H

Dynastat, Inc.
Research and Testing Services for the Voice Communications Community
2704 Rio Grande, Suite 4
Austin, TX 78705 USA
Phone: +1-512-476-5797 FAX: +1-512-472-2883
Contacts: Alan D. Sharpley (sharpley@dynastat.com)
Ira L. Panzer (ilpanzer@dynastat.com)

DYNASTAT – AUDIO TESTING METHODS AND PROCEDURES

Prepared for:

iBiquity Digital Corp.
20 Independence Blvd.
Warren, NJ

3 August 2001

III. Creating Comparable DAB-Off/DAB-On Samples for Each Condition

1) When selecting comparable DAB-Off and DAB-On samples for each condition (location, radio station, and radio), several criteria were used:
   - In keeping with laboratory samples, the DAB-Off and DAB-On samples should each be approximately 15 seconds long (12-17 seconds were considered acceptable).
   - The DAB-Off and DAB-On samples should be approximately the same length, within 0.5 seconds of each other. Parsing samples to match exactly was not possible due to the slight variations in musical phrasing.
   - The programmatic content of the DAB-Off and DAB-On samples should be as similar as possible. For example, if the DAB-Off sample consists of rock music, the DAB-On sample should also consist of rock music.
   - The musical characteristics of the DAB-Off and DAB-On samples should be as similar as possible (i.e. if the DAB-Off sample consists of a soloist with a light instrumental background, the DAB-On sample should consist of the same soloist with the same type of instrumental background). The nature of the DAB-Off and DAB-On samples should also be similar. For example, if the DAB-Off sample includes the chorus of a song, the DAB-On sample should also include the chorus of that song. Similarly, if the DAB-Off sample consists of male speech, the DAB-On sample should include male speech (spoken by the same male) as well.
2) To select an appropriate sample (based on the criteria listed above) from each 30-second DAB-Off or On wave file, two listeners played the wave files from beginning to end searching for a continuous segment that matched the above criteria.
3) Each identified segment was copied into a new wave file. Wave files were named using the following naming convention:

   NRSC_FT_Radio Station Call Letters_Receiver_Location Number_Host or First Adjacent_DAB-Off or On.wav

   For example, a sample taken from a recording taken of WETA at Location 5 using a Delphi receiver with DAB-On for the first adjacent station would have a file name as follows:

   NRSC_FT_WETA_Delp_Loc5_1st_On.wav
4) Each individual wave file was then edited to envelope the beginning and end of the file. Sound sample editing was performed identically for each radio.
5) Finally, all of the wave files were leveled. Please see the document entitled "Procedure for Editing and Leveling Samples" for additional details regarding the leveling process.

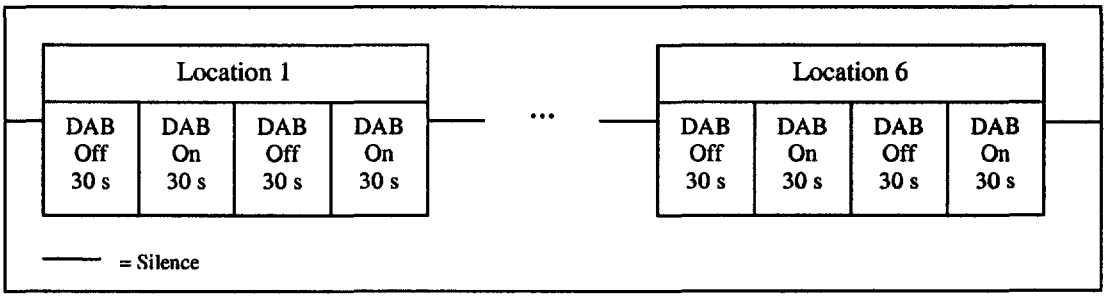
Table 1: Field Recordings

Station and Location Number	Host or First Adjacent
WETA	
Location 1	Host
Location 2	Host
Location 3	Host
Location 4a	First Adjacent
Location 4b	First Adjacent
Location 5	First Adjacent
Location 6	First Adjacent
WPOC	
Location 1	Host
Location 2	Host
Location 3	Host
Location 4	First Adjacent
Location 5	First Adjacent
Location 6	First Adjacent
WNEW	
Location 1	First Adjacent
Location 2	First Adjacent
Location 3	First Adjacent

II. Dividing the Field Recordings into DAB-Off/On Wave Files

1) Recordings for each radio station were included in the same wave file, with the recordings at different locations being separated by silence. **Figure 1** shows the structure of each wave file for each radio station.

Figure 1: Wave file for each radio station



2) Based on time-code records, each 30-second "DAB-Off" or "DAB-On" section of all of these wave files were edited and copied to new wave files. This process resulted in at least one "DAB-Off" and one "DAB-On" wave file for each location, at each radio station, and on each radio.

INTRODUCTION

Dynastat has performed the data collection for the subjective testing of the In-Band On-Channel (IBOC) hybrid system developed by iBiquity Digital Corporation. The testing effort to date has included both field and laboratory testing in the areas of compatibility and performance in FM mode. Dynastat has completed data collection for eleven experiments in this phase of the test program. These experiments have involved a variety of test methodologies and procedures which are described in the following sections.

BACKGROUND

Dynastat, Inc., Austin, Texas was formed in 1974 by Dr. William Voiers and his colleagues Alan Sharpley, and Ira Panzer. Dynastat has a long history of involvement with the evaluation of voice communication systems using subjective testing methods. Over the past two decades this has consisted of the development and implementation of methods for measuring speech intelligibility, speech quality, and speaker recognizability. Dynastat personnel have developed the Diagnostic Rhyme Test (DRT), the Diagnostic Acceptability Measure (DAM), and the Diagnostic Speaker Recognizability Test (DSRT). The DRT is one of the ANSI standards for measuring speech intelligibility (ANSI S3.2-1989) and the DAM has become a *de facto* standard for measuring speech quality at the Department of Defense. In addition, Dynastat has implemented most other methods currently in use for assessing performance of speech coding systems. For assessing speech intelligibility, these include the Modified Rhyme Test (MRT) and the Phonetically Balanced Word Test (PB), the other two ANSI standards along with the DRT. For measuring speech quality, Dynastat has also implemented all the International Telecommunications Union (ITU) standards (P.800), including the Absolute Category Rating (ACR) method which yields the Mean Opinion Score (MOS), the Degradation Category Rating (DCR) method from which the Degradation Mean Opinion Score (DMOS) is derived and the Comparison Category Rating (CCR) method which yields the CMOS.

It is Dynastat's policy to work with all our customers in determining their testing needs and to be equipped to provide virtually all subjective testing methods. Over the past two decades Dynastat has contracted with various international and national standard bodies to conduct subjective listener tests as an independent testing laboratory. These groups have included the International Telecommunication Union-Telecommunications Sector (ITU-T), the European Telecommunications Standards Institute (ETSI), and the Telecommunication Industry Association (TIA) as well as both Third-Generation Partnership Projects, 3-GPP and 3-GPP2. Dynastat is active in all of these standards groups providing expertise on test design, test implementation, test evaluation, and data analyses. Dynastat was approached by iBiquity to perform the evaluations for the NRSC effort. Except for the creation of the new test lab as explained in the body of this report, all systems were in place to undertake this effort.

SUBJECTIVE TESTING FACILITIES

Dynastat designed and built an Audio Testing Laboratory for this project. The laboratory is contained in a quiet interior room that will accommodate the testing of up to four listeners at a time. Each listener is seated in an individual sound-treated Tremetrics audio testing booth with a measured ambient noise level < 35dBA. Audio samples are presented to listeners binaurally over Sennheiser HD-600 open-backed headphones. The subjective rating scales are displayed on a flat-screen Viewsonic VG150 LCD monitor located on a table opposite the window of each booth and the listeners enter their responses using a PC mouse. The only equipment inside the booth is a chair, a laptop desk, a pair of headphones, and a PC mouse.

Each listening station includes the sound-treated booth, an HP Vectra VL400 PC, a high-quality Lucid DA9624 digital to analog converter, and the Sennheiser headphones. Figure 1 shows two views of listening stations contained in Dynastat's Audio Testing Laboratory. The PC's, A/D converters, and headphones were provided to Dynastat by iBiquity. Sound samples are stored on the hard-disk of each PC and are presented to the listeners under program control using a software package developed by iBiquity. The software also displays the appropriate rating scale(s) on the monitor and collects and stores the listener's responses. Each listening station is independent and self-contained and requires no experimenter control or interaction once the listener has started an experiment. Dynastat's Audio Laboratory includes four listening stations with the individual PC's networked to a server PC used for loading audio files and compiling listener responses.

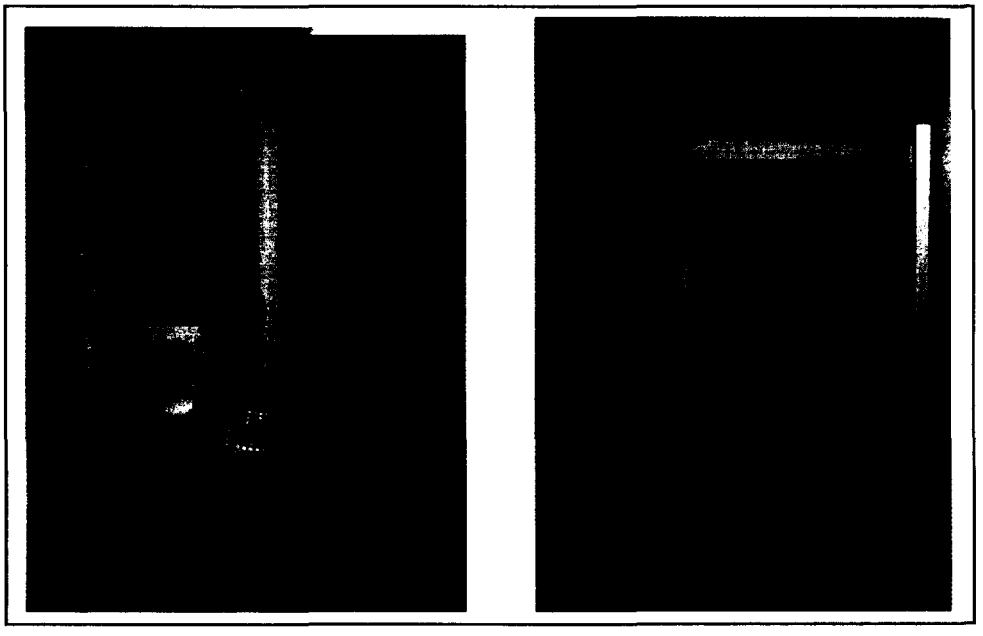


Fig. 1 Two views of a listening station contained in Dynastat's Audio Testing Laboratory.

LISTENER SAMPLE

The sample of listeners for the NRSC subjective experiments was stratified both for listener gender and age-group. For each experiment listeners were recruited to represent approximately equal representation in eight categories: four *Age-Groups* (16-24, 25-32, 33-42, 43-50) for each *Gender* (male, female). In general, each experiment required Dynastat to deliver the subjective data from 40 *qualified* listeners, where *qualification* was based on performance on an initial screening-test developed by iBiquity and a post-hoc screening test designed to eliminate obvious outliers. To achieve balance in the stratification of the sample and at the same time account for disqualifications due to failures of the screening-tests, approximately 50 listeners were recruited for each experiment. Listeners were recruited primarily from a pool of more than 2000 listeners contained in Dynastat's subjective testing database. This database is a continually evolving and expanding pool of listeners that Dynastat has maintained for use in subjective evaluation of speech-coding and voice-communications systems. Membership in Dynastat's subjective database is largely dictated by guidelines specified by ITU-T[1] and other standardization bodies.

PROCEDURES

Upon arrival at Dynastat, listeners completed a brief biographical data-sheet and received verbal instructions on the specific tasks to be performed in the experiment. Exhibit A shows the instructions that were read to listeners for a typical experiment[2]. Each listener was assigned a unique eight-character listener ID (i.e., **Eeeesai**) coded for experiment (**Eeee**), gender (**s** = 1 for male, 2 for female), age-group (**a** = 1 for 16-24, 2 for 25-32, etc.), and individual (i.e., individual within the category, **i** = 1, 2, etc.). For example, the ID "*E01a112*" would identify the 2nd individual listener who was a male listener in age-group 16-25 participating in Experiment 01a. The test administrator entered the listener's ID and biographical information into an Excel *Participant* file specific to the experiment. The overall duration of each experiment was approximately 2½ hours and typically included three phases: a training phase, a screening phase, and a testing phase consisting of one or more test sessions. The overall test duration is within the maximum testing time recommended by the ITU-T's recommendation P.800.

Training Phase

During the training phase listeners were presented a range of audio impairments typical of those involved in the testing phase of the experiment. The training phase was developed and provided to Dynastat by iBiquity and was used to expose and familiarize the listeners to the variety and range of conditions they were likely to hear in the subsequent screening and testing

[1] ITU-T Recommendation P.800, Methods for subjective determination of transmission quality, Aug., 1996.

[2] All experiments conducted by Dynastat in this phase of the testing effort used the ACR for at least a portion of the experiment. Some experiments also used other methodologies for a portion of the experiment. In those cases the ACR was always run first. Both the ACR and the other methodologies will be described in sections that follow.

phases. The impairments presented in the training phase ranged from subtle to extreme and served to train the listeners to listen carefully for potential impairments in the audio samples. The training materials were presented at a group listening station in the Audio Laboratory equipped with a Rane HC-6 distribution amplifier, which allowed the test administrator, and up to four listeners to hear the training materials over the Sennheiser HD-600 headphones. There were seven training samples, each involving multiple cuts. In each training sample, the first cut was a "clean" cut followed by two or more "impaired" cuts of the same materials. Each listener was asked if he could tell the difference between the cuts, i.e., if he could hear the impairments. The sample was replayed until all listeners acknowledged that they could hear the impairments. The experimenter asked only if the listeners could "hear the differences between the cuts." The experimenter never discussed the specific types of impairments involved in the training samples or how the listeners should judge or value those impairments.

Screening Phase

Immediately after the training phase, listeners participated in a pre-test screening phase to ensure that they were able to reliably distinguish between "clean" and "impaired" samples. The listener's task in the screening phase was a "Reference-A-B" comparison in which the listener was required to decide which of two "test" samples (A or B) was the same as the reference sample. In each trial one of the test samples was the same as the clean or unimpaired reference sample and the other sample was an impaired sample. Figure 2 shows the PC response display that was used for the screening task. Playback of samples was under the individual listener's control, but the screening software required him to listen to all three samples, reference and two test samples, before the response options were available. Listeners were free to replay any or all of the three samples until they were ready to enter their response and proceed to the next trial. The screening phase consisted of one practice trial and ten test trials. Listeners were provided no feedback on the "correctness" of their responses during the screening test. After completion of the screening phase, the listeners exited their booth for a short rest-break during which the test administrator scored their screening responses. Listeners were not informed of their specific performance in the screening phase, but depending on their score, were placed in one of three categories. If a listener scored 50% or less (i.e., 5 of 10 correct or at the "chance" level) he was paid a partial fee for his participation and was not allowed to proceed to the test phase of the experiment. If a listener scored 60% or 70% he was allowed to proceed to the test phase but his data was not used in the final set of ratings delivered to iBiquity (i.e., his data was disqualified based on screening test performance). Listeners who scored 80-100% proceeded to the test phase as a "qualified" listener and their rating data was used in "post-hoc" screening designed to provide the most reliable data possible. Description of the "post-hoc" data screening is provided in a later section.

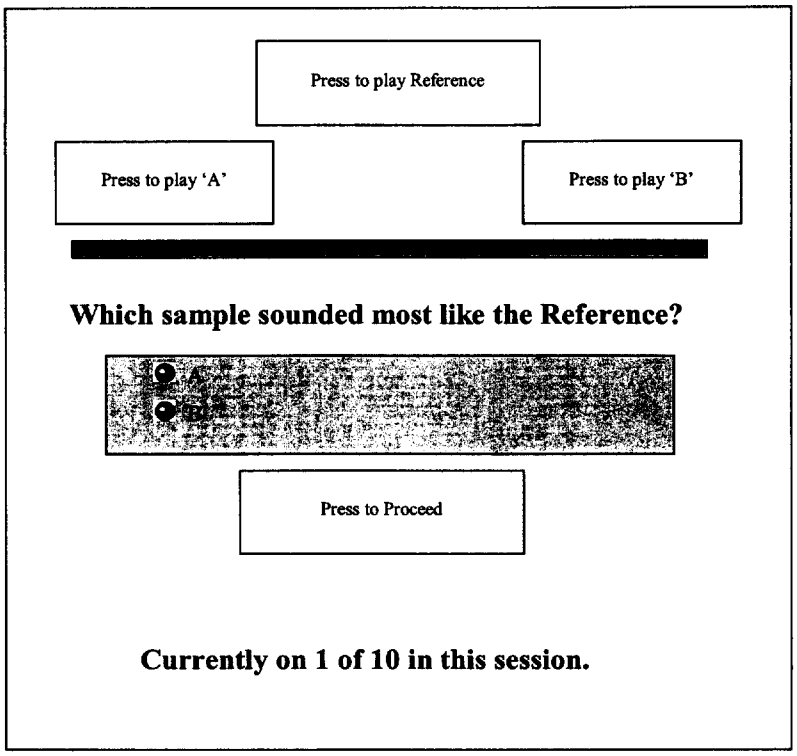


Fig.2 Response display for the Ref/A/B task in the screening phase.

Testing Phase

Table I presents a summary of the eleven experiments that Dynastat has conducted to date for this testing effort. The Absolute Category Rating (ACR) method was the primary subjective evaluation tool in these experiments and every experiment included the ACR as one portion of the experiment. For those experiments that also included another evaluation method (Exps. 5b, 7a, 9b), the ACR portion of the experiment was always run first. The ACR method yields the Mean Opinion Score (MOS), a measure of overall audio quality. The ACR requires the listener to judge the quality of an audio sample using a five category rating scale where: Excellent=5, Good=4, Fair=3, Poor=2, and Bad=1. The category judgments are reported as a measure of overall audio quality, an MOS, on a scale of 1 to 5. A response display for the ACR testing task is shown in Fig. 3. The listener controlled playback of the audio samples but on each trial he could enter his response only after listening to the entire sample. Typically, the testing phase consisted of two practice trials followed by approximately 200 test trials. The listener could adjust the playback volume during the practice trials. The playback volume set by the listener during the practice trials was then maintained throughout the remainder of the experiment. Test trials were grouped into sessions of approximately 50 trials each, separated by rest-breaks. During the rest-breaks listeners were required to remove the headphones and leave the booth.

Exp.	Test Methodology	# Audio Samples	# Listeners Retained	# Listeners Excluded	Min. FoM	FoM Index
E01a-f	ACR	195	60	7	0.720	0.814
E02	ACR	180	40	5	0.752	0.876
E03	ACR	222	40	9	0.701	0.804
E04	ACR	297	40	5	0.819	0.896
E05a	ACR	136	40	8	0.515	0.698
E05b	Point to Blend	72	"	"	0.629	0.746
E06	ACR	195	40	10	0.605	0.731
E07b	ACR	122	40	7	0.582	0.782
E07a	Interpret.of MOS	84	"	"	0.558	0.666
E08	ACR	214	40	5	0.682	0.787
E09a	ACR	108	60	5	0.709	0.720
E09b	Ticker	9	"	"	---	---
E10a	ACR	148	40	5	0.754	0.883
E10b	ACR	72	"	"	0.445	0.673
E11	ACR	268	40	9	0.807	0.860
Total		**2322**	**480**	**75**	**0.663**	**0.781**

Table I. Summary of experiments conducted by Dynastat.

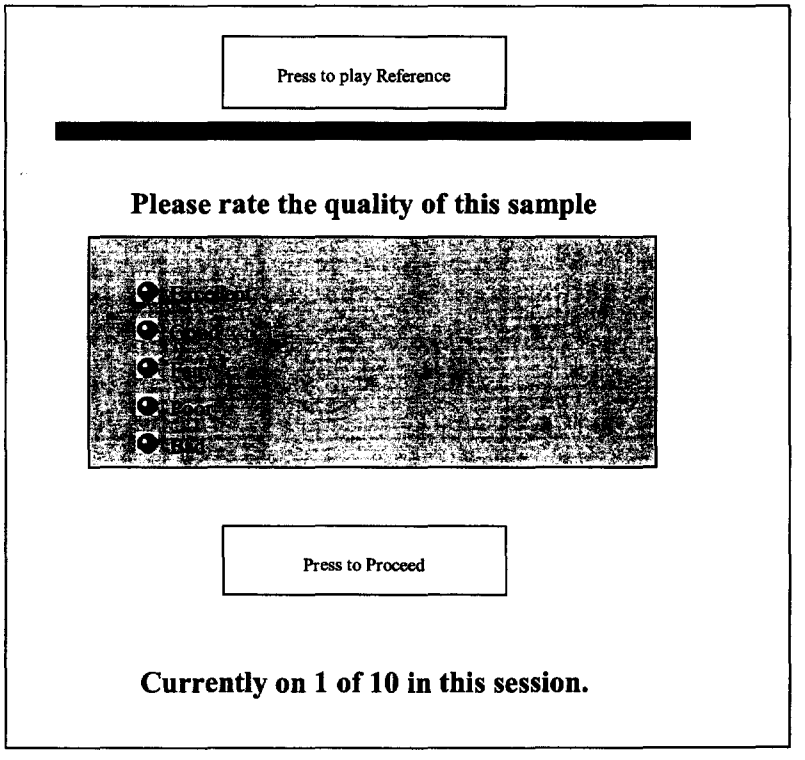


Fig. 3. Response display for the ACR task in the testing phase.

AUDIO MATERIALS

For each experiment approximately 200 processed audio samples were supplied to Dynastat by Advanced Television Technology Center (ATTC) for the *laboratory* tests and directly by iBiquity for the *field* tests. The audio materials were delivered to Dynastat either on CD-ROM or via Internet FTP. The files were provided in digital format (44.1KHz, 16 bit linear WAV). The digital files were loaded onto the hard-disk of the server PC and then distributed to the hard-disks of the individual PC's though a local area network.

AUDIO FILE PRESENTATION AND DATA COLLECTION

For each experiment Dynastat prepared an Excel file that controlled the audio file presentation and data collection software. During this process the audio files were loaded and verified, file order randomizations were created, and the overall layout of the experiment was established (i.e., number of test sessions, number of trials per session, and number of rest-breaks). The iBiquity software package automatically accumulated the listener responses into an experiment specific Excel *Response* spreadsheet. Once a test session had been initiated, the iBiquity software required no input from the test administrator. File presentation and data collection were controlled by the interaction of the listener and the software.

POST-HOC DATA ANALYSIS AND LISTENER SCREENING

At the conclusion of the data collection for an experiment, the total set of listener data (i.e., the experiment *Response* Excel file) was subjected to a post-hoc analysis to ensure the validity and the reliability of the data for each individual listener. For each experiment a "Figure of Merit" (FoM) was calculated for each listener participating in the experiment. The FOM was the "coefficient of correlation" between the individual listener's vector of ratings and the vector containing the average ratings for the remainder of the listeners involved in the experiment. Thirty years of experience with subjective rating data has shown this FoM to be a valuable screening measure to remove clear "outliers" from the rating data (i.e., listeners who either can't or won't perform the rating task). A practical lower threshold of 0.70[3] for the FoM was generally used to classify listeners as "outliers" and remove their data from the experiment. The last two columns in Table I show, for each experiment, the minimum value of FoM for the listeners that were retained in the final set of data delivered to iBiquity as well as the FoM index for the experiment. After eliminating listeners from the data set on the bases of pre-test and post-hoc screening, it was sometime necessary to remove additional listeners in order to satisfy the sample

[3] Since the FoM is based on a correlation coefficient it is subject both to the amount of variation in the rating data as well as the range of that data. The criterion value of .70 was arbitrarily chosen on the basis of empirical evidence and past experience in subjective testing efforts. For individual experiments the criterion value was adjusted according to the variation and range of the observed data.

stratification requirements. In case where one or more *qualified* listeners had to be removed from a specific gender/age-group category, listeners were randomly selected for deletion.

DATA DELIVERY

For each experiment Dynastat compiled and delivered two Excel worksheets to iBiquity. The *Participant* worksheet contained the biographical and ID information for the 40 listeners contained in the final data set. The *Response* worksheet contained the raw response data for those listeners. Exhibit C shows the Participant worksheet for each of the 11 experiments conducted by Dynastat.

OTHER PROCEDURES AND METHODOLOGIES

Experiment 5b – "Point to Blend" Test

In Experiment 5, the first portion of the experiment (Exp.5a) used the standard ACR rating scale shown previously in Fig. 3. In Experiment 5b, each audio sample included an analog/digital transition, the "Point to Blend." After listening to each sample listeners were presented the standard ACR rating display followed by a second response screen asking them to describe the basis for their response to the ACR question. Figure 4 shows the response display for the second question in the "Point to Blend" test.

What most affected your rating of the audio sample?

- ☐ The contrast between the best and worst audio
- ☐ The Best sounding portion of the audio sample
- ☐ The Worst sounding portion of the audio sample
- ☐ The transition point between the best and the worst audio
- ☐ I equally weighted the good and bad audio to arrive at my rating\
- ☐ Other (Write your answer on the sheet provided)

Fig. 4 Response display for the second question in Experiment 5b – "Point of Blend".

Experiment 7a – "Interpretation of MOS" Test

In Experiment 7, the first portion of the experiment (Exp.7b) used the standard ACR methodology and the rating scale shown previously in Fig.3. Note that for each listener Exp.7b was run first followed by Exp.7a. Experiment 7a involved two questions on each trial. The first question was a modification of the standard ACR rating scale using a continuous slider instead of the standard radio buttons. The sliding scale registered the listener's response within a tenth of an MOS point. After responding to the ACR question, the listener was then presented with a

response display posing the question: "Considering the quality of the sample you just heard, would you turn the radio off?" Listeners responded "Yes" or "No" using their mouse.

Experiment 9b – "Ticker" Test

In Experiment 9, the first portion of the experiment (Exp.9a) used the standard ACR methodology. In Exp.9b, listeners performed a different task, the "Ticker Test." Listeners were given a set of written instructions (reproduced in Exhibit B) explaining the task to be performed in Exp.9b. A special "Ticker" software program was provided by iBiquity to present the audio files and collect the listener's responses. For this experiment listeners used the PC keyboard instead of the mouse to register their responses. Exp.09b involved 10 trials - a single 15-sec. practice trial followed by nine 4-min. test trials. The listener's task was to detect impairments in the audio stream and to press one of two response keys depending on the judgment of the level of impairment: the "F1" key for a "small impairment;" the "F2" key for a "large" impairment. The criterion for distinguishing between "small" and "large" impairments was left up to the individual listener's discretion. Dynastat collected the raw "ticker" response data and forwarded it to iBiquity. There was no post-hoc analysis performed on the "ticker" data.

DATA AUDITING PROCEDURES

Dynastat was requested to perform an independent audit of the data analyses procedures that iBiquity was using to summarize and report the results of the subjective testing conducted by Dynastat. These results included Excel worksheets and "pivot-tables" derived from raw rating data delivered to iBiquity by Dynastat. The head of Dynastat's Evaluation Services section, Alan Sharpley, performed this task during a two-day visit to iBiquity's New Jersey facility. Mr. Sharpley developed a data-stream audit process to track the iBiquity procedures from raw rating data to final "pivot-table" output. The data-stream audit process involved the following steps:

1. Verify that the raw data files (i.e., Excel "Response" worksheets) for each experiment were the same files that Dynastat had previously delivered to iBiquity.
2. Decode the audio sample label using "Text to Columns" parsing to arrange raw listener rating data by conditions and experimental parameters.
3. Copy parameters from existing iBiquity pivot table (Data, Rows, Cols).
4. Construct new pivot table using Response Worksheet from step 2 and the parameters from step 3.
5. Compute "Difference Table" comparing values (i.e., rating means and confidence intervals) in iBiquity pivot-table with values in auditor's pivot-table.
6. Verify that the two-pivot tables, iBiquity's and auditor's, are identical.

Dynastat verified that all pivot-table results that were examined during the audit process were identical to the pivot-tables constructed by the auditor.

Exhibit A - Instructions for ACR Audio Rating Experiments

Overview

Welcome to this audio testing session. Today, you will be participating in a listening experiment which should last about two and a half hours. You will be listening to music and speech samples over headphones. We are studying how various radios sound under different transmission conditions. There are three parts to this study. The first part is training, where you will listen to the music you will be encountering in your tests. The second part is a discrimination test. The third part is an opinion test.

Training Task

In the training session, you will hear a variety of sound samples. These sound samples include typical transmission "impairments" you might hear during the discrimination and opinion tests. These impairments should be noticeable. During the course of each sample you will hear varying degrees of the "impairment". You will indicate to the administrator if differences are heard.

Discrimination Task

In the discrimination task we will be testing your ability to hear different impairments. In this task your job is to decide which of two samples (A or B) is most nearly the same as the reference sample. The response display is shown in Fig. 1. To begin click on the box labeled "Press to Play Reference". The complete reference sample will be played. Similarly, you will click on "Press to Play A" and "Press to Play B" to play these complete samples. The program will not let you enter a response until you have heard all three samples completely. After listening to the complete Reference, A, and B samples you can enter your response to the question "Which sample sounded most like the reference?". After indicating your response click on the box labeled "Press to proceed". If you would like to play any of the samples again, you can press the appropriate box and do so as much as needed until you have made your decision. Once you have indicated your response and clicked on the "Press to proceed" you will be ready to start your next trial. During the course of your practice trial for this task you can set the volume level my moving the slider box. Once this level is set it cannot be changed for the rest of the session.

The discrimination session will consist of one practice trial and 10 test trials. When you complete the task open the door and proceed to the waiting room for a 10-minute break. During the break the administrator will score your data and let you know if you passed the test. If you passed the test then you are eligible to participate in the opinion test. If you did not pass you will be paid $20 for your efforts.

Opinion Task – The ACR-MOS Test

In this part of the experiment we are evaluating systems that might be used for the radio transmission of sound samples. You are going to hear a number of recorded samples and rating how good you think they sound.

On each trial a single sample will be presented. Each sample will consist of a 10-15 second music or voice passage. Please listen to the complete sample, then indicate your opinion of the overall sound quality of the sample using the following 5-point scale: Exellent, Good, Fair, Poor, Bad. Figure 3 shows the response display.

This task is different from the discrimination task. There is no stated reference against which to compare the samples you are hearing. You simply hear a passage and then make a rating. You will have to use an internal reference to judge "the goodness" of the sample. By that we mean, when you are listening to a particular sample, think about how a very good radio station would sound in your car and over your home radio. Judge the sample in relation to your memory of those two references.

Many things go into a quality rating. You'll be listening for impairments as well as the overall aesthetic quality. By aesthetic we mean beauty, musicality, character, sound quality, etc. Try to judge each sample in an overall sense. This is especially hard to do if a big impairment happens to occur at the end of the sample. So, before you rate each sample, take a few seconds to think about the entire sample you just heard. In that way, it won't be just your last impression that carries the most weight.

The experiment will involve four test sessions separated by short rest periods. In the first session you will have a practice block of 2 trials to familiarize you with the rating task and adjust your listening volume. The practice block will be followed by 4 test sessions of 50 trials each. If you have any questions, please feel free to ask the test administrator.

Please do not discuss your opinions with any other listeners participating in the experiment. Thank you in advance for your participation.

Exhibit B – Instructions for Ticker Test

You will be participating in a listening experiment, which will last approximately 35 minutes. You will be listening to speech and music samples that were recorded off the radio at different radio stations. There will be a total of 10 trials during this experiment. The first trial is a Training Trial, which consists of a 15-second sample. Each of the remaining 9 trials are approximately 4 minutes long.

While listening to each sample (on headphones), your task is to record each time you hear an impairment. There are a variety of impairments that you might hear. For example, you may hear interference from another radio station, pops, clicks, hiss, distortions in the music or speech, changes in stereo-imagery (from stereo to mono), and loss of fidelity. Please be sure to record all of these impairments.

It is important to note that you may hear some amount of background noise or ringing, as you probably have heard when listening to the radio in your car. This background noise, which lasts for the entire sample, should not be recorded. If you were to record the background noise, you would probably end up holding one of the buttons down for the entire sample and would be unable to record any additional impairments as they happen. Instead, you should record only the impairments that occur in addition to the background noise or ringing. If you do not detect any additional impairments for a particular sample or portion of a sample, you need not press the buttons.

In addition to identifying impairments, you are also asked to decide whether each impairment is a "small" or "large" impairment. To indicate this, you should press one of the following keys:

F1: for a Small Impairment
F2: for a Large Impairment

You should press the appropriate button each time you hear an impairment. This may mean that you are pressing a button repeatedly within a short period of time. If the impairments are occurring too close together to record them with individual presses of the buttons, you may hold down the appropriate button for the duration of that impairment cluster.

When you are ready to begin, you may press the "F5" button on your keyboard. The computer will start playing the sample and you should begin to record all impairments that might exist. Be sure that you are actually ready to begin the task when you press the "F5" button because there is no way of pausing the sample once it has begun. A horizontal progress bar will provide information regarding the length of the sample that has already been played. When the progress bar is completely filled, the sample has been played in its entirety.

The two LEDs (labeled "F1" and "F2") will provide visual feedback indicating which button you have pressed and how many times you have pressed it. Each time you press the F1 key, the LED labeled "F1" will turn green, and each time you press the F2 key, the LED labeled "F2" will turn green.

You are required to take short breaks when you have completed a trial. When you have completed the last trial (Trial 9), please open the door and check with your experimenter.

Exhibit C

Participant Worksheets for Experiments 1-11.

Participant Worksheet for Experiment 1 (Parts a - f).

	Exp.Code	Gender	Age group	FoM	Screen %
1	E01a111	Male	1	0.852	100
2	E01a112	Male	1	0.724	90
3	E01a121	Male	2	0.868	100
4	E01a131	Male	3	0.901	90
5	E01a141	Male	4	0.739	80
6	E01a211	Female	1	0.849	90
7	E01a212	Female	1	0.771	100
8	E01a221	Female	2	0.826	100
9	E01a231	Female	3	0.771	90
10	E01a241	Female	4	0.814	100
11	E01b111	Male	1	0.894	100
12	E01b121	Male	2	0.878	100
13	E01b131	Male	3	0.811	100
14	E01b132	Male	3	0.910	80
15	E01b141	Male	4	0.826	80
16	E01b211	Female	1	0.867	90
17	E01b221	Female	2	0.867	90
18	E01b231	Female	3	0.783	100
19	E01b232	Female	3	0.850	100
20	E01b242	Female	4	0.812	80
21	E01c111	Male	1	0.850	100
22	E01c121	Male	2	0.895	100
23	E01c131	Male	3	0.821	100
24	E01c132	Male	3	0.792	90
25	E01c141	Male	4	0.858	90
26	E01c211	Female	1	0.724	100
27	E01c212	Female	1	0.849	100
28	E01c221	Female	2	0.779	100
29	E01c231	Female	3	0.852	100
30	E01c243	Female	4	0.816	90
31	E01d111	Male	1	0.840	90
32	E01d112	Male	1	0.832	80
33	E01d121	Male	2	0.852	100
34	E01d132	Male	3	0.843	100
35	E01d141	Male	4	0.861	100
36	E01d211	Female	1	0.843	80
37	E01d212	Female	2	0.862	100
38	E01d221	Female	2	0.785	100
39	E01d231	Female	3	0.780	90
40	E01d241	Female	4	0.894	90
41	E01e111	Male	1	0.793	80
42	E01e121	Male	2	0.834	100
43	E01e122	Male	2	0.821	100
44	E01e131	Male	3	0.886	100
45	E01e141	Male	4	0.864	100
46	E01e211	Female	1	0.828	90
47	E01e212	Female	1	0.860	100
48	E01e221	Female	2	0.815	100
49	E01e232	Female	3	0.743	90
50	E01e241	Female	4	0.754	100
51	E01f111	Male	1	0.818	90
52	E01f121	Male	2	0.847	100
53	E01f122	Male	2	0.864	100
54	E01f131	Male	3	0.831	90
55	E01f142	Male	4	0.748	100
56	E01f211	Female	1	0.720	90
57	E01f221	Female	2	0.861	90
58	E01f232	Female	3	0.741	90
59	E01f241	Female	4	0.861	90
60	E01f242	Female	4	0.838	100

Excluded listeners

	Exp.Code	Gender	Age group	FoM	Screen %	Reason
1	E01b241	Female	4	*	70	Screening
2	E01c241	Female	4	0.677	80	FoM
3	E01c242	Female	4	0.849	70	Screening
4	E01d131	Male	3	*	70	Screening
5	E01e231	Female	3	0.630	90	FoM
6	E01f141	Male	4	0.615	90	FoM
7	E01f231	Female	3	0.679	90	FoM

Participant Worksheet for Experiment 2.

	Exp.Code	Gender	Age-group	FoM	Screening (%)
1	E02x111	Male	1	0.909	90
2	E02x112	Male	1	0.905	100
3	E02x114	Male	1	0.792	80
4	E02x115	Male	1	0.916	100
5	E02x116	Male	1	0.942	100
6	E02x121	Male	2	0.872	100
7	E02x122	Male	2	0.923	100
8	E02x123	Male	2	0.902	100
9	E02x125	Male	2	0.908	100
10	E02x126	Male	2	0.875	90
11	E02x131	Male	3	0.909	80
12	E02x132	Male	3	0.890	90
13	E02x133	Male	3	0.853	100
14	E02x134	Male	3	0.830	100
15	E02x135	Male	3	0.899	90
16	E02x136	Male	3	0.874	90
17	E02x142	Male	4	0.752	90
18	E02x143	Male	4	0.891	100
19	E02x144	Male	4	0.904	100
20	E02x145	Male	4	0.802	80
21	E02x211	Female	1	0.868	90
22	E02x213	Female	1	0.907	100
23	E02x214	Female	1	0.847	100
24	E02x215	Female	1	0.891	100
25	E02x216	Female	1	0.903	100
26	E02x221	Female	2	0.912	90
27	E02x222	Female	2	0.871	100
28	E02x223	Female	2	0.869	100
29	E02x224	Female	2	0.897	100
30	E02x225	Female	2	0.877	90
31	E02x231	Female	3	0.901	90
32	E02x232	Female	3	0.891	100
33	E02x233	Female	3	0.882	100
34	E02x234	Female	3	0.862	100
35	E02x235	Female	3	0.912	100
36	E02x241	Female	4	0.933	100
37	E02x242	Female	4	0.862	100
38	E02x243	Female	4	0.898	90
39	E02x244	Female	4	0.894	100
40	E02x245	Female	4	0.807	90

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E02x113	Male	1	0.870	70	Screening
2	E02x124	Male	2	0.850	80	Random
3	E02x141	Male	4	0.874	40	Screening
4	E02x212	Female	1	0.844	100	Random
5	E02x236	Female	3	0.881	80	Random

Participant Worksheet for Experiment 3.

	Exp.Code	Gender	Age-group	FoM	Screening (%)
1	E03x111	Male	1	0.782	80
2	E03x112	Male	1	0.873	90
3	E03X113	Male	1	0.891	100
4	E03X115	Male	1	0.800	90
5	E03X116	Male	1	0.880	90
6	E03x117	Male	1	0.849	100
7	E03X122	Male	2	0.830	90
8	E03x123	Male	2	0.864	90
9	E03x124	Male	2	0.888	100
10	E03X125	Male	2	0.852	100
11	E03X126	Male	2	0.865	90
12	E03X127	Male	2	0.819	100
13	E03X132	Male	3	0.867	80
14	E03x133	Male	3	0.856	90
15	E03X135	Male	3	0.898	80
16	E03X136	Male	3	0.890	100
17	E03x141	Male	4	0.879	90
18	E03X142	Male	4	0.809	90
19	E03x143	Male	4	0.737	90
20	E03x145	Male	4	0.701	90
21	E03x211	Female	1	0.890	100
22	E03x213	Female	1	0.842	100
23	E03X214	Female	1	0.804	100
24	E03X215	Female	1	0.820	100
25	E03X218	Female	1	0.892	100
26	E03X221	Female	2	0.868	100
27	E03x222	Female	2	0.845	90
28	E03X224	Female	2	0.851	90
29	E03x225	Female	2	0.830	80
30	E03x226	Female	2	0.780	90
31	E03x232	Female	3	0.776	100
32	E03x233	Female	3	0.861	90
33	E03x234	Female	3	0.816	90
34	E03X235	Female	3	0.829	100
35	E03x236	Female	3	0.803	80
36	E03X241	Female	4	0.841	80
37	E03x242	Female	4	0.879	90
38	E03X244	Female	4	0.859	100
39	E03x245	Female	4	0.824	90
40	E03x246	Female	4	0.810	100

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E03x114	Male	1	0.723	80	FoM
2	E03X121	Male	2	0.140	100	FoM
3	E03x131	Male	3	0.671	90	FoM
4	E03x134	Male	3	0.711	70	Screening
5	E03x144	Male	4	0.669	90	FoM
6	E03x216	Female	1	0.649	100	FoM
7	E03X217	Female	1	0.797	100	FoM
8	E03x223	Female	2	0.826	50	Screening
9	E03x243	Female	4	0.765	60	Screening

Participant Worksheet for Experiment 4.

	Exp.Code	Gender	Age-group	FoM	Screening (%)
1	E04x112	Male	1	0.867	80
2	E04x113	Male	1	0.914	100
3	E04x114	Male	1	0.926	100
4	E04x115	Male	1	0.914	100
5	E04x117	Male	1	0.923	100
6	E04x118	Male	1	0.891	100
7	E04x121	Male	2	0.923	100
8	E04x122	Male	2	0.855	100
9	E04x125	Male	2	0.902	100
10	E04x126	Male	2	0.913	80
11	E04x127	Male	2	0.890	100
12	E04x131	Male	3	0.889	100
13	E04x132	Male	3	0.908	80
14	E04x133	Male	3	0.899	100
15	E04x134	Male	3	0.879	90
16	E04x141	Male	4	0.866	90
17	E04x142	Male	4	0.878	90
18	E04x143	Male	4	0.895	100
19	E04x144	Male	4	0.924	100
20	E04x145	Male	4	0.913	90
21	E04x211	Female	1	0.943	80
22	E04x212	Female	1	0.934	100
23	E04x214	Female	1	0.931	100
24	E04x215	Female	1	0.921	80
25	E04x216	Female	1	0.943	90
26	E04x218	Female	1	0.903	90
27	E04x221	Female	2	0.869	100
28	E04x222	Female	2	0.922	100
29	E04x223	Female	2	0.879	100
30	E04x225	Female	2	0.937	100
31	E04x226	Female	2	0.895	100
32	E04x231	Female	3	0.894	90
33	E04x232	Female	3	0.879	100
34	E04x233	Female	3	0.935	90
35	E04x234	Female	3	0.937	100
36	E04x236	Female	3	0.904	100
37	E04x241	Female	4	0.819	80
38	E04x242	Female	4	0.833	100
39	E04x244	Female	4	0.858	90
40	E04x245	Female	4	0.906	100

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E04x123	Male	2	0.912	70	Screening
2	E04x124	Male	2	0.852	100	Random
3	E04x217	Female	1	0.850	100	Random
4	E04x235	Female	3	0.872	100	Random
5	E04x243	Female	4	0.885	70	Screening

Participant Worksheet for Experiment 5 (Parts a and b).

	Exp.Code	Gender	Age-group	FoM(5a)	FoM(5b)	Screening (%)
1	E05x111	Male	1	0.646	0.712	100
2	E05x113	Male	1	0.697	0.814	100
3	E05x115	Male	1	0.764	0.692	100
4	E05x116	Male	1	0.785	0.788	100
5	E05x117	Male	1	0.772	0.629	100
6	E05x121	Male	2	0.622	0.741	80
7	E05x122	Male	2	0.792	0.861	100
8	E05x123	Male	2	0.754	0.701	100
9	E05x124	Male	2	0.775	0.840	100
10	E05x125	Male	2	0.728	0.791	100
11	E05x127	Male	2	0.778	0.777	90
12	E05x132	Male	3	0.867	0.731	90
13	E05x134	Male	3	0.682	0.690	80
14	E05x135	Male	3	0.676	0.725	90
15	E05x136	Male	3	0.846	0.791	100
16	E05x137	Male	3	0.760	0.759	100
17	E05x140	Male	4	0.694	0.747	100
18	E05x141	Male	4	0.679	0.706	100
19	E05x143	Male	4	0.564	0.710	90
20	E05x145	Male	4	0.628	0.724	100
21	E05x212	Female	1	0.668	0.707	100
22	E05x213	Female	1	0.823	0.742	100
23	E05x214	Female	1	0.791	0.809	100
24	E05x215	Female	1	0.806	0.832	90
25	E05x216	Female	1	0.711	0.796	90
26	E05x218	Female	1	0.703	0.718	100
27	E05x221	Female	2	0.515	0.760	80
28	E05x222	Female	2	0.753	0.799	90
29	E05x223	Female	2	0.673	0.845	90
30	E05x225	Female	2	0.766	0.761	90
31	E05x226	Female	2	0.748	0.795	90
32	E05x227	Female	2	0.815	0.764	100
33	E05x231	Female	3	0.697	0.735	100
34	E05x233	Female	3	0.692	0.722	100
35	E05x234	Female	3	0.732	0.820	100
36	E05x235	Female	3	0.722	0.847	90
37	E05x241	Female	4	0.688	0.705	80
38	E05x243	Female	4	0.780	0.686	100
39	E05x244	Female	4	0.704	0.843	100
40	E05x245	Female	4	0.770	0.808	90

Excluded listeners

	Exp.Code	Gender	Age-group	FoM(5a)	FoM(5b)	Screening (%)	Reason
1	E05x114	Male	1	0.760	0.800	60	Screening
2	E05x126	Male	2	0.743	0.603	80	FoM
3	E05x131	Male	3	0.513	0.670	90	FoM
4	E05x133	Male	3	0.469	0.700	100	FoM
5	E05x142	Male	4	0.721	0.660	70	Screening
6	E05x144	Male	4	0.193	0.571	90	FoM
7	E05x232	Female	3	0.662	0.514	100	FoM
8	E05x242	Female	4	0.573	0.612	100	FoM

Participant Worksheet for Experiment 6.

	Exp.Code	Gender	Age-group	FoM	Screening (%)
1	E06X111	Male	1	0.708	100
2	E06x112	Male	1	0.723	100
3	E06x114	Male	1	0.814	100
4	E06x115	Male	1	0.729	90
5	E06x116	Male	1	0.801	100
6	E06x117	Male	1	0.761	100
7	E06X121	Male	2	0.811	100
8	E06X122	Male	2	0.722	100
9	E06x123	Male	2	0.750	100
10	E06x124	Male	2	0.761	100
11	E06x125	Male	2	0.825	100
12	E06x127	Male	2	0.779	100
13	E06X131	Male	3	0.721	90
14	E06X132	Male	3	0.802	100
15	E06x133	Male	3	0.734	100
16	E06x134	Male	3	0.746	80
17	E06X143	Male	4	0.787	90
18	E06x144	Male	4	0.768	90
19	E06x145	Male	4	0.805	90
20	E06x147	Male	4	0.605	80
21	E06X211	Female	1	0.751	80
22	E06X212	Female	1	0.777	100
23	E06X213	Female	1	0.824	100
24	E06X217	Female	1	0.760	90
25	E06x218	Female	1	0.754	100
26	E06x221	Female	2	0.735	100
27	E06X222	Female	2	0.852	80
28	E06x223	Female	2	0.785	90
29	E06X224	Female	2	0.706	90
30	E06x225	Female	2	0.864	90
31	E06x227	Female	2	0.725	90
32	E06X231	Female	3	0.793	100
33	E06x233	Female	3	0.775	90
34	E06x234	Female	3	0.790	80
35	E06x235	Female	3	0.756	90
36	E06x236	Female	3	0.790	100
37	E06x242	Female	4	0.716	90
38	E06X243	Female	4	0.730	90
39	E06X244	Female	4	0.875	90
40	E06x245	Female	4	0.683	90

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E06x113	Male	1	0.619	100	FoM
2	E06X126	Male	2	0.701	80	FoM
3	E06x135	Male	3	0.676	100	FoM
4	E06x137	Male	3	0.678	100	FoM
5	E06x141	Male	4	0.383	80	FoM
6	E06X142	Male	4	0.561	70	Screening
7	E06x146	Male	4	0.588	90	FoM
8	E06x215	Female	1	0.694	100	FoM
9	E06X232	Female	3	0.647	100	FoM
10	E06x241	Female	4	0.512	90	FoM

Participant Worksheet for Experiment 7 (Parts a and b).

	Exp.Code	Gender	Age-group	FoM(7a)	FoM(7b)	Screening (%)
1	E07x111	Male	1	0.767	0.773	100
2	E07x112	Male	1	0.867	0.696	100
3	E07x113	Male	1	0.856	0.761	90
4	E07x114	Male	1	0.798	0.729	100
5	E07x115	Male	1	0.844	0.673	100
6	E07x117	Male	1	0.674	0.667	100
7	E07x121	Male	2	0.582	0.656	90
8	E07x122	Male	2	0.847	0.779	100
9	E07x123	Male	2	0.850	0.771	100
10	E07x125	Male	2	0.856	0.662	90
11	E07x126	Male	2	0.759	0.656	90
12	E07x133	Male	3	0.782	0.652	100
13	E07x134	Male	3	0.708	0.577	100
14	E07x135	Male	3	0.907	0.681	100
15	E07x136	Male	3	0.825	0.786	100
16	E07x137	Male	3	0.808	0.788	100
17	E07x141	Male	4	0.674	0.558	80
18	E07x142	Male	4	0.773	0.758	90
19	E07x143	Male	4	0.818	0.638	100
20	E07x144	Male	4	0.903	0.631	90
21	E07x146	Male	4	0.871	0.765	100
22	E07x211	Female	1	0.870	0.734	90
23	E07x212	Female	1	0.919	0.777	100
24	E07x213	Female	1	0.921	0.758	100
25	E07x214	Female	1	0.866	0.780	90
26	E07x215	Female	1	0.905	0.815	100
27	E07x218	Female	1	0.832	0.777	100
28	E07x221	Female	2	0.866	0.820	100
29	E07x222	Female	2	0.646	0.616	90
30	E07x224	Female	2	0.728	0.748	90
31	E07x226	Female	2	0.888	0.791	90
32	E07x231	Female	3	0.713	0.620	90
33	E07x232	Female	3	0.887	0.732	90
34	E07x235	Female	3	0.860	0.722	100
35	E07x237	Female	3	0.882	0.632	100
36	E07x238	Female	3	0.738	0.618	90
37	E07x241	Female	4	0.609	0.578	100
38	E07x242	Female	4	0.852	0.683	100
39	E07x243	Female	4	0.707	0.621	90
40	E07x244	Female	4	0.718	0.735	100

Excluded listeners

	Exp.Code	Gender	Age-group	FoM(7a)	FoM(7b)	Screening (%)	Reason
1	E07x124	Male	2	0.622	0.525	100	FoM
2	E07x131	Male	3	0.753	0.185	80	FoM
3	E07x145	Male	4	0.707	0.447	100	FoM
4	E07x216	Female	1	0.710	0.586	100	FoM
5	E07x225	Female	2	0.877	0.580	60	Screening
6	E07x228	Female	2	0.346	0.454	80	FoM
7	E07x234	Female	3	0.743	0.410	90	FoM

Participant Worksheet for Experiment 8.

	Exp.Code	Gender	Age-group	FoM	Screening (%)
1	E08x111	Male	1	0.853	90
2	E08x112	Male	1	0.800	80
3	E08x113	Male	1	0.789	100
4	E08x114	Male	1	0.871	90
5	E08x121	Male	2	0.801	80
6	E08x122	Male	2	0.692	80
7	E08x124	Male	2	0.784	100
8	E08x125	Male	2	0.692	90
9	E08x126	Male	2	0.787	100
10	E08x127	Male	2	0.807	100
11	E08x131	Male	3	0.844	100
12	E08x133	Male	3	0.729	90
13	E08x134	Male	3	0.870	100
14	E08x135	Male	3	0.887	90
15	E08x136	Male	3	0.877	90
16	E08x137	Male	3	0.864	90
17	E08x141	Male	4	0.794	80
18	E08x142	Male	4	0.843	100
19	E08x143	Male	4	0.848	100
20	E08x144	Male	4	0.874	100
21	E08x211	Female	1	0.856	90
22	E08x212	Female	1	0.775	100
23	E08x213	Female	1	0.870	80
24	E08x214	Female	1	0.806	100
25	E08x215	Female	1	0.834	100
26	E08x216	Female	1	0.759	100
27	E08x221	Female	2	0.825	100
28	E08x222	Female	2	0.834	100
29	E08x223	Female	2	0.855	100
30	E08x225	Female	2	0.811	90
31	E08x226	Female	2	0.795	100
32	E08x227	Female	2	0.795	90
33	E08x231	Female	3	0.812	90
34	E08x232	Female	3	0.828	100
35	E08x233	Female	3	0.758	100
36	E08x235	Female	3	0.791	80
37	E08x241	Female	4	0.694	80
38	E08x243	Female	4	0.682	100
39	E08x244	Female	4	0.815	90
40	E08x245	Female	4	0.752	80

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E08x132	Male	3	0.807	70	Screening
2	E08x224	Female	2	0.819	70	Screening
3	E08x234	Female	3	0.455	80	FoM
4	E08x236	Female	3	0.520	90	FoM
5	E08x242	Female	4	0.644	70	Screening

Participant Worksheet for Experiment 9 (Parts a and b).

	Exp.Code	Gender	Age group	Ticker group	FoM	Screen %
1	E09x112	Male	1	2	0.836	80
2	E09x113	Male	1	3	0.708	100
3	E09x114	Male	1	4	0.847	100
4	E09x115	Male	1	5	0.774	100
5	E09x116	Male	1	6	0.789	90
6	E09x117	Male	1	7	0.738	100
7	E09x118	Male	1	8	0.739	100
8	E09x121	Male	2	1	0.770	80
9	E09x122a	Male	2	2	0.748	100
10	E09x122b	Male	2	2	0.793	100
11	E09x123	Male	2	3	0.849	100
12	E09x126	Male	2	6	0.820	90
13	E09x127	Male	2	7	0.772	100
14	E09x128	Male	2	8	0.824	100
15	E09x133	Male	3	3	0.816	90
16	E09x134	Male	3	4	0.835	100
17	E09x135	Male	3	5	0.818	90
18	E09x136	Male	3	6	0.793	80
19	E09x137	Male	3	7	0.713	90
20	E09x143	Male	4	3	0.709	90
21	E09x146	Male	4	6	0.731	80
22	E09x147	Male	4	7	0.746	80
23	E09x211	Female	1	1	0.823	90
24	E09x212	Female	1	2	0.787	90
25	E09x213	Female	1	3	0.843	100
26	E09x214	Female	1	4	0.724	100
27	E09x215	Female	1	5	0.855	90
28	E09x216	Female	1	6	0.808	90
29	E09x217	Female	1	7	0.750	100
30	E09x218	Female	1	8	0.837	80
31	E09x222	Female	2	2	0.870	100
32	E09x223	Female	2	3	0.846	100
33	E09x224	Female	2	4	0.780	100
34	E09x225	Female	2	5	0.785	90
35	E09x227	Female	2	7	0.860	100
36	E09x228	Female	2	8	0.832	100
37	E09x232	Female	3	2	0.722	80
38	E09x233	Female	3	3	0.871	100
39	E09x235	Female	3	5	0.771	90
40	E09x237	Female	3	7	0.777	80
41	E09x238	Female	3	8	0.827	100
42	E09x241	Female	4	1	0.742	60
43	E09x242	Female	4	2	0.724	90
44	E09x243	Female	4	3	0.864	100
45	E09x245	Female	4	5	0.843	100
46	E09x246	Female	4	6	0.822	100
47	E09x247	Female	4	7	0.749	80
48	E09x248	Female	4	8	0.829	100

Excluded listeners

	Exp.Code	Gender	ge-gro	cker G	FoM	eening	Reason
1	E09x111	Male	1	1	0.634	90	FoM
2	E09x124	Male	2	4	0.589	80	FoM
3	E09x131	Male	3	1	0.115	100	FoM
4	E09x132	Male	3	2	0.672	90	FoM
5	E09x138	Male	3	8	0.636	70	Screening
6	E09x141	Male	4	1	0.413	90	FoM
7	E09x142	Male	4	2	0.652	100	FoM
8	E09x144	Male	4	4	0.565	100	FoM
9	E09x145	Male	4	5	0.464	90	FoM
10	E09x148	Male	4	8	0.554	80	FoM
11	E09x221	Female	2	1	0.409	90	FoM
12	E09x231	Female	3	1	0.200	70	Screening
13	E09x234	Female	3	4	0.296	100	FoM
14	E09x234	Female	3	4	0.678	90	FoM
15	E09x236	Female	3	6	0.683	90	FoM
16	E09x238	Female	3	8	0.547	100	FoM
17	E09x244	Female	4	4	0.625	80	FoM

Participant Worksheet for Experiment 10 (Parts a and b).

	Exp10a	FoM	Screening %
1	E10x111	0.859	100
2	E10x112	0.894	90
3	E10x113	0.811	100
4	E10x114	0.949	100
5	E10x115	0.924	100
6	E10x116	0.844	100
7	E10X121	0.961	100
8	E10x122	0.845	90
9	E10x123	0.786	100
10	E10x124	0.875	100
11	E10x125	0.754	100
12	E10x126	0.939	90
13	E10x131	0.950	100
14	E10x132	0.866	100
15	E10x133	0.828	90
16	E10x135	0.934	80
17	E10x142	0.931	80
18	E10x144	0.927	80
19	E10x145	0.800	80
20	E10x146	0.904	100
21	E10x211	0.902	80
22	E10x212	0.899	90
23	E10x213	0.950	100
24	E10x214	0.882	100
25	E10x215	0.851	90
26	E10x216	0.926	90
27	E10x221	0.863	90
28	E10x222	0.893	90
29	E10x223	0.912	90
30	E10x224	0.913	100
31	E10x225	0.890	100
32	E10x232	0.925	80
33	E10x233	0.907	100
34	E10x234	0.858	100
35	E10x235	0.878	100
36	E10x242	0.953	100
37	E10x243	0.939	90
38	E10x244	0.928	90
39	E10x245	0.934	100
40	E10x246	0.930	90

	Exp10b	FoM	Screening %
1	E10x111	0.693	100
2	E10x112	0.708	90
3	E10x113	0.737	100
4	E10x114	0.756	100
5	E10x115	0.627	100
6	E10x116	0.788	100
7	E10X121	0.551	100
8	E10x123	0.619	100
9	E10x124	0.714	100
10	E10x125	0.445	100
11	E10x126	0.797	90
12	E10x131	0.708	100
13	E10x132	0.706	100
14	E10x133	0.652	90
15	E10x135	0.735	80
16	E10x142	0.647	80
17	E10x144	0.553	80
18	E10x145	0.642	80
19	E10x146	0.748	100
20	E10x211	0.752	80
21	E10x212	0.771	90
22	E10x213	0.733	100
23	E10x214	0.752	100
24	E10x215	0.838	90
25	E10x216	0.793	90
26	E10x221	0.708	90
27	E10x222	0.779	90
28	E10x223	0.772	90
29	E10x224	0.818	100
30	E10x225	0.647	100
31	E10x231	0.616	90
32	E10x232	0.778	80
33	E10x233	0.617	100
34	E10x234	0.755	100
35	E10x235	0.646	100
36	E10x242	0.614	100
37	E10x243	0.712	90
38	E10x244	0.799	90
39	E10x245	0.653	100
40	E10x246	0.762	90

Excluded listeners

	Exp.Code	FoM	Screening (%)	Reason
1	E10x134	0.835	70	Screening
2	E10x137	*	40	Screening
3	E10x226	0.923	90	FoM(b)
4	E10x231	0.799	90	Random
5	E10x241	0.695	100	FoM(b)

* no data collected

Excluded listeners

	Exp.Code	FoM	Screening (%)	Reason
1	E10x122	0.199	90	FoM
2	E10x134	0.761	70	Screening
3	E10x137	*	40	Screening
4	E10x226	0.450	90	FoM
5	E10x241	0.306	100	FoM

* no data collected

Participant Worksheet for Experiment 11.

	Exp11	Gender	Age-group	FoM	Screening %
1	E11x111	Male	1	0.889	100
2	E11x112	Male	1	0.871	90
3	E11x114	Male	1	0.896	100
4	E11x115	Male	1	0.902	90
5	E11x116	Male	1	0.860	100
6	E11x122	Male	2	0.881	100
7	E11x123	Male	2	0.888	100
8	E11x124	Male	2	0.907	100
9	E11x125	Male	2	0.914	90
10	E11x126	Male	2	0.886	90
11	E11x127	Male	2	0.904	100
12	E11x131	Male	3	0.893	80
13	E11x132	Male	3	0.896	100
14	E11x134	Male	3	0.913	100
15	E11x135	Male	3	0.846	90
16	E11x137	Male	3	0.892	90
17	E11x141	Male	4	0.890	80
18	E11x142	Male	4	0.846	80
19	E11x144	Male	4	0.807	90
20	E11x145	Male	4	0.848	100
21	E11x211	Female	1	0.862	100
22	E11x212	Female	1	0.883	80
23	E11x213	Female	1	0.912	100
24	E11x215	Female	1	0.870	100
25	E11x217	Female	1	0.878	100
26	E11x221	Female	2	0.881	100
27	E11x223	Female	2	0.838	100
28	E11x224	Female	2	0.894	100
29	E11x226	Female	2	0.854	80
30	E11x232	Female	3	0.890	100
31	E11x233	Female	3	0.873	90
32	E11x235	Female	3	0.911	100
33	E11x236	Female	3	0.842	100
34	E11x237	Female	3	0.869	90
35	E11x241	Female	4	0.892	80
36	E11x242	Female	4	0.818	100
37	E11x244	Female	4	0.861	90
38	E11x245	Female	4	0.820	80
39	E11x246	Female	4	0.868	100
40	E11x247	Female	4	0.839	80

Excluded listeners

	Exp.Code	Gender	Age-group	FoM	Screening (%)	Reason
1	E11x113	Male	1	0.815	80	Random
2	E11x117	Male	1	0.835	90	Random
3	E11x121	Male	2	0.849	100	Random
4	E11x143	Male	4	0.792	80	Random
5	E11x146	Male	4	0.780	70	Screening
6	E11x214	Female	1	0.818	100	Random
7	E11x216	Female	1	0.824	100	Random
8	E11x222	Female	2	0.663	100	FoM
9	E11x243	Female	4	0.823	80	Random