Supplementary information for "Elevated moral condemnation of third-party violations in multiple sclerosis patients"

Indrajeet Patil[1], Liane Young[2], Vladimiro Sinay[3], Ezequiel Gleichgerrcht[4,*]

[1]Scuola Internazionale Superiore di Studi Avanzati, Neuroscience Sector, Trieste, Italy.

[2]Boston College, Chestnut Hill, MA, United States

[3]Institute of Cognitive Neurology, Buenos Aires, Argentina

[4]Department of Neurology, Medical University of South Carolina, Charleston, SC

*Corresponding Author:

Ezequiel Gleichgerrcht, MD
Department of Neurology
Medical University of South Carolina
96 Jonathan Lucas St. Suite 301
Charleston, SC 29425
Email: gleichge@musc.edu

Supplementary Text S1: Empathy and Alexithymia measurement

Empathy

The Interpersonal Reactivity Inventory (IRI) (Davis, 1983) is a widely used self-report questionnaire to assess specific aspects of dispositional empathy. Participants reported agreement with statements on a 5-point Likert scale (1: *never true for me*, 5: *always true for me*). The four subscales consisted of: (1) fantasy (F) scale, which measures the propensity to identify with fictional characters (e.g., "I daydream and fantasize, with some regularity, about things that might happen to me."); (2) perspective taking (PT) scale, which indexes the dispositional tendency to consider others' perspective in interpersonal interactions (e.g., "I sometimes find it difficult to see things from the "other guy's" point of view."; reverse-scored); (3) empathic concern (EC) scale, which measures the *other-oriented* tendency to experience feelings of warmth, compassion, and concern for other people (e.g., "I often have tender, concerned feelings for people less fortunate than me."); (4) personal distress (PD) scale, which measures the *self-oriented* tendency to feel personal unease and discomfort in reaction to the emotions of others (e.g., "Being in a tense emotional situation scares me."). We used the version of IRI validated for Spanish population (Pérez-Albéniz et al., 2003). One noteworthy departure in this version of IRI from the standard one is that the EC scale has 8 items, while the PD scale has 6 items because one item ("When I see someone get hurt, I tend to remain calm."; reverse-scored) from original PD scale was found to load on EC scale in the validation study. As different subscales had different number of items, we used mean (instead of sum) of items as a measure of empathy.

Alexithymia

To assess levels of trait alexithymia, we used the validated Spanish version of the Toronto Alexithymia Scale-20 (TAS-20) questionnaire (Bagby, Taylor, & Parker, 1994; Spanish version: Martínez Sánchez, 1996), which has been argued to be the best current measure overall for assessing alexithymia due to its sound reliability, validity, and broad

generalizability (Timoney & Holder, 2013). The TAS-20 is a 20-item scale that consists of three subscales: Difficulty Describing Feelings (DDF, 5 items, e.g., "I am often puzzled by sensations in my body."), Difficulty Identifying Feelings (DIF, 7 items, e.g., "I often don't know why I am angry."), and Externally-Oriented Thinking (EOT, 8 items, e.g., "I can feel close to someone, even in moments of silence."; reverse-scored). Items were rated using a 5-point Likert scale (1: *strongly disagree*, 5: *strongly agree*).

Internal reliability

Cronbach's alpha for subscales of TAS and IRI questionnaires split by group. (For notes on abbreviations, see Table 1 from main text). Group differences in Cronbach's alphas were investigated using `cocron` package in R (http://comparingcronbachalphas.org/) which implements inferential statistics on alphas.

Variable	Cronbach's alpha		Item count	$\chi^2(1)$	*p*
	HC	MS			
DDF	0.637	0.402	5	1.4986	0.2209
DIF	0.781	0.848	7	0.9113	0.3398
EOT	0.495	0.536	8	0.0512	0.8210
TAS-20	0.783	0.819	20	0.2742	0.6005
EC	0.805	0.531	8	5.3088	**0.0212**
PD	0.797	0.589	6	3.1787	0.0746
PT	0.725	0.541	7	1.7825	0.1818
F	0.843	0.615	7	5.3328	**0.0209**

As can be seen from the above table, although internal reliability was consistently lower for the MS group as compared to the HC group, it was within acceptable range. Internal reliability differed significantly between groups only for fantasy and empathic concern scale, but not on PT which was more crucial for the effects observed.

The only unacceptable value for alpha was observed for DDF subscale of TAS for the MS group. We ensured that the significant difference between groups was not due to poor internal reliability by re-computing DDF after dropping the item loading poorly ("I am able to describe my feelings easily"; reverse-scored) on this subscale and redoing the test. This

reanalysis upheld the original result: MS patients showed higher levels of DDF than HC group (MS = 9.97, HC = 8.05, $t(63.16) = 2.207$, $p = 0.031$, $BF_{10} = 1.885$).

Supplementary Text S2

The scenario-by-condition breakdown. The exact wording of the English version of scenarios can be found in Young et al. (2010) using the corresponding label. The Spanish translations of the scenarios are available from authors on request.

Scenario no.	Name	Condition
SC1	Popcorn	intentional
SC2	River	attempted
SC3	Coffee	neutral
SC4	Logan Airport	attempted
SC5	Vitamin	accidental
SC6	Safety town	intentional
SC7	Harness	attempted
SC8	Laboratory	neutral
SC9	Sesame	accidental
SC10	Meatloaf	attempted
SC11	Jellyfish	neutral
SC12	Veternerian	intentional
SC13	Latex	accidental
SC14	Mushrooms	neutral
SC15	Motorboat	accidental
SC16	Tracks	attempted
SC17	Igloo	intentional
SC18	Bike	accidental
SC19	Ham	intentional
SC20	Pool	neutral
SC21	Asthma	accidental
SC22	Chairlift	intentional
SC23	Iron	attempted
SC24	Bridge	neutral

Supplementary Text S3

Internal reliability for each scenario condition and for each type of moral judgment.

MS group and healthy controls showed poor internal reliability for appropriateness judgments on neutral and intentional harm conditions, respectively, and results from these conditions

should be interpreted while keeping in mind this limitation. All other cells showed satisfactory internal reliability.

Judgment	Group	neutral	accidental	attempted	intentional
Appropriateness	HC	0.678	0.612	0.682	0.398
	MS	0.395	0.549	0.810	0.516
Punishment	HC	0.575	0.624	0.860	0.849
	MS	0.623	0.553	0.779	0.727
Egocentric	HC	0.807	0.836	0.798	0.767
	MS	0.783	0.811	0.848	0.861

Note that it can be argued that the lack of difference between appropriateness of neutral and accidental scenarios was due to the low internal reliability for appropriateness judgments (see Table 3 in the main text). This explanation seems unlikely since the same pattern was observed not only for appropriateness judgment but also for punishment judgments, which were associated with high internal reliability. We therefore maintain that the result for appropriateness judgments may reflect genuine overreliance on intent information in the current sample, which led to especially lenient judgments of accidental harms.

Supplementary Text S4

Average moral judgments about appropriateness, punishment, and moral relativism for MS and HC groups for individual items belonging to neutral, accidental, attempted, and intentional harm conditions.

Scenario	Group	Appropriateness	Punishment	Egocentric
SC1	HC	6.632	4.842	68.58
	MS	6.605	5.974	79.66
SC2	HC	6	4.079	56.92
	MS	6.053	5.289	76.68
SC3	HC	2.368	1.526	77.29
	MS	2.263	1.842	83.24
SC4	HC	6.421	5.658	67.79
	MS	6.447	5.895	85.5
SC5	HC	2.263	1.368	65.89
	MS	3.158	2.553	79.92
SC6	HC	6.921	5.789	79.58
	MS	6.763	6.368	90.97
SC7	HC	6.5	5.053	68.71
	MS	6.263	5.842	86.42
SC8	HC	1.868	1.474	84.92
	MS	1.974	1.921	87.08
SC9	HC	1.711	1.684	77.97
	MS	1.816	1.526	86.97
SC10	HC	6.5	5.447	71.95
	MS	6.553	5.842	87.5
SC11	HC	1.868	1.368	82.13
	MS	2.632	1.974	81.18
SC12	HC	6.974	6.579	88.47
	MS	6.789	6.368	92.21
SC13	HC	1.368	1.263	81.29
	MS	1.816	1.579	90.39
SC14	HC	3.053	2.079	71.24
	MS	3.842	3.053	73.11
SC15	HC	2.395	1.737	76.5
	MS	3.579	2.947	72.37
SC16	HC	6.5	5.711	73.76
	MS	6.447	5.921	82.61
SC17	HC	5.763	4.868	68.53
	MS	6.368	5.684	84.47
SC18	HC	2.237	2	66.89
	MS	2.184	2.053	77.71
SC19	HC	6.026	4.763	70.03
	MS	6.658	5.789	87.74
SC20	HC	2.289	1.447	77.63
	MS	2.605	1.947	77.16
SC21	HC	3.211	2.053	59.74
	MS	3.605	2.895	71.45
SC22	HC	6.737	6.211	80.82
	MS	6.684	6.895	92.53
SC23	HC	5.421	4.079	67.37
	MS	5.974	4.5	76.42
SC24	HC	2.105	1.368	81.53
	MS	2.395	1.895	79.71

Supplementary Text S5

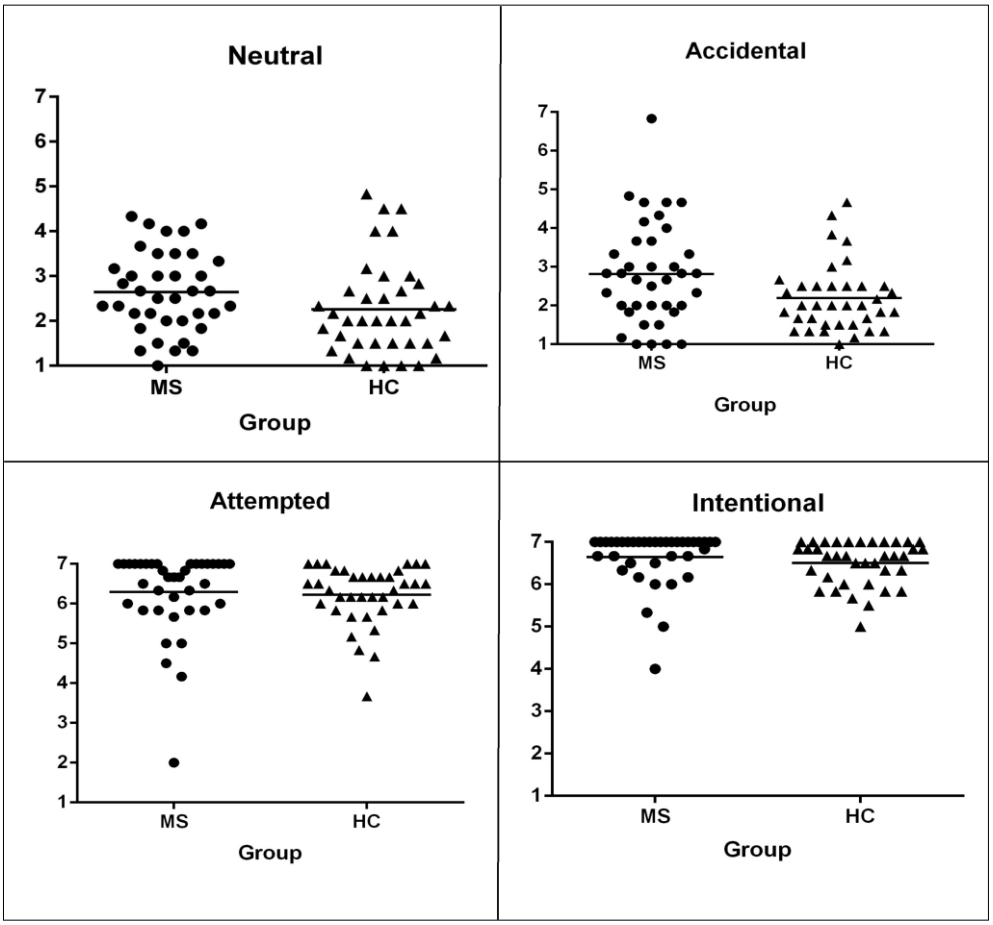
In this section, we provide:

(1) Scatter-plots for moral judgments for MS and HC groups. Higher score on the Y-axis represents more inappropriateness ratings, more severe punishment, and more morally egocentric moral judgments. The horizontal bar in the midst of the data represents mean value for the group.

(2) Details from planned comparisons carried out to decompose the main effect of group for each type of moral judgment.

- *Appropriateness judgments*

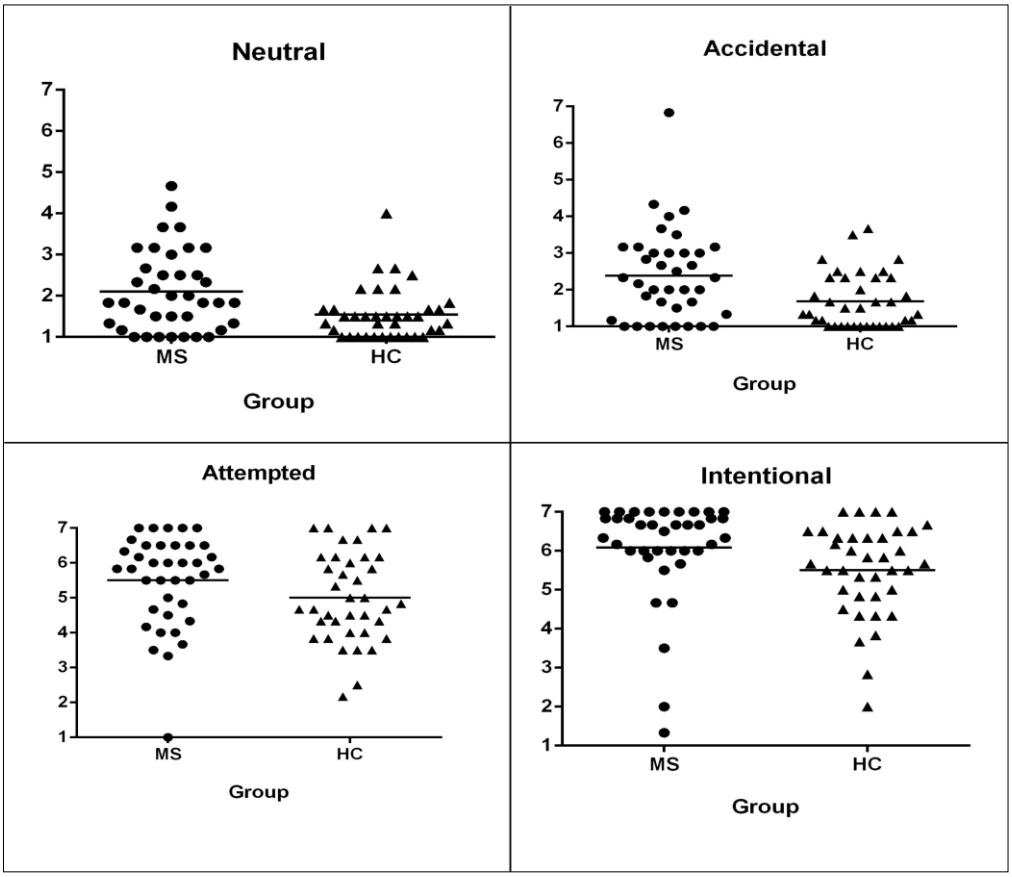
Bonferroni-corrected post-hoc comparisons to assess main effect of group revealed no significant differences.



- *Punishment judgments*

Bonferroni-corrected post-hoc comparisons to assess main effect of group revealed significant differences for all but attempted harm scenarios (p = 0.280, BF_{10} = 0.996). MS patients endorsed more severe punishment as compared to the control group not only for accidental (mean difference = 0.575, 95% CI [0.172, 0.977], p = 0.024, BF_{10} = 7.131) and intentional (mean difference = 0.671, 95% CI [0.163, 1.180], p = 0.040, BF_{10} = 4.383) harm cases, but also for neutral (mean difference = 0.561, 95% CI [0.186, 0.936], p = 0.016, BF_{10} = 9.828) cases.

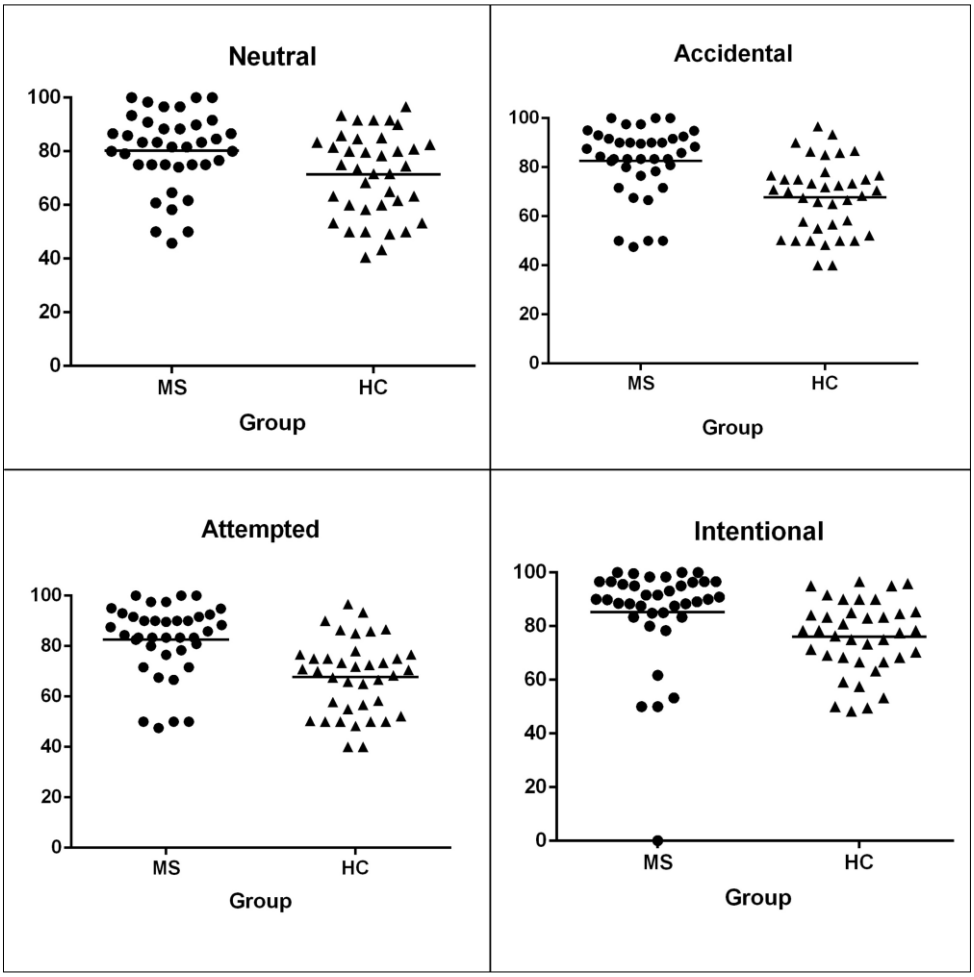
To explore the source of this effect, we repeated this mixed ANOVA analysis separately for each independent variable that differed between two groups, viz. EC, PD, DDF, DIF, and EOT (see Table 1). All group differences remained significant in ANCOVAs with EC, PD, DDF, and DIF as covariates. But in ANCOVA with EOT as the covariate, all significant group differences vanished, as shown by Bonferroni-corrected post-hoc *t*-tests for neutral (p = 0.344, BF_{10} = 1.378), accidental (p = 0.684, BF_{10} = 0.928), attempted (p = 0.244, BF_{10} = 1.278), and intentional (p = 0.196, BF_{10} = 1.747) harm cases. In other words, once the differences in EOT were accounted for, MS patients as a group were no more punitive than the controls.



- *Egocentricity judgments*

Planned comparisons to explore the main effect of group for egocentricity judgments were carried out with Bonferroni-corrected comparisons. This analysis showed that groups comparisons were not significant for neutral cases and only marginally significant for accidental harm cases (neutral cases: $p = 0.744$, $BF_{10} = 0.249$; accidental harm: $p = 0.060$, $BF_{10} = 3.210$), but both group comparisons were significant for the negative belief (attempted harm: mean difference = 14.772, 95% CI [8.148, 21.396], $p < 0.001$, $BF_{10} = 634$; intentional harm: mean difference = 11.930, 95% CI [5.811, 18.049], $p < 0.001$, $BF_{10} = 111$).

As with punishment judgments, we carried out mixed ANCOVA separately for each covariate of interest (EC, PD, DDF, DIF, and EOT) to explore possible source of this effect. This analysis showed that significant group differences in egocentric moral judgments for attempted ($p = 0.175$, $BF_{10} = 1.083$) and intentional ($p = 0.314$, $BF_{10} = 0.714$) harm cases disappeared only when EOT was added as a covariate to the model.


Neutral
100
80
60
40
20
0
MS
HC
Group
Accidental
100
80
60
40
20
0
MS
HC
Group
Attempted
100
80
60
40
20
0
MS
HC
Group
Intentional
100
80
60
40
20
0
MS
HC
Group

Supplementary Text S6: Correlation analysis between independent and dependent variables

- ***Correlation analysis: Alexithymia and moral judgments***

Correlation analysis between TAS scores and moral judgments about appropriateness and punishment ($\alpha = 0.05/4 = 0.0125$; for each judgment separately).

Variable	**statistic**	**HC** (n = 38)	**MS** (n = 38)
Punishment for neutral cases	ρ	0.056	-0.041
	p	0.737	0.809
Punishment for accidental harm	ρ	0.167	-0.102
	p	0.315	0.541
Punishment for attempted harm	ρ	-0.052	-0.137
	p	0.758	0.411
Punishment for intentional harm	ρ	-0.174	-0.134
	p	0.297	0.423
Appropriateness of neutral cases	ρ	0.191	0.063
	p	0.250	0.709
Appropriateness of accidental harm	ρ	0.123	-0.260
	p	0.463	0.115
Appropriateness of attempted harm	ρ	-0.169	0.034
	p	0.312	0.840
Appropriateness of intentional harm	ρ	-0.261	-0.291
	p	0.114	0.076

- ***Correlation analysis: Empathy and moral judgments***

Spearman's rho between IRI subscales for PT, EC, and PD and moral judgments about appropriateness and punishment for each group ($\alpha = 0.05/4 = 0.0125$; for each judgment separately). This analysis revealed that higher empathic concern was associated with higher punishment and inappropriateness ratings for attempted harm scenarios in controls, but not in MS population.

Variable	statistic	PT		EC		PD	
		HC (*n* = 38)	**MS** (*n* = 38)	**HC** (*n* = 38)	**MS** (*n* = 38)	**HC** (*n* = 38)	**MS** (*n* = 38)
Punishment for neutral cases	*ρ*	0.026	-0.066	0.286	0.041	0.042	0.165
	p	0.879	0.693	0.081	0.806	0.801	0.322
Punishment for accidental harm	*ρ*	-0.064	-0.044	-0.254	-0.034	-0.144	-0.070
	p	0.701	0.795	0.124	0.838	0.387	0.675
Punishment for attempted harm	*ρ*	-0.012	0.111	**0.499**	-0.040	0.295	-0.171
	p	0.943	0.509	**0.001**	0.812	0.073	0.304
Punishment for intentional harm	*ρ*	0.007	0.008	0.322	-0.143	0.129	0.108
	p	0.965	0.960	0.049	0.392	0.440	0.519
Appropriateness of neutral cases	*ρ*	-0.106	-0.158	0.121	0.126	0.067	0.134
	p	0.526	0.343	0.470	0.451	0.691	0.421
Appropriateness of accidental harm	*ρ*	-0.106	-0.207	-0.015	-0.188	-0.049	-0.192
	p	0.525	0.213	0.930	0.259	0.769	0.249
Appropriateness of attempted harm	*ρ*	0.092	-0.046	**0.477**	-0.046	0.400	0.119
	p	0.582	0.786	**0.002**	0.782	0.013	0.475
Appropriateness of intentional harm	*ρ*	0.045	-0.335	0.119	-0.391	-0.043	-0.184
	p	0.787	0.040	0.477	0.015	0.799	0.269

- ***Correlation analysis: Clinical variables, alexithymia, empathy, and moral judgments***

Spearman's rho between clinical variables denoting disease severity with alexithymia scores, empathy subscales, and moral judgments. Since this analysis was restricted only to the MS group (*n* = 38), we did not correct for multiple comparisons to avoid too stringent thresholds for a small sample size.

The only significant correlated observed was between EDSS and personal distress, i.e. higher scores on self-reported disability was associated with greater distress in dealing with anxiety-provoking social situations.

variable	statistic	MSSS	EDSS	No. Of Relapses	Disease duration
DDF	ρ	.115	.187	.163	.160
	p	.491	.260	.330	.338
DIF	ρ	.171	.315	.262	.286
	p	.304	.054	.112	.082
EOT	ρ	-.032	-.071	.058	-.002
	p	.850	.674	.729	.990
TAS-20	ρ	.095	.176	.204	.213
	p	.569	.289	.220	.200
IRI_PT	ρ	.122	.149	-.136	.012
	p	.467	.372	.417	.942
IRI_EC	ρ	.168	.239	.113	.126
	p	.315	.148	.500	.452
IRI_PD	ρ	.247	**0.326**	-.022	.187
	p	.135	**.046**	.898	.260
Punishment for neutral cases	ρ	-.203	-.126	-.069	.278
	p	.221	.453	.680	.091
Punishment for accidental harm	ρ	-.289	-.278	-.234	.083
	p	.079	.091	.157	.619
Punishment for attempted harm	ρ	.154	.008	-.217	-.212
	p	.357	.963	.192	.200
Punishment for intentional harm	ρ	.180	.077	-.215	-.249
	p	.280	.647	.194	.131
Appropriateness of neutral cases	ρ	-.220	-.156	-.075	.148
	p	.185	.349	.653	.374
Appropriateness of accidental harm	ρ	-.302	-.289	-.090	.098
	p	.066	.079	.590	.557
Appropriateness of attempted harm	ρ	.035	-.020	-.267	-.045
	p	.836	.907	.106	.787
Appropriateness of intentional harm	ρ	.068	-.020	-.032	-.079
	p	.684	.907	.851	.639

Supplementary Text S7: Correlation analysis between wrongness and punishment judgments

Spearman's *rho* between appropriateness and punishment moral judgments separately for healthy controls and MS patients for each condition. The difference in correlation values across groups was assessed using Fisher's *Z*-test (2-tailed). The difference was considered significant if $Z > 1.96$.

Correlation pair for condition	**statistic**	**HC** (n = 38)	**MS** (n = 38)	**Fisher's *Z***
neutral-neutral	ρ	0.738	0.66	0.641
	p	< 0.001	< 0.001	
accidental-acccidental	ρ	0.658	0.802	1.317
	p	< 0.001	< 0.001	
attempted-attempted	ρ	0.654	0.577	0.520
	p	< 0.001	< 0.001	
intentional-intentional	ρ	0.692	0.413	1.726
	p	< 0.001	.010	

References

Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The Twenty-Item Toronto Alexithymia Scale: II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, *38*, 33–40. doi:10.1016/0022-3999(94)90006-X

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, *44*(1), 113. doi: 10.1037/0022-3514.44.1.113

Pérez-Albéniz, A., De Paúl, J., Etxeberría, J., Montes, M. P., & Torres, E. (2003). Adaptación de interpersonal reactivity index (IRI) al español [Adaptation of Interpersonal Reactivity Index (IRI) in Spanish]. *Psicothema*, *15*(2), 267-272.

Timoney, L. R., & Holder, M. D. (2013). *Emotional Processing Deficits and Happiness: Assessing the Measurement, Correlates, and Well-Being of People with Alexithymia* (pp. 35-39). Dordrecht: Springer.

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*, 6753-6758. doi:10.1073/pnas.0914826107