Data and Methods + Table

II. Data Description

Design/Materials:

This is a 2x2 factorial design, where the first factor is the developer of the media at 2 levels (Human, AI) and the second factor is the type of media at 2 levels (Audio Clip, Image).

The experimental unit is 1 full-time Cal Poly student. Part-time students and students from other campuses are excluded from the experiment, so that information collected about a person's major and year stay consistent.

A completely randomized design structure is used, where participants are randomly assigned one of two types of media - images or audio clips.

AI-generated audio clips were generated from the resemble.ai platform. An original audio clip was extracted as a .mp4 file and passed through the resemble.ai software, which develops an artificially generated voice-over from the clip. This was done for all audio clips tested so that each original audio clip has an AI-generated version.

Images were sourced from the Which Face is Real project, a website that generates images and asks the user to identify which image is real or developed by AI.

During the survey, all materials were shown to each participant on my MacBook Air laptop, on maximum brightness and volume, so that the image and audio clips could be displayed as clearly as possible.

Methods:

56 full-time Cal Poly students were assigned to the audio or image treatment and asked to distinguish between real and AI duplicates.

Participants were gathered through a convenience sample, which was taken on Dexter Lawn from April 20, 2023 to April 21, 2023, between the hours of 12-6 pm PDT. Dexter Lawn was chosen because the area serves as a common meeting spot for students across all majors; it is usually crowded during the afternoon when vendors and clubs set up their booths.

Students were first asked if they were enrolled full-time at Cal Poly to confirm that they were eligible to participate. Then, participants were asked to click an online spin-the-wheel that I customized with labels of "Yes" and "No". If the spinner landed on "Yes", the student was assigned to the image treatment; if the spinner landed on "No", the student was assigned to the audio treatment. Based on the response, I presented one of two slideshows containing either the image or audio clips.

For participants assigned to images:

> Each slide has an image pair with the caption "Select the person who is real." Participants were given 15 seconds to look at each image pair and point to the person they thought was real. The following slide was then revealed, showing which person is real by highlighting their profile in green.
>
> The number of correct responses was tracked throughout the process, until all image pairs and the correct answers were revealed. The last slide credits the Which Face Is Real project, and provides a link to their website if participants were interested in learning more.

For participants assigned to audio clips:

> Each slide is embedded with one audio clip, which plays upon being clicked on. After each clip was played, the participant was given 15 seconds to decide if it was real or AI-generated. Once their answer was recorded, they were shown the next slide, which read "FAKE" in red if the clip was AI-generated or "REAL" in green if the clip was real.
>
> The number of correct responses was tracked throughout the process, until all audio clips and the correct answers were revealed. The last slide gives context about where the audio was sourced from, as well as link to the resemble.ai platform for more information.

Table:

The mean number of responses answered correctly, as well as the standard deviation, were computed for both the treatment groups as shown below.

Table X. Means and Standard Deviations of the Number of Correct Responses for Audio and Image Treatment Groups.

Treatment	Mean (Correct Responses)	Standard Deviation	N
Audio	3.71	1.12	24
Image	3.22	0.87	32