Three variables need to be coded by a second rater in order to check for inter-rater reliability:

1. Deepfake\_video\_check: This indicates whether participants were aware that the video they were watching in the experiment was a Deepfake or not.
2. Deepfake\_concept\_check: This indicates whether they were aware of the *concept* of Deepfakes prior to taking part in the study.
3. Recognised\_video\_manipulation: This is a third variable I was interested in. I wanted to know if people recognized that there was something “off” about the video/audio in the sense that it had been edited or manipulated in some way.

---------------

For the first variable (Deepfake\_video\_check) I coded an individual as “yes” if they did the following:

* If their response contained “Yes”
* If their response indicated that they recognized that it was a Deepfake to some extent (i.e., they said something to this effect; e.g., “I did”, “I recognized”, “I had my suspicions”)
* If their response indicated that the content was computer or AI generated or was an artificial system such as a robot.

I coded an individual as “no” if they did not meet the above criteria

---------------

For the second variable (Deepfake\_concept\_check) I coded someone as “yes” if they did the following:

* Indicated “Yes” in their response
* If their response indicated that they were aware of the concept of Deepfakes prior to taking part in the study (e.g., “I was aware”, “I heard about it”, “I knew”, “To a certain extent”)
* If they had not heard about one type of Deepfake (e.g., audio) but had heard about another (e.g., video)

I coded an individual as “no” if they did not meet the above criteria

---------------

For the third variable (Recognised\_video\_manipulation) I coded someone as “yes” if they did the following:

* Indicated that the video had been edited, contained evidence of manipulation (e.g., “there was something off about the video/audio”, “lighting was off and changed”, “lipsynth was off”). Anything that indicated that the person recognised that the content was altered, was not normal, had been tampered with, or contained video/audio artefacts.
* If the participant indicated that the video was a Deepfake

I coded an individual as “no” if they did not meet the above criteria