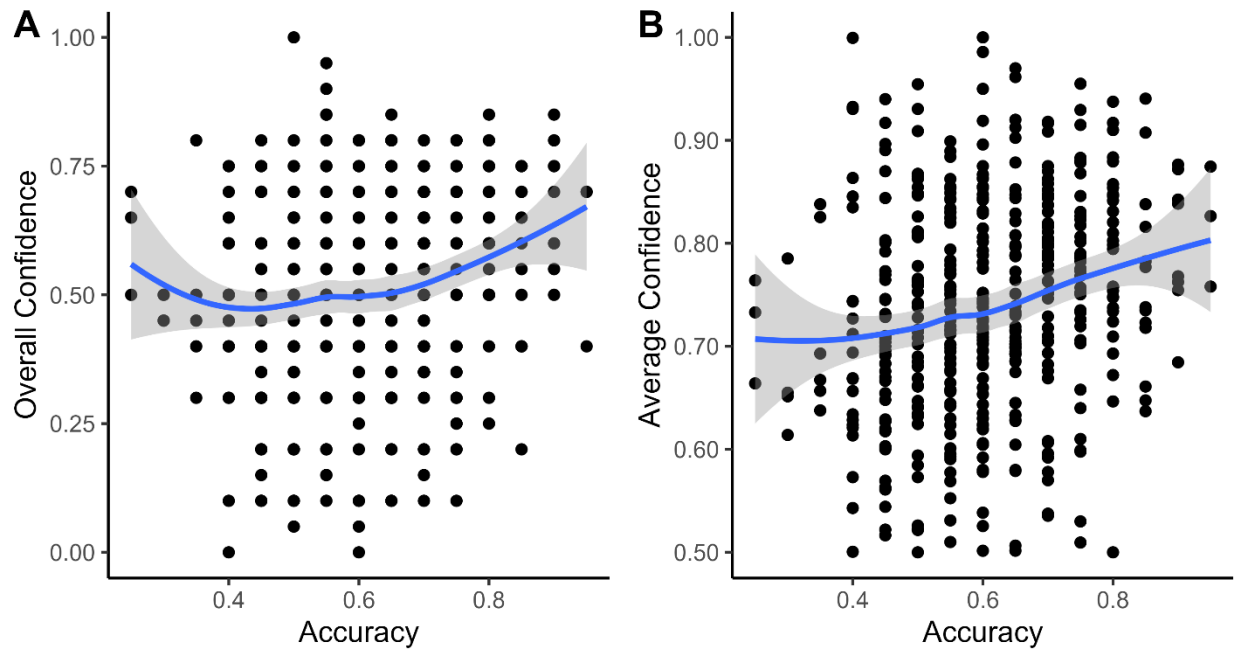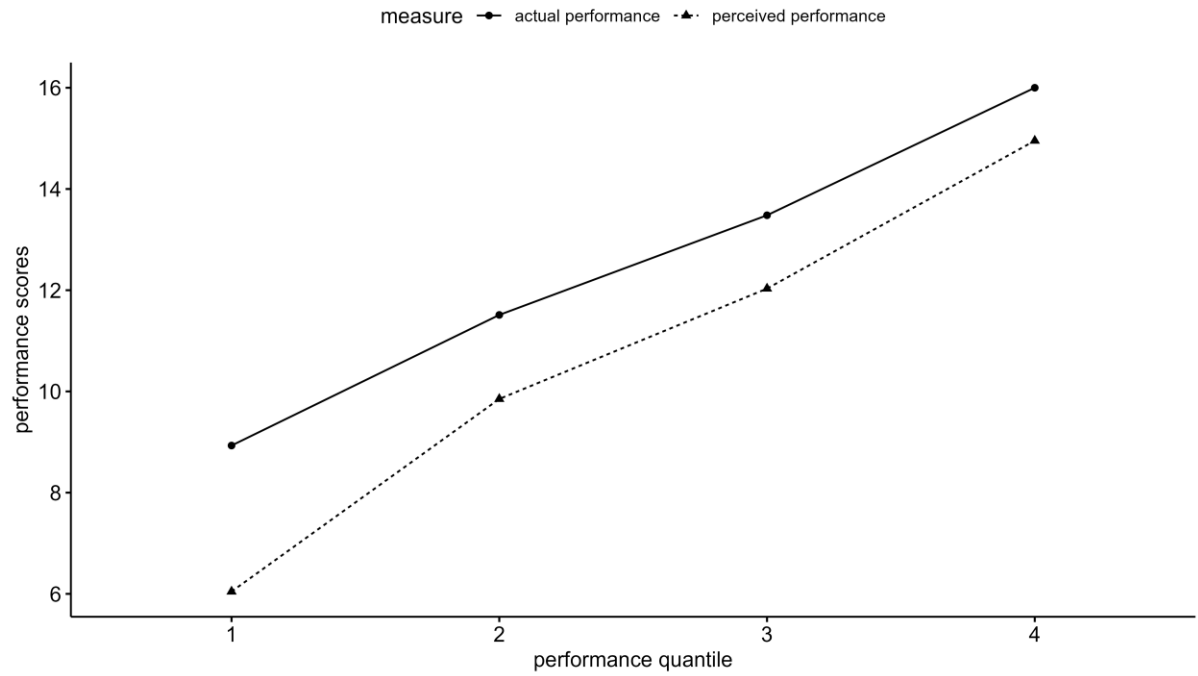**Supplementary Materials**

Figure S1

Relationship between Accuracy on the Detection Task and Overall Detection Confidence (A) and Average Detection Confidence (B)
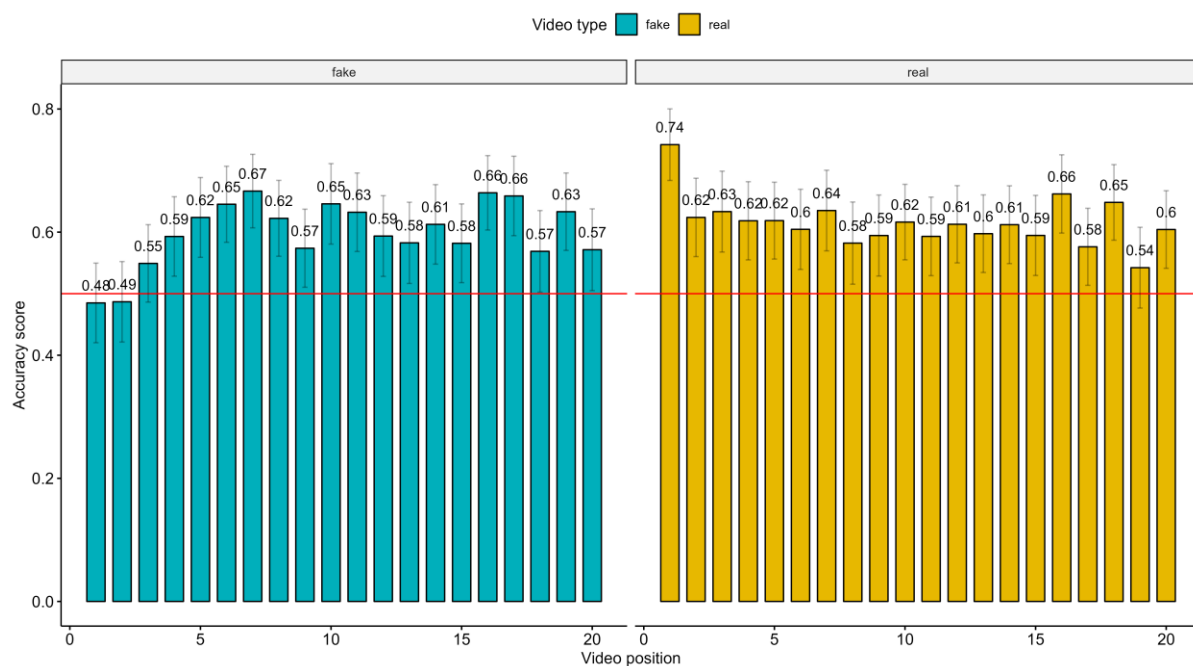

Note. A loess line is depicted in blue. Created with the ggplot2 package.

Figure S2

Relationship between Actual Detection Performance (Accuracy) and Perceived Performance (Overall Detection Confidence)
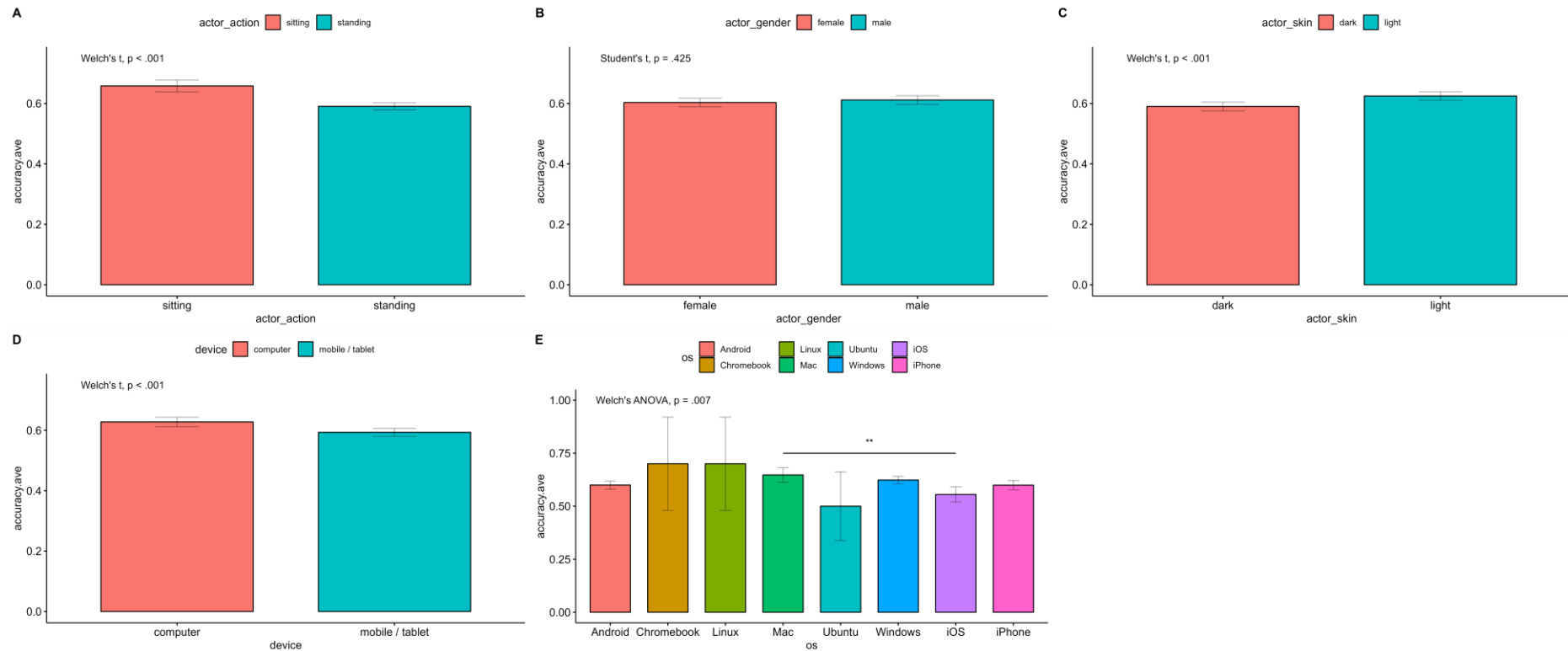

Note. Created with the ggpubr and ggplot package.

Figure S3

Accuracy by Video Position

Note. The numbers presented in the plots are the average accuracy scores for deepfake (blue) and real videos (yellow) by video position. Error bars represent 95% confidence intervals. Created with the ggpubr and ggplot package.

Figure S4

Accuracy scores by Actor/Actress' Action, Gender, Skin Color, Device Used, and Operating System of Devices Used

Note. The numbers presented in the plots are the average accuracy scores by the actor/actress' action (A), gender (B), skin color (C), device used (D) and the operating system of the devices used by participants (E). Error bars represent 95% confidence intervals. Welch's test was used when the Levene's test of homogeneity was violated. Created with the ggpubr and ggplot package.