**Supplementary Material**

Sample Size Calculation

The sample size was calculated so that, for a given facial emotion recognition task (e.g., sadness at an intensity of 75 %), the fraction of correct recognitions with at least 50% confidence can be estimated with an acceptable precision. The precision is defined by the length of the 95% confidence interval. The sample size was determined by a resampling method. Each sample size, ni=1,…,91 = 10, …, 100, was evaluated by simulating R = 999 times ni individual patients thereby assuming a Bernoulli distribution with p = 50 % for providing the correct answer. In each round, the 95% confidence interval for the recognition rate was calculated according to Blaker ([Blaker, 2000](#_ENREF_1)), and it is assessed whether the length of the 95% confidence interval was below a pre-defined threshold. Based on these calculations, 45 patients should be recruited for the study. This sample size allows in more than 90% of 999 hypothetical repetitions of the study (i.e. with a power of 90 %) estimating the recognition rate with a 95% confidence interval smaller than 30%.

**Reference**

Blaker, H. (2000). Confidence curves and improved exact confidence intervals for discrete distributions. *Canadian Journal of Statistics, 28*, 783–798.