

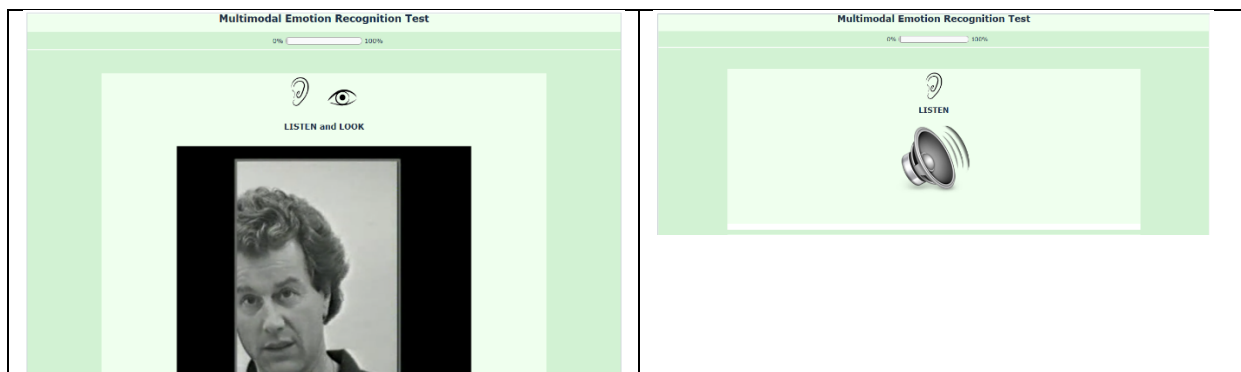
Appendix: Multimodal Emotion Recognition Test (MERT) User Manual

1 Goal, features, and versions of the MERT

The MERT (Bänziger, Grandjean, & Scherer, 2009) consists of 30 video clips of actors (three for each of 10 emotions: irritation, anger, anxiety, fear, happiness, elated joy, disgust, contempt, sadness, despair) that are presented in four modalities (still picture, video only, audio only, audio with video), yielding a total of 120 items (Fig 1. a). These recordings were taken from the GVEESS corpus (Banse & Scherer, 1996). A standard pseudo-linguistic sentence (i.e., a sentence without meaning) is used as verbal content.

After each of the 120 stimuli, participants are asked to choose which of the 10 emotions had been expressed by the actor (Fig 1. b). The total duration of the MERT is 45 minutes.

a)



b)

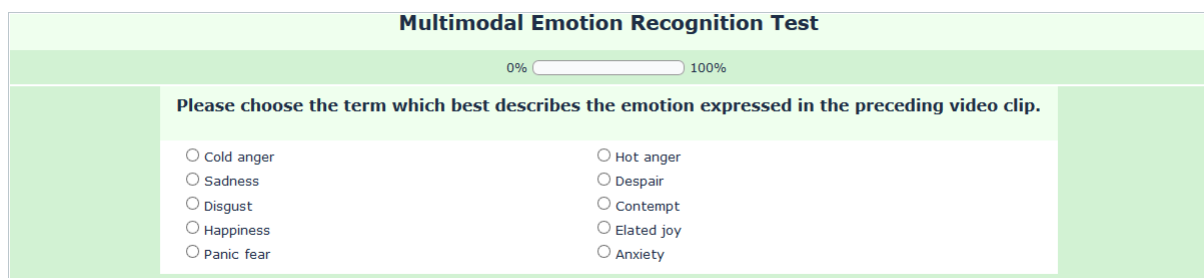
The image shows a screenshot of the MERT response format. It has a light green background and a progress bar at the top showing 0% completion. The title 'Multimodal Emotion Recognition Test' is at the top. Below the progress bar, the text 'Please choose the term which best describes the emotion expressed in the preceding video clip.' is displayed. Below this text, there are two columns of radio buttons, each followed by an emotion term. The terms are: Cold anger, Hot anger, Sadness, Despair, Disgust, Contempt, Happiness, Elated joy, and Panic fear, Anxiety.

Figure 1: Screenshots illustrating the test for an audiovisual and an audio item (a) and the response format after stimulus presentation (b)

2 MERT variables

For each emotion 3 variables are preserved in the file (eg videoU22212V, U22212V, EvalU22212V).

- The first variable indicates whether a video, sound or picture was presented. This is indicated in the first part of the variable name. The variable as such is always blank. The second part of the variable name starts with a letter indicating which emotion is shown (e.g. U for elated joy). Table 1, in the appendix, can be used to deduce which emotions the letters represent. The last or last two letters indicate which format is used (audiovisual clip (AV), visual clip (V), audio record (A) or a picture (P)). For a complete overview of the MERT variables see Appendix, Table 2.
- In the second variable the emotion selected by the respondent is saved.
- In the third variable (starting with Eval) the question is scored (1=correct; 0=wrong).

At the end of the file sumscores are calculated for each format as well as a total score for the whole test (expressed in % correct).

SUMA → audio; SUMV → visual clip; SUMAV → audiovisual clip; SUMP → picture

SUMA = round(sum(EvalH12209A, EvalZ21204A, EvalV11112A, EvalT12105A, EvalZ12103A, EvalZ21105A, EvalU22212A, EvalF22205A, EvalF12112A, EvalT11204A, EvalK11212A, EvalP21210A, EvalU21202A, EvalA11205A, EvalE11201A, EvalP12103A, EvalE11111A, EvalU22107A, EvalT22203A, EvalK22207A, EvalK12209A, EvalA22203A, EvalV11104A, EvalA21202A, EvalH11211A, EvalE11203A, EvalV12109A, EvalP12205A, EvalH21212A, EvalF21203A) / 30 * 100, 0)

SUMV = round(sum(EvalT12105V, EvalV12109V, EvalF12112V, EvalU22212V, EvalH21212V, EvalF21203V, EvalT11204V, EvalF22205V, EvalA22203V, EvalH12209V, EvalZ21105V, EvalE11201V, EvalE11111V, EvalK12209V, EvalK11212V, EvalP21210V, EvalU22107V, EvalP12205V, EvalU21202V, EvalV11112V, EvalT22203V, EvalA11205V, EvalP12103V, EvalE11203V, EvalK22207V, EvalV11104V, EvalA21202V, EvalZ12103V, EvalZ21204V, EvalH11211V) / 30 * 100, 0)

SUMAV = round(sum(EvalT22203AV, EvalZ12103AV, EvalZ21204AV, EvalH12209AV, EvalV12109AV, EvalZ21105AV, EvalK11212AV, EvalH11211AV, EvalU21202AV, EvalF12112AV, EvalE11201AV, EvalA11205AV, EvalT12105AV, EvalK22207AV, EvalP12103AV, EvalF21203AV, EvalP21210AV, EvalV11104AV, EvalP12205AV, EvalA22203AV, EvalH21212AV, EvalE11111AV, EvalU22212AV, EvalU22107AV, EvalA21202AV, EvalE11203AV, EvalT11204AV, EvalV11112AV, EvalF22205AV, EvalK12209AV) / 30 * 100, 0)

SUMP = round(sum(EvalK11212P, EvalF12112P, EvalE11201P, EvalA21202P, EvalF21203P, EvalF22205P, EvalP12205P, EvalE11111P, EvalA11205P, EvalU22212P, EvalT22203P, EvalE11203P, EvalZ21204P, EvalU21202P, EvalH21212P, EvalK22207P, EvalH11211P, EvalA22203P, EvalP21210P, EvalU22107P, EvalZ21105P, EvalV11112P, EvalP12103P, EvalV12109P, EvalK12209P, EvalH12209P, EvalT11204P, EvalV11104P, EvalZ12103P, EvalT12105P) / 30 * 100, 0)

SUMTOT = round(sum(EvalH12209A, EvalZ21204A, EvalV11112A, EvalT12105A, EvalZ12103A, EvalZ21105A, EvalU22212A, EvalF22205A, EvalF12112A, EvalT11204A, EvalK11212A, EvalP21210A, EvalU21202A, EvalA11205A, EvalE11201A, EvalP12103A, EvalE11111A, EvalU22107A, EvalT22203A, EvalK22207A, EvalK12209A, EvalA22203A, EvalV11104A, EvalA21202A, EvalH11211A, EvalE11203A, EvalV12109A, EvalP12205A, EvalH21212A, EvalF21203A, EvalT22203AV, EvalZ12103AV, EvalZ21204AV, EvalH12209AV, EvalV12109AV, EvalZ21105AV, EvalK11212AV, EvalH11211AV, EvalU21202AV, EvalF12112AV, EvalE11201AV, EvalA11205AV, EvalT12105AV, EvalK22207AV, EvalP12103AV, EvalF21203AV, EvalP21210AV, EvalV11104AV, EvalP12205AV, EvalA22203AV, EvalH21212AV, EvalE11111AV, EvalU22212AV, EvalU22107AV, EvalA21202AV, EvalE11203AV, EvalT11204AV, EvalV11112AV, EvalF22205AV, EvalK12209AV, EvalK11212P, EvalF12112P, EvalE11201P, EvalA21202P, EvalF21203P, EvalF22205P, EvalP12205P, EvalE11111P, EvalA11205P, EvalU22212P, EvalT22203P, EvalE11203P, EvalZ21204P, EvalU21202P, EvalH21212P, EvalK22207P, EvalH11211P, EvalA22203P, EvalP21210P, EvalU22107P, EvalZ21105P, EvalV11112P, EvalP12103P, EvalV12109P, EvalK12209P, EvalH12209P, EvalT11204P, EvalV11104P, EvalZ12103P, EvalT12105P, EvalT12105V, EvalV12109V, EvalF12112V, EvalU22212V, EvalH21212V, EvalF21203V, EvalT11204V, EvalF22205V, EvalA22203V, EvalH12209V, EvalZ21105V, EvalE11201V, EvalE11111V, EvalK12209V, EvalK11212V, EvalP21210V, EvalU22107V, EvalP12205V, EvalU21202V, EvalV11112V, EvalT22203V, EvalA11205V, EvalP12103V, EvalE11203V, EvalK22207V, EvalV11104V, EvalA21202V, EvalZ12103V, EvalZ21204V, EvalH11211V) / 120 * 100, 0)

Publications

- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, 70(3), 614.
- Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*, 9(5), 691–704.

Appendix

Table 1: Values for each emotion in data file

value	emotion
K	irritation
H	hot anger
T	sadness
Z	despair
E	disgust
V	contempt
F	happiness
U	elated joy
P	panic fear
A	anxiety

Table 2:

Overview of items, format and emotion presented

order	variable	format	target answer	
			value	emotion
1	H12209A	A	H	hot anger
2	T22203AV	AV	T	sadness
3	Z21204A	A	Z	despair
4	V11112A	A	V	contempt
5	T12105V	V	T	sadness
6	T12105A	A	T	sadness
7	V12109V	V	V	contempt
8	F12112V	V	F	happiness
9	Z12103A	A	Z	despair
10	U22212V	V	U	elated joy
11	H21212V	V	H	hot anger
12	Z21105A	A	Z	despair
13	U22212A	A	U	elated joy
14	Z12103AV	AV	Z	despair
15	K11212P	P	K	irritation
16	F21203V	V	F	happiness
17	F12112P	P	F	happiness
18	T11204V	V	T	sadness
19	Z21204AV	AV	Z	despair
20	F22205A	A	F	happiness
21	E11201P	P	E	disgust
22	H12209AV	AV	H	hot anger
23	F22205V	V	F	happiness
24	A21202P	P	A	anxiety
25	F21203P	P	F	happiness
26	A22203V	V	A	anxiety
27	V12109AV	AV	V	contempt

28	Z21105AV	AV	Z	despair
29	F22205P	P	F	happiness
30	H12209V	V	H	hot anger
31	F12112A	A	F	happiness
32	K11212AV	AV	K	irritation
33	P12205P	P	P	panic fear
34	T11204A	A	T	sadness
35	Z21105V	V	Z	despair
36	E11111P	P	E	disgust
37	H11211AV	AV	H	hot anger
38	E11201V	V	E	disgust
39	A11205P	P	A	anxiety
40	E11111V	V	E	disgust
41	U21202AV	AV	U	elated joy
42	K11212A	A	K	irritation
43	P21210A	A	P	panic fear
44	K12209V	V	K	irritation
45	U22212P	P	U	elated joy
46	T22203P	P	T	sadness
47	F12112AV	AV	F	happiness
48	K11212V	V	K	irritation
49	E11203P	P	E	disgust
50	U21202A	A	U	elated joy
51	A11205A	A	A	anxiety
52	P21210V	V	P	panic fear
53	Z21204P	P	Z	despair
54	U21202P	P	U	elated joy
55	U22107V	V	U	elated joy
56	E11201AV	AV	E	disgust
57	P12205V	V	P	panic fear
58	H21212P	P	H	hot anger
59	U21202V	V	U	elated joy
60	K22207P	P	K	irritation
61	V11112V	V	V	contempt
62	A11205AV	AV	A	anxiety
63	E11201A	A	E	disgust
64	H11211P	P	H	hot anger
65	T12105AV	AV	T	sadness
66	T22203V	V	T	sadness
67	P12103A	A	P	panic fear
68	A22203P	P	A	anxiety
69	E11111A	A	E	disgust
70	K22207AV	AV	K	irritation
71	P21210P	P	P	panic fear
72	U22107A	A	U	elated joy
73	P12103AV	AV	P	panic fear
74	F21203AV	AV	F	happiness
75	T22203A	A	T	sadness

76	P21210AV	AV	P	panic fear
77	K22207A	A	K	irritation
78	A11205V	V	A	anxiety
79	V11104AV	AV	V	contempt
80	P12205AV	AV	P	panic fear
81	U22107P	P	U	elated joy
82	A22203AV	AV	A	anxiety
83	P12103V	V	P	panic fear
84	Z21105P	P	Z	despair
85	V11112P	P	V	contempt
86	K12209A	A	K	irritation
87	P12103P	P	P	panic fear
88	V12109P	P	V	contempt
89	E11203V	V	E	disgust
90	A22203A	A	A	anxiety
91	K12209P	P	K	irritation
92	V11104A	A	V	contempt
93	K22207V	V	K	irritation
94	H12209P	P	H	hot anger
95	A21202A	A	A	anxiety
96	V11104V	V	V	contempt
97	A21202V	V	A	anxiety
98	H11211A	A	H	hot anger
99	H21212AV	AV	H	hot anger
100	E11203A	A	E	disgust
101	T11204P	P	T	sadness
102	E11111AV	AV	E	disgust
103	V12109A	A	V	contempt
104	U22212AV	AV	U	elated joy
105	U22107AV	AV	U	elated joy
106	P12205A	A	P	panic fear
107	Z12103V	V	Z	despair
108	A21202AV	AV	A	anxiety
109	E11203AV	AV	E	disgust
110	V11104P	P	V	contempt
111	H21212A	A	H	hot anger
112	Z12103P	P	Z	despair
113	T11204AV	AV	T	sadness
114	V11112AV	AV	V	contempt
115	F21203A	A	F	happiness
116	Z21204V	V	Z	despair
117	F22205AV	AV	F	happiness
118	T12105P	P	T	sadness
119	K12209AV	AV	K	irritation
120	H11211V	V	H	hot anger

AV=audio visual, A=Audio, V=Visual, P=Picture