**Appendix S1: Supporting Methods and Results**

**Emotional experience in patients with clinically isolated syndrome and early multiple sclerosis**

Thomas D. Hälbig, MD PhD1,2

Torsten Wüstenberg, PhD3

René M. Giess, MD1

Hagen Kunte, MD PhD4

Judith Bellmann-Strobl, MD1,5

Klemens Ruprecht, MD PhD6

Friedemann Paul, MD PhD1,5

1 NeuroCure Clinical Research Center, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

2 Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

3 Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

4 Faculty of Natural Sciences, MSB, Medical School Berlin, Berlin, Germany

5 Experimental and Clinical Research Center, Max Delbrueck Center for Molecular Medicine and   
Charité – Universitätsmedizin Berlin, Berlin, Germany

6 Department of Neurology, Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany Germany

**Methods**

*Subjects*

MS disease modifying drugs were prescribed as follows in patients: glatiramer acetate (45% of patients), dimethyl fumarate (14%), interferon beta-1b (10%), natalizumab (10%), fingolimod (7%), teriflunomid (7%) subcutaneous interferon beta-1a (44 μg) (3%). 7% of patients were untreated.

*Study design*

The study was designed as a combined behavioral-fMRI-study (see ClinicalTrials.gov; NCT02695394). Altogether 58 participants (29 patients with MS; 29 HC) were evaluated in the present study.16 patients and 16 HC underwent fMRI during the emotional processing test. Based on previous fMRI-studies on emotional processing in MS, this number of participants was assumed to be sufficient to demonstrate potential fMRI differences between patients and controls (Jehna et al. 2011; Krause et al. 2009; Passamonti et al. 2009). The remaining participants (13 patients, 13 HC) were tested without fMRI and conducted the emotional processing experiment in a behavioral laboratory with a personal computer (PC). To account for this, we used the experimental setting as an additional factor in our statistical analyses (see main manuscript).

The participants were randomized to undergo the testing in either the MRI- or the PC- lab-setting. The setting and procedures for all other assessments were identical for fMRI-subgroup and PC laboratory-subgroup. The fMRI-data currently are in the process of being analyzed and will be presented separately in a different manuscript.

*Neuropsychological background testing*

The following tests from the Brief Repeatable Battery of Neuropsychological Tests (BRB-N; Scherer et al. 2004) were used to test for neuropsychological deficits in verbal learning and memory (Selective Reminding Test; SRT), visuospatial learning and recall (10/36-Spatial Recall Test; SPART), attention (Paced Auditory Serial Addition Test; PASAT), sustained attention and speed of information processing (Symbol Digit Modalities Test; SDMT) and verbal fluency/executive functioning (Semantic Word List Generation; WLG, WORD). To test for fatigue, the Modified Fatigue Impact Scale (MFIS) (Fisk et al.1994) was used and to probe for depression, the Beck Depression Inventory (BDI-2; Hautzinger et al. 2009) was employed.

*Emotional processing testing*

For the recognition test that followed the initial presentation and rating of the “old” pictures, a second set of “new”, distracting stimuli was compiled. This set was composed according to the same criteria as described in the main manuscript. Valence and arousal values of all pictures used in this study are reported in Table S1. Means for the different valence and arousal levels for “old” and “new” pictures are reported in S2.

Further, in order to avoid a selection bias, two sets of pictures (“A”, “B”), each comprising “old” and “new” pictures was compiled according to the same picture selection criteria (Table S1). The use of sets “A” and “B” and the order of the sets “old” and “new” alternated randomly across participants and was fully counterbalanced. Mean valence and arousal values of sets “A” and “B” are reported in Table S2.

*Statistical analysis*

All analyses were conducted using the Statistical Package for Social Science version 23 (<https://www.ibm.com/analytics/spss-statistics-software>).

*Correlation analysis*

Associations between valence and arousal ratings (corrected for age, sex, BDI-2 and experimental setting) were analyzed by means of Spearman correlation analysis. Analyses were performed for HC and Patients separately for the sum scores of MFIS and SWLS, the positive and negative affect subscales of the PANAS and all eight subscales of the SF36 as well as the associated two sum scales of the SF36. Findings were considered as significant, if they passed a Bonferroni corrected p-threshold of p = .05/number of performed tests ([1(MFIS)+1(SWLS)+2(PANAS)+10(F36 sub- and sum-scales)] \* 2(groups) = 28 🡪 pBonf = .0018).

To test for group differences, we used a Bootstrap data resampling approach with 9999 resampling steps or 9999 boot samples. Based on the resulting distribution of Spearman correlation coefficients, we estimated the 95% CI borders. In case of non-overlapping CI’s, a group difference was assumed.

**Results**

*Emotional experience - Correlation analysis*

Table 3 shows the results of the Spearman correlation analysis. If only one of the two group-specific tests indicated a correlation at a threshold of p = .05 (uncorrected for multiple tests) or in case of non-overlapping 95% confidence intervals, the full report was generated for this subscale (also for the group without a significant correlation). Findings that fulfill the significance criteria are highlighted in bold.

Our correlation analyses revealed only for the **patient group** a **positive association** between **valence** ratings for **highly arousing, positive** pictures and **SWLS sum score**.

In addition, we found significant differences between **group specific associations** between **valence ratings** for

1. **high arousal, neutral** pictures and **SF36** subscale **emotional role function**(Pat: negatively associated; HC: positively associated),
2. **low arousal,** **negative** pictures and **SF36** subscale **physical functioning**(Pat: positively associated; HC: no association), and
3. **low arousal,** **negative** pictures and **SF36** sum score **physical health**(Pat: positively associated; HC: negatively associated).

These findings, suggest a link between life satisfaction and quality of life associated measures and emotional experience induced by emotional visual stimuli.

*Recognition Memory Performance*

Our analyses revealed no effects for “old” pictures but an interaction AROUSAL x EXPERIMERNTAL SETTING for “new” pictures (F(1,51) = 4.10, p = .048, 2 = .074). However, in post-hoc tests, this interaction did not become significant (Table 2, Table S3).

*Modulation of Response time*

Response times were modulated by experimental manipulations only for “old” but not for “new” pictures (Table 2, Table S3):Analyses revealed for ”old” pictures a main effect of EXPERIMENTAL SETTING (F(1,51) = 8.58, p = .005, 2 = .144), an interaction AROUSAL by EXPERIMENTAL SETTING (F(1,51) = 6.53, p = .014, 2 = .113) as well as an interaction AROUSAL x GROUP x EXPERIMENTAL SETTING (F(1,51) = 4.15, p = .047, 2 = .075).

*Post-hoc* tests show allover longer response times in the MRI compared to the laboratory PC-setting (p = .004). Moreover, the observed response time increase in MRI depends additionally on picture AROUSAL (Low: p = 0.28, High: p < .001).