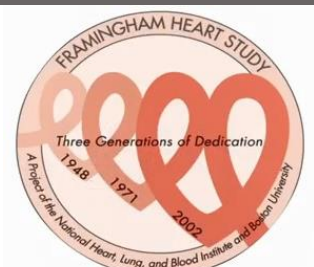


# DIGITAL VOICE CAPTURE MANUAL FOR IN-PERSON COGNITIVE TESTING

This manual is published by the ADRC Clinical Task Force Cognitive Working Group and represents a collaboration between the Framingham Heart Study Brain Aging Program at Boston University and the Indiana Alzheimer's Disease Research Center at Indiana University.

Sherral Devine, PhD  
Cody Karjadi, MS  
Hannah Craft, MPH



## Table of Contents

Introduction .....	2
About this Manual .....	3
<b>Part 1: Basic Digital Voice Capture.....</b>	<b>4</b>
Language for IRB and Informed Consent.....	4
Selecting Audio Recording Equipment.....	5
Audio Setup and Calibration.....	6
Recording .....	7
Saving Unedited Audio Files.....	7
<b>Part 2: Generating Analysis-Ready Files.....</b>	<b>9</b>
Personally Identifiable Information .....	9
Types of PII to Flag .....	9
Procedure for Testers to Flag PII.....	10
Processing Audio Recording .....	10
Labeling and Silencing PII in Audacity.....	10
Using Labels to Save Cognitive Tests as Individual Files .....	11
Saving Processed Audio Files .....	12
Quality Control .....	15
Manual QC for PII.....	15
QC for File Labels and Locations.....	16
Appendix A: Frequently Asked Questions (FAQs).....	17
Appendix B: Recording with the ZOOM H4N DVR .....	18
Appendix C. Digital Voice Recorder Alternatives.....	23
References .....	29

# Introduction

As research on Alzheimer's disease and related dementias (ADRD) centers on detecting symptoms earlier in the insidious onset process, there is an increasingly pressing need to develop methods for detecting them, presumably before the impact of the underlying pathological changes are irreversible. While there have been great advances in developing imaging and blood-based biomarkers of AD at the clinically pre-symptomatic level, the presumption of "pre-symptomatic" is currently being driven by the tools used to detect them. Measuring longitudinal changes in cognitive function is one of the core clinical indicators of pathological onset.

Surprisingly the research efforts to build more sensitive tools of cognitive function have not kept pace with that of PET/MRI imaging or fluid (blood/CSF) biomarkers. Yet cognitive function remains the primary outcome against which all these biomarkers as well as clinical trial treatment impact are measured. Thus, the AD research community is investing in methods for detecting AD pathology emergence at its earliest point but has not made concomitant investment in methods for detecting AD-related cognitive symptoms that might be emerging in parallel.

The implementation of digitally recording participant responses to neuropsychologist tests is the easiest, most cost-effective way to detect early changes in cognition. Speaking is a cognitively complex task and thus embedded in the spoken responses are acoustic and linguistic features that likely map onto the multiple cognitive domains implicated by neuropathological changes. As our cognitive capabilities shift, we express them through vocal responses in subtle ways, such as switching up word choices or sentence structures because of word finding problems, pausing, hesitating, and shifting as memory, attention, and executive functions are compromised.

Currently, there are no gold standards in methods for analyzing voice recordings, but just as with blood-based biomarkers, there is a growing, albeit still limited, literature suggesting that analysis of digital voice recordings as a method for differentiating those with and without cognitive impairment is promising. Thus, to facilitate the opportunities of using digital voice as a novel method for assessing cognition, we provide a manual of operations that describes how to collect digital voice recordings for research purposes.

Rhoda Au, PhD

Co-Principal Investigator; Framingham Heart Study Brain Aging Program

Boston University Professor of Anatomy & Neurobiology, Neurology, Medicine & Epidemiology

# About this Manual

This document is intended to guide Alzheimer's Disease Research Centers (ADRCs) and other interested groups in audio recording the UDS4 cognitive exam for research analysis. At this time, digital voice capture is encouraged but not required.

This manual is divided into two parts which provide ADRCs with two options for implementation. Part 1 contains best practices recommendations for basic digital voice capture. This involves recording and storing a digital audio file of the UDS4 cognitive exam with a prescribed file naming convention and data log for later research analysis. Centers can choose to only implement Part 1 at this time, with the option to pursue Part 2 when more resources are available.

Part 2 guides centers through generating analysis-ready files that can be shared with outside investigators. It involves audio editing for deidentification, dividing cognitive tests into individual files, and quality control measures. Part 2 is resource intensive, and we are currently exploring ways to financially support implementation. In addition, the National Alzheimer's Coordinating Center (NACC) is presently developing options for receipt and dissemination of deidentified digital speech files.

# Part 1: Basic Digital Voice Capture

## Language for IRB and Informed Consent

The Framingham Heart Study Brain Aging Program (FHS-BAP) records the consent process in addition to the battery, however centers can choose to only record the testing battery. The excerpts below can help centers develop their own language to include in IRB and consent forms.

FHS-BAP includes the following statement in its IRB application.

“Prior to beginning the consent process, participants will be informed that the examination will be audio recorded. The audio recording will also be stored for future data analysis. If the participant refuses, audio recording will not be done; otherwise, recording will begin and the consent process will commence. The consent process is being audio recorded to facilitate quality control, to ensure appropriate consenting. Participants have the right to refuse being recorded at any time during the exam.”

The FHS-BAP IRB application also contains the following language specifically regarding virtual visits.

“Use of this method of interview is expected to have minimal risks. The risk involves potential for breach of confidentiality. The risk is minimized by using a secure web platform that is used by many hospitals and clinics for doctors to communicate with patients, hence we believe it will be secure for this research interview. The visits will be digitally recorded (1x participant/month will be video recorded via Zoom recording and each participant will be audio recorded) and kept on FHS servers similar to our audio recordings of in person cognitive testing. The video recordings will be retained for short term storage, long enough for QC purposes, and will then be destroyed/removed. The FHS forms are filled and stored at the main FHS facilities along with all other FHS records. Participants can always decline to answer any question or decline to complete any test even if the participant consents and completes to the rest of the questions or test within the examinations. Specifically regarding use of smartphone-based cognitive tests: Use of the smartphone-based cognitive applications is expected to have minimal risks. The smartphone assessments are generally shorter in length (approximately 45 minutes in total) as compared to the standard NP in-exam (about 60 minutes in total), so participant burden should be reduced. Breaks will be provided between tests in order to reduce potential fatigue, which is likely to be transient and have very little overall impact. The risk involves potential loss of confidentiality. While the investigators are confident in the secure nature of data storage at FHS, such as locked filing cabinets for storing paper files and use of password protection and encryption for storing electronic files in computer systems, the information stored on the application developers' servers is out of the study team's control. However, Linus Health and other application developers will be blinded to any identifiable information using system-generated study IDs. Participants can always decline to answer any question or decline to complete any test even if the participant consents and completes to the rest of the questions or test within the smartphone-based cognitive assessment.”

The FHS-BAP Informed Consent Form contains the following statements about audio recordings.

Within the description of what will happen in the study:

“This session will be recorded using a digital audio recorder. Recordings will be analyzed in conjunction with other study information. We will also use recordings to make sure that your responses are accurately documented.”

For virtual visits:

“During Your Virtual Visit: You will be asked similar questions and administered the same tests that you would encounter during an in-person neuropsychological exam. The tests given virtually will be as close as possible to the tests given in person, modified only for use virtually. Later on, you may be offered a second in-person evaluation at your convenience, in your home. The tests will be audio and video recorded for data integrity purposes and further analysis.”

Within the Confidentiality section:

“We will store electronic files in computer systems with password protection and encryption. Access to these records is limited to authorized FHS staff. However, we cannot guarantee complete confidentiality. The files will be kept indefinitely, and there are no plans to destroy any of the records. Coded data and digital recordings from all tests will be stored in a repository and will be shared with qualifying investigators...”

“...Coded digital (video/audio) recording information will be analyzed by qualifying collaborators inside and outside of BU Medical Campus/BMC. Your name and other personally identifying information will not be shared with these entities.”

## Selecting Audio Recording Equipment

The audio equipment selected by a center will greatly affect the quality of recordings. In this section we have outlined the important details to consider in this selection process.

FHS-BAP uses the Zoom H4N recorder, which meets the criteria below. If your center also chose this recorder, see Appendix A for detailed instructions on its use. Additional recording device recommendations can be found in Appendix B.

Factors to consider when choosing a recorder:

1. Portability (if off-site testing is done)
2. Compatibility with lab’s computers (Mac/PC/Linux)
3. Both AC Adapter and Battery options for power (rechargeable battery preferred)
4. Microphone: We recommend choosing a recording device that can lay on the table in front of the participant and that has two microphones (for better sound quality than a single mic) as part of the recorder.

Lapel mics, which clip to the participant's shirt, may pick up too much rustling noise if the participant moves. Headsets might be too cumbersome or uncomfortable. Hanging (ceiling mounted) mics are not portable. If using external mics, avoid using condenser microphones (they pick up more room reverb); use cardioid mics instead.

5. Recording capacity: Depending on lab conditions, determine whether the recorder can collect multiple sessions before downloading the files to a computer, or whether it needs to be downloaded after each testing session.
6. Sampling rate: The sampling rate is one measure of audio quality, expressed in Hertz (Hz). The lowest acceptable sampling rate is 16000 Hz and we recommend capturing audio at a sampling rate of 44100 Hz.
7. Format/encoding: We highly recommend collecting voice data in the **WAV format with LINEAR16 PCM encoding with at least 16000 Hz sampling rate.** (For more detailed information on audio recording format and encoding, see references 2-3.)
8. Quantity: If multiple testing sessions are scheduled simultaneously, testers will need more than one recording device. If testers go off-site for sessions (i.e. to the participant's residence), they will need to take the device with them. Centers should consider acquiring an extra device in case one is broken or lost.
9. Additional equipment: if the chosen recorder requires an SD card, the center should ensure that multiple SD cards are purchased in case some are lost.

## Audio Setup and Calibration

Recorder placement and the sound quality of the testing room will greatly impact the recording quality. Sound reverberating off bare walls & floors, humming equipment, external noises (e.g. loud colleagues, birds chirping, etc.), and microphone direction are some of the factors that can diminish sound quality of voice recordings.

Sometimes these features are not under the tester's control, and we must work with what we have. This is especially true when recording is done in the participant's home. However, if you have the option to make changes to the testing environment, consider the following:

### Ideal Testing Room

1. Small to medium size
2. Multiple soft surfaces like carpet, couches, pillows, etc.
3. Avoid rooms that have a lot of hard surfaces that will make sound bounce around, such as windows, bare walls, and hard floors
4. Minimal exposure to external sounds (e.g., street noise, a conference room, loud colleagues, ringing phones, plumbing, weather)
5. Turn off noisy things in the room (e.g., fan, phone, air conditioner, computer in overdrive)
6. Lay a towel or piece of cloth under the recorder

### Placement of recorder

1. Point the mic(s) of the recorder toward the participant (and away from tester)
2. Place it in a location where it will be out of the way of testing (once you start recording, you don't want to be moving the recorder around)

3. If possible, place the recorder on furniture that is NOT the desk/table you are working on (because sounds such as pages turning, bangs on the table, etc. get picked up), but be sure it is close to the participant

The recording quality can be improved by sound-treating the testing room(s):

1. Floors
  - a. Carpet/rug
2. Ceiling/walls
  - a. [Bass Traps](#)
  - b. [Acoustic Panels](#)
  - c. Alternatively, can use [packing blankets](#) or mattress foam
  - d. Or, do it yourself (DIY)
    - i. [How to build a sound absorbing panel in 5 easy steps](#)
    - ii. [How to build your own acoustic panels](#)
    - iii. [Budget Audio Treatment](#)
    - iv. [Cheap Sound Treatment Tests in a Commercial Office](#)
    - v. [How to install acoustic foam without damaging your walls](#)
    - vi. [Tips for DIY](#)

The quality of audio can also be affected if the participant and/or examiner are wearing masks. We ask centers to track whether masks were worn for each testing session and if the participant, examiner, or both were wearing a mask.

A test recording in every room that is currently known to be a testing room should be sent to FHS-BAP at Boston University to confirm the quality of sound recording for data extraction. Email [adcvoice@bu.edu](mailto:adcvoice@bu.edu) for instructions on uploading test audio files. Please ensure that test recordings contain no PII. Make sure that every tester is trained in how to use the recording equipment and what to do with the audio recording after the testing session.

## Recording

We recommend that centers record the entire UDS4 cognitive battery in a single recording, because it might be distracting to the participant if the tester is repeatedly starting and stopping the recorder. However, centers have the flexibility to record individual tests if they prefer to do so. In a later section, we provide instructions on how to split a single recording into multiple audio files so that each test is in a separate audio file.

Some tests are virtually silent, such as Trails, but we encourage centers to continue recording even during these silences. Audio recordings can be used to analyze many different angles of testing, so it can be useful to record during quiet tests to catch if people speak or make noise during the test (such as verbalizing during Trails).

## Saving Unedited Audio Files & Data Log

After testing, the examiner will download the recording from the recorder to a computer and save it. **All files should be saved with the naming convention: [ADRC ID]\_[Accession#].wav**

It is imperative that centers maintain a data log with the following data variables for each recording:



- ADRC ID
- Accession #
- Visit date
- Visit number
- PTID/NACC ID
- Cognitive tests (NACC Code)
- Interviewer initials
- Whether the participant and/or interviewer were wearing masks
  - 1=interviewer
  - 2=participant
  - 3=both interviewer and participant
  - 4=no mask
- Where the recording took place
  - 1=clinic
  - 2=home
  - 3=nursing home or assisted living
  - 4=other
- Name of the recording device used

If your center is completing Part 2 below, these additional variables should be tracked in the data log:

- PHI removed (Y/N)
- Date of processing
- Initials of processor
- Program used
- Quality check

# Part 2: Generating Analysis-Ready Files

## Personally Identifiable Information

Personally Identifiable Information (PII) is information that can be used to identify, locate, or contact a single individual. **It is essential that all PII is removed from the recording prior to sharing with outside investigators.** Examiners should try to avoid using a participant's name during testing; however it is not uncommon for a participant to say something in the middle of testing that would be considered PII and therefore must be removed.

### Types of PII to Flag

This is a comprehensive list of the types of PII. Some types may occur more frequently during testing than others.

1. 18 HIPAA identifiers<sup>1</sup> of the individual or of relatives, employers, or household members of the individual:
  - a. Name (including maiden name)
  - b. All geographic subdivisions smaller than a state, including street address, city, county, precinct or neighborhood area, ZIP code, and their equivalent geocodes.
  - c. All elements of dates (except year) for dates directly related to an individual:
    - i. Birth date
    - ii. Admission date
    - iii. Discharge date
    - iv. Date of death
    - v. All ages over 89 (as well as the year of birth for this age group)
  - d. Telephone numbers
  - e. Fax numbers
  - f. Email addresses
  - g. Social Security numbers
  - h. Medical Record numbers
  - i. Health plan beneficiary numbers
  - j. Account numbers
  - k. Certificate/license numbers
  - l. Vehicle identifiers (e.g., serial numbers, license plate numbers)
  - m. Device identifiers and serial numbers
  - n. Web URLs
  - o. Internet protocol (IP) address
  - p. Biometric identifiers, including finger and voice prints
  - q. Full face photographic images and any comparable images
  - r. Any other unique identifying number, characteristic, or code
2. Research-related Identifiers
  1. Start-of-Exam Recorded Identifiers: Participant ID, tester ID, and date
3. Regional Identifiers
  1. Schools attended
  2. Place of work
  3. City of birth

## Procedure for Testers to Flag PII

PII may not occur often during the recorded testing, and testers can limit PII by collecting all participant-related information before beginning the recording and being mindful of speaking PII such as referring to the participant by name.

It is extremely important that the examiner pay attention to *every time* PII is spoken by either themselves or the participant. Upon speaking/hearing any such information, the examiner should mark the active battery page at the time of the PII. If the examiner is not using paper for the test (for example, if the center uses a computer or tablet instead), the center should agree on a way to mark when PII is spoken. All examiners in a center should use the same way of marking PII, in case the examiner who conducted the tests is not the same person who removes PII from the recording.

## Processing Audio Recording

Centers will use an audio editing software to remove PII from each recording. Some centers might have testers process the audio that they recorded, or they might have someone else process the recordings using the tester's notes where they flagged PII.

Centers can choose their preferred software as long as it can silence PII in such a way that it cannot be reversed and can save audio files in the WAV format. After silencing PII in the recording, the edited audio file should be saved with the "[ADRC ID]\_[Accession#].wav" file naming convention and corresponding details should be entered in the data log for sharing with external investigators. (**Do not save over the unedited audio file, save as a new file**).

We recommend using the software Audacity, which is a free audio recording and editing software for Windows, Mac, and Linux. Centers can download Audacity here: <https://www.audacityteam.org/download/>


Audacity can be used to record audio with a connected microphone, or centers can use a separate device to record and then download the recording to a computer that has the Audacity program. To open an audio file in Audacity, you can drag-and-drop the file or open Audacity, go to **File > Open**, and navigate to the appropriate file.

There are many online tutorials and instructional videos for Audacity users, so looking up your questions online will usually yield a solution. **For an introduction to the program, watch the first 4 minutes of this video:** <https://www.audacityteam.org/download/> (after 4 minutes, the video describes editing and noise reduction that you will *not* do for this project). We recommend that new users become familiar with the Audacity program and practice the following steps with sample audio *before* working on research recordings.

## Labeling and Silencing PII in Audacity

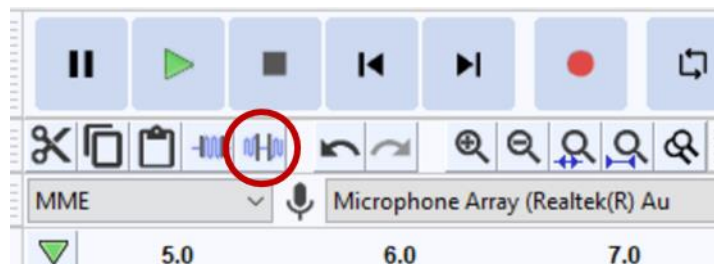
Audacity has a Silence tool that you will use to remove Personally Identifiable Information (PII) from the recording. To learn how to use the tool, watch this video (<https://youtu.be/Vgl6PUNv0fY>) and then follow the steps below.

You will follow these steps in conjunction with the steps in the next section (“Using Labels to Save Cognitive Tests as Individual Files”), so it is important to read and practice all the steps before working on recordings.

- After opening the appropriate audio recording in Audacity, go to the drop-down menu at the top of the program and find **Tracks**. Select **Add New > Label Track**. You will use the label track to mark when PII occurred in each recording.
- Listen to the audio to find the first occurrence of PII in the recording. If you click and drag on the audio track, you can highlight a portion of audio; if you press play, it will only play the highlighted section of audio. (This also works if you highlight part of the label track.)
- You can use the Zoom tools to zoom in on the audio: 
- Once you find PII, click and drag on the **label** track until you have highlighted the area with PII (you can press play to check). You can click and drag the start and end points of the highlighted section and keep replaying the segment until you have isolated the PII.

**Tip:** It might be hard to avoid including words on either side of the PII: for example, if someone says "Yep, my brother's name is John Smith and oh, um..." - It might be hard to not also grab when they say "is" and "and" on either side of "John Smith", depending on how fast they speak. We want to limit non-PII speech in the segments, but don't spend too much time trying to avoid capturing a word or two on either side of the segment. Overall, the idea is "do your best effort" on limiting non-PII speech.

- Press **ctrl + B** or **cmd + B**. This will create a label at the location you have highlighted. Type "PII".
- When you click on the "PII" label you made, it will highlight the selected audio. Click on the Silence icon or use **ctrl+L** to silence the selected audio:



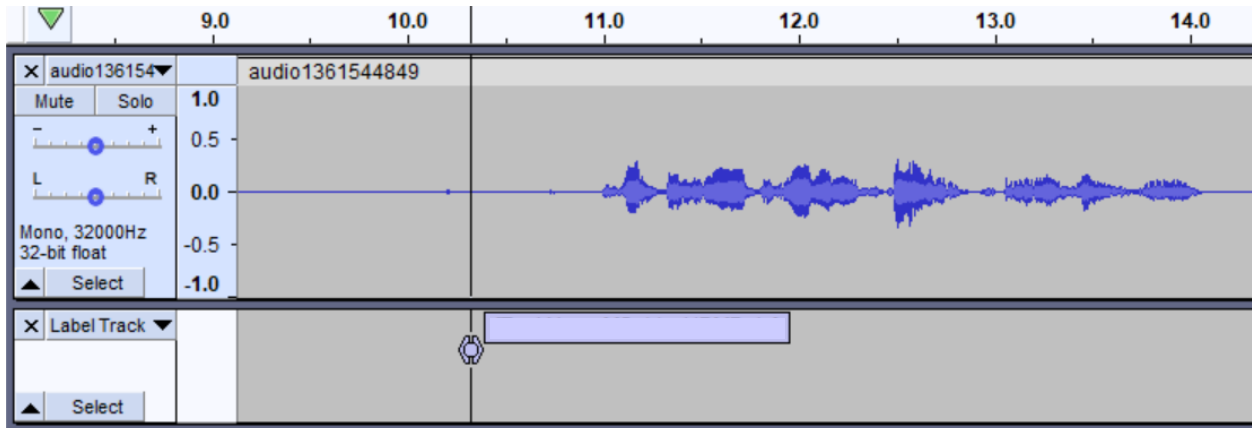
- Repeat these steps with each segment of PII until all are labeled and silenced.

## Using Labels to Save Cognitive Tests as Individual Files

Centers will most likely record all cognitive tests in a session as a single audio file. This is optimal because starting and stopping the recording for each test could be distracting to the participant. You can use Audacity to easily split and save each test as an individual audio file. Here is a video explaining this feature (<https://youtu.be/72ewbraaqi8>).

- At the top of the Audacity program, find the drop-down menu for **Tracks**. Select **Add New > Label Track**. (Centers will also use a label track to save timestamps of PII. This step will create a second label track. **It is important to use separate label tracks for the two tasks.**)

- Find the beginning of the first cognitive test. Click on the second label track so the vertical line is positioned before the first test begins.
- Press **ctrl + B** or **cmd + B**. This will create a label at the location you have selected. (If it creates a label on the PII label track instead of the cognitive test label track, it's because you need to click on the second label track before pressing **ctrl + B** or **cmd + B**.)
- Type in [ADRC ID]\_[Accession#] (this will eventually become the audio file name). Each test should have a different Accession # with the corresponding test noted in the data log using the NACC code.

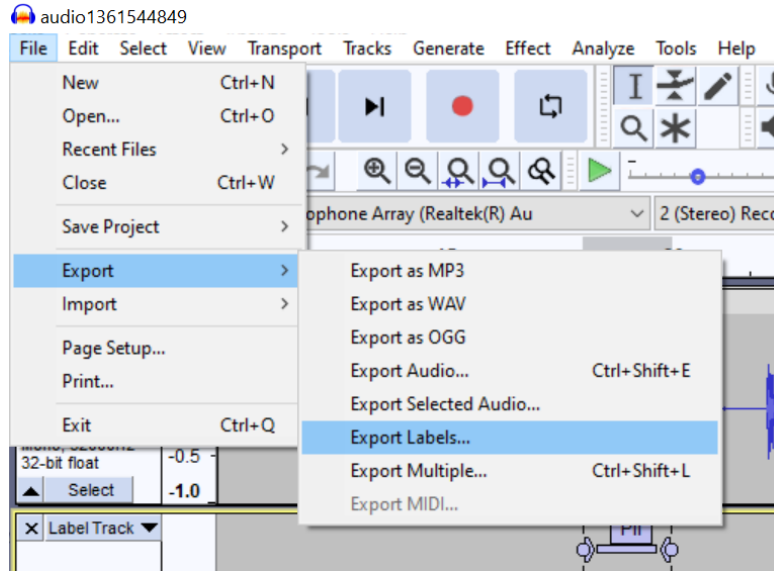


- Find the space between the end of the first test and the beginning of the second test. Click on the second label track at that location and use **ctrl + B** or **cmd + B** to create another label. Name it appropriately. Do this at the beginning of every test.

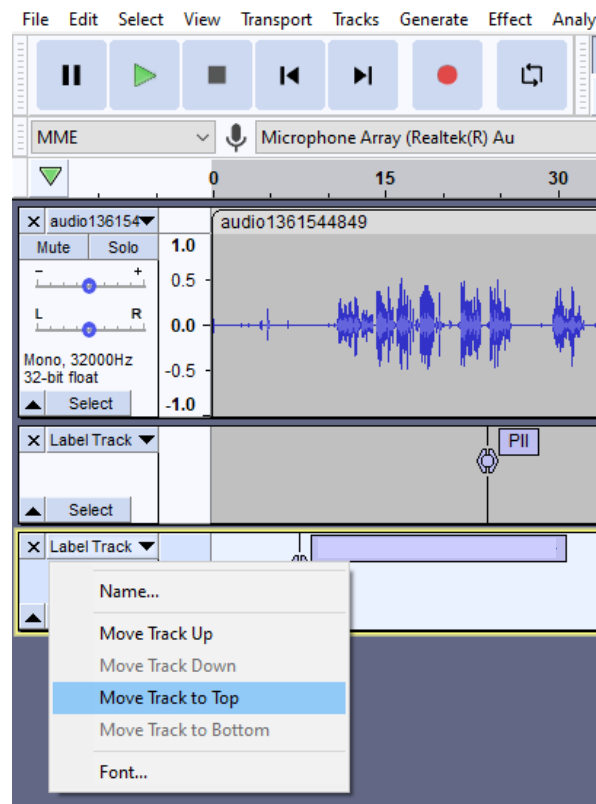
## Saving Processed Audio Files

After the PII and cognitive tests have been labeled and the PII is silenced, you need to save the labels and audio files separately. To do so, follow these steps:

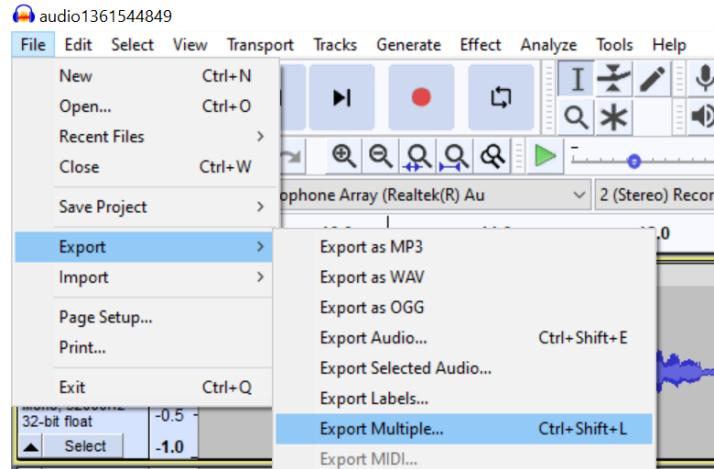
- To save the timestamps, click on **File > Export > Export labels**



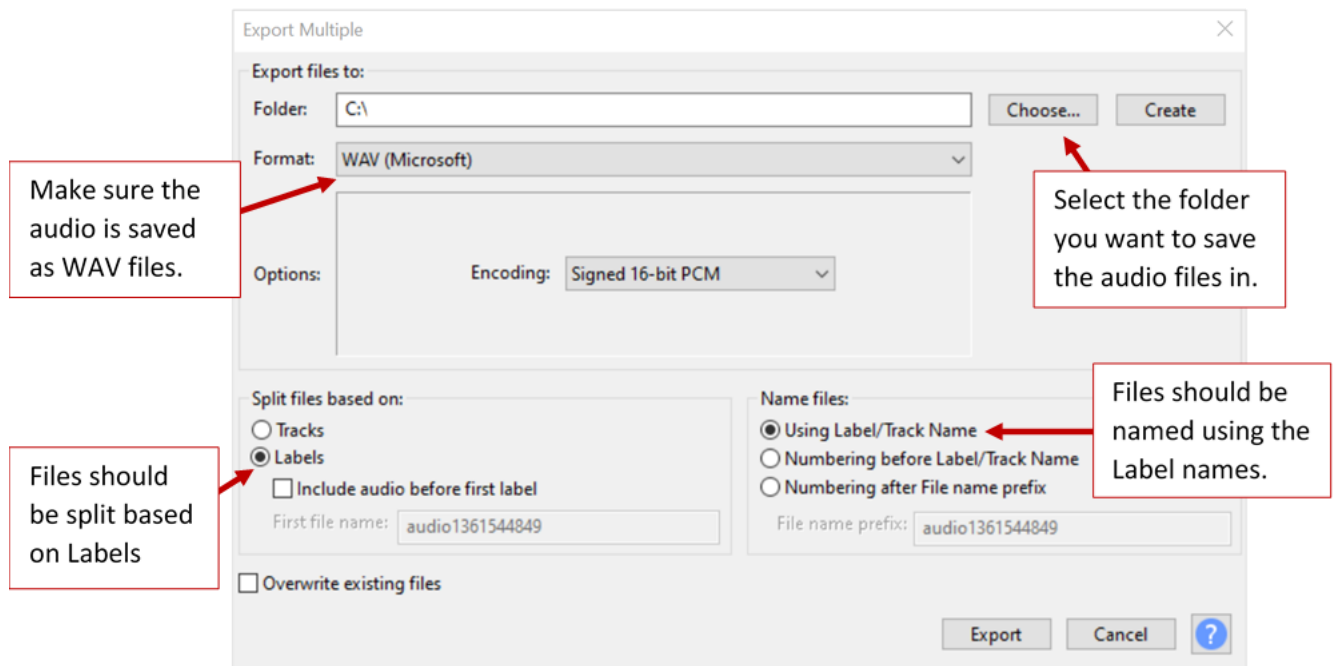
- Save the timestamp file (.txt is the default file type) in your center’s designated file location with the [ADRC ID]\_[Accession#] naming convention corresponding with the audio file. The text file will contain timestamps for both PII and the cognitive tests.
- Now it is time to save the audio files. Go to the label track that has labels for the cognitive tests. On the far right side of the track, click on the black triangle next to “Label Track” and select **Move Track to Top**:



- At the top of the program, click on **File > Export > Export Multiple**:



- The following menu will appear. Make sure you have these options selected:



- Click **Export**. Several windows will pop up, click **Ok** for all of them. The program will save the audio as individual files.
- Finally, save the Audacity project file in the center's designated file location. To do this, click on **File > Save Project > Save Project As...** Your center will use the Audacity project file to conduct QC and make any necessary changes to the labels or audio.

All audio files should be stored in a secure location and backed up regularly.

# Quality Control (QC)

The exact process for ensuring data quality is subject to each center's protocol. We recommend integrating efforts with the center's existing QC procedures and documentation. Below we offer suggested best practices that centers may want to adopt if feasible.

Centers should develop a system for tracking QC activities. This can be done using the center's preferred program such as REDCap or Excel. It should, at minimum, track the following variables:

1. Participant ID
2. Tester ID
3. Recording date
4. Name of audio file
5. ID of person conducting QC
6. Date of QC activity
7. What type of QC is being done (as outlined below: Supervisor, Peer, Intra, Data Integrity, etc.)
8. Whether the QC passed or failed
9. Why the QC failed, if applicable

We encourage centers to create a feedback loop for the QC process. This means that the people who conducted testing and processed the audio (this might be the same person or different people) are sent the QC results so they can tell if they made any mistakes. The testers and audio processors can sign off on the QC to confirm that they reviewed any errors.

## Manual QC for PII

Given the importance of maintaining the confidentiality of research participants, it is essential that QC measures are implemented to ensure consistent and accurate removal of PII from the recordings. While we provide a framework for QC below, we encourage centers to adapt and/or develop a process that works best with their existing infrastructure. Our recommended format includes three levels of QC: supervisor, peer, and self (intra).

**Supervisor QC:** A supervisor should regularly choose recordings at random to undergo QC. We recommend selecting at least one recording from each tester every month. They should listen to the entire recording to check that all PII has been labeled and silenced. If they find PII, they should flag it and ask a staff member proficient with Audacity to label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.

**Peer QC:** A Peer Reviewer should review 5-10% of the completed exams done by each tester on a regular basis. The recordings should be chosen at random. This does not have to be a trained NP tester; it just needs to be someone trained to listen for PII. The Peer Reviewer should listen to the recordings and make sure all PII has been labeled and silenced. If they find PII, they should flag it and ask a staff member proficient with Audacity to label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.

**Intra QC:** Each person processing the audio should review their own test recordings, chosen at random, to ensure all PII was accurately identified (we recommend reviewing one recording a



quarter). If they find PII, they should label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.

Each recording that undergoes QC should be logged in the center's designated QC tracking system. On a regular basis (we recommend quarterly), the tracking system entries should be reviewed to ensure that there are not (1) common pitfalls or (2) problems with the accuracy of any particular examiner or person processing the audio. Any common pitfalls or differences of opinion will be discussed with the team and steps taken to resolve them. If the person processing the audio is not consistently accurate, they should be given feedback and additional recordings from that person should be reviewed. The number of additional recordings that will be reviewed will be decided upon by a supervisor based on the given circumstances.

## QC for File Labels and Locations

We strongly recommend that centers implement a QC process that will ensure data files are labeled correctly and located in the appropriate folders. This is essential because the recording should contain no PII, which means the file name will be the only way to identify the recording. Audio files must be labeled according to the prescribed naming convention ([ADRC ID]\_[Accession#].wav) that corresponds with the correct row in the data log. Without a correct file name and match in the data log, it will be extremely difficult to align the voice data with participant phenotypic data. The file name and location should correspond with the data log with no discrepancies.

# Appendix A: Frequently Asked Questions (FAQs)

[inserts FAQs and answers here]

# Appendix B: Recording with the ZOOM H4N DVR

## Initial Set-up

1. Turn the power on by moving the power switch on the left panel of the device to "ON".
2. Press the menu button on right side of recorder.
3. Scroll to "SYSTEM" and enter it.
4. Enter "DATE/TIME" and set the date and time. The recorder uses military time. Press "OK".
5. Return to the menu and enter "REC".
6. Change "REC FORMAT" to "**WAV48kHz/24bit**". Exit out of the menu completely by repeatedly pressing the menu button.
7. Change the recording level to **50** using the "REC LEVEL" rocker on the right side of the recorder.
8. Set microphones to **120°**.
9. Initial set-up is complete. You may turn off the device.

## Loading the SD Card

1. Be sure the power is OFF when inserting or removing the SD card to avoid destroying data.
2. Insert the SD card into the slot on the left panel of the device
  - a. If "Format Card" appears on the display screen after inserting the card, it means that the SD card has not been formatted in the H4n Pro device. To format it, use the dial to select "YES".
3. To check the remaining capacity of the SD card, press "MENU" and select "SD CARD". Select "REMAIN" which will then display the remaining capacity meter, remaining space, and remaining recording time using the current settings.

## Recording Instructions

1. The recorder should always have functioning batteries installed, regardless of whether an AC adapter is being used.

2. Plug recorder into the AC adapter in the testing room (or, if testing elsewhere, have an AC adapter with you and try to arrange the testing location so you can plug in the recorder).
3. Turn the power on by moving the power switch on the left panel of the device to “ON”
4. Be sure the “Stereo Mode” indicator is lit.
5. Put the recorder in “Recording Standby Mode” by pressing the “REC” button.
  - a. Recording standby means the mic is on but is not yet recording.
  - b. The red light on the recorder blinks when in standby mode.
6. Confirm all settings are correct (recording level = 50; recording format = WAV48kHz/24bit; microphones are at 120°). See Initial Set-up section above for instructions.
7. Make sure that the MIC button is pushed on the front of the recorder (NOT the “1” or “2” buttons); see the [ZOOM H4N DVR Image](#) below.
8. Start recording by pressing the Play/Pause [▶/||] button. The time counter on the screen will advance, the recording symbol [●] will appear next to it, and the red light will stop blinking and remain on.
9. Record the following information: [Participant ID](#), [Date of testing](#), and [Examiner ID](#).
10. Press the Play/Pause [▶/||] button again to pause recording until ready to start recording the participant.
11. Lay the recorder down with the head of the recorder pointed directly at where the participant will be seated. In testing rooms, place the recorder on the file cabinet *next to* the testing table, as close to the participant as possible. It’s absolutely essential to place your paper holder on the opposite side of the table relative to the recorder, because the recorder is very sensitive and paper shuffling will muddle audio. If you are not in a testing room, try to arrange to place the recorder on a different surface, but still close to the participant, so it does not pick up all the paper shuffling, table jarring, etc.
12. After the participant has been consented and has signed the consent form, you may begin recording the examination. For our NP studies at FHS, however, we have IRB approval to audio record the consent process itself. In this case, first tell the participant, “We will be audio recording this session for analysis and quality control purposes” (or something along those lines), then begin recording. NOTE: If the participant reports that they do not want to be audio recorded, turn off the recorder, remove it from the table, and proceed with consenting/testing (unrecorded).

13. Since you are currently in “Standby” mode, press the Play [▶/||] button. Again, the time counter on the screen will advance, the recording symbol [●] will appear next to it, and the red light will stop blinking and remain on. **MAKE SURE THIS IS ALL HAPPENING BEFORE PROCEEDING WITH TESTING.**

14. Optionally, you may now slide the power switch toward “HOLD” on the left panel of the device to disable button operation during recording (although preferably you will not be touching the recorder at any time during testing, so this should not be necessary).

15. After all testing is complete, stop the recording by pressing the stop button [■].

16. **YOU MUST TURN OFF THE RECORDER BEFORE UNPLUGGING THE A/C ADAPTER OR ELSE THE RECORDING MAY BE LOST.** (This is only true if the batteries in your recorder are dead, but you should always follow this procedure to ensure data is not lost.

Although you are unlikely to need to play the recording back on the DVR device itself, because you will be using the ELAN software, this can be done by pressing the, [▶/||] button to play and the, [■] button to stop.

To play an older recording back, press “MENU” then select “FILE” using the dial. Select the file to play and press. Select “SELECT” and press. Press the [▶/||] button to start playback.

## Using USB to Transfer Files

1. Connect device to computer with USB cable.
2. Press the “MENU” button on the right panel of the device.
3. Select “USB” using the dial and press.
4. Select “STORAGE” and press.
5. The device is now connected to the computer and the files can be transferred
6. Save the file in the appropriate file location with the file naming convention for unedited recordings

## Dividing or Deleting a File

It is unlikely you will need to use these features; however, in the rare case that it may be necessary (e.g., two participants were accidentally recording in the same file), follow these directions:

1. Press the “MENU” button on the right panel of the device.
2. Select “FOLDER” using the dial and press.

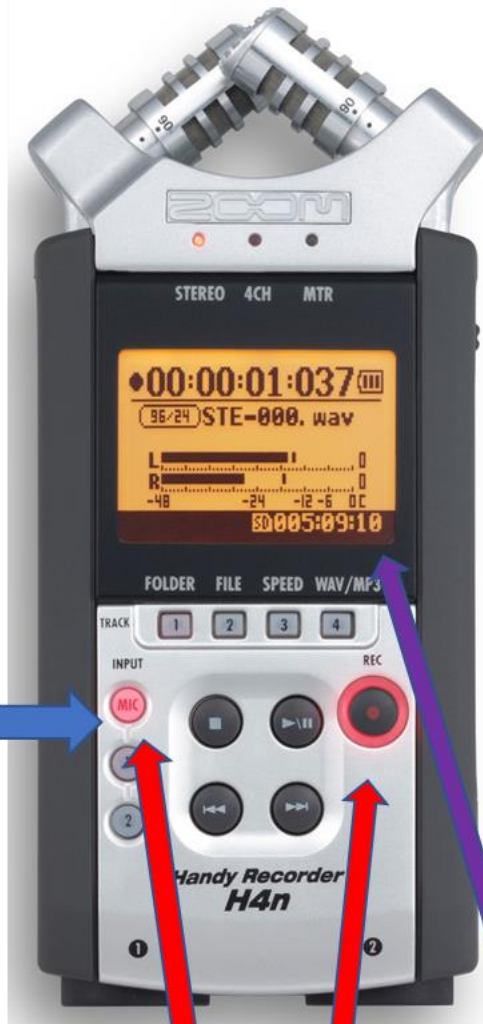
3. Select a folder using the dial and press.
  - a. To divide a file and a desired position, select “DIVIDE” and press. Press to start the playback and press again at the division point. Select “YES” to confirm the divide.
  - b. To delete a file, select “DELETE” using the dial and press. Select “YES” to confirm deleting. **\*\*Never delete files from the recorders until you are 100% certain they are correctly stored on the N drive\*\***

## Battery Type

1. To display the remaining battery life when using batteries, press “MENU”
2. Select “SYSTEM” using the dial and press.
3. Select “BATTERY” using the dial and press.
4. Select the battery type: Alkaline or Ni-MH.

## Software Update

1. To download the most recent system software, the device with an SD card must be connected to a computer with access to the internet.
2. Open the ZOOM website (<http://www.zoom.co.jp>)
3. Connect the H4n Pro to the computer with the USB cable
4. Copy the downloaded software to the root directory of the SD card
5. Disconnect the H4n Pro
6. Turn it on while holding down the [▶/||] button. Select “OK” when prompted to upgrade the version.



Check to be sure that for Input the “MIC” button is pressed!!

If either the “1” or “2” button is pressed, the recorder thinks there is an external mic, and it will not record anything.

A visual check to be sure that the recorder is actively recording:

1. MIC and REC buttons are both on and solid (NOT blinking)
2. The time counter is running

## Appendix C. Digital Voice Recorder Alternatives

This list was compiled by FHS-BAP and was last updated in June 2020. It can serve as a resource for centers that are “shopping around” to find a recording device that best suits them. Please note that some details may become outdated over time as device specifications and models change.

### Sony ICD-PX370 Mono Digital Voice Recorder with Built-In USB Voice Recorder

- Record in MP3 audio quickly and easily.
- MP3 files are compressed but require less memory, making them better for recording long lectures or meetings.
- Record up to 57 hrs of audio (MP3 128 kbps) with exceptional battery life that makes it possible to record for long periods of time.
- Transferring files to or from your computer is fast and convenient. Just plug the ICD-PX370 straight into a free USB port for an immediate connection—no USB cable needed.
- Turn on Auto Voice Recording and the ICD-PX370 will optimize audio capture settings for vocal frequencies. The result is a purer recording with reduced background noise. And when you listen back to the recording, Clear Voice technology cleans up the signal even more for improved clarity.
- The 4GB2 memory stores up to 59hrs35m of recording (MP3 128kbps stereo).
- Choose from four 'scene' presets (music, meeting, interview, dictation) to optimize the audio settings.
- User Manual [PDF]

### 32GB Digital Voice Recorder, Homder Voice Activated Recorder

- Dynamic noise reduction chip & dual microphones to capture sound clearly, gives you a really clear and natural audio.
- All recordings are named with a timestamp, convenient to find the file you are looking for.
- Built-in 32gb flash memory stores up to 2,000+ hours of maximum recording time
- utilizes DSP digital & AGC noise reduction technology to enhance human speech recordings and filter out background noise, to give a really full clear and warmer vocal recording.
- A single full charge (about 4 hrs) could be continuously used 60+ hrs.
- Multiple high-fidelity speakers ensure a crispy & loud enough playback even without headphones.



- Password function keeps your files far away from leaking.

#### EVISTR 16GB Digital Voice Recorder Voice Activated Recorder with Playback

- Voice Activated Record
- Reduce blank and whispering snippet
- Voice Recorder USB Rechargeable
- File name with Year, Month, Day, Hour, Seconds
- Dynamic noise cancellation microphone, capture 1536kpbs crystal clear audio
- Voice Recorder MAC Compatible (WIN Compatible)
- Easy to figure out, press REC: starts to record; press STOP, save the recordings safely. Small Voice Recorders with A-B repeat, fast forward, rewind function during playback, a helpful recorder for lectures, meetings, interviews, speeches, class
- Voice Activated Recorder: set the AVR voice activated function, record only when the teacher is speaking, reduce blank and whispering snippets, save space and time. Recording your appointment, meetings, interviews, speeches, lectures easily.
- Easy File Management: recordings with time stamp, easy to find out when you recorded, what it recorded.
- #1 best-seller on amazon
- User Manual [PDF]
- **Does not have Autosave feature**
  - Do not shut down the device until you press STOP to confirm the file saved properly, it will show "Saved!"
  - Do not shut down the device, while you are formatting, wait until it shows "format completed"

#### 16GB Digital Voice Activated Recorder - aiworth 1160 Hours Sound Audio Recorder Dictaphone

- E36 voice recorder equipped with dual sensitive microphone and professional recording IC, support up to 1536Kbps PCM recording, provide a super clear recorded voice
- Built-in 800mAh rechargeable battery, support up to 45 hours continuous recording. 16Gb flash memory could save 1160 hours recording files at most,
- in addition to this can support up to 32GB TF card (In addition to purchase) expansion and voice activated recording.
- The most user friendly voice recorder designed by aiworth, all operation buttons on the front side, operational logic like smart phone.

- With graphic user guide
- Lifetime software update
- Power-on password protection- 3-digit password,8000 combinations; without the password, no one could turn on the device and overheard your recorded files.After three trial and errors, device will auto turn off.
- 16 levels to adjust the play speed; play faster, jump to the point you exactly want to playback;play slowly let you hear every single word clearly.
- User Manual [PDF]

#### Aomago 8GB Audio Recorder Mini Portable Tape Dictaphone with Playback, USB, MP3

- This recorder upgraded its higher sensitive microphones, meaning that you can enjoy premium quality sound.
- Simple three-click recording, saving and playing, make it super user friendly.
- Set the recorder to voice activated recording, catch the speaking words only.
- A-B REPEAT FUNCTION: This is a great feature to help you study language, review lessons from selected starting point A to ending point B. You don't have to go back or forward to listen to the words any more.
- Easy transfer files: Voice recorder mac compatible. It supports recording files in MP3 or WAV format. You can transfer files easily by connecting to a computer via supplied Micro USB cable.
- 8GB MEMORY CAPACITY
- High Quality (128 kbps) : 7680 Mins
- Short Play (64 kbps) : 16920 Mins
- Long Play (32 kbps) : 33120 Mins
- 7 EQ modes
- Different languages
- USB connection, for uploads and downloads
- Battery life expectancy: up to 12 hours continuous recording
- Warning:
- Do not use the right "POWER" button to totally shut down your voice recorder, or it will reset your voice recorder system time to default.

- We suggest you press the PLAY/PAUSE button for two seconds to power off your voice recorder, and next time you just need to press “PLAY/PAUSE” again to wake up your voice recorder.
- When Battery is almost exhausted or too weak, functions may be limited, please recharge!
- Charge time between 3 to 4 hours, turn on the voice recorder before charging.
- Press the REC button while recording to pause or resume recording. The LED will flash when the recording is paused.
- User Guide [PDF]

#### Wohlman. Digital Voice Recorder 16GB 1536kbps Touch Screen High Recording Quality Noise Reduction Easy Operation Auto Activation MP3 Voice Recorder

- Clear recording with a resolution of 1536 Kbps and microphones with dynamic noise reduction, higher bit rate, higher recording quality, crystal-clear recordings and MP3 player.
- The built-in 180mAh battery can record 12 hours continuously. With 16 GB of internal storage, you can save up to 145 hours of recordings or 1500 songs.
- Automatic recording is possible with the preset time. Simply record and save with the "REC / Save" button. Simple operations with touch buttons.
- With the automatic voice recognition function, the recorder automatically starts recording when the sound is recognized. Without sound, it will be in standby mode to reduce recording capacity and power consumption. The detecting distance can reach up to 10m.
- With the A-B repeat play function, the recording can play back within a certain period of time. You can also fast forward and rewind during playback, which is useful for reviewing lessons, meeting records, songs, interviews, etc. We offer a one year guarantee.
- With the USB cable, you can easily transfer the files to the computer as well as delete the files directly. Compatible with Windows and IOS systems.
- The password setting secures your recording data
- Tschisen V93 is embedded with AGC noise reduction design and will give you high quality recordings.
- Speech recognition automatically picks up detected sounds and stops recording when it is quiet.

#### Olympus Voice Recorder WS-853 with 8GB, Voice Balancer, True Stereo Mic

- High quality MP3 recording
- USB Direct connect with battery charge function

- 8 gb internal memory
- Micro SD card slot
- Playback speed control 0.5X to 2.0X
- The True Stereo Mic with two directional microphones positioned at a 90 degree layout, enables highest quality recording with an authentic stereo experience.
- By differentiating the position and the distance of the speakers in meetings and conferences the recording is highly precise, letting you feel as if you are actually in the recording scene.
- Auto Mode function makes it easier for users by automatically adjusting the microphone sensitivity according to the volume of the speaker. To set this function, simply select 'Auto' for the recording level from the menu.
- The Simple Mode supports beginners by having the recorder display only the necessary information in large font. It also limits the functions in the menu to those which are frequently used.
- For advanced users, the Normal Mode with full functionality is recommended.
- The WS-853 can connect directly to a computer via the built-in USB connector. This makes it possible to easily save data anytime, anywhere without the need to bring along a USB cable. Furthermore, WS-853 is equipped with a protective cover to keep dust out of the connectors.
- The built in stand placed on the back of the body is carefully designed to reduce the noise from the surface when the recorder is placed on a table. It works much like a kickstand and allows users to read the menu without having to look down at the recorder.
- When recordings contain multiple speakers, the Voice Balancer makes smaller voices louder and ensures that louder voices stay below a given level, providing playback where everyone can be heard clearly. This comes in handy when recording sound sources from multiple positions, such as at a meeting. The prominent noise produced when amplifying small sounds is reduced. By eliminating the lower and higher frequency, the voice is even more enhanced.
- The noise-cancellation function powerfully reduces unwanted ambient noise such as air-conditioner noise or projector fan noise enabling clear playback quality. The function is very effective when playing back meeting recordings.
- NYTIMES #2 pick
- User Guide [PDF]

Sony ICDUX560BLK Digital Voice Recorder 1" Black

- NYtime #1 pick for voice recorders

- Built in stereo microphone and voice operated recording
- Three recording options: wide/stereo, narrow/focus and normal
- Quick charge; up to 1 hour recording time, with 3 minute charge
- Easy to use user interface and recording level indicator
- Micro SD memory card slot, headphone jack & mic input. LCD backlight
- Record in MP3/LPCM with a high-sensitivity S-Microphone
- Up to 4 GB of built-in storage, expandable via MicroSD (SDHC/SDXC) cards
- Focus and wide microphone modes to suit lectures or meetings
- Direct USB built-in for easy connection to PC
- FM radio to listen to or record radio broadcasts
- Normal, focus, and wide-stereo recording provide you with the opportunity to record the audio that you need to capture in any environment, while the slim and lightweight build make it easy to take with you wherever you go and the easy to use up makes file searching simple.
- UX560 received the highest overall ratings from our panel of test listeners (nytimes). It produces clear, understandable audio in the classroom, quiet office, and noisy coffee shop settings. It also offers a better collection of features than the other models we tested, with an easy-to-navigate menu system, a bright backlit screen, 39 hours of recording time (in MP3 format), 27-hour battery life, voice-activated recording to pause and restart after silences, and a pop-out USB 3.0 connector that lets you recharge the recorder and transfer files to a computer easily. Like many of the other recorders we looked at, it comes with an adequate amount of onboard storage (4 GB) but accepts microSD cards, so you can record and store hundreds of hours of recorded audio should you need it. The UX560 is also the slimmest recorder we tested—at 0.43 inch thick it can easily fit in a shirt or pants pocket.
- User Manual [PDF]

#### SONY PCM-D10

- Reliable hi-res recordings of up to 192kHz/24-bit
- 3-way adjustable high-resolution 40K frequency response microphones
- 2 XLR-TRS combo jacks with 48V phantom power
- Digital dual-path limiter function prevents distortion
- Bluetooth capability for both remote control and playback via Sony's free REC Remote app

- More expensive than most (\$500)
- Commonly used for podcasters, radio, amateur film makers
- Capture flawless Hi-Res sound anytime, anywhere with the pcm-d10 portable recorder. Record professional sound with Hi-Res Audio at up to 192kHz/24-bit. Whether it's your live music set, new podcast episode or breaking news report, the pcm-d10 unlocks a new level of detail and texture. The three-way adjustable microphones adapt to your situation, while the twin XLR/TRS combo jack lets you plug in your choice of input. High-quality dual ADCs maximize S/N and independent analog volume dials give you precise control of your inputs.

#### Tascam DR-05 recorders

- The dual internal condenser microphones can handle anything from subtle to loud, with sensitivity to capture every detail
- A revamped layout means operations like recording, adjusting levels, deleting bad takes and adding Markers are quick and easy
- Uses only two AA batteries, but can record for an outstanding 17.5 hours; It can also be powered by a USB mobile battery
- Connect to a PC using USB Audio Interface Mode for voiceover work, live streaming, podcasting and songwriting with studio-quality audio
- Used in a study investigating pitch modulation in human mate choice
  - In this study the researchers used a sampling rate of 96 kHz and 24-bit amplitude quantization. Recordings were stored onto microSDHC media cards as uncompressed WAV files and later transferred to a laptop computer for editing and analysis. This method allowed us to obtain high-quality, directional voice recordings that would otherwise be difficult to obtain in a noisy environment using a stationary microphone.
  - Acoustic editing and analysis were performed in Praat v. 6.0.21 [32]. Fragments of silence, acute noise, non-verbal vocalizations (e.g. laughter) and multi-voicing (e.g. the voice of the dating partner) were first manually removed from audio files. Recordings were then segmented into multiple parts each corresponding to a given participant and a single speed date. We further split each sound file into three equal time segments (beginning, middle and end of the date; mean segment duration  $50.6 \pm 23$  s), resulting in a total of 726 voice clips for acoustic analysis.

## References

1. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (Office for Civil Rights) 7-8 (2012).

2. Best Practices. Google Cloud. 2022. Updated 2/10/2022. 2022.  
<https://cloud.google.com/speech-to-text/docs/best-practices>
3. Introduction to audio encoding. Google Cloud. 2022. Updated 2/10/2022. 2022.  
<https://cloud.google.com/speech-to-text/docs/encoding>