Encoding Instructions

Encoding is the process of watching the videos and recording data from those videos. This data is sometimes open to interpretation, and is thus considered "qualitative". The goal of encoding the videos is to extract data points for further statistical analysis in the next stage of the research.

The Overall Process

Three reviewers will each independently watch and encode all of the study videos. The three reviewers will then come together, and check for any discrepancies between what they recorded. They will then discuss these discrepancies, and will come to an agreement on a final version of the encoding. That final version will be used for statistical analysis. This process is similar in concept to taking multiple measurements of a value with a scientific instrument, and then averaging those measurements, to increase the overall accuracy of the measurement.

Your Part

Watch each of the videos, one at a time. It is ok to (and expected you will need to) pause the videos while watching, and possibly rewatch the same part several times. Accuracy is the most important part of this, so take whatever time you need to ensure you are accurately recording what you observe in the video.

Each video runs continuously across up to 12 trials. Each trial has a reset period, and possibly some discussion between them. Only the data in each of these trials is to be counted. Disregard any data from outside of the trials (including the instructor questions and discussion with the team after a trial finishes). The study instructor will clearly speak to indicate the start and stop of each trial. See the file "VideoTimes" for the video start and stop times of each trial. Record the team number (in the video file name), modality (Desktop or HoloLens), and trail number (refer to the "VideoTimes" to be sure, but with only one exception (where the video is split into two parts) all videos start at trial #1 and move up from there).

Three types of data are being encoded:
1. Utterances
2. Gestures
3. Looks

Utterances

Watch the videos, and transcribe the utterances (spoken words and sounds). We are trying to capture four things:

1. Purpose of the speaking (**purpose**)
2. When the speaking happens (**time**)
3. Who is speaking (**speaker**)
4. Use of deictic expressions (**deictic expressions**)

Observe the participant's starting positions in the first trial in the first video. The participant closer to the camera is participant #1. The other participant is participant #2. Each participant keeps the same number for all videos they are in.

Purpose
- Categorize what is being spoken into an "utterance type" according to the chart below.
  - Only utterances spoken by one participant to the other participant should be considered. Utterances spoken by a participant to themself, or by a participant to the instructor should be counted as "other".
  - Do note who is speaking (speaker) and who is being spoken to (target) when counting the "other" utterances.
- For utterances classified as "other", you do not need to create a row and you do not need to record the time. You only need to count them based on the speaker and target. See the second example chart below.

Utterance Type	Description
Referents (R)	References to and attempts to describe a specific mark in the chart. E.g. "The red line", or "I found a node with eight connections"
Position (P)	Attempts to describe the position of marks in the chart. E.g. "The last bar on the left" or "The point under here"
Acknowledgements (A)	Responses to statements, confirming an understanding or a behavior. E.g. "yes" or "no" or "mmm-hmm" or "ok" or "there you go".
Viewport (V)	Requests to or descriptions of how to adjust the view, or for the other participant to do so.
Other	Anything either not spoken from one participant to the other, or anything spoken from one participant to the other that does not fall into one of the other types.

Time
- Record the exact time in the video the speaking starts, in minutes and seconds (mm:ss)
- Start a new row for each utterance. An utterance is the expression of a complete thought. Use your best judgment in determining how to divide up the different utterances; it may be a full sentence, a few sentences, or just part of a sentence. Long pauses or changing subjects are both clues a new utterance may have begun.

Speaker
- Record the person doing the speaking.
- Each utterance should only have a single speaker. If the speaker changes, a new utterance has started.
- You must record one of the following possible values:
  - 1 - the participant you assigned #1
  - 2 - the participant you assigned #2

Deictic Expressions
- Count the number of each type of deictic expressions used in the utterance. Each type should have its own count (column).
- A "deictic expression" is a word that depends on information not contained in the sentence ("depends on the context in which it is used").
- There are two types of deictic expressions to count:
  - Deictic Pronouns - Utterances that use the deictic pronouns "this," "that," "there,", "it", and related terms to refer to specific objects (physical or intangible objects).
  - Spatial Deictic - Utterances that refer to terms using spatial position, such as "above," "below," "in front of," "on top of," "next to," "behind," "right," "left," "up," "down," "touching", "closest to", "furthest from"
- When deciding whether to count a deictic pronoun, consider whether the object the pronoun refers to is mentioned in the sentence by name. **Do not count a pronoun if it refers to something named in the same sentence.** Some words also may or may not be deictic depending on the usage. See the examples below.

Record all of this by making two charts like the ones shown on the next page. **The second chart should be a total count - even if a utterance is recorded in the first chart, still count it in the second chart.**

Chart Example

Time	Purpose	Speaker	Deictic Pronouns	Spatial Deictic
0:13	R	1	2	0
0:17	P	1	0	0
0:18	P	2	1	3
0:23	A	1	1	1
0:25	A	2	0	1

Chart Example - Count Utterance Types

Condition	Count
P1 speaking to self	4
P2 speaking to self	7
P1 speaking to P2	15
P2 speaking to P1	21
Either person speaking to instructor	2

Counting Deictic Expression Examples

When deciding whether to count a deictic pronoun, consider whether the object the pronoun refers to is mentioned in the sentence by name.

- **Do not count a pronoun if it refers to something named in the same sentence.**
  Example: "Can you rotate the chart so that it faces this way?".
  - "It" is **not** counted, because "it" refers to "chart", which is mentioned in the sentence. If "chart" was replaced with the word "it", so that "it" appears in the sentence twice, then both instances of "it" would be counted, since neither would then refer to something (a noun) named in the sentence.
  - "This" **is** counted, because the phrase "this way" is specifying a direction. The direction is not indicated in the sentence, so it is information that must be gained outside of the sentence (such as through a hand gesture).
  - "That" is **not** counted, because it is not being used to refer to an object in this sentence, rather it is used only as an expression of speech. An example of where "that" would be counted, is the sentence "can you look at that one". In that case, "that" is referring to a specific object not specified in the sentence.
- Some pronouns, such as "that", can make a phrase deictic. Examples:
  - "That bar is green." The use of the word "that" implies there are multiple green bars, and *which* green bar needs to be specified: "Which one?" "*That* one!". This **is** a deictic expression.
  - "The bar is green." The use of the word "the" implies there is only a single green bar. There may be multiple bars, but only one is a green bar: "which bar?" "the *green* one!" This is **not** a deictic expression**.**
  - A helpful way to distinguish between the two, is to consider whether you need to point at the same time as saying the words for someone to know which bar you are talking about. You would probably need to point when saying "*That* bar is green", but you probably would not need to point (nor even look up) when saying "*The* bar is green." If you need to point, it is a deictic expression.

Gestures

Watch the videos, and record the gestures. Gestures are nonverbal communication, such as pointing or hand-waving or head nodding. We are trying to capture five things:

1. When the gesture happens (**time**)
2. What gesture is happening (**action**)
3. Who is gesturing (**gesturer**)
4. Who or what is being gestured towards/at (**target**)
5. What the gesture is being used to convey (**intent**)

Time
- Record the exact time in the video the gesture starts at, in minutes and seconds (mm:ss)
- Start a new row if a different person gestures, no matter how soon after the previous person gestures. Each row should only have a single gesture from one person.
- Start a new row for each new gesture.
- Continuous motion should be counted as a single gesture. If the person pauses gesturing before resuming, count that as two separate gestures.Similarly, if the person clearly breaks the gesturing (such as putting their hand down in between pointing at two different things) that is a pause between gestures.

Action
- Record the gesture the person made. Describe the gesture. "Pointed" or "hand wave" or "pull" or "grab" are common examples. Describe the gesture in more detail if needed to clearly communicate what the gesture was.
- Do not count movements not intended for communication, such as reaching up to adjust a face mask or headset, or scratch one's head.

Gesturer
- Record the person making the gesture.
- You must record one of the following possible values:
    - 1 - the participant you assigned #1
    - 2 - the participant you assigned #2

Target
- Record the person or thing being gestured at/towards.
- You must record one of the following possible values:
    - Self (S) - the participant is gesturing at themself
    - Other (O) - the participant is gesturing to their teammate
    - Chart (C) - the participant is gesturing towards the chart*
    - Background (B) - anything else not included in the other possible values listed here. Add a description of what the participant is gesturing towards.

Intent
- Record the intent of the gesture based on what the participant making the gesture is trying to communicate. Record as one of the following values:
    - Reference (R) - a gesture intended to indicate a specific part of the chart, or the chart as a whole
    - Description (D) - describe a shape, position, direction, size, trend, pattern, or related property
    - View (V) - request or describe an instruction to a teammate to change their viewport (point of view)
    - Conversational (C) - hand gestures used as part of the conversation, that do not fall into the above categories
- Use the participant's speaking, body language, and any other context clues to determine their intent with the gestures. Make a best guess as to the intent based on any available information in the video. Only record the intent of the gesture, however, even if the spoken words would fall into a different category. For example:
    - If the participant points at a mark and says "the round one", the gesture would be categorized as "Reference". The participant is using words to describe the shape, but the gesture itself is indicating a specific part of the chart.
    - By contrast: If the participant cups their hands and says "the round one", the gesture would be categorized as "Description". The participant is using their hands to visualize their description of the mark, and is not pointing at or motioning towards a specific mark.
- A participant using a hand to trace a trend line, or outline a shape on the chart itself, should be counted as a "Description" gesture, even though they are directly using the chart itself for the context of their gesture.

Gestures Chart Example

Time	Action	Gesturer	Target	Intent
0:11	Point	2	Chart	Reference
1:24	Hand wave	1	Other	Conversational
2:17	Two hands to show size	1	Background	Description

Looks

Watch the videos, and record the looks. Looks are when a participant makes an intentional effort to focus their gaze on the face of their teammate. We are trying to capture three things:

1. When the look happens (**time**)
2. Who is looking (**initiator**)

Time
- Record the exact time in the video the look starts at, in minutes and seconds (mm:ss)
- Start a new row if a different person looks, no matter how soon after the previous person looks. Each row should only have a single look from one person.
- Start a new row for each new look.
- Continuous looking should be counted as a single look. If the person stops looking to return their focus to something else (such as the chart) before resuming looking at their teammate, count that as two separate looks.
- If both teammates look at each other at the same time, record two separate looks, one for each teammate at the initiator.

Initiator
- Record the person looking at their teammate.
- You must record one of the following possible values:
  - 1 - the participant you assigned #1
  - 2 - the participant you assigned #2

Looks Chart Example

Time	Initiator
0:34	1
0:53	2
1:19	2
7:35	1