On Using Social Signals to Enable Flexible Error-Aware HRI: Supplementary Materials

Maia Stiber
mstiber@jhu.edu
Johns Hopkins University
Baltimore, Maryland, USA

Russell H. Taylor
rht@jhu.edu
Johns Hopkins University
Baltimore, Maryland, USA

Chien-Ming Huang
chienming.huang@jhu.edu
Johns Hopkins University
Baltimore, Maryland, USA

A SIGNIFICANT AU PREDICTOR WELCH'S T-TESTS

In order to determine the *Significant AU Predictors*, we evaluated the data set using Welch's t-tests. Table 1 displays the numerical values of the Welch's t-tests ran to evaluate the *Significant AU Predictors* metric for the complete dataset, as described in Section 4.4.1. Table 2 displays the numerical values of the Welch's t-tests ran to evaluate the *Significant AU Predictors* metric for the HRC-A data collection scenario. Table 3 displays the numerical values of the Welch's t-tests ran to evaluate the *Significant AU Predictors* metric for the HRC-C data collection scenario. Table 4 displays the numerical values of the Welch's t-tests ran to evaluate the *Significant AU Predictors* metric for the PbD data collection scenario.

Table 1: Welch's t-test values used to evaluate the *Significant AU Predictors* metric over the entire dataset. The t-tests were run on the AU intensities exhibited during error and no error timesteps. The bolded AUs are the statistically significant ones.

	Complete Dataset		
	Error Timestep AU Mean	No Error Timestep AU Mean	Welch's t-test
AU_01	0.41 ± 0.66	0.58 ± 0.93	t(1087.40) = -1.22, p = .22
AU_02	0.46 ± 0.81	0.59 ± 0.95	t(2721.40) = -7.67, p <.001
AU_04	0.45 ± 0.61	0.40 ± 0.58	t(2642.55) = 4.32, p <.001
AU_05	0.13 ± 0.33	0.15 ± 0.38	t(2709.39) = -0.75, p = .08
AU_06	0.42 ± 0.59	0.18 ± 0.39	t(2568.20) = 20.22, p <.001
AU_07	0.39 ± 0.58	0.37 ± 0.64	t(2693.89) = 1.99, p = .047
AU_09	0.34 ± 0.52	0.25 ± 0.50	t(2646.65) = 7.57, p <.001
AU_10	0.58 ± 0.71	0.42 ± 0.63	t(2624.82) = 1.57, p <.001
AU_12	0.57 ± 0.71	0.17 ± 0.35	t(2539.51) = 27.27, p <.001
AU_14	0.80 ± 0.78	0.39 ± 0.63	t(2603.79) = 25.82, p <.001
AU_15	1.40 ± 0.89	0.63 ± 1.22	t(2799.72) = -12.91, p <.001
AU_17	0.57 ± 0.85	0.73 ± 1.00	t(2819.04) = -9.09, p <.001
AU_20	0.43 ± 0.67	0.36 ± 0.73	t(2690.10) = 4.94, p <.001
AU_23	0.19 ± 0.50	0.30 ± 0.72	t(2820.94) = -10.30, p <.001
AU_25	0.34 ± 0.65	0.33 ± 0.69	t(2681.68) = 0.48, p = .63
AU_26	0.25 ± 0.48	0.32 ± 0.63	t(2767.09) = -7.16, p <.001
AU_45	0.43 ± 0.55	0.25 ± 0.47	t(2614.21) = 16.24, p <.001

Table 2: Welch's t-test values for the HRC-A scenario to evaluate the *Significant AU Predictors* metric for the open-source dataset analysis. The t-tests were run on the AU intensities exhibited during error and no error timesteps. The bolded AUs are the statistically significant ones.

	HRC-A		
	Error Timestep AU Mean	No Error Timestep AU Mean	Welch's t-test
AU_01	0.89 ± 1.06	1.10 ± 1.20	t(204.86) = -2.81, p = .005
AU_02	1.16 ± 1.04	1.29 ± 1.13	t(204.55) = -1.82, p = .07
AU_04	0.33 ± 0.56	0.45 ± 0.63	t(204.77) = -3.23, p = .002
AU_05	0.28 ± 0.48	0.25 ± 0.49	t(204.23) = 0.83, p = .41
AU_06	0.22 ± 0.37	0.21 ± 0.41	t(204.89) = 0.58, p = .57
AU_07	0.40 ± 0.61	0.45 ± 0.68	t(204.84) = -1.14, p = .25
AU_09	0.66 ± 0.72	0.52 ± 0.67	t(203.57) = 2.77, p = .006
AU_10	0.62 ± 0.74	0.54 ± 0.65	t(203.37) = 1.57, p = .12
AU_12	0.20 ± 0.34	0.16 ± 0.30	t(203.35) = 1.50, p = .14
AU_14	0.35 ± 0.61	0.48 ± 0.69	t(204.86) = -2.83, p = .005
AU_15	1.25 ± 1. 63	0.97 ± 1.45	t(203.38) = 2.39, p = .018
AU_17	1.11 ± 1.25	0.98 ± 1.24	t(203.46) = 1.47, p = .14
AU_20	0.42 ± 0.69	0.53 ± 0.81	t(205.16) = -2.37, p = .012
AU_23	0.73 ± 1.04	0.50 ± 0.95	t(203.53) = 3.06, p = .003
AU_25	0.088 ± 0.29	0.043 ± 0.23	t(202.91) = 2.23, p = .027
AU_26	0.13 ± 0.44	0.17 ± 0.50	t(204.90) = -1.35, p = .18
AU_45	0.19 ± 0.36	0.34 ± 0.56	t(208.25) = -5.61, p <.001

Table 3: Welch's t-test values for the HRC-C scenarios to evaluate the *Significant AU Predictors* metric for the open-source dataset analysis. The t-tests were run on the AU intensities exhibited during error and no error timesteps. The bolded AUs are the statistically significant ones.

	HRC-C		
	Error Timestep AU Mean	No Error Timestep AU Mean	Welch's t-test
AU_01	0.13 ± 0.33	0.17 ± 0.38	t(1088.44) = -3.72, p <.001
AU_02	0.24 ± 0.42	0.17 ± 0.38	t(1066.10) = 4.56, p <.001
AU_04	0.35 ± 0.47	0.30 ± 0.46	t(1073.37) = 3.35, p <.001
AU_05	0.20 ± 0.39	0.24 ± 0.43	t(1084.44) = -3.28, p = .001
AU_06	0.22 ± 0.41	0.072 ± 0.25	t(1045.88) = 11.19, p <.001
AU_07	0.43 ± 0.50	0.32 ± 0.48	t(1072.76) = 7.03, p <.001
AU_09	0.16 ± 0.37	0.063 ± 0.24	t(1048.12) = 8.61, p <.001
AU_10	0.29 ± 0.45	0.12 ± 45	t(1053.53) = 11.76, p <.001
AU_12	0.25 ± 0.43	0.010 ± .43	t(1050.34) = 11.61, p <.001
AU_14	0.59 ± 0.50	0.25 ± .44	t(106.75) = 12.40, p <.001
AU_15	0.23 ± 42	0.25 ± .47	t(1087.40) = -1.22, p = .22
AU_17	0.37 ± 0.48	0.33 ± 0.51	t(1082.23) = 2.83, p = .0048
AU_20	0.24 ± 0.42	0.18 ± 0.39	t(1068.08) = 4.71, p <.001
AU_23	0.22 ± 0.41	0.16 ± 0.37	t(1065.52) = 5.12, p <.001
AU_25	0.34 ± 0.47	0.18 ± 0.41	t(1063.52) = 10.65, p <.001
AU_26	0.24 ± 0.42	0.12 ± 0.34	t(1059.06) = 9.53, p <.001
AU_45	0.30 ± 0.45	0.21 ± 0.40	t(1065.28) = 6.34, p <.001

Table 4: Welch's t-test values for the PbD scenarios to evaluate the *Significant AU Predictors* metric for the open-source dataset analysis. The t-tests were run on the AU intensities exhibited during error and no error timesteps. The bolded AUs are the statistically significant ones.

	PbD		
	Error Timestep AU Mean	No Error Timestep AU Mean	Welch's t-test
AU_01	0.74 ± 0.78	0.98 ± 1.02	t(847.96) = -6.51, p <.001
AU_02	0.93 ± 1.06	1.03 ± 1.12	t(737.24) = -2.28, p = .023
AU_04	0.50 ± 0.67	0.52 ± 0.68	t(721.86) = -0.41, p = .68
AU_05	0.15 ± 0.32	0.21 ± 0.40	t(829.21) = -4.11, p <.001
AU_06	0.47 ± 0.56	0.14 ± 0.30	t(594.00) = 13.38, p <.001
AU_07	0.39 ± 0.49	0.30 ± 0.52	t(739.73) = 3.46, p = .001
AU_09	0.66 ± 0.63	0.41 ± 0.49	t(647.53) = 8.84, p <.001
AU_10	0.47 ± 0.50	0.38 ± 0.57	t(766.56) = 4.05, p <.001
AU_12	0.75 ± 0.66	0.19 ± 0.33	t(586.28) = 19.57, p <.001
AU_14	1.16 ± 0.75	0.50 ± 0.69	t(688.95) = 19.53, p <.001
AU_15	0.28 ± 0.56	0.58 ± 0.91	t(1023.25) = -10.54, p <.001
AU_17	0.30 ± 0.55	0.46 ± 0.81	t(931.05) = -6.03, p <.001
AU_20	0.55 ± 0.58	0.43 ± 0.58	t(721.58) = 4.27, p <.001
AU_23	0.07 ± 0.28	0.12 ± 0.42	t(950.60) = -3.54, p <.001
AU_25	0.15 ± 0.43	0.06 ± 0.21	t(583.40) = 4.95 p <.001
AU_26	0.05 ± 0.21	0.14 ± 0.43	t(1325.30) = -8.04, p <.001
AU_45	0.81 ± 0.67	0.50 ± 0.67	t(696.22) = 10.04, p <.001