The supplementary material includes description of scenarios that we developed to record emotionally charged person-person interactions for the MPIIEmo dataset as well as a summary of existing datasets for analysis of emotional expressions.

The details on scenarios and sub-scenarios are shown in Fig. 1 - 7. Description of each scenario includes a short background of the situation that led to a scene and a description of each sub-scenario that correspond to possible developments. These descriptions were given to subjects prior to recording and are used to guide their conversation. Fig. 1 - 7 also include example snapshots to illustrate each of the recordings. For the purpose of illustration we choose images from recordings of different pairs of subjects for every scenario.

Emotion Recognition from Embedded Bodily Expressions and Speech during Dyadic Interactions

Supplementary Material
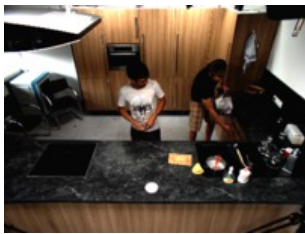
1. B is **happy** that A reminded him and throws away the garbage.

2. B is **angry** about himself, because A has needed to remind him. A tries to calm him down, its ok to forget once.

3. B is **angry** about the As irony. A thinks that his remarks were ok and becomes **angry** himself.

4. B is **bewildered** because he knows that today is not his turn. A forgot about it. A is **surprised** and apologizes himself.

Scenario 1: "MuellWe". Scenario description: *A is in the kitchen, B comes in. A asks B with irony if he is perhaps the next one to throw away the garbage.* (A: white shirt, B: black shirt)

1. A is coming along and is **happy** about the party. B is **happy**, that A will join.

2. A was invited, but has a lot of work to do and can't come. She is **sad** about it. B **sympathizes** with her.

3. A is **surprised**. She didn't know about the party but she would like to come along. B is **happy** about it.

4. The party is organized by the company where A and B are both together. A was not invited. She is **angry**. She has the impression that she is a second-class employee in the company. B is **surprised** and **bewildered** that A was not invited.

Scenario 2: "PartyPe". Scenario description: *A is in the kitchen, B comes in. B asks A if she is also coming to the party.* (A: grey pullover, B: green pullover)

1. A is **happy** and says thank you. B is **happy** that she could do A a favor.

2. A is **surprised** and **happy** that B brought a new flavour. B is **proud** that she discovered it.

3. A ist **angry** because B didn't bring what he asked for. B is **sad** that she made a mistake. Actually, she was looking forward to do a favour for A.

4. A reacts indifferently. B is **irritated** and **angry**, she was expecting to be thanked for making an effort.

Scenario 3: "Mitbrin". Scenario description: *B comes back from shopping and gives A a pack of candy that A asked her to buy. A reacts in various ways.* (A: male person, B: female person)

1. A is **happy** together with B.

2. A applied for the same job, but didn't get it. A is **angry** at B, because B took away the job from A and on top of that talks about it happily.

3. A is **sad** because it means that B will be moving away soon. They were very good friends. B also becomes **sad** when he thinks about it.

4. A is **surprised** because B told him before that the job was very difficult to get. B is **proud** of hermself.

Scenario 4: "JobOffe". Scenario description: *B has just learned that she passed a job interview for a position that she wanted for a long time. She enters the kitchen and talks about it happily.* (A: white shirt, B: black shirt)

1. B's girlfriend broke up with him. B is very **sad** and talks about it. A has **sympathy** for B and tries to solace him.

2. B's girlfriend broke up with him. B is very **sad** and talks about it. A is **angry** at B's ex-girlfriend.

3. B broke up with his girlfriend. He is **happy** about it, because she got on his nerves since a long time. A is **bewildered** that B is not sad because of the break up.

4. B broke up with his girlfriend. He is **happy** about it, because she got on his nerves since a long time. A is **sad** because he was getting along with B's ex-girlfriend very well.

Scenario 5: "SplitUp". Scenario description: *B and his girlfriend broke up. B comes into the kitchen. A is there too.* (A: female person, B: male person)

1. A is **angry** that B arrives so late. B is sorry.

2. A is **angry** and points out with irony how late B is. B is textbfangry because of the irony.

3. B apologizes immediately, A is fine with it.

4. A is **surprised** that B finally arrived after all. B is sorry.

Scenario 6: "ZuSpaet". Scenario description: *A und B were supposed to meet at home to go climbing together. B is half an hour late. A is waiting in the kitchen, B comes in a hurry.* (A: male person, B: female person)

1. A is **surprised** and very **happy** that he can go on a trip with B for free.

2. A has exams during the time of the trip and can't go. Both are **disappointed** and **sad**.

3. A hates trips on a ship. B has already registered A for the trip. This makes A **angry**. A thinks that B should have known that he would rather not go on the sea trip.

4. B is **surprised** and **bewildered** that B wants to go on this trip, because it will be his exam time. However, A is still **happy** that B takes him on the trip.

Scenario 7: "ReiseGe". Scenario description: *B just learned that he won two tickets for a journey on a cruise ship. He is extremely **happy** about it and since he has an extra ticket he would like to take A with him. B goes to the kitchen to tell A about it.* (A: person on the left, B: person on the right)

Name	modalities	# views	# people	emotion annotations	naturalness of interaction	size	nature of data	language
USC Creative IT [7]	audio, video, mocap (full body)	3D positions	2	valence, activation, dominance	acted	19 actors, ca. 9 h	actors improvising on theatrical exercises and excerpts of plays	English
IEMOCAP [1]	audio, video, mocap (face and wrists)	3D positions (mocap)	2	valence, activation, dominance + discrete	acted	10 actors, ca. 12 h	scripted and improvisation-based, actors sitting	English
Vera am Mittag [5]	audio, video	1	2-5	none for full-body video part of dataset	natural	12 h	unscripted, authentic discussions between guests of a talk show	German
SEMAINE [6]	audio, video	2	2 (1 avatar)	valence, activation, power, anticipation, overall emotion intensity + optional dimensions	natural	100 recordings, approx 5 min each	user interacting with avatar in front of computer	English
Belfast naturalistic dataset [3]	audio, video	1	2	valence, activation and categorial annotations	natural	125 speakers, 298 clips (10-60 s each)	Interactive unscripted discourse, from TV shows and lab interactions	English
Metanet [2]	audio, video	5	4	categories: neutral, emphasized, negative (frustrated), enthusiastic	Spontaneous, guide-lined conversation	6 people, 5h	interactions in a meeting-room environment	English, Spanish and Catalan
MPIIEmo (ours)	audio, video (full body)	8	2	valence, activation, power, anticipation and categorial annotations	acted	14 actors, 196 clips, 125 min	Improvisation of coarsely scripted daily-life situations in a kitchen	German

Table 8: Comparison of different available datasets for studying affective bodily expressions in interactions. The datasets presented here are a subset of the ones listed in [4].

References

[1] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan. Iemocap: Interactive emotional dyadic motion capture database. *Language resources and evaluation*, 42(4):335–359, 2008. 5

[2] T. Butko, C. Nadeu, and A. Moreno. A multilingual corpus for rich audio-visual scene description in a meeting-room environment. In *ICMI workshop on multimodal corpora for machine learning: Taking Stock and Roadmapping the Future*, 2011. 5

[3] E. Douglas-Cowie, R. Cowie, and M. Schröder. A new emotion database: considerations, sources and scope. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 2000. 5

[4] A. Ěerekoviæ. An insight into multimodal databases for social signal processing: acquisition, efforts, and directions. *Artificial Intelligence Review*, pages 1–30. 5

[5] M. Grimm, K. Kroschel, and S. Narayanan. The vera am mittag german audio-visual emotional speech database. In *Multimedia and Expo, 2008 IEEE International Conference on*, pages 865–868. IEEE, 2008. 5

[6] G. McKeown, M. F. Valstar, R. Cowie, and M. Pantic. The semaine corpus of emotionally coloured character interactions. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 1079–1084. IEEE, 2010. 5

[7] A. Metallinou, C.-C. Lee, C. Busso, S. Carnicke, and S. Narayanan. The usc creativeit database: a multimodal database of theatrical improvisation. *Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality 18 May 2010*, page 55, 2010. 5