Proceedings of 20th International Congress on Acoustics, ICA 2010
23-27 August 2010, Sydney, Australia

Effect of Experimental Design on the Results of Clarity-Index Just-Noticeable-Difference Listening Tests

Michelle C. Vigeant (1) and Robert D. Celmer (1)

(1) Acoustics Prog. and Lab., Dept. of Mech. Engr., Univ. of Hartford, 200 Bloomfield Ave, W. Hartford, CT 06117, USA

PACS: 43.55.Hy

ABSTRACT

The just noticeable differences (JND) of room acoustics parameters are useful quantities in design and research, as these values provide a guideline as to when a design change will result in a subjectively noticeable difference. The clarity index for music (C_{80}) JND has been studied previously by Cox *et al* (1993) and Bradley *et al* (1999), who found C_{80} JNDs of 0.7 dB and 0.9 dB, respectively. These studies had limitations in that Cox *et al* had a relatively small subject pool and Bradley *et al*'s study used speech signals rather than music, as the focus was C_{50}. Two new studies have been conducted to further investigate the C_{80} JND. In Study 1, 51 musically trained subjects were exposed to a total of 54 AB paired comparisons producing results suggesting a higher JND of 1.6 dB. A pilot study, Study 2, was conducted to compare two testing methods: Test Method 1, which was used in Study 1, required the 11 subjects to listen to all of signal A and then all of signal B before giving a response, while Test Method 2 allowed subjects to switch in real-time between the two signals, as it was hypothesized that Test Method 2 would yield a lower JND similar to previous work. However, this study yielded an even higher C_{80} JND of 3.8 dB averaged over both test methods. In particular, an interaction effect was found with test method and the order in which the subjects received each test method. The results that most closely matched the predicted trendline were obtained when the subjects completed the first half of the test using Test Method 1 and the second half using Test Method 2. If the first half of the test was considered a training period, then the results in this case from only Test Method 2 gave a C_{80} JND of 4.4 dB, which was much higher than found in previous work.

INTRODUCTION

Many parameters are used to quantify the acoustics of a concert hall, such as clarity index (C_{80}), early decay time (EDT), and early lateral energy fraction (LF). However, limited work has been conducted to determine the smallest difference, or just-noticeable-difference (JND) of these and other common parameters. Knowing the JND for a given parameter, allows for error-bars to be added to both measured and computer model predicted results, and provides an indication to both designers and researchers if a given change in the design or a variable of interest will make a large enough difference to be perceived subjectively [1].

The focus of this paper is the JND of clarity index for music, C_{80}. The clarity index is defined as the ratio, expressed in decibels, of the early sound energy, during the first 80 ms, to the late energy, after 80 ms, as derived from the impulse response, *p(t)* [2]:

$$C_{80} = \frac{\int_0^{80\,ms} p^2(t)dt}{\int_{80\,ms}^{\infty} p^2(t)dt} \quad (1)$$

The 80 ms upper integration limit was defined according to the human ear's ability to integrate direct sound with reflected sound arriving up to 80 ms later [3]. C_{80} provides an indication of the overall clarity of the sound expected in a hall. If the ratio of the early to late sound energy is high, then there will be a sensation of definition or clarity, i.e., how much it is possible to distinguish between notes played by individual instruments. If however, the ratio is small, then there be will a decrease in definition and an increase in fullness of tone, but if the ratio is too small, then the sound is perceived as muddy [3].

The earliest well-documented work in the field of JND's or difference limens of room acoustic parameters was conducted by Reichardt *et al* [4,5]. This work was conducted in the 1960's with the current technology of the time to generate the simulated impulse response (IR), however the technology was limited and the IR may not have been accurate [1]. More recently, only two major studies have been carried out to investigate the C_{80} JND.

The first was conducted by Cox *et al* [1], with the goal to identify the JNDs for both C_{80} and LF. A system of delays and effect units were used to create impulse responses and were played back over eight loudspeakers in an anechoic chamber. Two motifs were used in the study, a short section from Handel's *Water Music Suite*, and a motif from the fourth movement of Mendelssohn's *Symphony no. 3 in A minor*, Op.56. A total of 7 to 10 subjects participated in the series of tests conducted, who had normal hearing and some musical training. While clarity was being changed, reverberation time was held constant. A disparity was found between the difference limens for the two motifs; the Handel motif was found to have a C_{80}JND of 0.44 dB, while the Mendelssohn motif was found to have a C_{80} JND of 0.92 dB. The two values were averaged to give a difference limen for clarity of 0.67 dB.

The second study was carried out by Bradley *et al* [6], but with the focus on the JND of the clarity index of speech, C_{50}.

The simulated impulse responses were generated in a similar manner to Cox *et al*, with updated equipment, and also played back in an anechoic chamber over eight loudspeakers. Rather than musical motifs, speech was used as the test signals. Three C_{50} base cases (-3.0 dB, +1.0 dB, +5.0 dB) were used to represent low, mid, and high levels of clarity, and these signals were compared to other signals with small changes in C_{50}. The reverberation time was held constant for all signals in each base case. Ten subjects participated in the study with normal hearing. The results were averaged over the three base cases, and the JND for C_{50} was found to be 1.1 dB, and based on the linear relationship between C_{50} and C_{80}, the C_{80} JND was predicted to be 0.9 dB.

Two new studies have been conducted to further investigate the C_{80} JND. Based on the findings by Cox *et al* and Bradley *et al* that the C_{80} JND is approximately 1 dB, Study 1 was designed to contain pairs of signals with C_{80} differences around this value. A total of 51 musically trained subjects participated in this study. The test was carried out using Test Method 1, which required subjects to listen to all of the first signal, then all of the second signal, and then answer if the two signals were the same or different. The results of Study 1 indicated that the C_{80} JND might be higher than the previously reported results. In Study 2, a pilot study was carried out to compare results from two different testing methods, Test Method 1, as described, and Test Method 2 where subjects can toggle back and forth in real-time between the two signals, A and B, before giving their answer, with the hypothesis that Test Method 2 would yield a C_{80} JND closer to the previously found values. Test Method 2 was based on the work by Bradley *et al* [6].

STUDY 1 – C_{80} JND USING TEST METHOD 1

Experimental set-up

The simulated impulse responses (IR's) were generated using a Yamaha digital mixing engine, type DME64n, and the associated software, Yamaha DME Designer v3.0. A total of eight loudspeakers, Genelec 1030A, were used to playback the IR's in a 100 Hz full anechoic chamber, with six loudspeakers located directly in front of the listener and two in the rear, as shown in Figures 1 and 2.

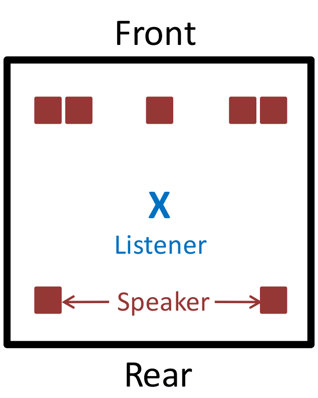

Figure 1. Top view of the speaker arrangement in the anechoic chamber. Six speakers were at the front, with two centre speakers aligned in the vertical plane, as shown in Figure 2.

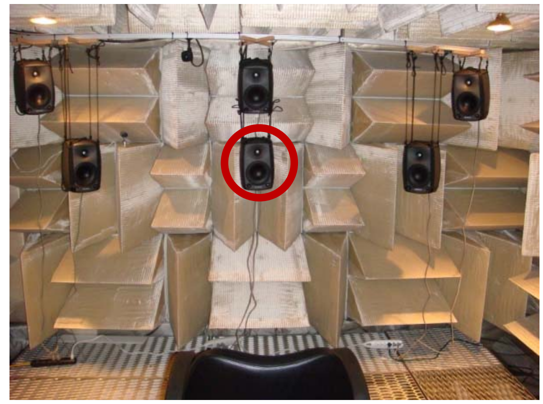

Figure 2. View of the front six loudspeakers from behind the listener's position (top of chair is shown). The front centre speaker, which plays the direct sound, is circled.

The DME Designer software allows for the generation of specific early reflections to be sent out to the individual loudspeakers, and control of the reverberation time settings. The signal input, in this case a musical motif, was processed in three simultaneous paths (Figure 3). The first was the direct sound path, which sends the sound out to the centre speaker, circled in Figure 2. The early reflections were generated in the second path, with a total of 31 early reflections in the first 80 ms, with three sent to the centre speaker in addition to the direct sound, and four sent to each of the remaining seven loudspeakers. The final path was the reverberation processing, which was sent out to each of the eight loudspeakers after the first 80 ms, each with slight time delays to avoid any coherence problems. The time delays and levels coming from each loudspeaker were first normalized so that all sound would reach the listener at the same time and level before any signal processing was carried out. Equalizers were used to adjust the levels of the early reflections and/or late energy as needed to achieve the desired C_{80}, while keeping the reverberation time held constant. The properties of the generated signals were measured using the sine sweep impulse response method with WinMLS 2004 software and a Brüel and Kjær type 4190 microphone.

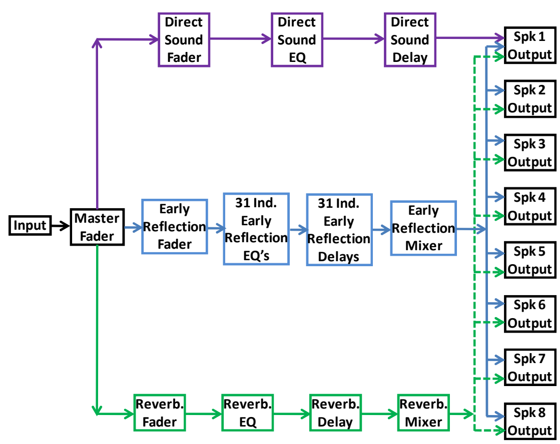

Figure 3. Overview of signal processing applied using Yamaha DME Designer Software and DME64n digital mixing engine.

Test signals in Study 1

Similar to Bradley *et al*'s [6] experimental design, two C_{80} base cases were generated and subsequently measured, along with signals with various small C_{80} differences. Base Case 1 had a C_{80} value of -3 dB at 1 kHz, with a reverberation time (T30) of 2.1 s, to be in the range of the ideal values for a large concert hall with a symphonic repertoire [7]. To model a smaller space meant for chamber music, Base Case 2 had a C_{80} value of +1 dB at 1 kHz, with a T30 of 1.6 s [7]. Each of the base case signals were compared to signals with the same T30, respectively, with positive differences in C_{80} of 0.5, 0.8, 1.0, 1.2, 1.5, 2.0, and 3.0 dB, as shown in Table 1. In addition, for each base case, the lowest and highest C_{80} value was compared to itself to give a total of four controls with a 0.0 dB difference in C_{80}.

Table 1. All AB pairs used in Study 1. Note that two 0.0 dB comparisons were made for each base case, with a "low" 0 dB difference, comparing the base case signal with itself, and and a "high" 0 dB difference, comparing the upper end of the range of the C_{80} differences for each base case.

	C_{80} of Signal A @ 1 kHz (dB)	C_{80} of Signal B @ 1 kHz (dB)	Δ C80 (dB)	Com-ment
Base Case 1: C80 = -3 dB (T30 = 2.1s)	-3.0	-3.0	*0.0*	*"Low" 0 dB Diff.*
	-3.0	-2.5	*0.5*	-
	-3.0	-2.2	*0.8*	-
	-3.0	-2.0	*1.0*	-
	-3.0	-1.8	*1.2*	-
	-3.0	-1.5	*1.5*	-
	-3.0	-1.0	*2.0*	-
	-3.0	0.0	*3.0*	-
	0.0	0.0	*0.0*	*"High" 0 dB Diff.*
Base Case 2: C80 = +1 dB (T30 = 1.6s)	+1.0	+1.0	*0.0*	*"Low" 0 dB Diff.*
	+1.0	+1.5	*0.5*	-
	+1.0	+1.8	*0.8*	-
	+1.0	+2.0	*1.0*	-
	+1.0	+2.2	*1.2*	-
	+1.0	+2.5	*1.5*	-
	+1.0	+3.0	*2.0*	-
	+1.0	+4.0	*3.0*	-
	+4.0	+4.0	*0.0*	*"High" 0 dB Diff.*

Motifs in Study 1

The motifs were selected from the limited available high-quality anechoic recordings. Three different motifs were chosen for several reasons. First, it was desirable to have motifs that were both performed by a large ensemble as well as a solo instrument. Second, the motifs should have somewhat quick-moving notes, and be relatively short. A 10.9 s passage from the third movement of Bizet's *L'Arlésienne Suite No. 2* was chosen as a large ensemble motif, Motif 1. For a solo piece, *Theme* by Weber, a solo cello piece, was chosen as Motif 2 with a length of 10.3 s. Motif 3, a 10.3 s passage from the beginning of Handel's *Water Music Suite*, was chosen as the same motif that was also used in the Cox *et al* [1] study. The orchestral anechoic recordings, Motifs 1 and 3 were taken from the DENON *Anechoic Orchestral Music Recordings* CD [8], while Motif 2 was taken from the Bang & Olufsen *Music for Archimedes* CD [9].

Test subjects in Study 1

A total of 51 test subjects participated in Study 1, with 30 males and 21 females. The subjects were required to have hearing thresholds of 25 dB HL or lower between 250 – 8,000 Hz. In addition, all subjects needed to have a minimum of five years of musical training, with the subjects having an average of 10.0 years of formal training and 12.1 years of experience.

Testing procedure in Study 1

The subjects were presented signal pairs, A and B, and were instructed to indicate if the signals sounded the same or different. In particular, the participants were instructed to *focus on how clear each individual note sounds, and also how clear the note sounds relative to the subsequent note*. The subjects were allowed to listen to the pair as many times as they wanted before giving a response.

For Study 1, Test Method 1 was employed, where subjects had to listen to each signal in its entirety before hearing the next signal. In other words, subjects had to listen to all of signal A, and then all of signal B, before making a decision if the signals were the same or different. In addition, if they wanted to hear the pair repeated, the subjects were again presented all of signal A and then all of signal B.

The subjects were presented a total of seven sets of AB pairs. The first set, which used a different motif, contained only four AB pairs, and was meant as a practice set, unbeknownst to the test subjects. The remaining six sets each contained nine AB pairs, with both the motif and base case held constant in each set (3 motifs X 2 base cases). The order of the individual sets and the order of the presented pairs were randomized for each test subject. The testing lasted an average of approximately 1 hour.

Study 1 results and discussion

The goal of this study was to determine an overall C_{80} JND, while two additional independent variables were introduced to examine their effects on the C_{80} JND, base case and motif. For brevity, only C_{80} JND averaged over all conditions will be discussed.

When the data were compiled over all 51 test subjects and averaged over the variables of base case and motif, small differences were found in the percentages of subjects who reported hearing a difference when the difference in C_{80} (ΔC_{80}) increased, as shown in Figure 4, contrary to the hypothesis that a low percentage would report hearing a difference when ΔC_{80} = 0.0 dB and a relatively high percentage would report hearing a difference at ΔC_{80} = 3.0 dB. The JND is defined as the point where subjects report hearing a difference 50% of the time, and in general, subjects reported this difference for most of the signals, including the signals with no difference at all in C_{80}. The regression analysis, however, did reveal a significant trend with a R^2 value of 0.599 and $p < 0.007$.

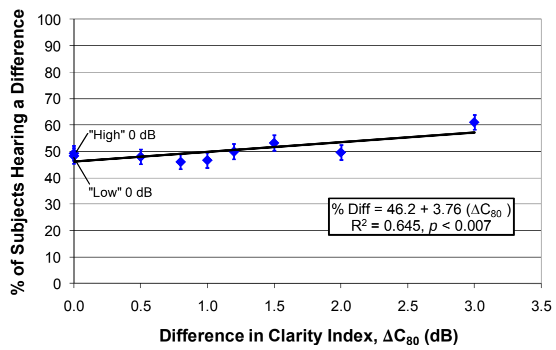

Figure 4. Data averaged over all conditions in Study 1. In general, subjects reported hearing a difference about 50% of the time for all ΔC_{80} conditions. The error bars in this plot and in ALL PLOTS indicate the standard error of the mean.

From these initial data analyses, it was concluded that the subjects found the test extremely difficult and that they were likely guessing for many of the AB pairs. The former conclusion was supported by the feedback obtained from the subjects on the post-test questionnaires. A further analysis was conducted by filtering the data. The data were filtered down to include only those subjects who were correct 65% of the time for the pairs presented with a 0.0 dB difference and for the pairs with the largest difference of 3.0 dB. This data reduction resulted in keeping 17 of the 51 test participants' data. A much steeper slope was obtained from this data set, as shown in Figure 5, with less than 30% of the subjects reporting hearing a difference when none existed, and close to 70% hearing a difference at the largest difference. This regression was also significant with a R^2 value of 0.885 and $p <$ 0.0001. Although the percentages at these two extremes could be smaller and larger, respectively, the C_{80} JND was extracted from these results and found to be 1.6 dB, which is more than 50% larger than the previous studies have shown.

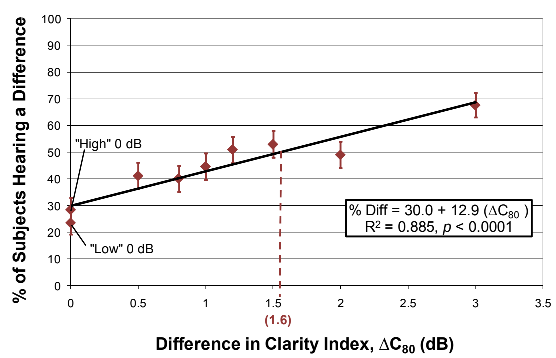

Figure 5. Filtered data averaged over all conditions in Study 1. Note the lower percentage who reported hearing a difference when ΔC_{80} = 0.0 dB and higher percentage who reported hearing a difference at ΔC_{80} = 3.0 dB, as compared to the unfiltered data shown in Figure 4.

Based on the results of Study 1, a pilot study was conducted to investigate if the testing method used had a significant effect on the results. In particular, it was hypothesized that Test Method 2, where subjects were able to change between signals A and B in real-time, would find the test easier, and therefore might yield a lower C_{80} JND.

STUDY 2 – A COMPARISON OF TEST METHODS 1 & 2 FOR C_{80} JND

Test signals in Study 2

The experimental set-up for this study was identical to Study 1. The same signal generation and playback equipment were used in this study.

The same two C_{80} base cases, where Base Case 1 had a C_{80} value of -3 dB at 1kHz with a T30 of 2.1 s, and Base Case 2 had a C_{80} value of +1 dB at 1kHz with a T30 of 1.6 s, were used as in Study 1. However, based on the relative difficulty experienced by the test participants in Study 1 and also the results that indicated that the C_{80} JND might be above 1.0 dB, higher positive differences in C_{80} between signals were used: 3.0, 5.0 and 7.0 dB. Only four differences, including 0.0 dB, were used in this study to both reduce the overall testing time and also as this study was intended as a pilot study. All comparisons for each base case for this study are summarized in Table 2.

Table 2. All AB pairs used in Study 2.

	C_{80} Signal A @ 1 kHz (dB)	C_{80} Signal B @ 1 kHz (dB)	Δ C80 (dB)
Base Case 1: C80 = -3 dB (T30 = 2.1s)	-3.0	-3.0	*0.0*
	-3.0	0.0	*3.0*
	-3.0	+2.0	*5.0*
	-3.0	+4.0	*7.0*
Base Case 2: C80 = +1 dB (T30 = 1.6s)	+1.0	+1.0	*0.0*
	+1.0	+4.0	*3.0*
	+1.0	+6.0	*5.0*
	+1.0	+8.0	*7.0*

In an additional effort to reduce the overall testing time, only two motifs were used in this study. The first motif, Motif 1, was the same as in Study 1, Bizet's *L'Arlésienne Suite No. 2*, while Motif 2, Debussy's *Prelude to the Afternoon of a Faun*, was a different motif. Motif 2 was chosen to provide more contrast to Motif 1, as it has a flowing string passage as compared to the Bizet, which has a nicely articulated woodwind passage.

Test subjects in Study 2

As Study 2 was meant as a pilot study to investigate two different testing methods, only 11 test subjects were used, with 7 males and 4 females. However, unlike in Study 1, two different subject types were used in this study. The first subject type consisted of students and faculty from the University of Hartford's The Hartt School, Music Production Technology (MPT) program, who train (and teach) in the field of live music recording and mixing. The MPT subjects averaged 11.5 hours of *critical listening* per week, where for the purposes of this study, *critical listening* is defined as recording and/or mixing music. In addition, these subjects were also musically trained and had on average 12.5 years of formal instructions on their instrument. The second subject type consisted of music performance majors also from the University of Hartford's The Hartt School. These subjects averaged 10 years of formal instruction on their instrument. All subjects reported regular attendance at live classical concerts at least once per month.

A second difference for this set of test participants was that the hearing threshold criteria were made stricter by lowering the requirement to 15 dB HL (from 25 dB HL) or lower between 250 – 8,000 Hz based on work by Bech [10].

Testing procedure in Study 2

As in Study 1, the subjects were presented a pair of signals and were instructed to indicate if the signals sounded the same or different. They were given the same specific instructions about what differences to listen for and were again allowed to listen to the pair as many times as they wanted before giving a response. However, in this study two different testing methods were used.

In addition, to Test Method 1, where subjects had to listen to the entire signal for both signals without interruption, subjects were also required to use a second method. Test Method 2 allowed subjects to change between signals A and B in real-time to allow them to compare the signals at specific points during the melodies [6]. Subjects were able to control the signals and enter their responses using the custom testing box shown in Figure 6.

Figure 6. Custom testing box used in Study 2. Note toggle switch at the bottom of the box for the subjects to use with Test Method 2 to switch between the signals in real-time.

For Study 2, the subjects were presented a total of four sets of AB pairs: Set A–Training Set for Test Method 1, Set B–Test Method 1 Set, Set C–Training Set for Test Method 2, and Set D–Test Method 2 Set, as shown in Table 3. The training sets, Sets A and C, consisted of six AB pairs and contained different motifs than those used in the actual sets to give the subjects some variety. The actual test sets, Sets B and D, contained a total of 24 questions, but the first eight questions were additional practice questions, which the subjects were unaware were practice only. The training sets contained two pairs with 0.0 dB differences, and two pairs each with the largest differences of +5.0 and +7.0 dB, while the additional practice questions in each of the actual test sets, contained only the largest two differences, but with the actual motifs being used in the test.

Table 3. Testing orders for Study 2. Half of the participants received Testing Order 1, and the remainder received Testing Order 2.

	First Half of Test		Second Half of Test	
Testing Order 1	Test Method 1, *no switch*		Test Method 2, *with switch*	
	Set A	Set B	Set C	Set D
	Training (6)	Practice (8) & Test (16)	Training (6)	Practice (8) & Test (16)
Testing Order 2	Test Method 2, *with switch*		Test Method 1, *no switch*	
	Set C	Set D	Set A	Set B
	Training (6)	Practice (8) & Test (16)	Training (6)	Practice (8) & Test (16)

All subjects were required to complete the test using both test methods. Two testing orders were used: (1) Testing Order 1 had Test Method 1, *no switch*, Sets A and B, and then Test Method 2, *with switch*, Sets C and D; while (2) Testing Order 2 had the reverse, *with switch* first and then *no switch* (Table 3). All subjects received the training questions, practice questions, and actual test questions in random order. The odd-numbered subjects were assigned Testing Order 1, and the even were given Testing Order 2.

Study 2 results and discussion

The main independent variable of interest in Study 2 was the effect of test method on the C_{80} JND, so the results will focus on this variable, and for brevity's sake, the effects of motif and base case, and also the interaction effects with subject type will not be discussed.

The results were averaged over all conditions for the 11 test subjects and are shown in Figure 7. The range of percentages of people hearing a difference was much larger from this study, with just over 20% at the no difference extreme (ΔC_{80} = 0.0 dB) and close to 75% at the highest difference of ΔC_{80} = 7.0 dB. The correlation coefficient, R^2, was found to be 0.977 and $p < 0.06$. The regression curve was not found to be significant at the 0.05 level, which was likely due to the small number of test subjects. These data produced a JND of 3.8 dB. Possible reasons for the steeper slope in Study 2, as compared to Study 2, could be due to the additional training and practice questions, and using a larger range of C_{80} differences.

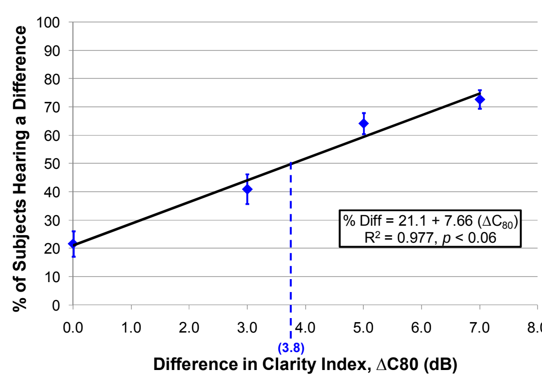

Figure 7. Data averaged over all conditions in Study 2. Note that the slope is much steeper than for the results in Study 1. The calculated C_{80} JND is 3.8 dB.

It was hypothesized that the Music Production Technology (MPT) students and faculty would give results with a steeper slope, and thus a lower JND than the music performance majors, but this was not found to be the case, as shown in Figure 8. For the MPT subjects, the relationship was found to have a R^2 value of 0.964 and $p < 0.09$, while for the music performance majors the results gave a R^2 value of 0.991 and $p < 0.003$. The JND for both subject types is similar, with 3.7 dB for the MPT subjects, and 3.9 dB for the performance subjects.

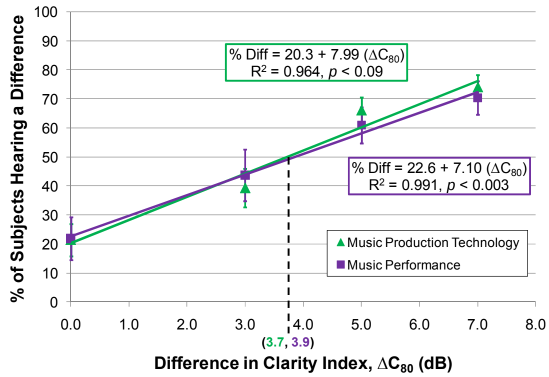

Figure 8. Data averaged over all conditions in Study 2, separated by subject type. Note that the slopes for the two subject types are very similar, with a calculated C_{80} JND of 3.7 dB for the MPT subjects and C_{80} JND of 3.9 dB for the performance subjects.

An initial analysis comparing Test Method 1 to 2, revealed a JND of 2.8 and 4.6, respectively, as shown in Figure 9. These results were contrary to the hypothesis that Test Method 2, *with switch*, would yield a lower JND. The Test Method 1 results, however, have a higher percentage of subjects reporting hearing a difference when the signals were the same, nearly 30%, as compared to Test Method 1 where less than 20% reported hearing a difference. The overall trendline for the results of Test Method 2 is much steeper as was hypothesized.

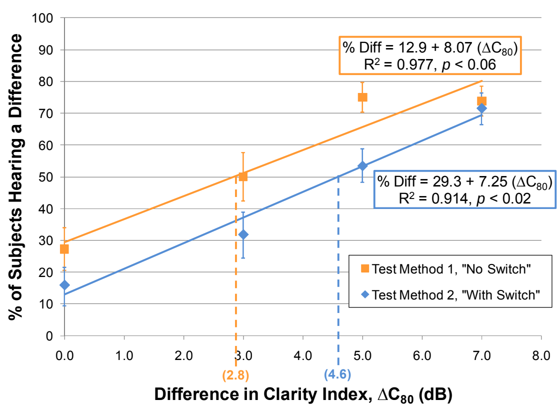

Figure 9. Data averaged over all conditions in Study 2, separated by test method. Note the steeper slope for the results using Method 2, over Method 1. A higher C_{80} JND of 4.6 dB was found for Method 2, as compared to C_{80} JND of 2.8 dB for Method 1.

A further analysis was conducting examining the interaction effect of testing order with test method, as shown in Figure 10. Half of the subjects were given Testing Order 1 (red data), where they began the test using Test Method 1, *no switch* (dashed lines), and then completed the second half using Test Method 2, *with switch* (solid lines), while the other half of the subjects had Testing Order 2 (blue data), which was the reverse. The subjects who were given Testing Order 1, yielded the steepest slope overall in the results for Test Method 2, *with switch*, after they had spent the first half of the test using Test Method 1, *no switch*. These results imply that by giving the subjects an extended training period using Test Method 1, *no switch*, and then giving them the actual test set using Test Method 2, *with switch*, will perhaps yield the most accurate results. Subjects given Testing Order 2 (blue data), performed better in the first half of the test with Test Method 2, *with switch*, and their performance declined in the second half of the test with Test Method 1, *no switch*. However, these subjects reported hearing a difference more often when there wasn't a difference between the two signals that the subjects who were assigned Testing Order 1. In particular, the results for the subjects who completed Testing Order 1, with Test Method 2 second, achieved the lowest percentage at the 0.0 dB difference of 8%, and above 70% reporting hearing a difference at the maximum difference of 7.0 dB (thickest solid line in plot). These data, which have the steepest slope and are most like the hypothesized results, gave a C_{80} JND of 4.4 dB

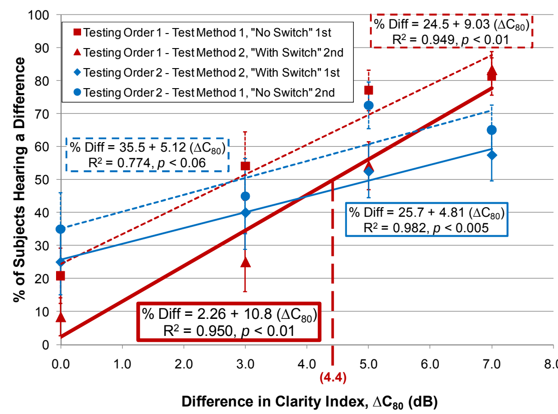

Figure 10. Data averaged over all conditions in Study 2, separated by testing order (Testing Order 1 in red and Testing Order 2 in blue) and test method (Test Method 1 in solid lines and Test Method 2 in dashed lines). Subjects who were given Testing Order 1, with *no switch* first and then *with switch second*, yielded the lowest percentage of subjects reporting a difference when no difference existed (thick red line).

SUMMARY

Two studies were conducted to investigate the clarity index just noticeable difference, C_{80} JND. In Study 1, the test was designed to contain C_{80} differences close to 1.0 dB based on previously published data. Test Method 1 was used in the first study, which forced subjects to listen to all of Signal A and then all of Signal B, before responding if the two signals were the same or different. The test, however, proved exceedingly difficult for the test subjects, with subjects reporting hearing a difference 50% of the time for most of the signal pairs presented, even though the differences in C_{80} ranged between 0.0 – 3.0 dB.

A pilot study, Study 2, was carried out to evaluate the effect of test method on the C_{80} JND, as it was hypothesized that Test Method 1 might have contributed to the test difficulty. A second testing procedure, Test Method 2, was added to this study, which allowed subjects to change between Signals A and B in real-time (*with switch*). It was hypothesized that Test Method 2 would yield a much lower percentage of subjects reporting hearing a difference when none existed and a much higher percentage would report hearing a difference at the maximum difference included in the study than Test Method 1 (*no switch*). Given that the C_{80} differences seemed too small in Study 1, much larger differences of 3.0, 5.0 and 7.0 dB were used in Study 2. The results from this second study gave a much higher C_{80} JND of 3.8 ±1.1 dB.

In particular, an interesting interaction effect was found between testing order and test method. The results that most conformed to the hypothesis were yielded from the subjects who began the test using Test Method 1, *no switch*, and then in the second half used Test Method 2, *with switch*. Therefore, it can be concluded that subjects require an extended training period using Test Method 1, before given the actual test using Test Method 2. The results from this pilot study, using this proposed procedure, gave a C_{80} JND of 4.4 dB. This result differs dramatically from the previously published results in the literature, which may be due to the fact that a more technologically advanced impulse response simulation system was used than in Cox *et al* [1], and that music signals were used to directly obtain the C_{80} JND, rather than from interpolation from C_{50} JND results obtained using speech signals [6]. Further work will be conducted using a larger subject population to investigate the proposed procedure.

ACKNOWLEDGMENTS

This work was supported by the Paul S. Veneklasen Research Foundation (PSVRF). The authors thank John Lo Verde, PSVRF President, for his input on these studies. In addition, the authors wish to acknowledge John S. Bradley, principal research officer at the National Research Council of Canada, for his extensive help with the experimental set-up of the apparatus to produce and playback the generated impulse responses, and for his feedback and suggestions on the experimental design. Finally, the authors thank the students who worked on these studies for their senior projects: Meghan Ahearn and Matthew Schaeffler (Study 1) and Clothilde Giacomoni and Christopher Jasinski (Study 2).

REFERENCES

1. T.J. Cox, W.J. Davies, and Y.W. Lam, “The sensitivity of listeners to early sound field changes in auditoria” *Acustica* **79**, 27-41 (1993)
2. W. Reichardt, O. Abdel Alim, and W. Schmidt, “Abhängigkeit der grenzen zwischen brauchbarer und unbrauchbarer durchsichtigkeit von der art des musikmotives, der nachhallzeit und der nachhalleinsatzzeit”, *Appl. Acoust.* **7**, 243-264 (1974)
3. M. Mehta, J. Johnson, J. Rocafort, *Architectural Acoustics: Principles and Design* (Prentice-Hall, Inc., New Jersey, 1999) pp. 265
4. W. Reichardt and W. Schmidt, “The audible steps of spatial impression in music performances” (in German) *Acustica* **17**, 175-179 (1966)
5. W. Reichardt and W. Schmidt, “The detectability of changes in sound field parameters for music” (in German) *Acustica* **18**, 274-282 (1967)
6. J.S. Bradley, R. Reich, and S.G. Norcross, “A just noticeable difference in C_{50} for speech” *Appl. Acoust.* **58** 99-108 (1999)
7. L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*, 2nd Edition (Springer-Verlag, Inc., New York, 2004) pp. 536
8. Bang & Olufsen, *Music for Archimedes*, Audio CD (1992)
9. Denon, *Anechoic Orchestral Music Recording*, Audio CD (1988)
10. S. Bech, “Selection and training of subjects for listening tests on sound-reproducing equipment” *J. Aud. Eng. Soc.* **40** 590-609 (1992)