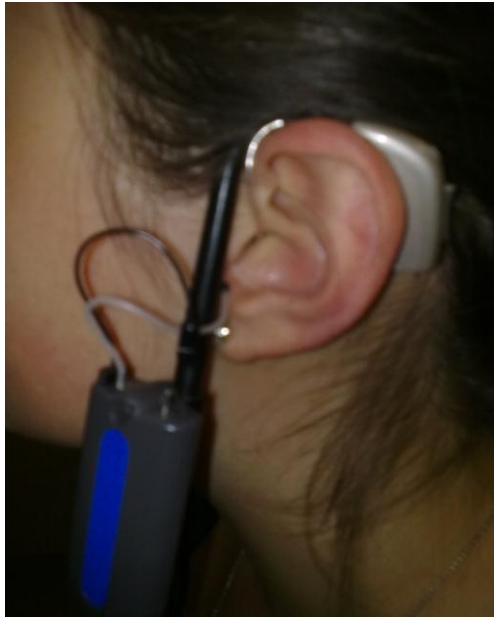


Figure 8: Probe tube placement during the REUG measure.

After the REUG was measured, the participant's conventional earmould/open-fit was placed and the hearing aid was turned on to conduct the REIG measurement. Figure 9 shows the placement during the REIG using open-fit (a) and using conventional earmould (b). Then both REUR and REIG measures were repeated after removing the hearing aid, the conventional earmould/open-fit and the probe tube and then replacing them.



(a)



(b)

Figure 9: Probe tube and hearing aid placement during REIG measure for open-fit (a) and conventional earmould (b).

Each participant underwent all conditions in the order which was determined by the modified Latin Square. At the end of testing otoscopy was performed again. The entire screening and test session for each participant lasted for approximately 45 minutes. To minimize the procedural variability, a single tester performed the whole experiment.

2.10 Pilot Study

Prior to the experiment a pilot study was performed on one participant, in order to check the equipment and test procedure and ensure a reasonable test time. No changes on the test procedures or equipments were made as a result of the pilot study.

2.11 Data Management

All raw data gathered from the experiment was exported to Excel. Mean difference values between the first and the second REIG measures were calculated. The difference values were then entered into the statistical program SPSS, version 17.0 in order to carry out the data analysis.

3.0 RESULTS

3.1 Introduction

The aim of this project was to investigate the short-term REIG test-retest reliability using open-fit ear-tips and to compare it with that using conventional earmould. REIG measurements were conducted for all 8 test conditions on all 20 participants (5 males and 15 females). All participants completed the experiment. In all participants otoscopy and tympanometry screening detected no abnormality in their external and middle ears. The REIG was measured and calculated using REM equipment and software. The results were collated in an Excel spreadsheets and then analyzed using SPSS 17.0. Statistical analysis was performed in order to determine whether or not the results demonstrated a statistically significant difference between the test-retest reliability for open-fits and conventional earmould REM.

3.2 Descriptive Statistics

The test-retest difference was obtained by subtracting the test measurement from the retest one and averaging the data regardless of sign.

Table 3 shows means and standard deviations of open-fit REIG test-retest difference in dB at 0° and 45° loudspeaker-to-head azimuth for each frequency. Results indicated that the standard deviation does not exceed 1.92 dB up to 6 kHz. Open-fit mean test-retest difference values for REIG at 0° and 45° are displayed in Figure 10.

Table 3: Mean and SD in dB for REIG of open-fit test-retest difference at 0° and 45° across frequency.

	<u>250 Hz</u>		<u>500 Hz</u>		<u>1000 Hz</u>		<u>2000 Hz</u>		<u>3000 Hz</u>		<u>4000 Hz</u>		<u>6000 Hz</u>		<u>8000 Hz</u>	
	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°
REIG																
Mean	.83	.53	.83	.79	1.7	.95	1.75	.81	1.52	2.29	1.03	2.61	1.65	1.80	3.17	2.97
SD in dB	.49	.30	.52	.66	1.56	.72	.98	.49	.82	1.87	.85	1.92	1.13	1.35	2.21	1.85

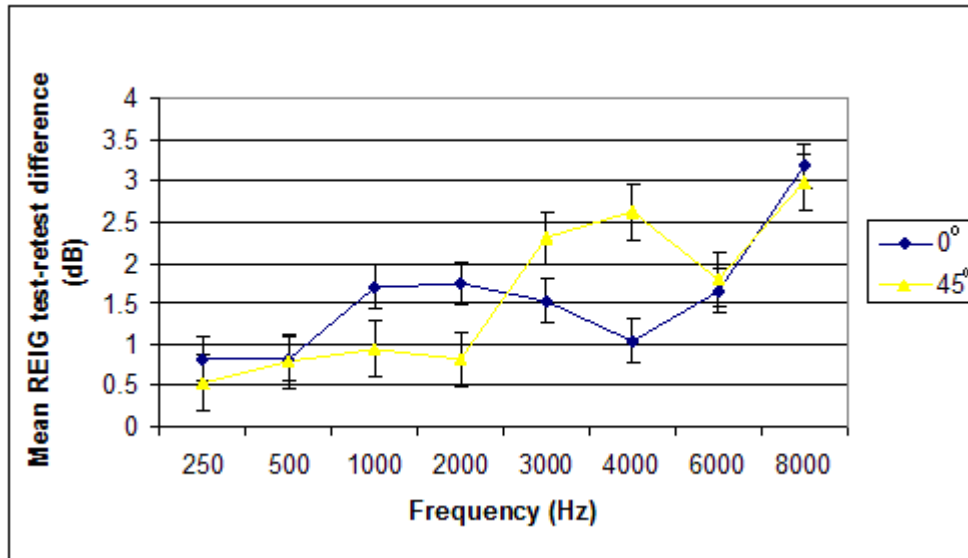


Figure 10: Mean REIG test-retest difference for open-fit at 0° and 45° azimuths. Error bars show +/- 1.0 SD.

Table 4 shows means and standard deviations of conventional earmould REIG test-retest differences in dB at 0° and 45° loudspeaker-to-head azimuth for each frequency. Results indicated that the standard deviation does not exceed 2.45 dB up to 6 kHz. Conventional earmould mean test-retest difference for REIG at 0° and 45° are displayed in Figure 11.

Table 4: Mean and SD in dB for REIG of conventional earmould test-retest differences at 0° and 45° across frequency.

	<u>250 Hz</u>		<u>500 Hz</u>		<u>1000 Hz</u>		<u>2000 Hz</u>		<u>3000 Hz</u>		<u>4000 Hz</u>		<u>6000 Hz</u>		<u>8000 Hz</u>	
	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°	0°	45°
REIG																
Mean	1.71	1.56	2.41	1.86	1.52	1.23	1.40	.83	1.12	1.64	1.38	2.64	2.64	4.50	4.18	3.30
SD																
In dB	1.21	1.36	2.13	1.05	.86	.96	1.25	.85	.76	1.70	1.32	1.74	1.87	2.45	3.58	1.59

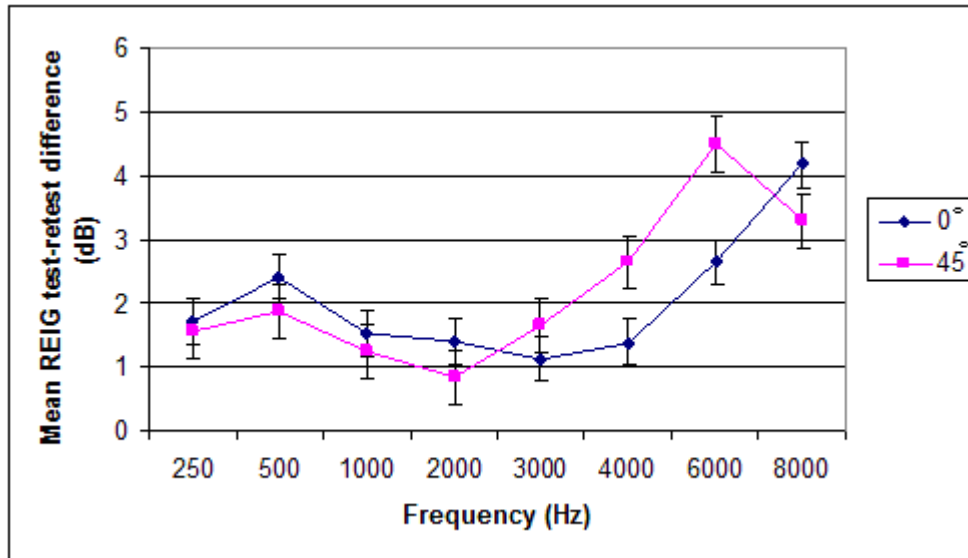


Figure 11: Mean REIG test-retest difference for conventional earmould at 0° and 45° azimuths. Error bars show +/- 1.0 SD.

3.3 Data Distribution

In order to establish whether the data were normally distributed, the Shapiro-Wilk test of normality was used, in addition to examining the normal probability Quartile-Quartile (Q-Q) plots. The repeated measure of the four test conditions across the eight different frequencies resulted in 32 variables. The Shapiro-Wilk test of normality showed that 26 of 32 variables were normally distributed. However, the normal probability (Q-Q) plots showed that the non-normal distributed variables were not significantly deviated from normal distribution. Appendix E shows the normal probability (Q-Q) plots for two of the normally distributed conditions and two abnormally distributed conditions according to Shapiro-Wilk test. Parametric tests were chosen to perform statistical analysis, since the majority of the data were normally distributed and the non-normally distributed variables were not severely deviated, besides analysis of variance (ANOVA) is not significantly sensitive to moderate deviations from normality (Glass et al., 1972; Harwell et al., 1992; Lix et al., 1996).

3.4 Inferential Statistics

3.4.1 Significance of difference between test and retest measures.

Thirty-two paired-sample t-tests (eight frequencies, each tested for open-fit and conventional earmould at 0° and 45° azimuths) were carried out to examine if there is a significant difference between the test and the retest REIG measurements. Table 5 shows the results of the 32 paired-sample t-tests. As a multiple t-tests were conducted, the significance level was adjusted using a Bonferroni correction for 32 t-tests by dividing the significance level of .05 by 32, resulting in a significance level of $p < .001$. The results indicated that there was no significant difference between test and retest REIG measurements as significance values were larger than .001 for all conditions across all frequencies.

Table 5: Results of 32 paired-sample t-tests examining the difference between test and retest REIG measurements for all conditions across frequency.

Paired Samples Test												
	Open-fit						Conventional earmould					
Azimuth	0°			45°			0°			45°		
Frequency (Hz)	t	df	Sig.	T	df	Sig.	t	df	Sig.	T	df	Sig.
250	-3.737	19	.002	2.350	19	.030	1.868	19	.077	1.029	19	.316
500	-1.356	19	.191	.408	19	.688	1.118	19	.277	-.375	19	.712
1000	1.707	19	.104	1.212	19	.240	-.150	19	.883	1.544	19	.139
2000	.902	19	.378	-.345	19	.734	-1.176	19	.254	.969	19	.345
3000	.201	19	.843	-.067	19	.947	.584	19	.566	.882	19	.389
4000	-.715	19	.483	-.387	19	.703	-.748	19	.464	.435	19	.669
6000	.132	19	.897	1.355	19	.191	-.443	19	.663	.677	19	.507
8000	-.544	19	.593	-1.923	19	.070	1.982	19	.062	-.158	19	.876

3.4.2 Analysis of variance

A five-way mixed repeated measures ANOVA was carried out initially to investigate the effect of the independent variables; ear measured, gender, earmould type, loudspeaker-to-head azimuth and frequency, on the REIG test-retest difference. The

dependant variable was the amount of test-retest difference of REIG in dB SPL. The within-subjects factors were earmould type (open-fit and conventional earmould), loudspeaker-to-head azimuth (at 0° and 45°) and frequency (.25, .5, 1, 2, 3, 4, 6, 8 kHz). The between-subjects factors were gender (female and male) and ear that was tested (right and left). The Mauchly's test of Sphericity indicated that for all factors sphericity could be assumed except for frequency, earmould-type*frequency and earmould-type*azimuth*frequency conditions, as a result for these three conditions the Greenhouse-Geisser epsilon was used to adjust the degrees of freedom. The effects of gender and ear were firstly analysed to examine if they could be collapsed across conditions for further analyses.

3.4.3 Effect of gender on the test-retest difference

The data was analysed using the 5-way ANOVA to examine whether the test-retest difference that was measured for each condition differed between male and female participants. Table 6 shows an overall mean and SD for females and males.

Table 6: Mean and SD for the males and females data averaged across conditions.

Gender	Mean test-retest difference (dB)	SD (dB)
Females	1.80	1.30
Males	2.00	1.32

It can be seen from Table 6 that the test-retest difference averaged means and SD are similar for both genders. ANOVA results indicated that no significant difference was found between males and females $F(1, 18) = 1.475$, $p = .242$. Therefore, further analyses in the following sections were not examined separately for females and males and a 4-way ANOVA was carried out.

3.4.4 Effect of ear on the test-retest difference

The data was analysed using a 4-way ANOVA to examine whether the test-retest difference that was measured for each condition differed between the right and left ears. Table 7 shows an overall mean and SD for the right and left ears.

Table 7: Mean and SD for the right and left ears data averaged across conditions.

Ear	Mean test-retest difference (dB)	SD (dB)
Right	1.83	1.38
Left	1.86	1.28

It can be seen from Table 7 that the test-retest difference averaged means and SD are similar for both ears. ANOVA results indicated that no significant difference was found between right and left ears $F(1, 18) = .143$, $p = .710$. As a consequence, further analyses in the subsequent sections were not examined separately for right and left ears and a 3-way ANOVA was carried out.

3.4.5 Effect of loudspeaker-to-head azimuth on the test-retest difference

REIG test-retest difference was examined in relation to loudspeaker-to-head azimuth. Table 8 shows an overall mean and SD for 0^0 and 45^0 azimuths.

Table 8: Mean and SD for 0^0 and 45^0 azimuths data averaged across conditions.

Azimuth	Mean test-retest difference (dB)	SD (dB)
0^0	1.80	1.34
45^0	1.89	1.30

It can be seen from Table 8 that the test-retest difference averaged means and SD are similar for both azimuths. Figure 12 shows open-fit mean test-retest difference for both 0^0 and 45^0 across frequency.

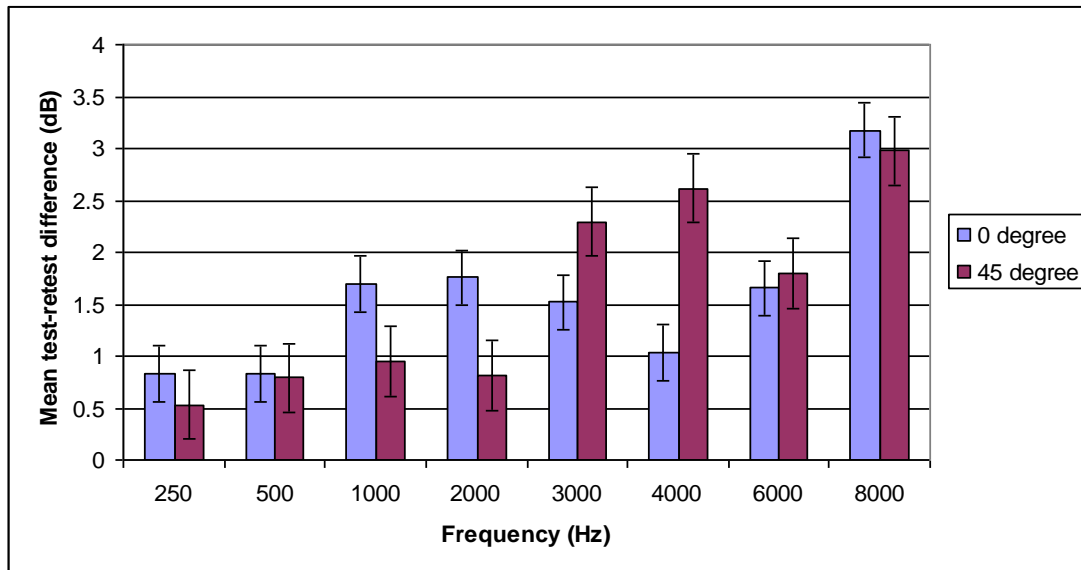


Figure 12: Open-fit mean test-retest difference at 0° and 45° across frequency. Error bars show +/- 1.0 SD.

Figure 13 shows conventional earmould mean test-retest difference for both 0° and 45° across frequency.

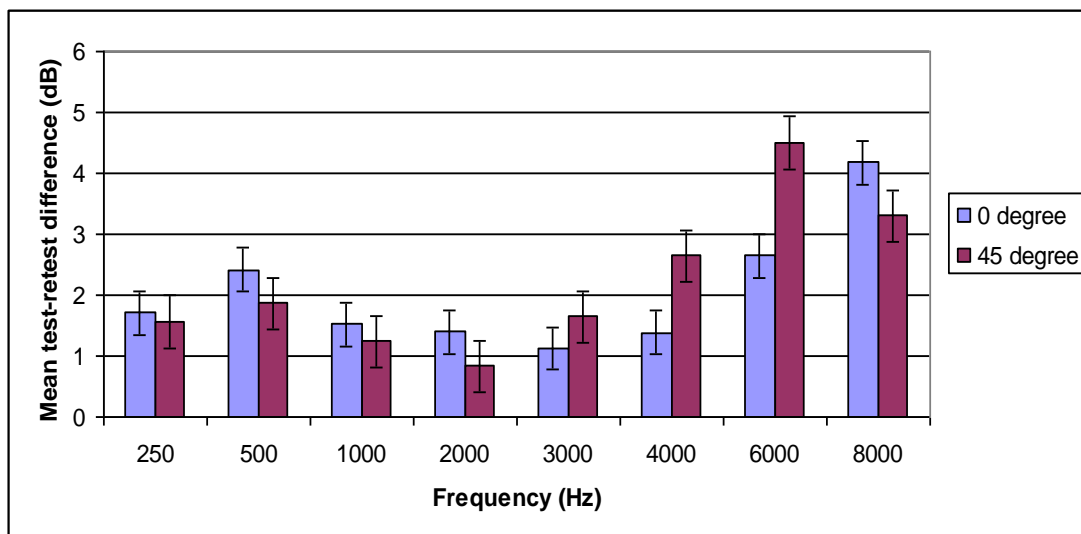


Figure 13: Conventional earmould mean test-retest difference at 0° and 45° across frequency. Error bars show +/- 1.0 SD.

Both Figures 12 and 13 shows that the mean test-retest difference was larger at 0° up to 2000 Hz, while at 3,4, and 6 kHz the mean test-retest difference was larger at 45°. However, ANOVA results indicated no significant test-retest difference between the two azimuths in all test conditions, $F(1, 19) = .248$, $p = .624$. Therefore, further analyses in the following sections were done with both azimuth conditions collapsed and a 2-way ANOVA was carried out.

3.4.6 Effect of earmould type on the test-retest difference

The effect of the earmould type on the REIG test-retest difference was examined averaged across azimuth. Table 9 shows an overall mean and SD for open-fit and conventional earmould.

Table 9: Mean and SD for open-fit and conventional earmould data averaged across conditions.

Earmould type	Mean test-retest difference (dB)	SD (dB)
Open-fit	1.57	1.10
Conventional earmould	2.12	1.54

It can be seen from Table 9 that the test-retest difference averaged means and SD are dissimilar for open-fit and conventional earmould. Figure 14 shows the mean test-retest difference between open-fit and conventional earmould across frequency.

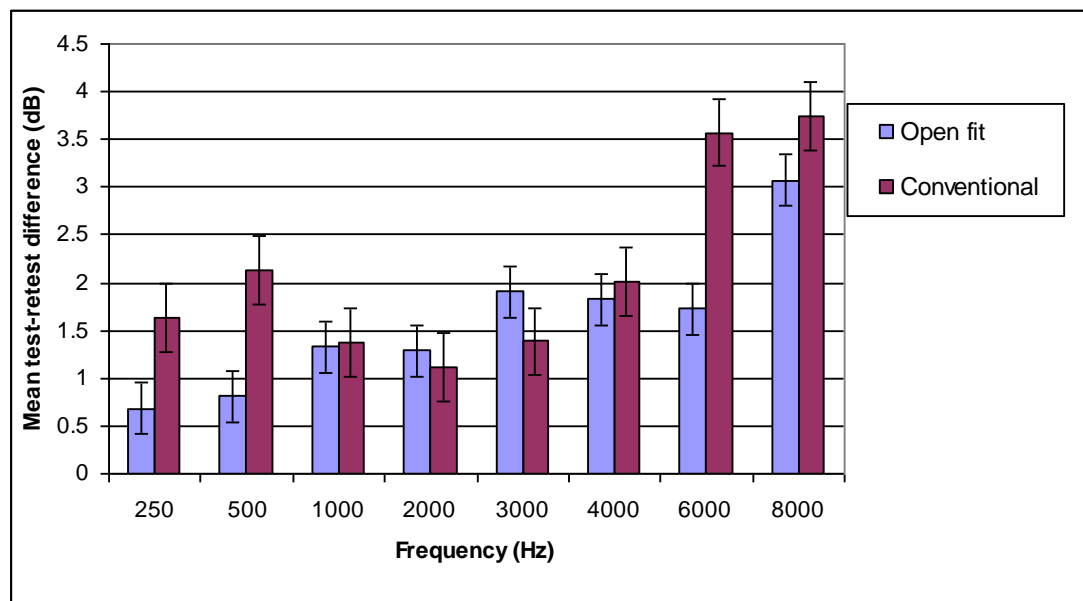


Figure 14: Mean REIG test-retest difference for open-fit and conventional earmould. Error bars show ± 1.0 SD.

Figure 14 reveals that for mid-frequencies (1, 2, 3, 4 kHz) the mean test-retest difference was almost the same for both open-fit and conventional earmould. However, at both low frequencies (.25 and .5 kHz) and high frequencies (6 and 8 kHz) the mean test-retest difference was larger for conventional earmould than for open-fit. The results of the ANOVA showed the effect of earmould type to be

significant $F(1, 19) = 13.690$, $p = .002$, with a small effect size (partial eta squared = .215) and a small observed power (.149).

3.4.7 Effect of the stimulus frequency on the test-retest difference

The effect of the stimulus frequency on the REIG test-retest difference was examined. Figure 14 in 3.4.6 section shows the mean test-retest difference for all frequencies for open-fit and conventional earmould. It can be seen from Figure 14 that the mean test-retest difference for 6 and 8 kHz were larger than that for the other frequencies for both open-fit and conventional earmould. The results were analysed using the ANOVA to establish whether there was significant of frequency. Using Greenhouse-Greisser epsilon as sphericity could not be assumed for frequency as Mauchly's test of sphericity had indicated. The ANOVA revealed a significant effect of frequency $F(3.9, 74) = 26.4$, $p = .000$, with a small effect size (partial eta squared = .219) and a small observed power (.153). Pairwise comparisons were then examined to establish which frequencies were significantly different from each other. This was done by comparing each frequency with all other frequencies. A Bonferroni correction was applied to the significant level to account for the multiple comparisons. The Pairwise comparisons results indicated that 6 and 8 kHz frequencies had a significantly higher mean test-retest difference values than all the other test frequencies (see Appendix F).

3.4.8 Interaction effects

The 2-way ANOVA examination of interaction effects was carried out on the collapsed data. The results revealed that the following interactions were statistically significant. Azimuth*Frequency interaction effect was significant $F(7,113) = 7.259$, $p = .000$, with a small effect size (partial eta squared = .207) and a small observed power (.141). This interaction effect suggests that the effect of azimuth varies with frequency. This interaction was not examined further because of the non significant overall effect of azimuth. A significant interaction effect was also found for Earmould-type*Frequency $F(3, 71) = 6.038$, $p = .000$, with a small effect size (partial eta squared = .204) and a small observed power (.139). Figure 14 in section 3.4.7 shows that the mean test-retest difference was larger for conventional earmoulds at

.25, .5, 6 and 8 kHz than that for open-fit. Sixteen paired-sample t-tests were carried out to examine at which frequencies the test-retest difference between open-fit and conventional earmould is statistically significant. Table 10 shows the results of the 16 paired-sample t-tests. As a multiple t-tests were conducted, the significance level was adjusted using a Bonferroni correction, resulting in a significance level of $p < .003$. The results indicated that there was significant difference between open-fit and conventional earmould at 250, 500 and 6000 Hz.

Table 10: Results of 16 paired-sample t-tests examining the test-retest difference between open-fit and conventional earmould at each frequency for both azimuths.

Paired Samples Test						
	0°			45°		
Frequency (Hz)	t	df	Sig.	T	Df	Sig.
250	-3.338	19	.003	-3.297	19	.003
500	-3.419	19	.003	-4.021	19	.001
1000	.453	19	.655	-1.018	19	.321
2000	.971	19	.344	-.093	19	.927
3000	1.414	19	.174	-1.223	19	.236
4000	-1.054	19	.305	-.052	19	.959
6000	-2.798	19	.003	-4.180	19	.001
8000	-1.036	19	.313	-.567	19	.578

4.0 DISCUSSION

The aim of this study was to examine the short-term test-retest reliability for REIG for open-fit in comparison to conventional earmoulds test-retest reliability. The test-retest difference for REIG was also examined with regards to azimuth, gender, the ear measured and stimulus frequency. The analysed results are discussed below in relation to previous findings and clinical relevance.

4.1 The effect of gender on the test-retest difference for REIG

For this experiment, the results were split by gender and analysed to establish whether there was a significant difference between the test-retest differences for REIG for females and males. The mean test-retest difference values displayed in Table 6 indicate that females and males have a very similar mean and SD values. Statistical analysis indicated no significant difference between females and males. Therefore, Hypothesis 6 (see section 2.3.1) which states that there will be no statistical significant difference between the test-retest differences for REIG for females and males can be accepted. This finding cannot be compared to the results of other studies as no examination of the differences between genders was conducted. For example, Ringdahl & Leijon (1984) and Valente et al. (1990) studies recruited both genders in their experiments, but did not mention whether or not there was a significant gender effect on test-retest REIG measurements.

4.2 The effect of ear on the test-retest difference for REIG

In this experiment, the effect of ear on the test-retest difference for REIG was analysed. This was to ascertain whether the test-retest difference values are similar and the data could be collapsed across ears. Table 7 shows that the mean test-retest difference values are the same for both right and left ears. Statistical analysis indicated no significant ear effect and therefore hypothesis 7 (see section 2.3.1) can be accepted. As there was no significant difference between ears, the further findings of

the present study could be applied to both ears. This finding cannot be compared to the results of other studies as no examination of the differences between ears was conducted. For example, Valente et al. (1990) tested both left and right ears but did not mention whether or not there was a significant effect of ear on test-retest REIG results.

4.3 The test-retest reliability for REIG using open-fit

The test-retest differences of REIG for open-fit were found to be not significant in this study, across all test conditions. Results in Table 3 revealed less than 1.4 dB mean SD of test-retest difference in .25 to 4 kHz frequency range averaged for both azimuths, which indicated good test-retest reliability relative to most behavioural measures of auditory performance (Feigin et al., 1989). This was in agreement with Ricketts and Mueller (2009) who found reliable test-retest REIG results for open-fit. However, the present study test-retest difference results were better than that found by Ricketts and Mueller (2009) in up to 1.3 dB SD difference for 1 to 4 kHz frequency range. The precise procedure used for controlling head movement in the present study may be the main reason for better test-retest reliability. In addition, Ricketts and Mueller (2009) small sample size (2 participants), and their unclear methodology (see section 1.5.4) might affect their reliability findings. The findings of the present study cannot be comprehensively compared to the results of Ricketts and Mueller's (2009) findings as in their study the averaged standard deviation in only 1 to 4 kHz frequency range was mentioned. Therefore, no high or low frequencies comparisons can be done.

Table 5 indicated that no significant test-retest difference was found using open-fit across .25 to 8 kHz frequency range for both azimuths. According to the findings of the present study, the experimental hypothesis (hypothesis 1; see section 2.3.1) can be accepted.

4.4 The test-retest reliability for REIG using conventional earmould

In this study, the test-retest differences of REIG for conventional earmould were also found to be not significant across all test conditions. Results in Table 4 revealed up to 1.6 dB mean SD of test-retest difference in .25 to 4 kHz frequency range averaged for both azimuths which indicated good test-retest reliability. This was expected as all previous studies indicated good test-retest reliability for conventional earmould (see section 1.4). The present study findings are constant with up to 2.2 dB SD findings for Killion and Revit (1987), Barlow et al. (1988), Feigen et al. (1989), Valente et al. (1990) and Hawkins et al. (1991) in .25 to 4 kHz range. The present study findings are also better than the up to 2.6 dB SD reported by Ringdahl and Leijon (1984) and Dillon and Murray (1987) in .25 to 4 kHz. Regarding the higher frequencies the present study findings were consistent with all the previous studies findings, as generally when the frequency increases the test-retest variability increases. Table 4 showed that the test-retest standard deviation difference for up to 8 kHz frequency did not exceed 2.5 dB averaged across both azimuths, compared to up to 2.1, 3.5, 4.7 dB standard deviation differences found by Feigen et al. (1989), Killion and Revit (1987) and Ringdahl and Leijon (1984) studies respectively. The present study lower test-retest variability findings might be mainly due to the accurate procedure used for controlling head movement and maintain it in the exact position relative to the loudspeaker. In addition to other procedural differences such as, the shorter loudspeaker-to-head distance (.5 m) used in this study comparing to the .6 and 1 m used in Ringdahl and Leijon (1984) and Dillon and Murray (1987) respectively. Moreover, particular care was taken to place and keep the probe tube in the same position relative to tympanic membrane for each measurement.

According to the findings of the present study, the experimental hypothesis (hypothesis 2; see section 2.3.1) can be accepted.

4.5 The effect of loudspeaker-to-head azimuth on the test-retest difference for REIG

In this experiment, participant's head azimuth relative to the loudspeaker was found to have no significant effect on REIG test-retest difference for both open-fit and conventional earmould.

Test-retest variability findings using open-fit were in agreement with Shaw (2010), as he found no significant test-retest difference between 45^0 and 0^0 for REUG measurements utilizing the modified pressure with stored equalization method. This agreement in findings is due to the similarity in both measurement procedures. Both Shaw (2010) and the present studies used same loudspeaker-to-head distance, same REM system (Aurical), same participants' number and inclusion criteria, and same probe tube insertion depth. In addition, in Shaw (2010) participants were instructed to focus on a specific point on the wall during the 45^0 azimuth measures and to not move their heads throughout the whole measurements.

In this experiment's findings, even though there was no significant difference between both azimuths, as the frequency increased above 2 kHz, the mean test-retest difference at 45^0 increased more than that for 0^0 . Test-retest variability findings using conventional earmould were consistent with the findings of Stone and Moore (2004). A possible explanation of better test-retest reliability found in the present study compared to Stone and Moore (2004) study that in the present study special care was taken to ensure no head movement during the measurements (by using the laser light attached to the participants' headband in order to keep the light at a specific point on the wall). In addition, the present study measurements were performed in anechoic environment condition compared to near-anechoic and different positions measurements utilized in the Stone and Moore (2004) procedure.

However, even though Killion and Revit (1987) did their experiment using conventional earmould in anechoic double-walled room and monitored their 10 participants head movements, their test-retest reliability findings were worse than that

found in the present study (up to 2 dB standard deviation difference at 8 kHz). Moreover, they found that 45° azimuth had better tests-retest REIG reliability compared to 0° azimuth. In their procedure in order to control head movement, they instruct each participant to look across the tip of his/her nose and to select a point on the wall for each eye's of sight, maintaining that head position throughout the measurements. In the present study, it was easy to monitor even small head movements as both the participant and the examiner were able to see the deviation of laser light from the specific point in the wall. The difference in head movement control procedures and their use of the foam eartip might be the reasons for inconsistency in findings between their study and the present study.

It was hypothesised that there would not be a significant test-retest difference between 45° and 0° loudspeaker-to-head azimuth. This was based upon Shaw (2010) findings. According to the analysed data results, the experimental hypothesis (hypothesis 4; see section 2.3.1) can be accepted.

4.6 The effect of earmould type on the test-retest difference for REIG

The findings of the present study have revealed a significant test-retest difference between open-fit and conventional earmould. Figure 14 and Table 10 indicated that the test-retest difference is more pronounced at both low frequencies (.2 and .5 kHz) and high frequencies (above 4 kHz) in conventional earmould. This observed difference may be due to the different sound equalization methods used for open-fit and conventional earmould. A modified pressure with stored equalization method was used for open-fit, while a modified pressure with concurrent equalization was used for conventional earmould. As a result, any low frequency sound leakage resulted from the conventional earmould 1mm vent or un-tight earmould sealing has reached the active reference microphone (Hawkins, 1987; Feigen et al., 1989). Therefore, higher test-retest variability in low frequency range using the conventional earmould was observed. Moreover, it was harder to place the conventional earmould in the ear canal without affecting the probe tube position than that for open-fit. Small probe tube placement difference relative to the tympanic membrane between aided and unaided measures mainly affect the high frequency range (see section 1.3.3). As a result,

higher test-retest variability in high frequency range using the conventional earmould was observed. This measurement variability at low and high frequency was previously noticed in the majority of the previous studies (see section 1.4).

It was hypothesised that open-fit and conventional earmould would not have significant test-retest difference. This was based upon Ricketts and Mueller (2009). They found similar test-retest variability results for open-fit, using modified pressure with stored equalization method, when compared to Hawkins et al. (1991) results for conventional earmould, using modified pressure with concurrent equalization method. Both these studies test-rest difference findings were analysed for 1 to 4 kHz frequency range. In agreement with Ricketts and Mueller (2009), the present study results found almost similar test-retest difference values for both open-fit and conventional earmould at 1 to 4 kHz frequency range (see Figure 14). However, based upon the present study statistical finding, the experimental hypothesis (hypothesis 3; see section 2.3.1) can be rejected.

4.7 The effect of stimulus frequency on the test-retest difference for REIG

The test-retest variability was inspected across frequency range to examine whether there were significant differences between the test-retest difference values for the frequencies compared with each other. Figure 14 displayed that the mean test-retest difference at each frequency increases for the frequencies above 4 kHz, with the most variability at 8 kHz. In addition, Tables 3 and 4 displayed that the SD generally increases as the frequency increased. Statistical analysis indicated a significant test-retest difference for REIG between frequencies. When examined using pairwise comparison, findings indicated that the test-retest reliability for 6 and 8 kHz frequencies were significantly less than that for other test frequencies. Therefore the experimental (hypothesis 5; see section 2.3.1) can be accepted. This frequency difference was expected prior to conducting the measurements. This was because the majority of the previous studies (i.e. Ringdahl & Leijon, 1984; Killion & Revit, 1987; Hawkins, 1987; Humes et al., 1988) found that greater test-retest differences at high frequencies (more than 3 kHz). This effect was due to the fact that high frequencies

are more sensitive to subtle changes in the probe tube position in the ear canal, as standing waves effect is more pronounced on frequencies greater than the quarter wavelength resonant frequency of patients' ear canal (more than 3000 Hz) (Hawkins & Mueller, 1986; Gilman & Dirks, 1986; Dirks & Kincaid, 1987; Sullivan, 1988).

4.8 Clinical implications

According to the findings of the present study linked to the previous studies' findings, some clinical implications can be suggested. In clinical practice, as the present study found no statically significant difference between 0^0 and 45^0 azimuths, it is recommended to use a 0^0 azimuth placement. Positioning the patient at 0^0 azimuth is easier than the 45^0 azimuth during the REM especially when bilateral hearing aid fittings is needed; as the later requires two patient positions if a 45^0 azimuth is used.

Moreover, according to the present study findings, open-fit REM has good test-retest reliability so this is a reliable test to be used clinically. It is recommended in clinical practice to use the modified pressure with stored equalization method for open-fit REM, in order to minimise the test variability across high and low frequency. In addition, a special care should be taken to control the head movement during the REM.

4.9 Limitations

In this study, a number of considerations were taken into account throughout testing to avoid any factors that would reduce the experiment validity. For example, maintaining consistent participant position relative to the loudspeaker and consistent probe tube insertion depth throughout testing were considered. Moreover, to limit errors sufficient sample was tested and this study sample power calculation indicated a 0.95 sample power, which is considerably high. The power analysis using the measured effective size (0.6 dB) and SD (1.3 dB) of this study indicated that 13 participants are needed to achieve a 0.8 statistical power and a 0.05 significance level. This sample power calculation is in agreement with the previously made power

calculation based on previous studies, which indicated that 15 participants are needed to achieve the same sample power (see section 2.4.1).

However, the experiment does have some limitations that may have affected the accuracy of the findings that were obtained. Firstly, due to time constraints only two repeated measurements were performed on each participant at each azimuth and for each ear tip. More repeated measurements would give more reliable results.

A second limitation is that the tester was not expert in probe tube measurements. Therefore, there is a possibility that more accuracy in probe tube and earmould placement progressed with the time. However, special care was taken in performing the test procedure throughout the experiment to ensure accurate results.

A further limitation of this study is the experiment was conducted on participants with a mean age of 26 years. The majority of hearing aid users are much older than this age. Moreover, the participants were instructed to minimise their head movement throughout the measurements utilizing the attached to head laser light. Elderly participants are more likely to have cognitive and health problems (Worrall & Hickson, 2003) which may affect their ability in maintaining their head in a fixed position throughout testing. Therefore, it is questionable whether the study findings can be generalised to population of hearing aid users.

Another limitation of this study is that a Bonferroni adjustment was used in many of the statistical analyses. As the number of multiple tests increases, the probability of finding an effect by chance increases. In order to overcome this, a Bonferroni correction was applied and the significant level made stricter, to make it harder to find a significant effect (Dancey & Reidy, 2004). Therefore, it is possible that a significant test-retest difference may have been found for several more variables if the significant level adjustment had not been so stringent.

4.10 Further research

The present study findings indicated that there are several areas where further research could be conducted. Firstly, it would be beneficial to repeat the same experiment using more representative sample from hearing aid users with wide age range. Moreover, repeating the same measurements with more repeats would seem more reliable.

A further study could be conducted in conditions more similar to the clinical situation. For example, monitoring the head movement using routine clinical instruction in usual clinical environment rather than anechoic booth would seem more reliable. Moreover, further study on the several clinically used loudspeaker-to-participant distances would seem to be clinically feasible.

More extensive research could also be performed examining the open-fit test-retest reliability such as; comparing the test-retest reliability using the modified pressure with concurrent sound equalization method with that using the stored sound equalization method and comparing the test-retest reliability across different examiners would seem to be clinically feasible.

5.0 CONCLUSIONS

Good test-retest reliability was found for both open-fit and conventional earmould, as no significant difference between the test and retest measurements was found for REIG using open-fit and conventional earmould. In agreement with the previously published findings, the test-retest difference for REIG was found to vary significantly with frequency, with the most variability being present at 6 and 8 kHz. The present study findings indicated that the test-retest difference for REIG varied significantly between open-fit and conventional earmould. It was also found that there was no significant difference between 45^0 and 0^0 loudspeaker-to-head azimuths for both open-fit and conventional earmould REM.

As it was expected no significant difference was found between the test-retest difference for REIG between right and left ears and between males and females.

6.0 REFERENCES

Ackley R, Decker T, & Limb C (2007). An essential guide to hearing and balance disorders. Lawrence Erlbaum Associates.

Barlow N, Auslander M, Rines D & Stelmachowicz P (1988) Probe-tube microphone measures in hearing-impaired children and adults. *Ear Hear.* 9, 243-247.

Brown J (2005) Experimental design generator and randomiser. [Online]. Available from: <http://www.edgarweb.org.uk>.

British Society of Audiology and British Academy of Audiology (BSA & BAA) (2007). Guidelines on the use of real ear measurements to verify the fitting of digital signal processing hearing aids. <http://www.thebsa.org.uk/docs/RecPro/REM.pdf>

British Society of Audiology (BSA) (1992). Recommended procedure for tympanometry. *British journal of Audiology.* 26, 255-257.

British Society of Audiology (BSA) (2010). Recommended procedure for ear examination. http://www.thebsa.org.uk/docs/RecPro/RecProc_EarExam_25Jan2010.pdf

BSEN60645-5: 2005 Electroacoustics – Audiometric equipment. Part 5: Instruments for the measurement of aural acoustic impedance/admittance. London: British Standards Institution.

Burkhard M & Sachs R (1975). Anthropometric manikin for acoustic research. *J Acoust Soc Am.* 58, 214–222.

Christensen L & Matsui G (2004). Hearing aid satisfaction with ReSoundAir. GN Resound White Paper.

Dancey C & Reidy J (2004). Statistics without maths for psychology 3rd, Prentice Hall , London

Dillon H & Murray N (1987). Accuracy of twelve methods for estimating the real ear gain of hearing aids. *Ear Hear.* 8, 2-1.

Dillon H (2001) *Hearing Aids*. Sydney: Boomerang Press, pp 93.

Dirks D & Kincaid G (1987). Basic acoustic considerations of ear canal probe measurements. *Ear Hear.* 8, 60S–67S.

Fabry D (2007). Open-fit hearing aids and the topics of directionality, omnidirectional performance in noise, and real-ear measurement. *The Hearing Review*. http://www.hearingreview.com/issues/articles/HPR_2007IN_03.asp.

Feigen J, Kopun J, Stelmachowicz P & Gorga M (1989). Probe-tube microphone measures of ear canal sound pressure levels in infants and children. *Ear Hear.* 10, 254–258.

Flattens A (2010) HJ Report. *Hearing journal.* 63, 7-8. Available from: http://journals.lww.com/thehearingjournal/Fulltext/2010/08000/Hj_Report.2.aspx

Flynn M (2003). Opening Ears: The scientific basis for an open ear acoustic system. *Hearing Review.* 10, 34-67.

Gilman S & Dirks D (1986). Acoustics of ear canal measurements of eardrum SPL in simulators. *Journal of the acoustical society of America* 80, 783-793.

Glass G, Peckham P & Sanders J (1972). Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. *Rev. Educ. Res.* 42, 237-288.

Gnewikow D & Moss M (2006). Hearing aid outcomes with open- and closed- canal fittings. *The Hearing Journal.* 59, 66-72.

Goode R & Krusemark J (1999). Advantages of a new miniature hearing aid for mild to moderate hearing loss. *Laryngoscope*. 109, 1919-1923.

Harrowven R (1998). Insertion gain versus median ear corrected coupler gain: a comparison of two fitting methods in new NHS hearing aid users. *Br J Audiol*. 32, 153–165.

Harwell M, Rubinstein E, Hayes W & Olds C (1992). Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. *J. Educ. Stat.* 17, 315-339.

Hawkins D & Mueller H (1986). Some variables affecting the accuracy of probe tube microphone measurements, *Hearing Instruments* 37: 8-12.

Hawkins D & Mueller H (1992). Procedural considerations in probe-microphone measurements. In Mueller H, Hawkins D & Northern J (eds.) *Probe microphones measurements. Hearing aid selection and assessment* . San Diego: Singular Publishing Group, pp. 67_89.

Hawkins D (1987). Variability in clinical ear canal probe microphone measurements. *Hear Instrum*. 38, 30-32.

Hawkins D, Alvarez E & Houlihan J (1991). Reliability of three types of probe tube microphone measurements. *Hear Instr* 42,14-16.

Hellstrom P & Axelsson A (1994). Miniature microphone probe tube measurements in the external auditory canal. *J. Acoust. Soc. Am.* 93, 907-919.
[Http://medgadget.com/archives/ent](http://medgadget.com/archives/ent).

Humes L Hipskind N & Block M (1988). Insertion gain measured with three probe tube systems. *Ear Hear*. 4, 108-112.

Ickes M, Hawkins D & Cooper W (1991). Effect of reference microphone location and loudspeaker azimuth on probe tube microphone measurements. *Journal of the American Academy of Audiology*. 2, 156-163.

Khanna M & Stinson R (1985). Specification of the acoustical input to the ear at higher frequencies. *J. Acoust. Soc. Am.* 77, 577-589.

Kiessling J, Brenner B, Thunberg C, Groth J & Drylund O (2005). Occlusion effect of earmolds with different venting systems. *J Am Acad Audiol*.16, 237–249.

Killion M & Revit L (1987). Insertion gain repeatability versus loudspeaker location: You want me to put my loudspeaker where? *Ear & hearing*. 8, 68-73

Killion M & Revit L (1993). CORFIG and GIFROC: real ear to coupler and back. In: G.A. Studebaker & I. Hochberg (eds.) *Acoustical Factors Affecting Hearing Aid Performance*. Boston: Allyn and Bacon, pp. 65–85.

Kochkin S & MarkeTrak (1994). Why 20 million in U.S. don't use hearing aids for their hearing loss. *The Hear Journal*. 46, 20-27.

Kuk K, Keenan D & Ludvigsen C (2005). Efficacy of an open-fitting hearing aid. *Hear Rev* 12, 26–32.

Lantz J, Jensen O, Haastrup A, et al. (2007). Real-ear measurement verification for open, non-occluding hearing instruments, *International Journal of Audiology*. 46, 11-16.

Lix L, Keselman J & Keselman H (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Rev. Educ. Res.* 66, 579-619.

Mackenzie D (2006). Open-canal fitting and the hearing occlusion effect. *The Hearing Journal*. 59, 50-56.

Mackenzie D, Mueller H, Ricketts T, et al. (2004). The hearing aid occlusion effect: Measurement devices compared. *The hearing Journal*.57, 30-39.

Mueller H & Ricketts T (2006). Open-canal fittings: Ten take-home tips. *Hear J* 59, 24-36.

Mueller H (1992). Terminology and procedures. In Mueller H, Hawkins D & Northern J (eds.) *Probe microphones measurements. Hearing aid selection and assessment*. San Diego: Singular Publishing Group, pp. 41_66.

Mueller H (2006). Open is in. *The Hearing Journal*. 59, 11-14.

Noble W, Sinclair S & Byrne D (1998). Improvement in aided sound localization with open earmolds: observation in people with high frequency hearing loss. *Am. Acad. Aud.* 9, 25-34.

Pumford J & Sinclair S (2001). Real-ear measurement: Basic terminology and procedures. *Audiology Online*, Article 285. <http://www.audiologyonline.com>.

Revit L (1992). Two techniques for dealing with the occlusion effect. *Hear Instr.* 43, 16-19.

Ricketts T & Mueller H (2009). Whose NAL-NL fitting method are you using? *Hearing Journal*. 62 (8), 10-17.

Ringdahl A & Leijon A (1984). The reliability of insertion gain measurements using probe microphones in the ear canal. *Scandinavian Audiol.* 13, 173-8

Royal National Institute for Deaf People (RNID) (2008). www.rnid.org.uk.

Sandlin R (2000). Textbook of hearing aid amplification – 2nd edition, San Diego: Singular Publishing Group.

Schweitzer C & Jesse K (2006). The value proposition of open-fit hearing aids. *Hearing Review*. 13, 10-19.

Shaw P (2010). Are real-ear measurements (REM) accurate when using the modified pressure with stored equalization (MPSE) method? *International Journal of Audiology*. 49, 463–466.

Staples C & Aiken S (2006). Where did my gain go? Thin Tube Open-fit BTE Verification.

http://www.bernafon.ca/eprise/main/ca_en/_downloads/Bernaфон_Gain_Tech_Topic.pdf.

Stone M & Moore B (1999). Tolerable hearing aid delays. I. Estimation of limits imposed by the auditory path alone using simulated hearing losses. *Ear and Hearing*. 20, 182–192.

Stone M & Moore B (2002). Tolerable hearing aid delays II. Estimation of limits imposed during speech production. *Ear Hear*. 23, 325-338.

Stone M & Moore B (2004). Estimated variability of real-ear insertion response (REIR) due to loudspeaker type and placement. *International Journal of Audiology*. 43, 271-275.

Strom K (2006). The HR 2006 Dispenser Survey. *The Hearing Review*. 13, 16-39.

Sullivan R (1988). Probe tube microphone placement near the tympanic membrane. *Hearing Instruments*. 39, 43-60.

Swan I & Gatehouse S (1995). The value of routine in-the-ear measurement of hearing aid gain. *Br J Audiol*. 29, 271–277

Taylor B (2006). Real-world satisfaction and benefit with open-canal fittings. *Hearing Journal*. 59, 74-82.

Tecca J (1990). Clinical application of real-ear probe tube measurement. In R. Sandlin (Ed.), *Handbook of hearing aid amplification, volume 2: Clinical consideration and fitting practice* (pp. 225-255). Boston, MA: College-Hill Press.

Valente M, Meister M, Smith P, et al. (1990). Intratester test-retest reliability of insertion gain measures *Ear & Hearing*. 11, 181-184.

Worrall L & Hickson L (2003). *Communication Disability in Aging*. Clifton Park, New York: Delmar Learning.

Yanz J (2006). Open-ear fittings: An entry in hearing care for mild losses. *The Hearing Review*. 13, 48-52.

7.0 APPENDICES

Appendix A – The randomisation of condition order for all participants

Participant number	Test-retest order								
	Ear	1 st order		2 nd order		3 ^{ed} order		4 th order	
1	Left	Open 0°	Open 0°	Open 45°	Open 45°	Conv 0°	Conv 0°	Conv 45°	Conv 45°
2	Right	Conv 0°	Conv 0°	Conv 45°	Conv 45°	Open 0°	Open 0°	Open 45°	Open 45°
3	Left	Open 45°	Open 45°	Open 0°	Open 0°	Conv 45°	Conv 45°	Conv 0°	Conv 0°
4	Left	Conv 45°	Conv 45°	Conv 0°	Conv 0°	Open 45°	Open 45°	Open 0°	Open 0°
5	Right	Open 0°	Open 0°	Open 45°	Open 45°	Conv 0°	Conv 0°	Conv 45°	Conv 45°
6	Left	Conv 0°	Conv 0°	Conv 45°	Conv 45°	Open 0°	Open 0°	Open 45°	Open 45°
7	Right	Open 45°	Open 45°	Open 0°	Open 0°	Conv 45°	Conv 45°	Conv 0°	Conv 0°
8	Right	Conv 45°	Conv 45°	Conv 0°	Conv 0°	Open 45°	Open 45°	Open 0°	Open 0°
9	Left	Open 0°	Open 0°	Open 45°	Open 45°	Conv 0°	Conv 0°	Conv 45°	Conv 45°
10	Right	Conv 0°	Conv 0°	Conv 45°	Conv 45°	Open 0°	Open 0°	Open 45°	Open 45°
11	Left	Open 45°	Open 45°	Open 0°	Open 0°	Conv 45°	Conv 45°	Conv 0°	Conv 0°
12	Left	Conv 45°	Conv 45°	Conv 0°	Conv 0°	Open 45°	Open 45°	Open 0°	Open 0°
13	Right	Open 0°	Open 0°	Open 45°	Open 45°	Conv 0°	Conv 0°	Conv 45°	Conv 45°
14	Left	Conv 0°	Conv 0°	Conv 45°	Conv 45°	Open 0°	Open 0°	Open 45°	Open 45°
15	Right	Open 45°	Open 45°	Open 0°	Open 0°	Conv 45°	Conv 45°	Conv 0°	Conv 0°
16	Right	Conv 45°	Conv 45°	Conv 0°	Conv 0°	Open 45°	Open 45°	Open 0°	Open 0°
17	Left	Open 0°	Open 0°	Open 45°	Open 45°	Conv 0°	Conv 0°	Conv 45°	Conv 45°
18	Right	Conv 0°	Conv 0°	Conv 45°	Conv 45°	Open 0°	Open 0°	Open 45°	Open 45°
19	Left	Open 45°	Open 45°	Open 0°	Open 0°	Conv 45°	Conv 45°	Conv 0°	Conv 0°
20	Right	Conv 45°	Conv 45°	Conv 0°	Conv 0°	Open 45°	Open 45°	Open 0°	Open 0°

*Open 0°: open-fit at 0° loudspeaker-to-head azimuth

*Open 45°: open-fit at 45° loudspeaker-to-head azimuth

*Conv 0°: Conventional earmould at 0° loudspeaker-to-head azimuth

*Conv 45°: conventional earmould at 45° loudspeaker-to-head azimuth

Appendix B – Screening form

Screening form

Name: Participant number:

Age: Gender:

Please answer and give details for the following questions:

1) Have you had any operations on your ears?

.....
.....

2) Have you recently had any ear pain, infection or discharge from your ears?

.....
.....

3) Have you ever had a perforation in either of your ear drums?

.....
.....

4) Do you suffer with tinnitus (noises) in either of your ears?

.....
.....

5) Do you feel you are particularly sensitive to loud sounds or find everyday sounds uncomfortably loud?

.....
.....

Participant's signature:

Appendix C – Consent form

Consent form to be completed by adult subjects taking part in an experiment

(Adults are 18 years of age or older.)

Exposure Number:

University of Southampton

Institute of Sound and Vibration Research

Before completing this form, please read the list of contra-indications which has been provided by the experimenter on the reverse of this form.

This consent form applies to a subject volunteering to undergo an experiment for research purposes. The form is to be completed before the experiment commences.

I,

of

(address or department)

consent to take part in TEST-RETEST RELIABILITY OF REAL EAR MEASUREMENTS FOR OPEN CANAL HEARING AID FITTINGS to be conducted by Wala' Alaqrabawi during the period June 30 to August 31 2010.

The purpose and nature of this experiment have been explained to me. I understand that the investigation is to be carried out solely for the purposes of research. I am willing to act as a volunteer for that purpose on the understanding that I shall be entitled to withdraw this consent at any time, without giving any reasons for withdrawal. My replies to the above questions are correct to the best of my belief, and I understand that they will be treated by the experimenter as confidential.

Date: Signed:

(Volunteer subject)

I confirm that I have explained to the subject the purpose and nature of the investigation which has been approved by the Human Experimentation Safety and Ethics Committee.

Date: Signed:

(Researcher in charge of experiment)

This form must be submitted to the Secretary of the Human Experimentation Safety and Ethics Committee on completion of the experiment.

Appendix D – Participant information sheet

Participant information sheet

Experiment title

Test-retest reliability of real ear measurements for open canal hearing aid fittings

Experiment aim

To investigate the short-term test-retest reliability for open-fit real ear measurements and to compare these test-retest reliability results with those for conventional earmould.

Instructions

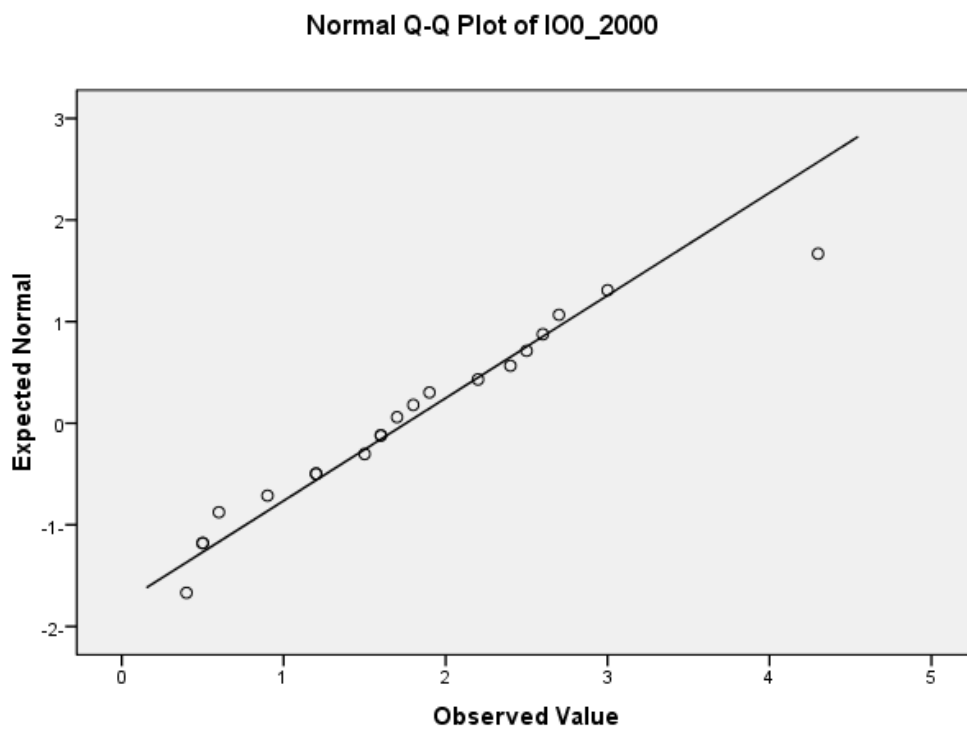
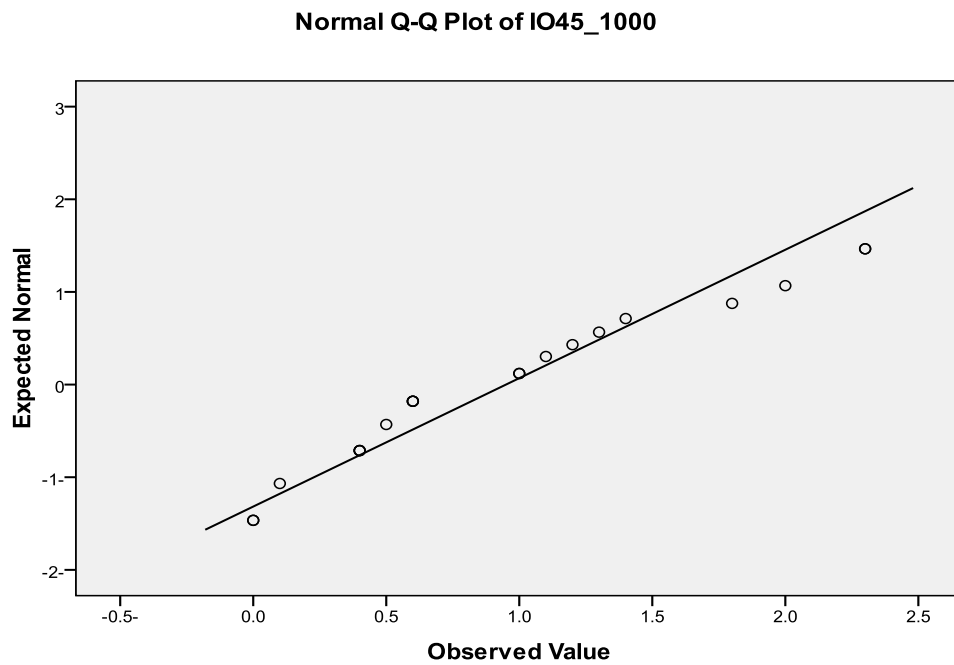
Before testing I will have a look into your ears using otoscopy to make sure that your ears are suitable for the measurements and do not have any contraindications and I will check your middle ear function using tympanometry.

Then you will be seated on the chair in front of the loud speaker and be fitted with a hearing aid. Two positions for each ear tip will be required. You should not move your head during the measurements. You will put this head-band to attach the laser light to it. This is to ensure accurate position throughout the measurements, as the laser light should be pointed at the 0° or 45° tap markers on the wall. I will then insert a thin probe tube into your ear for a certain depth you might feel a little tickling. This probe tube will help in measuring the sound levels near your ear drum. Measured sounds will be produced by the loudspeaker in front of you. Then I will place the hearing aid and its ear tip into your ear. And play a sound from the loud speaker again.

You do not need to respond in any way just stay still and quiet as possible. If you feel uncomfortable at any point let me know and I will stop the measurements.

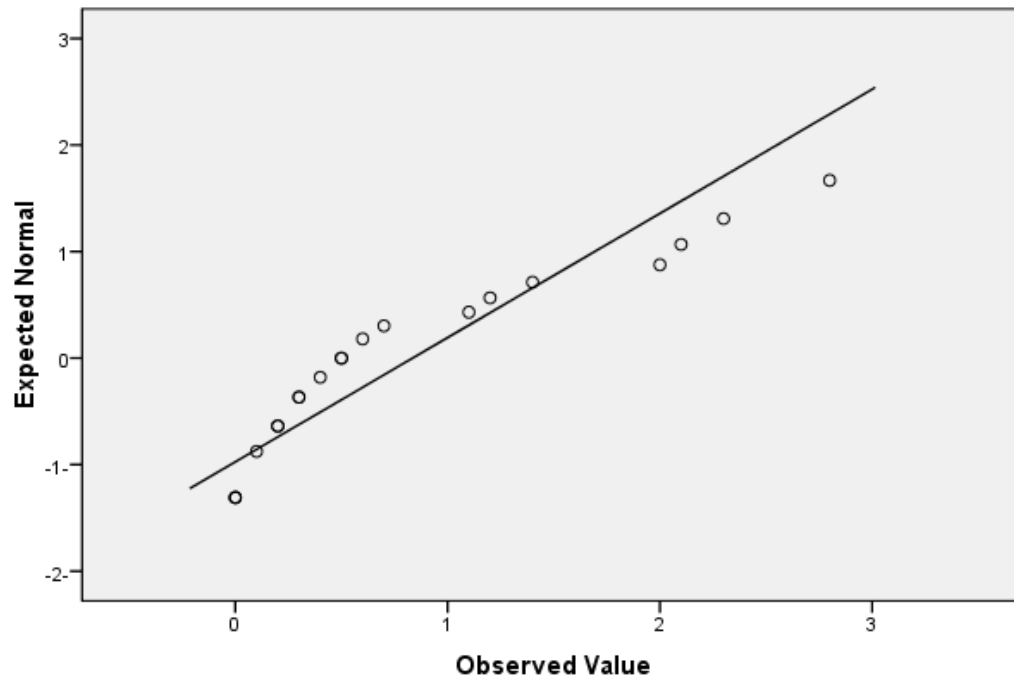
Appendix E - Examples of normally distributed and non-normally distributed normality probability (Q-Q) plots

Normally distributed normality probability (Q-Q) plots

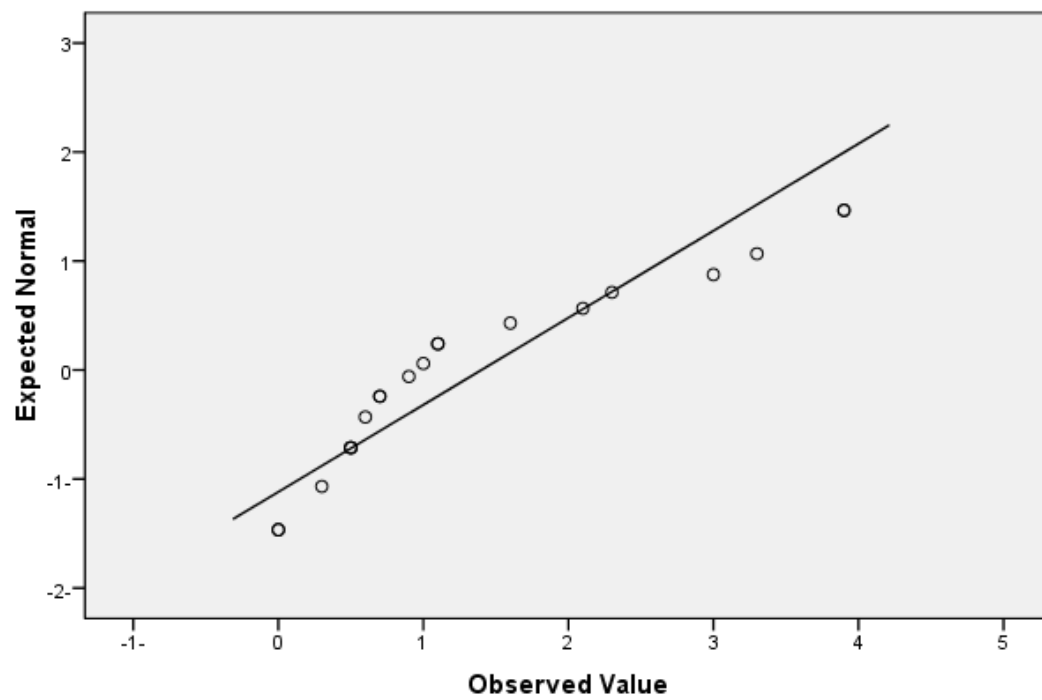


Non-normally distributed normality probability (Q-Q) plots

Normal Q-Q Plot of IC45_2000



Normal Q-Q Plot of IC0_2000



Appendix F – Pairwise Comparisons results when each frequency was compared with all other frequencies

Pairwise Comparisons

Measure: MEASURE_1

(I) frequency	(J) frequency	Mean Difference (I-J)	Std. Error	Sig.a	95% Confidence Interval for Differencea	
					Lower Bound	Upper Bound
250	500	-.314	.148	1.000	-.856	.229
	1000	-.193	.195	1.000	-.906	.521
	2000	-.043	.180	1.000	-.701	.616
	3000	-.488	.205	.809	-1.239	.264
	4000	-.760	.223	.087	-1.575	.055
	6000	-1.491	.183	.000	-2.163	-.820
	8000	-2.249	.267	.000	-3.227	-1.271
500	250	.314	.148	1.000	-.229	.856
	1000	.121	.197	1.000	-.598	.841
	2000	.271	.197	1.000	-.449	.991
	3000	-.174	.247	1.000	-1.079	.732
	4000	-.446	.263	1.000	-1.410	.517
	6000	-1.178	.217	.001	-1.974	-.381
	8000	-1.935	.240	.000	-2.813	-1.057
1000	250	.193	.195	1.000	-.521	.906
	500	-.121	.197	1.000	-.841	.598
	2000	.150	.135	1.000	-.343	.643
	3000	-.295	.188	1.000	-.984	.394
	4000	-.568	.220	.530	-1.373	.238
	6000	-1.299	.230	.001	-2.139	-.458
	8000	-2.056	.284	.000	-3.097	-1.016
2000	250	.043	.180	1.000	-.616	.701
	500	-.271	.197	1.000	-.991	.449
	1000	-.150	.135	1.000	-.643	.343
	3000	-.445	.160	.339	-1.029	.139
	4000	-.718	.160	.080	-1.303	-.132
	6000	-1.449	.152	.000	-2.007	-.891
	8000	-2.206	.261	.000	-3.162	-1.250
3000	250	.488	.205	.809	-.264	1.239
	500	.174	.247	1.000	-.732	1.079
	1000	.295	.188	1.000	-.394	.984
	2000	.445	.160	.339	-.139	1.029
	4000	-.273	.160	1.000	-.857	.312
	6000	-1.004	.176	.001	-1.649	-.359
	8000	-1.761	.329	.001	-2.967	-.555
4000	250	.760	.223	.087	-.055	1.575
	500	.446	.263	1.000	-.517	1.410

	1000	.568	.220	.530	-.238	1.373
	2000	.718	.160	.080	.132	1.303
	4000	.273	.160	1.000	-.312	.857
	6000	-.731	.201	.052	-1.467	.004
	8000	-1.489	.314	.005	-2.638	-.339
6000	250	1.491	.183	.000	.820	2.163
	500	1.178	.217	.001	.381	1.974
	1000	1.299	.230	.001	.458	2.139
	2000	1.449	.152	.000	.891	2.007
	3000	1.004	.176	.001	.359	1.649
	4000	.731	.201	.052	-.004	1.467
	8000	-.758	.310	.705	-1.894	.379
8000	250	2.249	.267	.000	1.271	3.227
	500	1.935	.240	.000	1.057	2.813
	1000	2.056	.284	.000	1.016	3.097
	2000	2.206	.261	.000	1.250	3.162
	3000	1.761	.329	.001	.555	2.967
	4000	1.489	.314	.005	.339	2.638
	6000	.758	.310	.705	-.379	1.894