Supplemental Material

Bänziger, T., Mortillaro, M., Scherer, K. R. (2011). Introducing the Geneva Multimodal Expression Corpus for experimental research on emotion perception. Manuscript submitted for publication.

The supplemental material is organized in three sections: in Section A, we report the mean ratings of intensity, plausibility and authenticity for the Emotion Expressions (EEs) included in the core-set rating study, organized per emotion; in Section B, we discuss issues related to response bias and chance accuracy and we report several indexes that can be used to estimate emotion recognition accuracy; in Section C, we report stimulus hit rates and unbiased hit rates of individual emotions computed for each encoder.

Section A - Intensity, authenticity and plausibility

Table A. *Mean Ratings for Intensity, Authenticity, and Plausibility (Core-set Rating Study) Compared to Believability Ratings (Full-set Rating Study) and Organized per Emotion*

	Intensity				Authenticity				Plausibility				Believability			
	A	V	AV	Total	A	V	AV	Total	A	V	AV	Total	A	V	AV	Total
Positive high arousal																
Amusement	4.27	4.16	4.24	4.23	6.88	6.40	7.69	6.99	6.65	7.85	7.59	7.36	7.30	7.12	7.24	7.22
Pride	3.29	3.48	3.41	3.39	6.42	5.64	7.41	6.49	5.84	7.17	6.66	6.55	6.60	6.72	6.72	6.68
Joy	3.99	4.20	4.17	4.12	6.06	6.30	7.49	6.62	5.79	7.66	7.32	6.92	6.38	7.21	7.01	6.87
Positive low arousal																
Relief	3.61	3.21	3.55	3.46	6.28	5.68	6.45	6.14	5.92	7.03	7.27	6.74	6.74	6.82	6.70	6.76
Interest	2.40	2.77	2.86	2.68	5.54	5.71	5.98	5.74	5.52	6.49	6.62	6.21	6.22	6.57	6.45	6.41
Pleasure	2.92	3.44	3.48	3.28	5.37	5.37	6.31	5.68	5.26	6.65	6.63	6.18	6.23	6.01	6.44	6.23
Negative high arousal																
Hot anger	4.28	4.25	4.33	4.28	7.23	6.67	8.19	7.36	7.07	8.03	8.01	7.71	7.45	7.59	7.47	7.50
Panic fear	4.13	3.83	4.06	4.00	7.31	6.01	6.49	6.61	6.93	7.45	7.67	7.35	7.13	6.45	6.69	6.76

Despair	3.55	3.08	3.27	3.30	6.28	5.64	6.47	6.13	5.45	6.57	6.45	6.15	6.57	6.66	6.64	6.62
Negative low arousal																
Irritation	3.02	3.05	3.13	3.07	6.45	5.85	7.18	6.49	6.09	7.08	7.26	6.81	6.70	6.76	6.75	6.74
Anxiety	2.70	2.82	2.77	2.76	6.45	5.42	6.32	6.06	5.64	6.53	6.72	6.30	6.45	6.77	6.52	6.58
Sadness	2.55	2.43	2.63	2.54	6.29	5.27	6.27	5.94	6.02	6.93	6.61	6.52	6.71	6.38	6.45	6.52
Total for 12 repeated emotions	3.38	3.38	3.48	3.41	6.36	5.81	6.83	6.33	5.99	7.10	7.04	6.71	6.69	6.74	6.75	6.73
Additional emotions																
Admiration	3.53	3.24	3.81	3.52	5.79	5.37	6.85	6.00	5.04	6.87	7.50	6.47	6.64	6.70	6.43	6.59
Tenderness	2.61	2.52	2.89	2.67	5.72	5.91	6.38	6.00	5.32	6.06	6.07	5.82	6.19	6.53	6.55	6.42
Disgust	3.16	3.57	3.87	3.53	6.02	5.33	6.30	5.88	5.23	6.65	6.61	6.16	5.81	5.97	6.30	6.02
Contempt	3.07	3.28	3.36	3.24	6.35	6.25	7.11	6.57	5.34	6.61	7.22	6.39	6.13	6.49	6.61	6.41
Surprise	2.73	2.68	3.08	2.83	6.24	5.51	6.70	6.15	6.77	6.42	6.61	6.60	6.60	6.80	6.69	6.69
Total for 5 additional emotions	3.04	3.06	3.41	3.17	6.02	5.66	6.67	6.12	5.52	6.54	6.83	6.30	6.29	6.50	6.51	6.43
Total for all emotions	3.32	3.32	3.47	3.37	6.30	5.79	6.81	6.30	5.91	7.01	7.00	6.64	6.63	6.70	6.71	6.68

Note. “A”: Audio only; “V”: Video only; “AV”: Audio-Video. N= 154.

Section B - Chance Expectation, Response Bias, and Confusions Between Emotions

For reasons of space economy, in our article we reported results for accuracy of emotion recognition only in the form of the classic accuracy index (stimulus hit rate), i.e., the number of emotion expression items for a given target category that was correctly classified by the raters. However, some authors pointed out a number of methodological issues involved in using this accuracy index (e.g., Wagner, 1993). In particular, accuracy estimates are affected by two kinds of biases - the relative number of items for each emotion category presented (presentation bias) and the relative utilization of different emotion categories by the raters (response bias). These biases affect accuracy indices. With respect to presentation bias, the more and the more difficult (ambiguous) stimuli are presented, the greater the chance of confusions. With respect to response bias, an overuse of a specific category label (i.e., assigning this label more often than there are stimuli in that class) will result in a higher probability that the items in this category will be correctly identified. Conversely, category underuse (i.e., assigning this label less often than there are stimuli in that class) will lead to a lower probability that stimuli can be correctly identified. These biases in the distribution of the ratings also affect the chance level, i.e., the criterion to adopt for judging to what extent an accuracy index exceeds chance effects that would result from random guessing (Scherer, Banse, Walbott, & Goldbeck, 1991).

There have been several attempts to propose indices that are corrected for chance guessing and/or response bias. However, apart from the fact that many of the implicit assumptions of such correction methods (e.g., the assumption of random guessing) are not met in EE recognition studies, the resulting indices (for example, the “unbiased hit rate” suggested by Wagner, 1993) generally mask the origin of the potential biases and do not allow to take these into account in interpreting the results. The appropriate and most complete way to provide all the relevant information about potential biases is to reproduce the detailed confusion matrices

for the data. In consequence, in this supplementary material we provide the confusion matrices for each of the three modalities.

For each modality, both a confusion matrix with the actual frequencies of the classifications by raters (raw frequencies; Tables B1, C1, and D1) as well as a version showing the respective percentages is provided (Tables B2, C2, and D2). Next to each of the raw matrices, a column shows the respective over- or underuse of each emotion category in the rater classifications. In addition to the matrices, we provide in each case six sets of derived indices (Tables B3, C3, and D3), consisting of the following information:

H_s*: stimulus hit rate;* this corresponds to the classic accuracy rate reported in Table 5 of the article; i.e., the proportion of the number of presented items that were correctly identified as the target category (the diagonal cell divided by the column margin in the confusion matrices below);

p_s: *stimulus-based chance level*; the probability that a category is correctly identified by chance (random guessing) given the number of available categories (18) but adjusted for the differential number of stimuli that were presented in each category (the column margin divided by the total number of items presented);

H_j*: judge hit rate*; representing the percentage of correct choices by the judges based on the total number of uses of a category label in their judgments (the diagonal cell divided by the row margin). This may reflect the judges' recognition ability for a given category, independently of their overall use of the category (provided they are not guessing randomly).

p_j: *judge-based chance level*; the probability that judges will correctly identify a category by chance (random guessing) adjusted for the differential number of times they used the respective category label (the row margin divided by the Grand total);

pi: *proportion index*; this index converts any mean accuracy that originates as a proportion or percentage, no matter how many response options each item had, to its equivalent proportion if there were only two response options, i.e., a binomial decision (Rosenthal & Rubin, 1989). Thus, performance can be expressed on the common metric *pi* allowing direct comparison between the results of studies with a different number of answer alternatives (see Hall, Andrzejewski, Murphy, Schmid Mast, & Feinstein, 2008);

H_u*: unbiased hit rate*; representing "[…] the joint probability that a stimulus category is correctly identified given that it is presented at all and that a response is correctly used given that it is used at all" (Wagner, 1993, p.3). It is computed for each rater by squaring the value of the cells on the diagonal of the confusion matrix and dividing the result by the product of the corresponding row and column marginal totals. We report this index, averaged across raters, because of its recent popularity in the literature. However, we feel that it is not ideal for recognition studies based on enacted corpora because of the fact that these corpora are not made of easily recognizable expression prototypes but rather realistic and idiosyncratic expressions based on memory recall or mental imagery. The H_u will severely reduce the estimated hit rate for any responses that do not fit the target category, assuming that only one single answer, the one corresponding to the intended category, can be "true".

In addition to the information about the frequency and types of confusions in the matrices, these additional indices allow to evaluate the relative effect of stimulus frequency and category use on the recognition accuracy. For example, if H_j is considerably higher than H_s, this may be due to an underuse of the respective category (possibly because of too many ambiguous stimuli or stable confusions for some raters). Conversely, if H_s is much higher than H_j this can be due to overuse of the category, indicating a response bias which artificially

boosts recognition accuracy. One might surmise that the best estimate might lie somewhere between these two values. Similarly, the best estimate of a chance level given the number of alternative might lie somewhere between p_s and p_j.

References

Hall, J. A., Andrzejewski, S. A., Murphy N. A., Schmid Mast, M., & Feinstein, B. A. (2008). Accuracy of judging other's traits and states: Comparing mean levels across tests. *Journal of Research in Personality, 42,* 1476-1489.

Rosenthal, R. & Rubin, D. B (1989). Effect size estimation for one-sample multiple-choice-type data: Design, analysis, and meta-analysis. *Psychological Bulletin, 106,* 332-337.

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15,* 123-148.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17,* 3-28.

Table B1. *Confusion Matrix (Raw Frequencies) – Audio-Video Modality*

		Intended Emotion																		
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad	Total	Use
Ratings	Adm	62	1	10						12	1	7			22	7	2	2	126	18
	Amu		146	2				6			1	16					1		172	-8
	Ten	2		55					1	1					25	1		1	86	-4
	Ang	1			129	1	9	3	4	2	23	1	4	5				1	183	3
	Dis	2		1	1	72	2		1	1	3		28		3		2	7	123	33
	Des		1	2	1	4	81		8		1	1	2	5	1	5	2	52	166	-68
	Pri	1	1	1				115		4	1	7			2			1	133	-47
	Anx			1			37		104	15	8	1	1	33		2	12	10	224	44
	Int	4		2			1	8	4	85	13	2	2		12		7	1	141	-39
	Irr				47	1	19	4	9	9	105		2	3	1	4		8	212	32
	Joy	2	28				1	11		1		112			2	2	1		160	-38
	Con	2		1	1	4	1	5	5	7	9		40	1	2		1	20	99	9
	Fea				1		16		12					126				1	156	-24
	Ple	5	1	7				11		1		16			126	15	1		183	-33
	Rel	3					1	1	1	1		10			13	137			167	-13
	Sur	21		1		1	11		16	28	5	24		7			53	5	172	82
	Sad		2	1		4	46		5				3		1		2	91	155	-61
	Neu	3		6		3	9	16	10	13	10	1	8		6	7	6	16	114	114
	Total	108	180	90	180	90	234	180	180	180	180	198	90	180	216	180	90	216	2772	

Note : The column « Use » refers to the overuse (positive values) or underuse (negative values) of the emotion label with respect to the number of target itmes. List of abbreviations : « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table B2. *Confusion Matrix (Percentages) – Audio-Video Modality*

		Intended Emotion																
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad
Ratings	Adm	57	1	11	0	0	0	0	0	7	1	4	0	0	10	4	2	1
	Amu	0	81	2	0	0	0	3	0	0	1	8	0	0	0	0	1	0
	Ten	2	0	61	0	0	0	0	1	1	0	0	0	0	12	1	0	0
	Ang	1	0	0	72	1	4	2	2	1	13	1	4	3	0	0	0	0
	Dis	2	0	1	1	80	1	0	1	1	2	0	31	0	1	0	2	3
	Des	0	1	2	1	4	35	0	4	0	1	1	2	3	0	3	2	24
	Pri	1	1	1	0	0	0	64	0	2	1	4	0	0	1	0	0	0
	Anx	0	0	1	0	0	16	0	58	8	4	1	1	18	0	1	13	5
	Int	4	0	2	0	0	0	4	2	47	7	1	2	0	6	0	8	0
	Irr	0	0	0	26	1	8	2	5	5	58	0	2	2	0	2	0	4
	Joy	2	16	0	0	0	0	6	0	1	0	57	0	0	1	1	1	0
	Con	2	0	1	1	4	0	3	3	4	5	0	44	1	1	0	1	9
	Fea	0	0	0	1	0	7	0	7	0	0	0	0	70	0	0	0	0
	Ple	5	1	8	0	0	0	6	0	1	0	8	0	0	58	8	1	0
	Rel	3	0	0	0	0	0	1	1	1	0	5	0	0	6	76	0	0
	Sur	19	0	1	0	1	5	0	9	16	3	12	0	4	0	0	59	2
	Sad	0	1	1	0	4	20	0	3	0	0	0	3	0	0	0	2	42
	Neu	3	0	7	0	3	4	9	6	7	6	1	9	0	3	4	7	7
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Note : Numbers represent percentages. « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table B3. *Measures of Accuracy - Audio-Video Modality*

Emotion	H_s	p_s	H_j	p_j	pi	H_u
Admiration	.57	.04	.49	.05	.96	.32
Amusement	.81	.06	.85	.06	.99	.70
Tenderness	.61	.03	.64	.03	.96	.43
Hot Anger	.72	.06	.70	.07	.98	.53
Disgust	.80	.03	.59	.04	.99	.52
Despair	.35	.08	.49	.06	.90	.18
Pride	.64	.06	.86	.05	.97	.58
Anxiety	.58	.06	.46	.08	.96	.30
Interest	.47	.06	.60	.05	.93	.31
Irritation (cold anger)	.58	.06	.50	.08	.96	.31
Joy (elation)	.57	.07	.70	.06	.95	.42
Contempt	.44	.03	.40	.04	.93	.23
Fear (panic)	.70	.06	.81	.06	.98	.59
Pleasure	.58	.08	.69	.07	.96	.42
Relief	.76	.06	.82	.06	.98	.64
Surprise	.59	.03	.31	.06	.96	.20
Sadness	.42	.08	.59	.06	.93	.27

Note : « H_s » : Stimulus Hit Rate ; « p_s » : chance-level adjusted for the number of stimuli presented ; « Hj » : Judge Hit Rate ; « p_j » : chance-level adjusted for the number of times a label was used ; « pi » : Proportion Index; «H_u» : Unbiased Hit Rate.

Table C1. *Confusion Matrix (Raw Frequencies) – Video-Only Modality*

		Intended Emotion																		
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad	Total	Use
Ratings	Adm	33	1	10			1	7		10	1	8	1		22	11	2		107	-7
	Amu	1	137	6				8		4		22			10	2			190	0
	Ten	4	2	33			1	1	2	1	1	2	1		32	3	1		84	-11
	Ang				151	1	3	6	6	6	32	1	2	20			2		230	40
	Dis	2		4	1	48	4		6	4	5		32	1		3	3	10	123	28
	Des	4	7	4	1	14	69		8	2	8	2	2	12		10	2	37	182	-65
	Pri	5		1				109	1	5		5			24	3		2	155	-35
	Anx	7		2		3	72	1	85	15	4		1	48		1	6	15	260	70
	Int	11		6			7	10	8	71	7	3			15	3	5	3	149	-41
	Irr	1		3	35	2	22	2	16	24	77	1	1	8	1	7	5	5	210	20
	Joy	3	23	2				17				111			5	1	1		163	-46
	Con	4		3	2	3	6	3	7	17	38		39	4	8	3	9	24	170	75
	Fea	2					7		24	1	1	3	1	88			3	1	131	-59
	Ple	6	8	4			1	11		3		9			79	22			143	-85
	Rel	3	6	5				5			1	16			27	113	3		179	-11
	Sur	25	2				14		20	13	5	26	3	7		3	42	9	169	74
	Sad	1	4	1		23	34		1	1	4		11	1			1	96	178	-50
	Neu	2		11		1	6	10	6	13	6		1	1	5	5	10	26	103	103
	Total	114	190	95	190	95	247	190	190	190	190	209	95	190	228	190	95	228	2926	

Note : The column « Use » refers to the overuse (positive values) or underuse (negative values) of the emotion label with respect to the number of target itmes. List of abbreviations : « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table C2. *Confusion Matrix (Percentages) – Video-Only Modality*

		Intended Emotion																
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad
Ratings	Adm	29	1	11	0	0	0	4	0	5	1	4	1	0	10	6	2	0
	Amu	1	72	6	0	0	0	4	0	2	0	11	0	0	4	1	0	0
	Ten	4	1	35	0	0	0	1	1	1	1	1	1	0	14	2	1	0
	Ang	0	0	0	79	1	1	3	3	3	17	0	2	11	0	0	2	0
	Dis	2	0	4	1	51	2	0	3	2	3	0	34	1	0	2	3	4
	Des	4	4	4	1	15	28	0	4	1	4	1	2	6	0	5	2	16
	Pri	4	0	1	0	0	0	57	1	3	0	2	0	0	11	2	0	1
	Anx	6	0	2	0	3	29	1	45	8	2	0	1	25	0	1	6	7
	Int	10	0	6	0	0	3	5	4	37	4	1	0	0	7	2	5	1
	Irr	1	0	3	18	2	9	1	8	13	41	0	1	4	0	4	5	2
	Joy	3	12	2	0	0	0	9	0	0	0	53	0	0	2	1	1	0
	Con	4	0	3	1	3	2	2	4	9	20	0	41	2	4	2	9	11
	Fea	2	0	0	0	0	3	0	13	1	1	1	1	46	0	0	3	0
	Ple	5	4	4	0	0	0	6	0	2	0	4	0	0	35	12	0	0
	Rel	3	3	5	0	0	0	3	0	0	1	8	0	0	12	59	3	0
	Sur	22	1	0	0	0	6	0	11	7	3	12	3	4	0	2	44	4
	Sad	1	2	1	0	24	14	0	1	1	2	0	12	1	0	0	1	42
	Neu	2	0	12	0	1	2	5	3	7	3	0	1	1	2	3	11	11
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Note : Numbers represent percentages. « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table C3. *Measures of Accuracy - Video-Only Modality*

Emotion	H_s	p_s	H_j	p_j	pi	H_u
Admiration	.29	.04	.31	.04	.87	.14
Amusement	.72	.06	.72	.06	.98	.54
Tenderness	.35	.03	.39	.03	.90	.18
Hot Anger	.79	.06	.66	.08	.99	.55
Disgust	.51	.03	.39	.04	.95	.23
Despair	.28	.08	.38	.06	.87	.12
Pride	.57	.06	.70	.05	.96	.42
Anxiety	.45	.06	.33	.09	.93	.17
Interest	.37	.06	.48	.05	.91	.20
Irritation (cold anger)	.41	.06	.37	.07	.92	.17
Joy (elation)	.53	.07	.68	.06	.95	.38
Contempt	.41	.03	.23	.06	.92	.12
Fear (panic)	.46	.06	.67	.04	.94	.34
Pleasure	.35	.08	.55	.05	.90	.23
Relief	.59	.06	.63	.06	.96	.41
Surprise	.44	.03	.25	.06	.93	.13
Sadness	.42	.08	.54	.06	.93	.25

Note : « H_s » : Stimulus Hit Rate ; « p_s » : chance-level adjusted for the number of stimuli presented ; « Hj » : Judge Hit Rate ; « p_j » : chance-level adjusted for the number of times a label was used ; « pi » : Proportion Index; «H_u» : Unbiased Hit Rate.

Table D1. *Confusion Matrix (Raw Frequencies) – Audio-Only Modality*

		Intended Emotion																		
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad	Total	Use
Ratings	Adm	37		9		2	6	5	3	19	1	15	9	1	21	15	2	4	149	29
	Amu	1	132	9			9	4		1	2	29	1		6	3	5	4	206	6
	Ten	9		23		1		3	3	11	3	1	4		26	17	2	10	113	13
	Ang		1		118	2	5	17		3	13	4	2	6	2	1	1	1	176	-24
	Dis	2		1	4	20	5	10	8	5	10	6	10	1	6	5	4	12	109	9
	Des	5	12	5		11	63	3	9	3	2	29	3	9	10	12	1	43	220	-40
	Pri	5			4	10	5	32	2	7	8	12	2	2	10			2	101	-99
	Anx	2	2		2	3	46	6	68	12	3	5	9	48	5	9	13	26	259	59
	Int	8		10		1	1	13	23	47	25	2	2	1	18		14	5	170	-30
	Irr	5			54	3	6	37	10	6	74	2	3	5	3	8	3	6	225	25
	Joy	2	27				13	4		1	1	62	1	2		4			117	-103
	Con	4	7	5	10	9	2	32	16	12	26	2	24		7	6	4	15	181	81
	Fea				3	2	22	1	12	1		9	2	118		1	1	5	177	-23
	Ple	4	4	8		8	4	6	1	2	1	3	3		65	18	1	5	133	-107
	Rel	13		10	1	7	5	5	9	10	3	5	3	1	27	87	2	19	207	7
	Sur	12	2	1	2	5	22	14	15	22	11	15	8	6	3	6	33	2	179	79
	Sad	2	12	8	1	8	43		5	4		16	4		4	4	1	44	156	-84
	Neu	9	1	11	1	8	3	8	16	34	17	3	10		27	4	13	37	202	202
	Total	120	200	100	200	100	260	200	200	200	200	220	100	200	240	200	100	240	3080	

Note : The column « Use » refers to the overuse (positive values) or underuse (negative values) of the emotion label with respect to the number of target itmes. List of abbreviations : « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table D2. *Confusion Matrix (Percentages) – Audio-Only Modality*

		Intended Emotion																
		Adm	Amu	Ten	Ang	Dis	Des	Pri	Anx	Int	Irr	Joy	Con	Fea	Ple	Rel	Sur	Sad
Ratings	Adm	31	0	9	0	2	2	3	2	10	1	7	9	1	9	8	2	2
	Amu	1	66	9	0	0	3	2	0	1	1	13	1	0	3	2	5	2
	Ten	8	0	23	0	1	0	2	2	6	2	0	4	0	11	9	2	4
	Ang	0	1	0	59	2	2	9	0	2	7	2	2	3	1	1	1	0
	Dis	2	0	1	2	20	2	5	4	3	5	3	10	1	3	3	4	5
	Des	4	6	5	0	11	24	2	5	2	1	13	3	5	4	6	1	18
	Pri	4	0	0	2	10	2	16	1	4	4	5	2	1	4	0	0	1
	Anx	2	1	0	1	3	18	3	34	6	2	2	9	24	2	5	13	11
	Int	7	0	10	0	1	0	7	12	24	13	1	2	1	8	0	14	2
	Irr	4	0	0	27	3	2	19	5	3	37	1	3	3	1	4	3	3
	Joy	2	14	0	0	0	5	2	0	1	1	28	1	1	0	2	0	0
	Con	3	4	5	5	9	1	16	8	6	13	1	24	0	3	3	4	6
	Fea	0	0	0	2	2	8	1	6	1	0	4	2	59	0	1	1	2
	Ple	3	2	8	0	8	2	3	1	1	1	1	3	0	27	9	1	2
	Rel	11	0	10	1	7	2	3	5	5	2	2	3	1	11	44	2	8
	Sur	10	1	1	1	5	8	7	8	11	6	7	8	3	1	3	33	1
	Sad	2	6	8	1	8	17	0	3	2	0	7	4	0	2	2	1	18
	Neu	8	1	11	1	8	1	4	8	17	9	1	10	0	11	2	13	15
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Note : Numbers represent percentages. « Adm » : Admiration ; « Amu » : amusement ; « Ten » Tenderness ; « Ang » : Hot anger (rage) ; « Dis » : Disgust ; « Des » : Despair ; « Pri » : Pride ; « Anx » : Anxiety ; « Int » : Interest ; « Irr » : Irritation (cold anger) ; « Joy » : Elated Joy ; « Con » : contempt ; « Fea » : (Panic) Fear ; « Ple » : Pleasure ; « Rel » : Relief ; « Sur » :Surprise ; « Sad » : Sadness ; « Neu » : Neutral.

Table D3. *Measures of Accuracy Audio-Only Modality*

Emotion	H_s	p_s	H_j	p_j	pi	H_u
Admiration	.31	.04	.25	.05	.88	.11
Amusement	.66	.06	.64	.07	.97	.45
Tenderness	.23	.03	.20	.04	.84	.10
Hot Anger	.59	.06	.67	.06	.96	.43
Disgust	.20	.03	.18	.04	.81	.06
Despair	.24	.08	.29	.07	.84	.09
Pride	.16	.06	.32	.03	.76	.08
Anxiety	.34	.06	.26	.08	.90	.10
Interest	.24	.06	.28	.06	.84	.09
Irritation (cold anger)	.37	.06	.33	.07	.91	.14
Joy (elation)	.28	.07	.53	.04	.87	.17
Contempt	.24	.03	.13	.06	.84	.06
Fear (panic)	.59	.06	.67	.06	.96	.42
Pleasure	.27	.08	.49	.04	.86	.15
Relief	.44	.06	.42	.07	.93	.23
Surprise	.33	.03	.18	.06	.89	.09
Sadness	.18	.08	.28	.05	.79	.07

Note : « H_s » : Stimulus Hit Rate ; « p_s » : chance-level adjusted for the number of stimuli presented ; « Hj » : Judge Hit Rate ; « p_j » : chance-level adjusted for the number of times a label was used ; « pi » : Proportion Index; «H_u» : Unbiased Hit Rate.

Section C – Hit Rates per encoder

Researchers interested in using stimuli taken from the GEMEP CS may be interested in knowing the recognition accuracy for each stimulus in each of the three modalities. Unfortunately, unbiased Hit Rates (H_u) cannot be computed for individual EEs. Nevertheless, it is possible to compute the H_u index per encoder for each emotion. We report in the following table the raw hit rates and the unbiased hit rates per encoder in the three perception modalities organized per emotion.

Table E. *Emotion Raw Stimulus Hit Rate (H_s) and Unbiased Hit Rate (H_u) per Encoder*

		AV			V			A		
Encoder	Emotion	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u
A01	Amu	.67	.38	.36	.37	.15	.20	.70	.25	.14
	Ten	.39	.23		.32	.11		.15	.04	
	Ang	.72	.55		.84	.64		.50	.36	
	Des	.33	.17		.34	.16		.08	.01	
	Pri	.78	.60		.26	.11		.30	.15	
	Anx	.61	.21		.53	.14		.20	.03	
	Int	.44	.25		.32	.19		.15	.02	
	Irr	.28	.13		.21	.08		.15	.03	
	Joy	.50	.32		.45	.29		.33	.22	
	Con	.28	.15		.11	.02		.10	.01	
	Fea	.94	.70		.58	.35		.75	.51	
	Ple	.75	.52		.39	.22		.45	.30	
	Rel	.83	.69		.53	.26		.25	.07	
	Sur	.72	.39		.74	.27		.25	.05	
	Sad	.28	.09		.21	.04		.15	.03	
A02	Adm	.72	.63	.56	.42	.24	.42	.15	.05	.21
	Amu	.89	.89		.89	.85		.20	.20	
	Ang	.83	.52		.89	.59		.75	.70	
	Dis	.78	.78		.58	.53		.05	.01	
	Des	.72	.63		.42	.28		.35	.06	
	Pri	.83	.69		.79	.74		.15	.05	
	Anx	.83	.48		.58	.23		.45	.16	
	Int	.56	.51		.37	.17		.00	.00	
	Irr	.50	.30		.37	.20		.85	.54	
	Joy	.44	.32		.37	.26		.55	.47	
	Fea	.50	.41		.74	.49		.35	.31	
	Ple	.67	.30		.42	.21		.15	.03	
	Rel	.72	.67		.89	.66		.50	.20	
	Sad	.83	.66		.74	.49		.60	.23	
A03	Amu	.94	.94	.34	.74	.64	.24	.50	.38	.11
	Ten	.67	.57		.53	.40		.20	.07	
	Ang	.44	.22		.47	.13		.25	.09	
	Des	.44	.21		.11	.02		.70	.35	
	Pri	.33	.18		.53	.31		.05	.01	
	Anx	.06	.00		.05	.01		.55	.16	
	Int	.67	.40		.58	.37		.10	.01	
	Irr	.33	.08		.11	.01		.00	.00	
	Joy	.94	.80		.79	.59		.05	.01	
	Con	.44	.15		.58	.12		.20	.03	
	Fea	.72	.63		.00	.00		.60	.40	
	Ple	.00	.00		.05	.01		.15	.05	
	Rel	.61	.61		.63	.63		.20	.05	
	Sur	.17	.13		.05	.05		.10	.05	
	Sad	.33	.13		.37	.29		.00	.00	

(*table continues*)

Table E (*continued*)

		AV			V			A		
Encoder	Emotion	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u
A04	Adm	.61	.56	.39	.11	.02	.22	.00	.00	.16
	Amu	.67	.62		.74	.47		.50	.42	
	Ang	.89	.57		.89	.63		.50	.28	
	Dis	.61	.48		.21	.17		.35	.16	
	Des	.08	.04		.08	.02		.05	.01	
	Pri	.28	.20		.37	.26		.05	.01	
	Anx	.67	.31		.32	.06		.30	.07	
	Int	.11	.03		.21	.06		.20	.04	
	Irr	.67	.27		.58	.23		.45	.19	
	Joy	.67	.44		.58	.49		.50	.22	
	Fea	.56	.43		.42	.31		.55	.30	
	Ple	.72	.52		.32	.15		.30	.18	
	Rel	.94	.73		.21	.06		.80	.32	
	Sad	.33	.33		.37	.15		.15	.06	
A05	Adm	.50	.26	.45	.21	.14	.32	.40	.12	.23
	Amu	.78	.73		.68	.47		1.00	.69	
	Ang	.83	.60		1.00	.59		.95	.67	
	Dis	.94	.85		.95	.90		.15	.06	
	Des	.56	.33		.58	.24		.20	.06	
	Pri	.72	.72		.26	.19		.30	.11	
	Anx	.56	.33		.37	.18		.65	.26	
	Int	.39	.23		.32	.16		.15	.06	
	Irr	.56	.33		.26	.10		.50	.29	
	Joy	.50	.23		.63	.24		.25	.14	
	Fea	.61	.45		.37	.17		.20	.07	
	Ple	.33	.15		.16	.03		.30	.20	
	Rel	.72	.55		.68	.49		.55	.29	
	Sad	.78	.60		.84	.56		.20	.16	
A06	Amu	.89	.68	.60	.74	.47	.35	.75	.47	.26
	Ten	.94	.85		.47	.27		.40	.23	
	Ang	.72	.72		.84	.56		.75	.63	
	Des	.39	.23		.32	.11		.35	.14	
	Pri	.61	.56		.63	.45		.00	.00	
	Anx	.78	.52		.74	.45		.60	.27	
	Int	.50	.45		.11	.05		.40	.15	
	Irr	1.00	.67		.47	.19		.45	.23	
	Joy	.94	.85		.95	.71		.50	.28	
	Con	.44	.36		.11	.03		.50	.28	
	Fea	.78	.64		.47	.39		.60	.38	
	Ple	.72	.67		.21	.08		.50	.33	
	Rel	.72	.72		.84	.71		.25	.25	
	Sur	.67	.57		.42	.42		.45	.17	
	Sad	.83	.46		.68	.32		.40	.15	

(*table continues*)

Table E (*continued*)

		AV			V			A		
Encoder	Emotion	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u
A07	Amu	.83	.78	.48	.84	.84	.31	.90	.81	.21
	Ten	.67	.67		.16	.06		.30	.09	
	Ang	.61	.42		.42	.26		.30	.11	
	Des	.33	.21		.03	.00		.40	.21	
	Pri	.94	.89		.89	.76		.00	.00	
	Anx	.72	.43		.42	.08		.25	.05	
	Int	.61	.40		.26	.09		.25	.08	
	Irr	.78	.42		.58	.19		.30	.05	
	Joy	.44	.36		.58	.58		.20	.16	
	Con	.44	.19		.68	.23		.30	.08	
	Fea	.78	.68		.37	.18		.70	.49	
	Ple	.78	.60		.58	.45		.30	.18	
	Rel	.83	.66		.89	.63		.45	.18	
	Sur	.83	.48		.58	.27		.85	.60	
	Sad	.22	.07		.13	.05		.10	.02	
A08	Amu	.83	.63	.37	.63	.58	.34	.80	.46	.15
	Ten	.39	.34		.26	.22		.10	.01	
	Ang	.72	.63		.95	.71		.60	.55	
	Des	.17	.03		.05	.00		.05	.00	
	Pri	.61	.61		.79	.59		.05	.01	
	Anx	.56	.28		.32	.13		.10	.01	
	Int	.33	.12		.32	.11		.35	.11	
	Irr	.17	.05		.26	.09		.05	.01	
	Joy	.50	.38		.58	.49		.15	.06	
	Con	.61	.40		.58	.34		.10	.02	
	Fea	.67	.53		.58	.53		.80	.67	
	Ple	.78	.54		.58	.49		.40	.12	
	Rel	.94	.73		.84	.45		.35	.14	
	Sur	.56	.17		.42	.11		.00	.00	
	Sad	.22	.16		.37	.27		.15	.06	
A09	Adm	.75	.72	.58	.50	.41	.42	.50	.27	.22
	Amu	.72	.59		.79	.47		.50	.24	
	Ang	.89	.71		.95	.81		.90	.74	
	Dis	.78	.73		.21	.09		.15	.04	
	Des	.61	.40		.79	.54		.35	.19	
	Pri	.89	.84		.84	.64		.50	.23	
	Anx	.67	.32		.63	.32		.00	.00	
	Int	.61	.52		.74	.52		.45	.20	
	Irr	.67	.50		.63	.40		.35	.16	
	Joy	.67	.44		.47	.33		.15	.03	
	Fea	.78	.68		.68	.52		.70	.54	
	Ple	.83	.66		.63	.47		.25	.18	
	Rel	.89	.65		.32	.17		.35	.19	
	Sad	.39	.34		.26	.13		.10	.03	

(*table continues*)

Table E (*continued*)

		AV			V			A		
Encoder	Emotion	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u	H_s	H_u	Mean H_u
A10	Adm	.11	.03	.32	.00	.00	.24	.30	.12	.17
	Amu	.89	.84		.79	.54		.75	.70	
	Ang	.50	.28		.68	.68		.40	.18	
	Dis	.89	.79		.58	.40		.30	.11	
	Des	.11	.01		.47	.13		.10	.01	
	Pri	.39	.34		.37	.17		.20	.07	
	Anx	.33	.11		.53	.14		.30	.11	
	Int	.50	.23		.53	.21		.30	.09	
	Irr	.89	.43		.58	.27		.60	.18	
	Joy	.11	.06		.00	.00		.10	.05	
	Fea	.67	.53		.42	.34		.65	.40	
	Ple	.33	.29		.21	.11		.00	.00	
	Rel	.39	.30		.11	.03		.65	.30	
	Sad	.39	.25		.58	.34		.10	.01	

Note. «AV»: Audio-Video modality; «V»: Video only modality; «A»: Audio only modality. Emotion labels: «Adm»: Admiration; «Amu»: amusement; «Ten»: Tenderness; «Ang»: Hot anger (rage); « Dis»: Disgust; «Des»: Despair; «Pri»: Pride; «Anx»: Anxiety; «Int»: Interest; «Irr»: Irritation (cold anger); «Joy»: Elated Joy; «Con»: contempt; «Fea»: (Panic) Fear; «Ple»: Pleasure; «Rel»: Relief; «Sur»:Surprise; «Sad»: Sadness.