Ⅲ. 研究成果の刊行物・別刷

Psychiatry and Clinical Neurosciences 2011; 65: 191–198 doi:10.1111/j.1440-1819.2011.02189.x

Regular Article

Functional magnetic resonance imaging study on the effects of acute single administration of paroxetine on motivation-related brain activity

Toshiyuki Marutani, MD,[1,2] Noriaki Yahata, PhD,[3,4] Yumiko Ikeda, MS,[3] Takehito Ito, PhD,[1] Manami Yamamoto, PhD,[1] Masato Matsuura, MD, PhD,[5] Eisuke Matsushima, MD, PhD,[2] Yoshiro Okubo, MD, PhD,[6] Hidenori Suzuki, MD, PhD[3] and Tetsuya Matsuda, PhD[1]*

[1]*Tamagawa University Brain Science Institute,* [2]*Section of Liaison Psychiatry & Palliative Medicine, Graduate School of Tokyo Medical & Dental University,* [3]*Department of Pharmacology, Nippon Medical School,* [4]*Department of Neuropsychiatry, Graduate School of Medicine, University of Tokyo,* [5]*Section of Biofunctional Informatics, Graduate School of Allied Health Sciences, Tokyo Medical and Dental University, and* [6]*Department of Neuropsychiatry, Nippon Medical School, Tokyo, Japan*

Aim: The aim of the present study was to investigate the effects of acute paroxetine administration on brain activity related to motivation.

Methods: Sixteen healthy subjects participated in a randomized, single-blind, no-drug/placebo-controlled, cross-over study. After administration of no drug, placebo or paroxetine (selective serotonin reuptake inhibitor; 20 mg), subjects underwent functional magnetic resonance imaging while performing a monetary incentive delay task. We analyzed the differences in brain activities of the reward anticipation/motor preparation period that are subject to motivational modulation. For this purpose, we subdivided the incentive trials on the basis of whether the reaction times (RT) were slower or faster than the subject's mean RT (slow RT and fast RT trials).

Results: No drug and placebo showed robust activation differences in the globus pallidus and putamen for the fast RT trials compared to the slow RT trials, whereas paroxetine showed none. Paroxetine showed significantly lower activations in the globus pallidus, insula, putamen and dorsolateral prefrontal cortex compared to no drug in the fast RT trials.

Conclusions: Paroxetine single acute administration diminished brain activity induced by motivation in healthy subjects. This may partially explain the increased lack of motivation seen in patients with relatively mild symptoms after taking a dose of paroxetine for the first time.

Key words: functional magnetic resonance imaging, motivation, paroxetine, reaction time, reward anticipation.

SELECTIVE SEROTONIN REUPTAKE inhibitors (SSRI) are first-line drugs for the treatment of major depressive disorder (MDD). MDD is characterized by disturbances in emotion, motivation and behavior in the presence of autonomic nervous symptoms.[1] A core symptom of MDD includes decreased motivation,[2,3] which SSRI sometimes rather aggravate in some patients.[4–6]

Motivational processing includes reward anticipation, motor preparation and related processes, including arousal and attention.[7,8] Several pharmacological functional magnetic resonance imaging (fMRI) studies have assessed the functions and/or mechanisms of SSRI related to motor, attention and reward. The effects of SSRI on motor function,[9,10] attention,[11] loss/no-loss comparison[12] and neural

Correspondence: Tetsuya Matsuda, PhD, Tamagawa University Brain Science Institute, 6-1-1, Tamagawa Gakuen, Machida, Tokyo 194-8610, Japan. Email: tetsuya@lab.tamagawa.ac.jp
Received 15 September 2010; revised 23 December 2010; accepted 23 December 2010.

processing of both rewarding and aversive stimuli[13] in healthy subjects have been studied by fMRI. McCabe reported that seven days of citalopram treatment diminished the brain activity induced by deliveries of rewards and aversive stimuli. They used primary rewards, chocolate taste and unpleasant strawberry taste as stimuli. Their conclusion indicated that the results could explain the experience of emotional blunting described by some patients during SSRI treatment.[13–15]

Decrease in motivation is also clinically observed after taking an initial dosing.[16] We have clinically observed some patients, especially patients with mild symptoms who reported decreased motivation after taking an initial dose of SSRI. Then, in the present study, we focused on the effects of an SSRI single acute administration on brain activity during motor preparation and reward anticipation, which are subject to motivational modulation. For this purpose, we used a monetary incentive delay (MID) task.[17] This task has been used in numerous reward-processing studies, and variations of the MID task have been used in a variety of other research.[18–20] Regardless of the details, the reward anticipation/motor preparation period and the subsequent button press during the task are essential. It is likely that the subject's motivations fluctuate over repeated trials of the MID task, and this is reflected in reaction time (RT). We expected that paroxetine would attenuate brain activity induced by motivation.

METHODS

Subjects

Sixteen healthy subjects participated in this study, but two were excluded because of an extremely low hit rate (less than 60%). Fourteen healthy subjects (eight men, mean age ± SD: 31 ± 3.8 years) were included in the final analysis. All subjects were native Japanese speakers and right-handed, as assessed by the Edinburgh Handedness Inventory. They filled out a questionnaire about their medical history and medications and were then interviewed by a medical staff member. They had no history of present or past psychiatric illnesses, neurological disorders, significant physical illnesses or head injuries, and no alcohol- or drug-related problems. They had not taken any types of medication for at least 1 day prior to scanning.

After a complete explanation of the study, including the possible side-effects of paroxetine, written informed consent was obtained from all the subjects and all the subject identifiers were removed. The protocol was approved by the local ethics committee.

Drug administration

We chose paroxetine as the SSRI for this study because it has the highest affinity for the human serotonin (5-HT) transporter among SSRI and other antidepressants according to radioligand binding assay studies[21–23] with a reported equilibrium dissociation constant (K_D) of 0.13 ± 0.01 nmol.

All subjects were examined after administration of paroxetine (S, 20 mg [minimally effective dose] paroxetine hydrochloride hydrate tablet), placebo (P, 12 mg lactobacillus bifidus tablet) or no drug (N) in a randomized, single-blind, no drug/placebo controlled, cross-over design. Three to 43 days (average 14.0 ± 13.7 days) passed between experiments. The order of drug administration was counterbalanced across subjects. The drug administration order consisted of six combinations (N-P-S, N-S-P, P-N-S, P-S-N, S-N-P, S-P-N) and we randomly assigned each combination to each subject.

The maximum drug concentration time (Tmax) of paroxetine 20 mg was reported to be 5.05 ± 1.22 h in healthy Japanese subjects.[24] Accordingly, placebo (P) and paroxetine (S) were given 5–5.5 h before initiating scanning to ensure maximum and stable plasma concentrations.

A previous positron emission tomography study suggests that 80% 5-HT transporter blockade is important for therapeutic effect of SSRI.[25] A single dosing of minimum therapeutic dose of an SSRI showed around 80% 5-HT transporter occupancy, which was almost the same as long-term dosing data.[26] Accordingly, a single dosing of paroxetine 20 mg of this study should have enough 5-HT transporter occupancy for therapeutic effect.

Reward task

Subjects performed an incentive task during functional scanning after a short pre-scanning training task. The task paradigm was an event-related design. The task was created with E-Prime 1.2 (Psychology Software Tools), which consisted of 98 7–8-s trials with 4-s inter-trial intervals (approx. 19 min. total). During each trial, subjects were shown one of three cue shapes (500 ms), a fixed crosshair during a variable delay (2500–3500 ms), and they responded with

a button press during the presentation of a gray square target (500 ms). They were then shown a fixed yellow crosshair (3000 ms) and this was followed by feedback (500 ms) notifying subjects if they had gained the points indicated by the cue, gained no points (= 0 point), or failed to press the button within 500 ms. The inter-trial interval was set to 4000 ms.

The cues signaled the possibility of no gain, 0 points ($n = 10$; denoted by a circle), 100 points ($n = 44$; denoted by a circle with one horizontal line) or 500 points ($n = 44$; denoted by a circle with three horizontal lines). There were three pseudorandom and predetermined orders of trials presented to subjects depending on experimental order, i.e. the combinations of medication and trial presentation order were counterbalanced.

Before scanning, subjects were instructed that the duration of target presentation was fixed to 500 ms but the button press limits differed from trial to trial. Fourteen 100-point cue and 500-point cue trials were predetermined to have a feedback of 0 points despite any efforts. In eight of these 28 trials, RT were not collected and were excluded from the analysis. The other trials required a fixed 500-ms time limit for the button press. If the subject did not respond in the appropriate interval, the message 'Press the button!' was displayed. We asked subjects to respond as quickly as possible to gain the maximum number of points, but the points earned were not reflected in the payment for participation in the study. Subjects were also asked to respond within the target presentation time even if the cue was a circle without line (potential 0 points). The total points earned were displayed at the end of the session.

During the original MID task,[17] RT were collected during the practice session so that the task difficulty level was set to achieve a success rate of 66%. However, we fixed the target duration to 500 ms so that the hit rate would reflect subjects' efforts more accurately. We also performed more trials to compare the effects of differences of RT in incentive trials. To maintain cue incentives, predetermined trials of gain cued with non-gain feedback were intermixed. Subjects were not told of their running point totals to minimize possible confounding effects.

fMRI data acquisition

The fMRI scans were acquired with a 3T Siemens MAGNETOM Trio Tim system scanner (Siemens, Erlangen, Germany). A total of 575 functional images were taken with a T2*-weighted gradient echo planer imaging sequence (TE = 25 ms; TR = 2000 ms; FA = 90°; matrix 64 × 64; FOV 192 × 192 cm) sensitive to the blood oxygenation level dependent (BOLD) contrast. Whole brain coverage was obtained with 34 axial slices (thickness 4 mm; in-plane resolution 3 × 3 mm).

Behavioral data analysis

For each drug condition for each subject, the mean RT to the target was calculated. Trials in which subjects did not press the button within the time limit were excluded from this calculation. Since the goal of this study was to investigate motivational motor preparation, we divided the RT of the incentive trials (100 and 500 points) on the basis of whether the RT were slower or faster than the subject's mean RT (RTslow and RTfast). For the purpose of this analysis, the 100- and 500-point trials were pooled to increase the sample set; there were no significant differences in hit rate or proportion of successful button presses among drug conditions for the different point trials. The mean RT of the slow RT and the fast RT trials were calculated, and these data were entered into a 3 (drug: non-drug, placebo, and paroxetine) × 2 (RT: slow and fast)-repeated-measures ANOVA using SPSS 16.0 J (SPSS Japan, Tokyo, Japan). The level of significance was set at 0.05.

fMRI data analysis

Image pre-processing and data analysis were performed with the statistical parametric mapping software package, SPM5 (Wellcome Department of Imaging Neuroscience, London, UK) running MATLAB 2007a (Mathworks, Natick, MA, USA). During pre-processing, the echo planer images were corrected for sequential slice timing, and all images were realigned to the first image to adjust for possible head movements. The realigned images were then spatially normalized to a standard Montreal Neurological Institute (MNI) template.[27] After normalization, all scans had a resolution of $2 \times 2 \times 2$ mm^3. Functional images were spatially smoothed with a 3-D isotropic Gaussian kernel (full width at half maximum of 8 mm). Low-frequency noise was removed by applying a high-pass filter (cut-off period = 192 s) and the default correction for AR1 auto correlation was performed for the fMRI time series at each voxel. A temporal smoothing function

was applied to the fMRI time series to enhance the temporal signal-to-noise ratio. Significant hemodynamic changes for each condition were examined using the general linear model with boxcar functions convoluted with a hemodynamic response function. Statistical parametric maps for each contrast of t-statistic were calculated on a voxel-by-voxel basis.

We then assessed the RT effect for each drug condition and the drug effect for the slow or fast RT during reward anticipation. We divided the trials into slow and fast RT trials, and we created the t-contrasts for the anticipation period between the offset of cue presentation and the onset of target presentation for the three different drug conditions in single-subject analysis (Nslow, Nfast, Pslow, Pfast, Sslow, Sfast).

A random effects analysis was performed to examine for population-wide effects. First, we used a 3 (medication: no drug, placebo and paroxetine) × 2 (RT: fast and slow) full factorial design to investigate brain activation between the different RT trials under each drug condition. There were significant activations for Nfast > Nslow in basal ganglia and primary motor cortex of which evident correlations have been revealed with reward anticipation[17,18,28,29] and motor preparation,[30] whereas activations for Pfast > Pslow and Sfast > Sslow were almost none. Then, to focus on regional activations in the reward anticipation and motor preparation-related areas in placebo and paroxetine, after paired *t*-tests were applied to Nfast > Nslow at the $P < 0.001$ level, uncorrected, with a voxel threshold of k = 10, we proceeded to a region-of-interest (ROI) analysis.

RESULTS

Behavioral data

The average hit rate of 90 trials was 92.9 ± 5.5%, 95.5 ± 4.2%, and 92.8 ± 6.7% for no drug, placebo, and paroxetine, respectively.

The average RT of all the trials and of the incentive (100 and 500 points) trials were 297.42 ± 38.69 ms and 294.02 ± 42.66 ms, 294.03 ± 40.31 ms and 290.45 ± 46.69 ms, 299.77 ± 38.96 ms and 298.08 ± 44.72 ms, under no drug, placebo and paroxetine conditions, respectively. There were no significant differences among three drug conditions.

Then we subdivided the RT of each incentive trial based on their relationship to the subject's mean RT, and the mean RT of each group, RTslow and RTfast, were compared for each drug treatment group.

A 3 (drug: non-drug, placebo, and paroxetine) × 2 (RT: slow and fast)-repeated-measures ANOVA revealed an effect of RT, $F_{1,13} = 398.73$, $P < 0.001$. Post hoc analyses with Bonferroni correction showed significant differences between RTslow and RTfast for each drug condition. RTslow and RTfast were 329.70 ± 27.04 ms and 258.34 ± 19.12 ms, 327.64 ± 31.61 ms and 253.25 ± 24.36 ms, 334.74 ± 30.04 ms and 261.43 ± 20.26 ms under no drug, placebo and paroxetine conditions, respectively. However, no significant differences were detected in the same RT (slow or fast) group among the different drug conditions.

fMRI data

The significantly activated areas for Nfast > Nslow were left primary motor cortex (T = 8.50), left globus pallidus (GP) (T = 5.95), right GP (T = 5.14), left dorsolateral prefrontal cortex (DLPFC) (T = 5.57), left transverse temporal gyrus (T = 5.26), right transverse temporal gyrus (T = 5.26), left thalamus (T = 4.87), right thalamus (T = 3.53), left insula (T = 4.71), right insula (T = 4.69), left putamen (T = 4.41), right putamen (T = 4.57), vermis (T = 4.50), right nucleus accumbens (NAcc) (T = 4.49) and left caudate (T = 4.27).

To investigate motivation-related areas under placebo and paroxetine conditions, we then performed a ROI analysis for the peak voxel of the regions significantly activated in Nfast > Nslow whole brain *t*-test. The ROI were selected based on previous fMRI studies of reward anticipation; GP,[28] insula,[27,29] putamen,[17,18,29] NAcc,[17] caudate,[29] DLPFC[31] and motor preparation; primary motor cortex.[30] The MNI coordinates [x y z] of ROI were left GP [−24 −10 0], right GP [20 −10 0], left insula [−38 −14 10], right insula [40 2 8], left putamen [−22 8 −2], right putamen [28 4 8], right NAcc [10 10 −14], left caudate [−6 12 4], left DLPFC [−36 32 26] and left primary motor cortex [−32 −22 54]. We collected beta values of each ROI and entered the data into 3 (drug conditions: N, P, S) × 2 (RT: slow, fast)-repeated-measures ANOVA using SPSS 16.0J. The level of significance was set at 0.05.

This ROI analysis using an ANOVA with repeated measures revealed a significant interaction between drug and RT in left insula ($F_{2,26} = 4.406$, $P = 0.022$), right insula ($F_{2,26} = 5.379$, $P = 0.011$), right NAcc ($F_{2,26} = 3.387$, $P = 0.049$), left primary motor cortex ($F_{2,26} = 4.016$, $P = 0.030$), a significant drug effect in

right insula ($F_{1,13} = 6.948$, $P = 0.021$) and left primary motor cortex ($F_{2,26} = 7.894$, $P = 0.002$), a significant RT effect in left GP ($F_{1,13} = 1.573$, $P < 0.0001$), right GP ($F_{1,13} = 37.957$, $P < 0.0001$), right insula ($F_{1,13} = 6.948$, $P = 0.021$), left putamen ($F_{1,13} = 45.757$, $P < 0.0001$), right putamen ($F_{1,13} = 13.968$, $P = 0.002$), right NAcc ($F_{1,13} = 5.755$ $P = 0.032$), left caudate ($F_{1,13} = 10.553$, $P = 0.006$), left DLPFC ($F_{1,13} = 10.568$, $P = 0.006$) and left primary motor cortex ($F_{1,13} = 38.179$, $P < 0.0001$).

Post hoc analysis with Bonferroni correction showed there were significant differences between paroxetine and placebo in only the left primary motor cortex, Pslow and Sslow ($P = 0.003$), Pfast and Sfast ($P = 0.008$), in which the activations were greater under placebo treatment (Fig. 1j). In the absence of drug or placebo treatment, the fast RT trials (Nfast) showed significantly higher activation than the fast RT trials of paroxetine condition (Sfast) in left GP ($P = 0.023$), left insula ($P = 0.008$), right insula ($P = 0.007$), right putamen ($P = 0.008$), left DLPFC ($P = 0.022$) and left primary motor cortex ($P = 0.003$) (Fig. 1a,c,d,f,i,j), which was not shown in comparison between the slow RT trials. There were no significant differences between placebo and no drug in any of the ROI.

Considering the way the ROI were defined, it was natural that there were significant differences between Nslow and Nfast in all the ROI. In paroxetine conditions, Sfast was significantly more activated than Sslow in only the left primary motor cortex (Fig. 1j). When subjects were given placebo, Pfast activation was greater than Pslow only in the left GP ($P = 0.045$), left putamen ($P = 0.007$) and left primary motor cortex ($P = 0.042$) (Fig. 1a,e,j).

DISCUSSION

Disturbances in motivation and motor activity are seen in MDD and these symptoms are sometimes exacerbated by SSRI in some patients. To investigate this paradoxical effect, we wish to use fMRI to monitor affected patients in response to drug therapy. However, as a first step, we studied normal subjects following a single dose of the SSRI paroxetine.

In this collection of normal subjects, there were no differences among the three drug conditions within each of the average RT of the whole, no incentive, incentive and subdivided RT trials. Thus, paroxetine administration did not affect subject behavior or performance globally. The average RT of no incentive and incentive trials showed no significance, which might be induced by the instruction for the subjects to press the button within the short duration of 500 ms even when the cue was no incentive.

Then, we investigated brain activations between the slow RT and the fast RT trials, which were behaviorally subdivided with significance, within treatment groups. The fast RT trials recruited greater activation in the GP, insula, putamen, NAcc, DLPFC, caudate and primary motor cortex than slow RT trials under no drug treatment. Under placebo conditions, the fast RT trials recruited greater activation in the GP, putamen and primary motor cortex. However, the paroxetine condition showed greater activations in the fast RT trials compared to the slow RT trials only in the primary motor cortex. These results indicated paroxetine desensitized RT influence on reward-anticipation-related brain activity, meanwhile no drug and placebo conditions reflected RT influence fully or partially in the reward-related areas.

In the next step, we looked into the activation differences in the same RT (slow or fast) group among the different drug conditions. Paroxetine significantly suppressed activation in the left GP, bilateral insula, right putamen and left DLPFC as reward-anticipation-related areas compared to no drug in the fast RT trials reflecting higher motivation, not in the slow RT trials reflecting lower motivation.

In the primary motor cortex, the activation under paroxetine administration was significantly weaker than no drug in the fast RT trials only, but weaker than placebo in both fast and slow RT trials. Besides, the fast RT trials were activated greater than the slow RT trials in all three drug conditions. Thus, the characteristics shown in the reward-related areas collapsed in the primary motor cortex, although paroxetine reduced activation compared to no drug and placebo in any case.

Taken together, paroxetine attenuated the brain activity in the reward-anticipation-related areas between the subdivided RT groups and compared to no drug in the more motivated fast RT trials. When anhedonia, one of the major symptoms of MDD, is considered as decreased motivation and sensitivity to rewarding experiences, our results suggest that a single dose of paroxetine may create a relatively anhedonic state in healthy subjects.

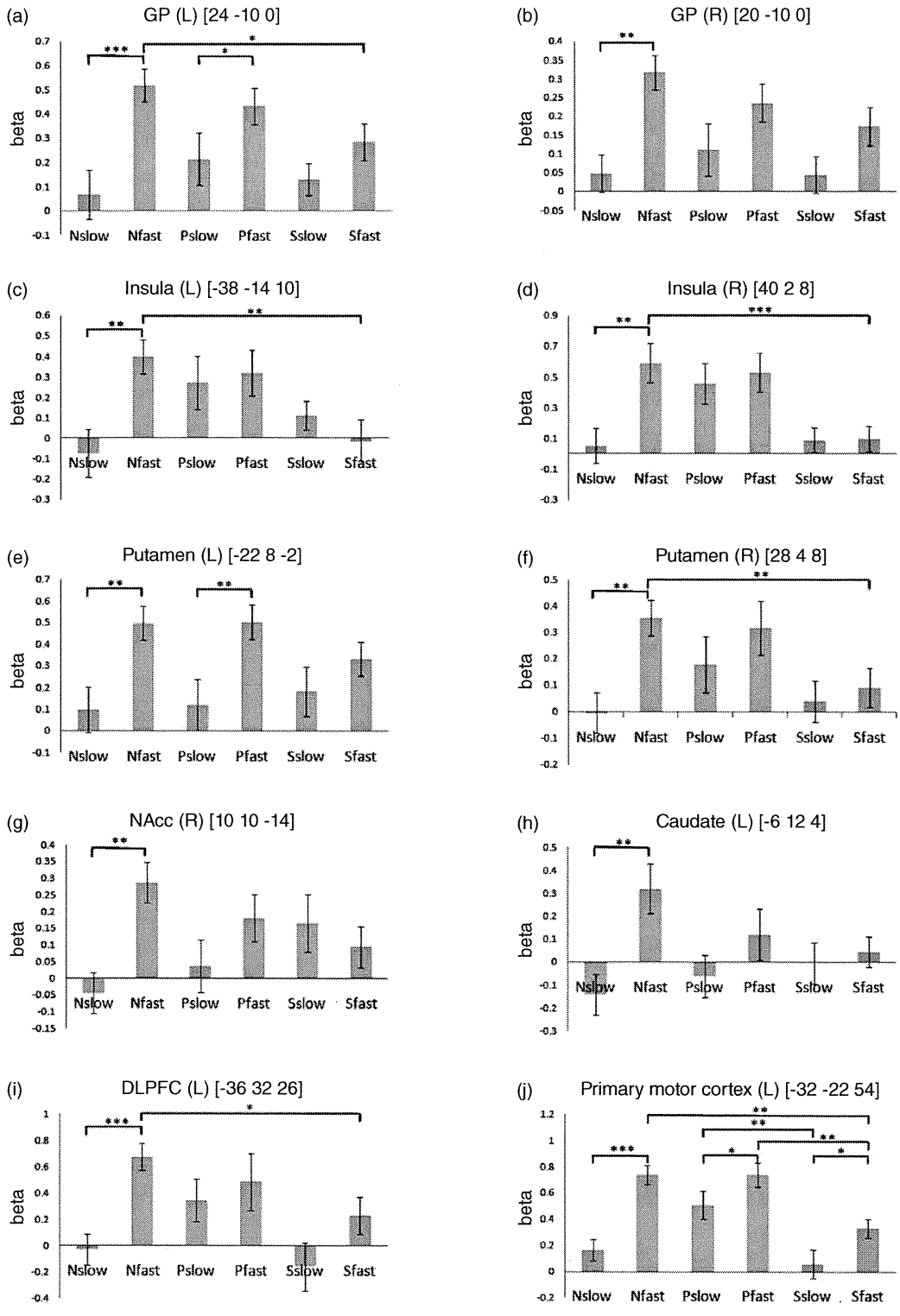

Figure 1. The mean beta values for the peak activation categorized by drug and reaction time type for the defined regions-of-interest. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Error bars show SEM. DLPFC, dorsolateral prefrontal cortex; GP, globus pallidus; L, left; NAcc, nucleus accumbens; R, right.

These results might partly come from the duration of drug administration because sufficient antidepressive effects of SSRI are not apparent normally until after 3–6 weeks of treatment. The increase in 5-HT produced by a single administration of SSRI not only stimulates the postsynaptic 5-HT receptors but also stimulates the somatodendritic inhibitory 5-HT_{1A} autoreceptors and presynaptic 5-HT_{1B} and 5-HT_{1D} autoreceptors. This varied activity could produce a net reduction in the activity of the 5-HT system.[32] Long-term treatment with SSRI induces desensitization/internalization of 5-HT autoreceptors, and this could lead to the downregulation of some postsynaptic receptors, such as the 5-HT_{2A} and 5-HT_{2C} subtypes. The end result of this process is thought to be a net activation of the 5-HT system.[32] Our results may partly arise from a net reduction of serotonin function by 5-HT autoreceptors produced by acute paroxetine administration.

We should briefly mention a relatively strong affinity of paroxetine for the norepinephrine transporter, $K_D = 40 \pm 2$ nmol[23] and muscarine receptor, $K_i = 72 \pm 3$ nmol/L,[22] but it is beyond the scope of the present study to examine the effects of paroxetine on these pathways.

In conclusion, paroxetine single acute administration diminished brain activity induced by motivation in healthy subjects. Our results may partially explain clinically observed decreased motivation seen in patients with relatively mild symptoms taking an initial paroxetine tablet dose of 10 or 20 mg for the first time. Further research is needed to clarify the effects of SSRI on brain activity with respect to cognitive and motor functions.

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of the members of the Tamagawa University Brain Science Institute. This work was supported by Grants-in-Aid for Scientific Research #19790838 and #18300275 for T. M. and Tamagawa University Global Center of Excellence (GCOE) program from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT).

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*, 4th edn. American Psychiatric Association, Washington, DC, 1994.
2. Naranjo CA, Tremblay LK, Busto UE. The role of the brain reward system in depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 2001; **25**: 781–823.
3. Treadway MT, Buckholtz JW, Schwartzman AN, Lambert WE, Zald DH. Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS ONE* 2009; 4: e6598.
4. Opbroek A, Delgado PL, Laukes C *et al.* Emotional blunting associated with SSRI-induced sexual dysfunction. Do SSRIs inhibit emotional responses? *Int. J. Neuropsychopharmacol.* 2002; **5**: 147–151.
5. Prince J, Cole V, Goodwin GM. Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *Br. J. Psychiatry* 2009; **195**: 211–217.
6. Shelton RC, Tomarken AJ. Can recovery from depression be achieved? *Psyichiatr. Serv.* 2001; **52**: 1469–1478.
7. Dickinson A, Balleine B. Motivational control of goal-directed action. *Anim. Learn. Behav.* 1994; **22**: 1–18.
8. Roesch MR, Olson CR. Neutral activity related to reward value and motivation in primate frontal cortex. *Science* 2004; **304**: 307–310.
9. Loubinoux I, Pariente J, Boulanouar K *et al.* A single dose of the serotonin neurotransmission agonist paroxetine enhances motor output: double-blind, placebo-controlled, fMRI study in healthy subjects. *NeuroImage* 2002; **15**: 26–36.
10. Loubinoux I, Pariente J, Rascol O, Celsis P, Chollet F. Selective serotonin reuptake inhibitor paroxetine modulates motor behavior through practice. A double-blind, placebo-controlled, multi-dose study in healthy subjects. *Neuropsychologia* 2002; **40**: 1815–1821.
11. Wingen M, Kuypers KP, van de Ven V, Formisano E, Ramaekers JG. Sustained attention and serotonin: a pharmaco-fMRI study. *Hum. Psychopharmacol.* 2008; **23**: 221–230.
12. Del-Ben CM, Deakin JFW, Mckie S *et al.* The effect of citalopram pretreatment on neuronal responses to neuropsychological tasks in normal volunteers: an fMRI Study. *Neuropsychopharmacology* 2005; **30**: 1724–1734.
13. McCabe C, Mishor Z, Cowen PJ, Harmer CJ. Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biol. Psychiatry* 2010; **67**: 439–455.
14. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: a clinical review. *J. Psychiatr. Pract.* 2004; **10**: 196–199.
15. Hasler G, Drevets WC, Manji HK, Charney DS. Discovering endophenotypes for major depression. *Neuropsychopharmacology* 2004; **29**: 1765–1781.
16. Beer MH, Porter RS, Jones TV. *The Merck Manual of Diagnosis and Therapy*, 18th edn. Merck Sharp & Dohme Corp, Whitehouse Station, 2006.
17. Knutson B, Adams CM, Fong GW, Hommer D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 2001; **21** (RC159): 1–5.

18. Dillon DG, Holmes AJ, Jahn AL, Bogdan R, Wald LL, Pizzagalli DA. Dissociation of neural regions associated with anticipatory versus consummatory phases of incentive processing. *Psychophysiology* 2008; **45**: 36–49.
19. Knutson B, Taylor J, Kaufman M, Peterson R, Glover G. Distributed neural representation of expected value. *J. Neurosci.* 2005; **25**: 4806–4812.
20. Scott DJ, Stohler CS, Egnatuk CM, Wang H, Koeppe RA, Zubieta JK. Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron* 2007; **55**: 325–336.
21. Owens MJ, Morgan WN, Plott SJ, Nemeroff CB. Neurotransmitter receptor and transporter binding profile of antidepressants and their metabolites. *J. Pharmacol. Exp. Ther.* 1997; **283**: 1305–1322.
22. Owens MJ, Knight DL, Nemeroff CB. Second-generation SSRIs: human monoamine transporter binding profile of escitalopram and R-fluoxetine. *Biol. Psychiatry* 2001; **50**: 345–350.
23. Tatsumi M, Groshan K, Blakely RD, Richelson E. Pharmacological profile of antidepressants and related compounds at human monoamine transporters. *Eur. J. Pharmacol.* 1997; **340**: 249–258.
24. Irie H, Fujita M, Inokawa Y, Narita H. Phase 1 clinical study of paroxetine HCl (Study 3): Pharmacokinetics after single oral administration of paroxetine HCl 10,20 and 40 mg to healthy adult male volunteers. *Jpn. Pharmacol. Ther.* 2000; **28**: S47–S68 (in Japanese).
25. Meyer JH, Wilson AA, Sagrati S *et al.* Serotonin transporter occupancy of five selective serotonin reuptake inhibitors at different doses: an [11C]DASB positron emission tomography study. *Am. J. Psychiatry* 2004; **161**: 826–835.
26. Suhara T, Takano A, Sudo Y *et al.* High levels of serotonin transporter occupancy with low-dose clomipramine in comparative occupancy study with fluvoxamine using positron emission tomography. *Arc. Gen. Psychiatry* 2003; **60**: 386–391.
27. Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. *Hum. Brain Mapp.* 1999; **7**: 254–266.
28. Tanaka S, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 2004; **7**: 887–893.
29. Knutson B, Fong GW, Adams CM, Varner JL, Hommer D. Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport* 2001; **12**: 3683–3687.
30. Carrillo-de-la-Peña MT, Galdo-Álvarez S, Lastra-Barreira C. Equivalent is not equal: primary motor cortex (MI) activation during motor imagery and execution of sequential movements. *Brain Res.* 2008; **1226**: 134–143.
31. Critchley HD, Mathias CJ, Dolan RJ. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* 2001; **29**: 537–545.
32. Cools R, Roberts AC, Robbins TW. Serotoninergic regulation of emotional and behavioural control processes. *Trends Cogn. Sci.* 2008; **12**: 31–40.

OPEN ACCESS Freely available online

PLoS one

Eye Gaze during Observation of Static Faces in Deaf People

Katsumi Watanabe[1,2,3]*, Tetsuya Matsuda[4], Tomoyuki Nishioka[5], Miki Namatame[5]

1 Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan, **2** Human Technology Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan, **3** Japan Science and Technology Agency, Saitama, Japan, **4** Brain Science Institute, Tamagawa University, Tokyo, Japan, **5** Department of Synthetic Design, Tsukuba University of Technology, Tsukuba, Japan

Abstract

Knowing where people look when viewing faces provides an objective measure into the part of information entering the visual system as well as into the cognitive strategy involved in facial perception. In the present study, we recorded the eye movements of 20 congenitally deaf (10 male and 10 female) and 23 (11 male and 12 female) normal-hearing Japanese participants while they evaluated the emotional valence of static face stimuli. While no difference was found in the evaluation scores, the eye movements during facial observations differed among participant groups. The deaf group looked at the eyes more frequently and for longer duration than the nose whereas the hearing group focused on the nose (or the central region of face) more than the eyes. These results suggest that the strategy employed to extract visual information when viewing static faces may differ between deaf and hearing people.

Citation: Watanabe K, Matsuda T, Nishioka T, Namatame M (2011) Eye Gaze during Observation of Static Faces in Deaf People. PLoS ONE 6(2): e16919. doi:10.1371/journal.pone.0016919

Editor: Jan Lauwereyns, Kyushu University, Japan

Received October 16, 2010; **Accepted** January 16, 2011; **Published** February 16, 2011

Funding: This research was supported by the Japan Science and Technology Agency and Grant-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) (21300301) and the Japan Science and Technology Agency. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kw@fennel.rcast.u-tokyo.ac.jp

Introduction

It has been hypothesized that deaf people may explore and see the visual world differently from hearing people because of their adaptation to hearing loss and/or consequential changes in communication strategy. Some studies have supported altered visual functions in deaf people, especially in the distribution and processes of visual attention [1]–[6].

Facial processing is considered to be one of the fundamental visual processes necessary for successful social interaction. This is because, for sighted people, facial processing constitutes a basic skill for detecting and recognizing other people's emotional states. A few studies have shown that facial processing in deaf people might differ from that of hearing people. For example, McCullough & Emmorey [7] showed that American deaf people are better at detecting subtle differences in facial features (particularly around the eyes and mouth) and suggested that long-term experience in discriminating grammatical facial expressions used with American Sign Language (ASL) and lip-reading may contribute to enhanced detection of nuances in relevant facial features (see also [8],[9]).

Since high spatial resolution visual processes are possible only at the fovea, humans produce a series of foveal fixations to extract visual information [10], which are closely linked with overt visual attention [11]. With regard to facial processing, studies investigating eye movements have consistently found a systematic fixation sequence in which the eyes are not directed equally to all regions of a face but only to selected parts; i.e., mainly the eyes and mouth [12]–[16].

Several studies that examined the eye movements of deaf people have found that they tend to look at facial regions in a similar magnitude as do hearing people [17]–[19]. For example, Muir & Richardson [17] conducted gaze-tracking experiments with deaf people watching sign language video clips and found that participants fixated mostly on the facial regions rather than on the hand movements of the signer, presumably to detect facial movements related to expression. In addition, Emmorey et al. [19] compared eye movements of beginning signers with experienced signers of ASL during ASL comprehension and found differences in fixation patterns: Beginning signers looked at facial regions around the signer's mouth while native signers fixated more on the areas around the eyes. Although these previous studies showed that there are minor differences in fixation patterns between certain groups, the sequence of fixation on the eyes and mouth has been considered to be a universal information extraction pattern.

Nevertheless, the idea of strictly universal facial processing has recently been challenged by several studies that investigated cultural influences on eye movements [20],[21]. Blais et al. [20] showed that Western Caucasian observers consistently fixated on the eye region and partially on the mouth area, confirming the triangular fixation pattern, whereas East Asian observers fixated more on the central region of the face (i.e., around the nose region). These results were interpreted by the authors in the context of cultural influences on visual environment affordance (analytic versus holistic processes [22]) and indicate that, even for face processing, strategies employed to extract visual information are shaped by experience (see also [23]–[27]).

Since hearing loss imposes significant constraints on everyday life, it is possible that deaf people use a visual strategy that is different from that used by hearing people, which might lead to differences in scan paths. The triangular pattern of scan path during observations of faces has not been examined quantitatively in deaf people. The purpose of the current study was therefore to report differential scan paths between deaf and hearing people. We chose an emotional valence evaluation task with static faces because it was easy to understand and perform by both deaf and hearing participants.

Because our participants were Japanese (i.e., East Asian), we expected to observe the generally dominant fixations on the central region of the face (i.e., around the nose region) [20]. Then, there were several possibilities, besides that of no difference in eye movement pattern between deaf and hearing participants. Firstly, since deaf people communicate with sign languages, manually signed languages, and/or lip-reading, the mouth region would be of importance for deaf people and therefore fixated more during face observation. Secondly, eye contact is an imperative component of communication and this is more so in a deaf community [28]. Hence, fixations in the eye region might be more pronounced in deaf people. Thirdly, visual processing in deaf individuals exhibit more emphasis on the peripheral visual field [2]–[6]. Therefore, in addition to the general tendency toward the nose region [20] [21], deaf individuals might make more eye movements in the parafoveal and peripheral regions, irrespective of whether the region is the main parts of faces (eyes and mouth) or not.

Methods

Ethics Statement

The procedures were approved by the internal review board of the Tsukuba University of Technology, and written informed consent was obtained from all participants prior to the testing.

Participants

We recruited 24 congenitally deaf Japanese people and 29 Japanese people with normal hearing function. Due to procedural failures during the experiment and/or spontaneous withdrawal from the study, data from 10 participants were excluded. The remaining participants comprised 20 congenitally deaf Japanese people (10 males and 10 females; mean age = 21.7 years, standard deviation = 0.75) and 23 Japanese people with normal hearing function (11 males and 12 females; mean age = 24.6 years, standard deviation = 3.11). All deaf participants were undergraduate students at Tsukuba University of Technology, where one of the entrance criteria is hearing loss of 60 dB or more. The deaf participants typically used manually signed Japanese and/or lip-reading for communication. None of the hearing participants were practiced in sign languages, manually signed languages, or lip-reading.

Stimuli

Stimuli were obtained from a commercially available database (the ATR face database DV99; ATR-Promotions, Inc.) and consisted of 4 male and 4 female Japanese identities expressing 10 different expressions (neutral [NE], fear [FE], happiness with the mouth opened [HO], happiness with the mouth closed [HC], sadness [SD], surprise [SP], anger with the mouth opened [AO], anger with the mouth closed [AC], disgust [DI], and contempt [CT]). The images were displayed on a 17-inch LCD monitor and viewed at a distance of about 55 cm, subtending 20 degrees of visual angle (27 cm) vertically and 36 degrees of visual angle (54 cm) horizontally. Each face image was centrally located and about 20 cm in height, which represents the size of a real face. Approximate positions of the eyes and mouth were aligned. Presentation of stimuli was controlled by Tobii Studio software (ver. 2.1.12, Tobii Technology, Stockholm, Sweden).

Eye tracking

Eye movements were recorded at a sampling rate of 60 Hz with the Tobii T-60 eye-tracker (Tobii Technology), which has an average gaze position error of 0.5 degrees and near-linear output over the range of the monitor used. Only the dominant eye of each participant was tracked although viewing was binocular. A manual calibration of eye fixations was conducted at the beginning of each session using a 9-point fixation procedure as implemented in the Tobii Studio software, and drift correction was performed for each trial.

Procedure

Participants were informed that they would be presented with a series of face pictures in order to evaluate the emotional valence of each face stimulus shown. Before each trial, participants were instructed to fixate on a cross at the center of the screen to perform an automatic drift correction. The participant initiated each trial by pressing a space bar. After a 2-s fixation period, a face was presented for 3 s. Then, the evaluation display appeared, and the participants used a computer mouse to click on the emotional valence of the face in the picture (out of a 7-point positive-negative scale with 1 being most positive). Evaluation was not speeded. Upon the participant's click of the mouse, the next trial began. A session consisted of 3 training trials with neutral expressions followed by 144 test trials. For the test trials, each combination of 8 identities and 9 expressions (except for NE) was presented once (72 trials), and each neutral expression of 8 identities was repeated 9 times (72 trials). The presentation of face stimuli was randomized.

Data analysis

The rating scores of emotional valence were first averaged for each expression by each participant. The mean rating scores were then grouped by the combination of hearing loss and participants' gender. A 3-way analysis of variance (ANOVA) was conducted to assess statistical significance, with hearing loss (deaf versus hearing) and participants' gender (male versus female) as between-group factors and facial expression of the stimulus as a within-group factor.

For each participant, we calculated the time that they fixated (fixation duration) and the number of fixations (fixation frequency) on the following areas of interests (AOIs): the eyes, the nose, and the mouth. AOIs were defined for each face (Figure 1). To control for differences in the sizes of AOIs, we normalized the fixation duration by the area of the AOI so that the sum of relative fixation duration would be 1 for each trial (relative fixation duration). The same normalization was performed for fixation frequency to calculate relative fixation frequency. Relative fixation duration and relative fixation frequency on the different AOIs were averaged separately for expressions within each participant. The averages were then grouped by combining hearing loss and participants' gender separately for AOIs. The relative fixation duration on AOIs was entered into a 4-way ANOVA, with hearing loss and participants' gender as between-group factors and facial expression and AOIs as within-group factors. The same ANOVA was conducted on the relative fixation frequency.

Results

Evaluation of emotional valence

The averaged rating scores of emotional valence are shown in Figure 2. Face stimuli with happy expressions (HO and HC) were

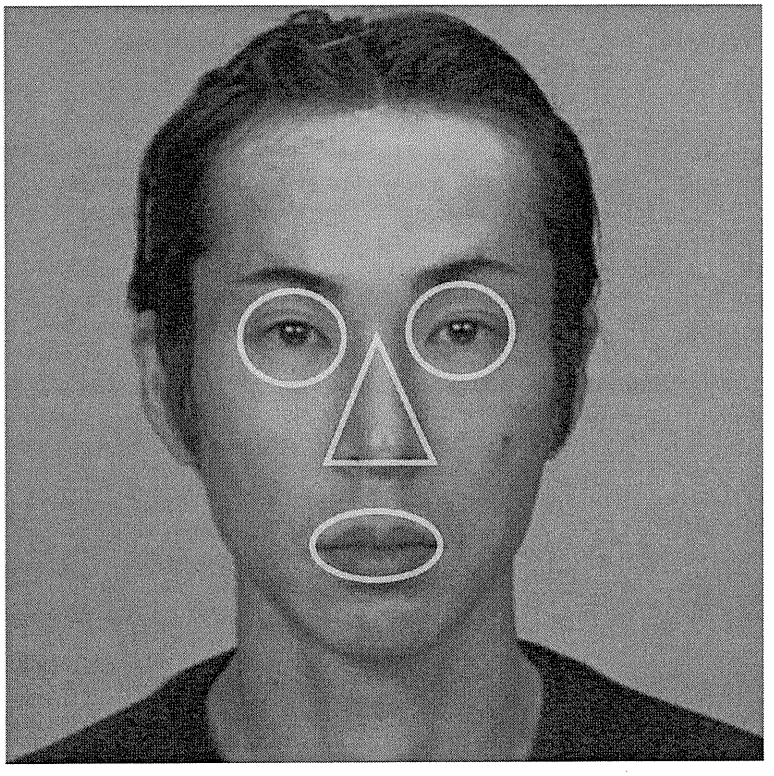

Figure 1. Example of areas of interest (AOIs). For each face stimulus, we defined AOIs: the eyes, nose, and mouth. In order to control for differences in sizes of AOIs, the fixation duration and fixation frequency were normalized by the AOI (relative fixation duration and relative fixation frequency) so that the sum of fixation duration and that of fixation frequency would be 1 for each trial.
doi:10.1371/journal.pone.0016919.g001

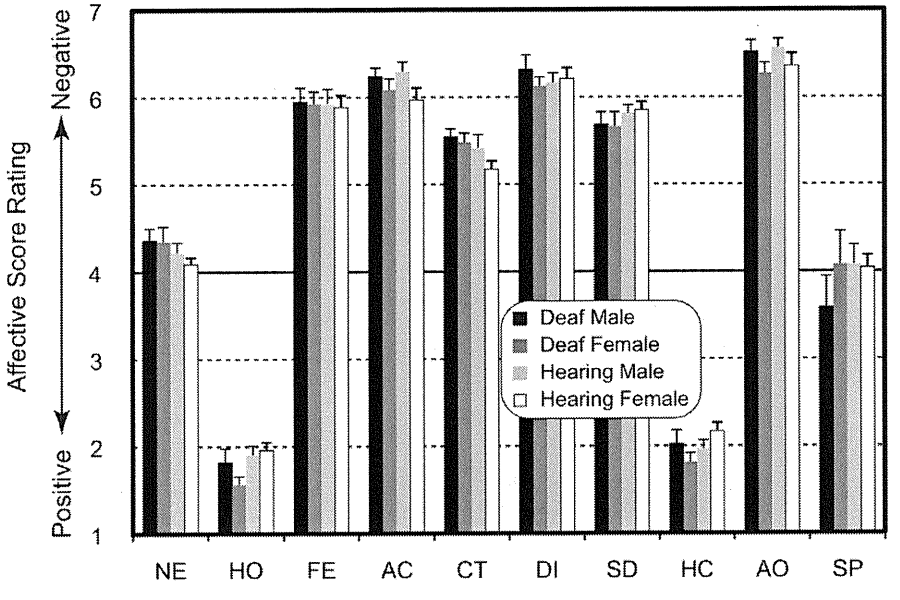

Figure 2. Mean rating scores of emotional valence as a function of expression in face stimuli. The face stimuli with happy expressions (HO and HC) were evaluated positively while the face stimuli with sad (SD), angry (AO and AC), disgust (DI), and contempt (CT) expressions were rated negatively. The faces with neutral (NE) and surprised (SP) expressions were evaluated neither positively nor negatively. NE = neutral; HO = happiness with the mouth opened; FE = fear, AC = anger with the mouth closed; CT = contempt; DI = disgust; SD = sadness; HC = happiness with the mouth closed; AO = anger with the mouth opened; and SP = surprise.
doi:10.1371/journal.pone.0016919.g002

evaluated positively while face stimuli with fear (FE), sad (SD), angry (AO and AC), disgust (DI), and contempt (CT) expressions tended to be rated negatively. Faces with neutral (NE) and surprised (SP) expressions were evaluated, on average, neither positively nor negatively. Three-way ANOVA showed that a main effect of expression ($F(9,351) = 597.7$, $P<0.001$) was significant while the main effects of hearing loss and participants' gender were not significant ($F(1,39 = 0.3$, $P = 0.60$, $F(1,39) = 1.2$, $P = 0.29$, respectively). No interaction reached a significant level ($F<1.3$, $P>0.23$). These results suggest that the participants evaluated the emotional valence of the faces presented as stimuli consistently, irrespective of hearing loss and participants' gender.

Eye movements

Data from trials where no gazes were directed at AOIs (i.e., the eyes, nose, or mouth) were excluded from the following analysis, which were 0.8%, 0.4%, 0.4%, and 0.5%, for deaf male, deaf female, hearing male, and hearing female groups, respectively. There was no significant difference in the number of discarded trials (Fisher exact test; $P = 0.32$). Figure 3 depicts the relative fixation duration mapped onto an example image of neutral face separately summed for deaf participants (Figure 3a), normal-hearing participants (Figure 3b), female participants (Figure 3c), and male participants (Figure 3d). This figure suggests that the gaze patterns differed among the participant groups.

Figure 4a shows relative fixation duration as a function of AOI, averaged over all facial expressions within different combinations of participant groups. In general, participants tended to fixate on the eyes and nose longer than on the mouth. In addition, deaf participants looked at the eyes longer than the nose whereas normal-hearing participants gazed at the nose longer than the eyes (Figure 4b). The tendency to fixate longer on the eyes appeared to be stronger in females compared with male participants (Figure 4c). Figure 5 shows the relative fixation duration for the different AOIs as a function of facial expression averaged over all participants. The differential relative fixation durations for different AOIs were apparent; i.e., fixation duration on the eyes and the nose was longer than on the mouth. In addition, the pictures of faces with neutral expressions appeared to lead to longer fixation on the eyes in exchange for shorter fixation duration on the mouth.

Four-way ANOVA revealed significant main effects of participants' gender ($F(1,39) = 8.7$, $P<0.01$; female > male), AOI ($F(2,78) = 26.8$, $P<0.001$; post hoc Ryan's method, eyes = nose > mouth, $P<0.001$), and expression ($F(9,351) = 5.6$, $P<0.001$).

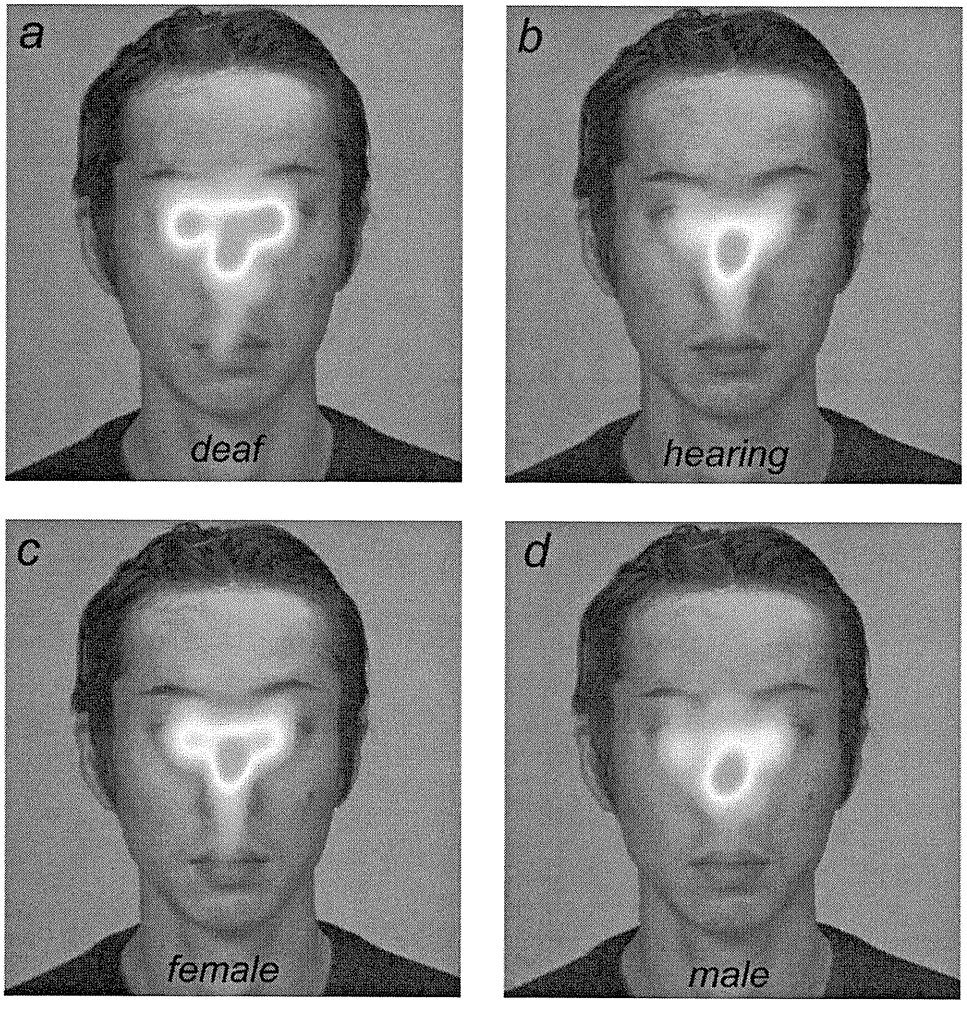

Figure 3. Total fixation duration mapped onto an example face image: (a) deaf participants, (b) hearing participants, (c) female participants, and (d) male participants. Red regions represent the places where the participants' eyes stayed longer. The fixation patters differed among the participant groups.
doi:10.1371/journal.pone.0016919.g003

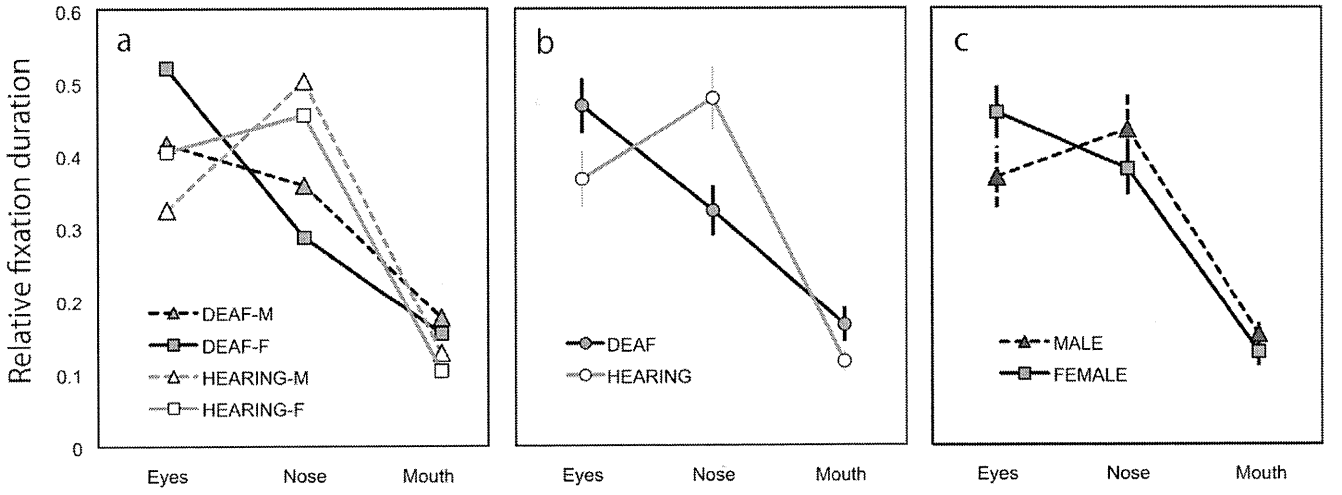

Figure 4. Relative fixation duration. (**a**) Relative fixation duration as a function of area of interest averaged over all facial expressions within different combinations of participant groups. Participants tended to fixate on the eyes and nose longer than on the mouth. (**b**) Relative fixation duration compared between deaf and hearing participants. The deaf participants looked at the eyes longer than the nose whereas the hearing participants gazed at the nose longer than the eyes. (**c**) Relative fixation duration compared between female and male participants. The female participants tended to fixate on the eyes longer than did the male participants.
doi:10.1371/journal.pone.0016919.g004

A significant interaction between hearing loss and AOI ($F(2,78) = 5.3$, $P<0.01$) and between expression and AOI ($F(18,702) = 2.8$, $P<0.001$) were found. There were also significant interactions between hearing loss and expression ($F(9, 351) = 2.0$, $P<0.05$) and among participants' gender, expression, and AOI ($F(18, 702) = 1.7$, $P<0.05$). Analyses of simple main effect indicated that the normal-hearing group looked at the nose longer than the eyes whereas the deaf group tended to look at the eyes more than the nose ($P<0.05$). The interaction between expression and AOI was mainly due to the fact that participants fixated longer on the eyes in pictures of neutral faces than in pictures of faces with other expressions ($P<0.05$). The data of relative fixation frequency corroborated the results of relative fixation duration (Figures 6 and 7).

An additional analysis was performed to test whether the fixation duration and fixation frequency outside the AOIs differ among participant groups. Whereas significant main effects of participants' gender were found (male > female; fixation duration, $F(1,39) = 13.6$, $P<0.01$; fixation frequency, $F(1,39) = 15.1$, $P<0.01$), no statistical difference was observed between deaf and hearing participants (fixation duration, $F(1,39) = 0.14$, $P = 0.7$; fixation frequency, $F(1,39) = 1.14$, $P = 0.29$), corroborating the results of 4-way ANOVA.

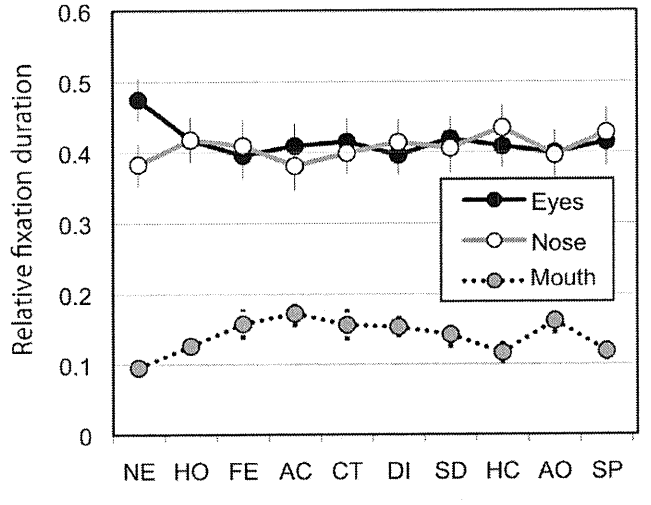

Figure 5. Relative fixation duration for the different areas of interest as a function of facial expression, averaged over all the participants. The faces with neutral expression led to longer fixation duration on the eyes. NE = neutral; HO = happiness with the mouth opened; FE = fear, AC = anger with the mouth closed; CT = contempt; DI = disgust; SD = sadness; HC = happiness with the mouth closed; AO = anger with the mouth opened; and SP = surprise.
doi:10.1371/journal.pone.0016919.g005

Discussion

In the present study we examined the possible difference in the pattern of eye movements between congenitally deaf and normal-hearing Japanese individuals while they evaluated the emotional valence of static faces. The results can be summarized as follows: (1) The emotional valence of face stimuli were evaluated consistently irrespective of hearing loss and participants' gender; (2) participants fixated (in terms of frequently and duration) on the eyes and nose more than on the mouth (the main effect of AOI), confirming overall fixation dominance on the eyes; (3) female participants tended to look at the main facial parts (i.e., the eyes, nose, and mouth) more than did male participants (the main effect of participant's gender); (4) faces with neutral expressions induced fixations on the eyes more than did faces with other expressions (the interaction between expression and AOI); and (5) deaf participants looked at the eyes more than the nose whereas normal-hearing participants tended to look more at the nose (the interaction between hearing loss and AOI).

It has been reported that females have an advantage in decoding nonverbal emotion [29]–[35] and that females look more at the main parts of the face than do males, with particular emphasis on the eyes [34],[35]. The main effects of participant's gender supported this notion. Although the interaction between participants' gender and AOI and the interaction among participants' gender, hearing loss, and AOI did not reach a significant level, our data clearly showed a tendency in the female

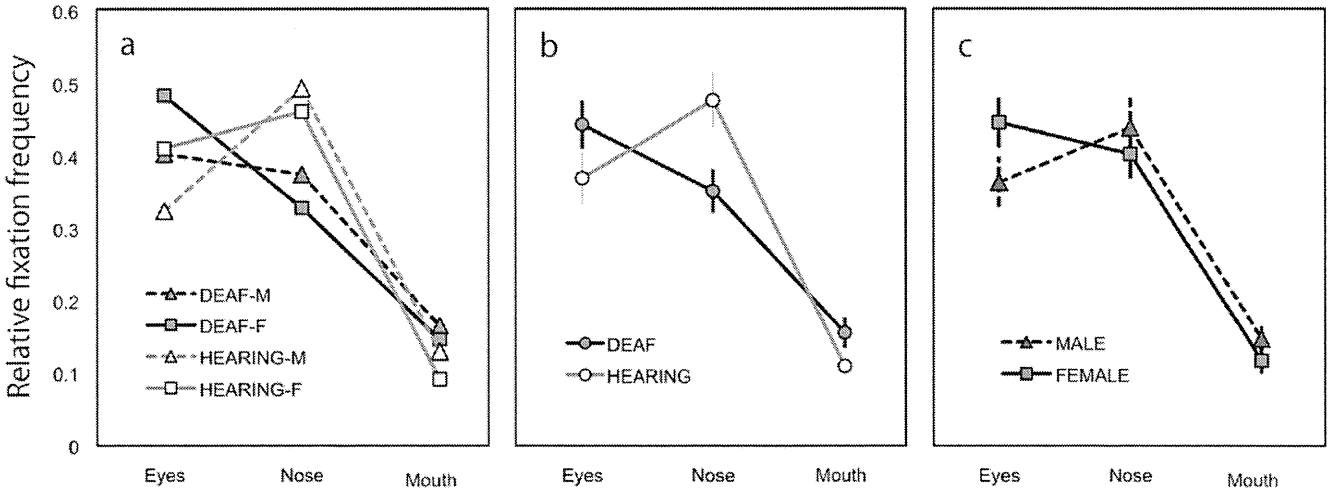

Figure 6. Relative fixation frequency. (**a**) Relative fixation frequency as a function of area of interest averaged over all facial expressions within different combinations of participant groups. (**b**) Relative fixation frequency compared between deaf and hearing participants. (**c**) Relative fixation frequency compared between female and male participants. The results for fixation frequency corroborated those of fixation duration.
doi:10.1371/journal.pone.0016919.g006

participants to fixate on the eyes (Figure 4c and Figure 6c). Thus, the present results may be taken as evidence supporting a gender difference in fixation pattern for faces with emotional expressions [34],[35].

Irrespective of participant group, faces with neutral expressions tended to produce more fixations on the eyes than did faces showing other expressions. Faces with emotional expressions have distinct features that help observers to interpret the expression. On the other hand, neutral faces are ambiguous and lack the visual cues for comprehension of emotion. It has been suggested that understanding and communication of emotion depends greatly on the visual processing of the eye region [36]–[39]. Therefore, it is possible that people fixate more on the eyes of ambiguous neutral faces in an attempt to discern emotional clues. However, it should

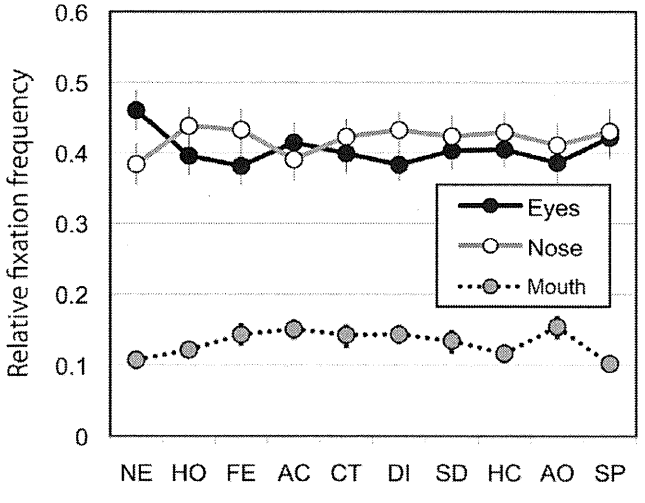

Figure 7. Relative fixation frequency for the different areas of interest as a function of facial expression averaged over all the participants. The results for fixation frequency corroborated those of fixation duration. NE = neutral; HO = happiness with the mouth opened; FE = fear, AC = anger with the mouth closed; CT = contempt; DI = disgust; SD = sadness; HC = happiness with the mouth closed; AO = anger with the mouth opened; and SP = surprise.
doi:10.1371/journal.pone.0016919.g007

be stated that there was a possible confound in the present experiment that the neutral faces were repeated 9 times while the others were presented once and the interaction between expression and AOI may be due to repetition rather than expression. Further investigations are warranted to examine whether less emotional facial expressions indeed lead to more fixation on the eye region. Specifically, a future study should avoid the possible confound between viewing less emotional face expressions and repeated viewing.

The main focus of the present study was to investigate potential differences in fixation patterns between deaf and hearing participants. The hearing participants in the present study looked at the nose (i.e., central) region most rather than at the eye region. Since all the participants in the current study were Japanese, this may be attributed to a cultural influence on eye movement. Blais et al. [20] reported that Western Caucasian observers consistently fixated the eye region, and partially the mouth, whereas East Asian observers fixated more on the central region of the face to extract information from faces. They hypothesized that this difference is due to the social norm in East Asian cultures that direct or excessive eye contact may be considered rude [40] and to the difference in cognitive strategy (holistic/analytic approach to visual information: [22],[41]). On the other hand, our Japanese deaf participants looked at the eye region most, closer to the fixation pattern of Western Caucasians in Blais et al [20]. It has been reported that in a deaf community, eye contact is vital for communication because avoiding eye contact disrupts communication more profoundly than it does in sighted communities [28]; this holds true for a Japanese deaf community. Therefore, it is possible that the increased fixation on the eye region in our Japanese deaf participants may reflect their communication strategy. In this sense, the present study may be taken as an extension of Blais et al. [20], showing that living in a specific community (more specifically, deaf community in Tsukuba University of Technology in Japan) might alter how we look at faces (also see [23]–[27]).

The underling mechanism for differential scan paths between deaf and hearing individuals remains to be clarified. However, one possible mechanism is the altered distribution and processes of visual attention [2]–[5]. Deaf individuals are more distracted by

visual information in the parafovea and periphery [5]. Since there was no difference in fixation duration and frequency outside the AOIs and no increase of fixation in the mouth region, the present finding cannot be explained solely by the attention emphasis on the peripheral processing. However, it is still possible that altered peripheral visual attention and scrutinizing strategy for faces may interact to produce the differential scan paths.

Limitations of the present study

Although the difference in fixation pattern was clear, it should be noted that the present study has considerable limitations. One limitation is that the stimuli used in the present study were static, rather than dynamic, stimuli. Many studies of emotional expression have used static face stimuli. Yet, facial expressions are highly dynamic, and thus, static stimuli represent unnatural snapshots of them. Recent studies on dynamic facial expressions have shown that visual processes for facial expressions are essentially tuned to dynamic information [36],[42],[43]. Evidence supporting this notion comes from facilitative effects of dynamic presentation on facial processing [44]–[50] and enhanced neural activities for dynamic, as opposed to static, face stimuli [51]–[53]. Therefore, it is likely that the pattern of results would be different if dynamic stimuli were used. In particular, the relatively less fixations in the mouth region might be due to the use of the static face stimuli. It has been shown that the mouth region conveys useful information for emotion discrimination [54]–[58], and this seems to be more so with dynamic face stimuli, e.g., [59].

Another limitation stems from the use of the evaluation task of emotional valence. Many previous studies have examined the scan paths during emotion discrimination and identification (e.g., [56],[57]) but little study has employed an evaluation task of emotional valence. Therefore, the present results may not be compared directly with those of the previous studies. Also, in order to elucidate the mechanism for valence evaluation and emotional processes, it is important to consider the relation between the time-course of evaluation processes and eye movement. The face stimuli used in the present study included some variations in visual information for emotional valence evaluation, which in turn would lead to different demands for different face stimuli. Since the decision was not timed, we did not know when the participants reached their decisions. Therefore, the eye movement pattern may reflect either pre-decision or post-decision processes or both.

The final limitation is the demographic peculiarity of the participants. It is possible that the use of sign language (Japanese Sign Language; JSL) leads to enhanced attention to the eye region because changes in eye configurations convey various syntactic distinctions and grammatical information in JSL as in ASL [60],[61]. However, until around 2002, most Japanese schools for the deaf emphasized oral education; i.e., teaching through lip-reading. Although manually signed Japanese (which is a signed form of the Japanese language) has recently started to be used in schools for the deaf, even now Japanese sign language is not officially taught. Therefore, it is difficult to infer whether the difference in fixation pattern is due to the hearing loss itself, to the extended use of sign language, and/or to the specific historical situation of Japanese deaf education.

Despite the above limitations, the present study showed the differential scan paths during observation of static face stimuli between deaf and hearing participants. Further investigations, preferably with speeded response or confidence/difficulty rating of decision, with dynamic stimuli, and with cross-cultural comparisons, will shed light on how and to what extent hearing loss influences how we look at faces and interpret others.

Acknowledgments

The use of materials generated through the use of the DB99 (Figure 1 and Figure 2) was licensed by the ATR Promotions.

Author Contributions

Conceived and designed the experiments: KW TM MN. Performed the experiments: MN. Analyzed the data: KW TN MN. Contributed reagents/materials/analysis tools: TM TN. Wrote the manuscript: KW MN.

References

1. Tharpe AM, Ashmead D, Sladen DP, Ryan HA, Rothpletz AM (2008) Visual attention and hearing loss: Past and current perspectives. Journal of the American Academy of Audiology 19: 741–747.
2. Stivalet P, Moreno Y, Richard J, Barraud PA, Raphel C (1998) Differences in visual search tasks between congenitally deaf and normally hearing adults. Cognitive Brain Research 6: 227–232.
3. Proksch J, Bavelier D (2002) Changes in the spatial distribution of visual attention after early deafness. Journal of Cognitive Neuroscience 14: 687–701.
4. Bavelier D, Dye MWG, Hauser PC (2006) Do deaf individuals see better? Trends in Cognitive Sciences 10: 512–518.
5. Dye MWG, Hauser PC, Bavelier D (2008) Visual skills and cross-modal plasticity in deaf readers: Possible implications for acquiring meaning from print. Annals of the New York Academy of Sciences 1145: 71–82.
6. Dye MWG, Hauser PC, Bavelier D (2009) Is visual selective attention in deaf individuals enhanced or deficient? The case of the useful field of view. PLOS One 4(5): e5640.
7. McCullough S, Emmorey K (1997) Face processing by deaf ASL signers: evidence for expertise in distinguishing local features. Journal of Deaf Studies and Deaf Education 2: 212–222.
8. Bettger J, Emmorey K, McCullough S, Bellugi U (1997) Enhanced facial discrimination: effects of experience with American sign language. Journal of Deaf Studies and Deaf Education 2: 223–233.
9. Kubota Y, Quérel C, Pelion F, Laborit J, Laborit MF, et al. (2003) Facial affect recognition in pre-lingually deaf people with schizophrenia. Schizophrenia Research 61: 265–270.
10. Yarbus AL (1965) Role of Eye Movements in the Visual Process. Moscow: Nauka.
11. Findlay JM, Gilchrist ID (2003) Active Vision - The Psychology of Looking and Seeing. Oxford, UK: Oxford University Press.
12. Walker-Smith G, Gale A, Findlay J (1977) Eye movement strategies involved in face perception. Perception 6: 313–326.
13. Janik SW, Wellens AR, Goldberg ML, Dell'Osso LF (1978) Eyes as the center of focus in the visual examination of human faces. Perceptual and Motor Skills 47: 857–858.
14. Groner R, Walder F, Groner M (1984) Looking at faces: local and global aspects of scanpaths. In: Gale AG, Johnson F, eds. Theoretical and applied aspects of eye movements research. Amsterdam: Elsevier. pp 523–533.
15. Henderson JM, Williams CC, Falk RJ (2005) Eye movements are functional during face learning. Memory & Cognition 33: 98–106.
16. Armann R, Bülthoff I (2009) Gaze behavior in face comparison: The roles of sex, task, and symmetry. Attention, Perception & Psychophysics 71(5): 1107–1126.
17. Muir LJ, Richardson IEG (2005) Perception of sign language and its application to visual communications for deaf people. Journal of Deaf Studies and Deaf Education 10: 390–401.
18. De Filippo CL, Lamsing CR (2006) Eye fixations of deaf and hearing observers in simultaneous communication perception. Ear & Hearing 27(4): 331–352.
19. Emmorey K, Thompson R, Colvin R (2009) Eye gaze during comprehension of American sign language by native and beginning signers. Journal of Deaf Studies and Deaf Education 14: 237–243.
20. Blais C, Jack RE, Scheepers C, Fiset D, Caldara R (2008) Culture shapes how we look at faces. PLoS One 3: e3022.
21. Jack RE, Blais C, Scheepers C, Schyns PG, Caldara R (2009) Cultural confusions show that facial expressions are not universal. Current Biology 19: 1543–1548.
22. Chua HF, Boland JE, Nisbett RE (2005) Cultural variation in eye movements during scene perception. Proceedings of the National Academy of Sciences 102: 12629–12633.
23. Tanaka JW, Kiefer M, Bukach CM (2004) A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. Cognition 93: B1–B9.
24. Michel C, Caldara R, Rossion B (2006) Same-race faces are perceived more holistically than other-race faces. Visual Cognition 14: 55–73.

25. Michel C, Rossion B, Han J, Chung CS, Caldara R (2006) Holistic processing is finely tuned for faces of our own race. Psychological Science 17: 608–615.
26. Yuki M, Maddux WW, Masuda T (2007) Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. Journal of Experimental Social Psychology 43: 303–311.
27. de Heering A, Rossion B (2006) Prolonged visual experience in adulthood modulates holistic face perception. PLoS ONE 3: e2317.
28. Mindess A (2006) Reading between the Signs: Intercultural Communication for Sign Language Interpreters. Boston, MA: Intercultural Press.
29. Hall JA (1978) Gender effects in decoding nonverbal cues. Psychological Bulletin 85: 845–857.
30. Hall JA (1984) Nonverbal sex differences: Communication accuracy and expressive style. Baltimore: Johns Hopkins University Press.
31. Thayer JF, Johnsen BH (2000) Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. Scandinavian Journal of Psychology 41: 243–246.
32. Hall JA, Matsumoto D (2004) Gender differences in judgments of multiple emotions from facial expressions. Emotion 4: 201–206.
33. Rahman Q, Wilson GD, Abrahams S (2004) Sex, sexual orientation, and identification of positive and negative facial affect. Brain and Cognition 54: 179–185.
34. Hall JK, Hutton SB, Morgan MJ (2009) Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? Cognition & Emotion 24: 629–637.
35. Vassallo S, Cooper SL, Douglas JM (2009) Visual scanning in the recognition of facial affect: Is there an observer sex difference? Journal of Vision 9: 1–10.
36. Ekman P, Friesen WV (1975) Unmasking the face: A guide to recognizing emotions from facial clues. Englewood Cliffs, NJ: Prentice Hall.
37. Adolphs R, Gosselin F, Buchanan TW, Tranel D, Schyns P, et al. (2005) A mechanism for impaired fear recognition after amygdala damage. Nature 433: 68–72.
38. Adolphs R (2006) Perception and emotion: How we recognize facial expressions. Current Directions in Psychological Science 15: 222–226.
39. Calvo MG, Nummenmaa L (2008) Detection of emotional faces: Salient physical features guide effective visual search. Journal of Experimental Psychology: General 13: 471–494.
40. Argyle M, Cook M (1976) Gaze and Mutual Gaze. Cambridge, UK: Cambridge University Press.
41. Nisbett RE, Miyamoto Y (2005) The influence of culture: holistic versus analytic perception. Trends in Cognitive Sciences 9: 467–473.
42. Humphreys GW, Donnelly N, Riddoch MJ (1993) Expression is computed separately from facial identity, and it is computed separately for moving and static faces: Neuropsychological evidence. Neuropsychologia 31: 173–181.
43. Roy C, Blais C, Fiset D, Gosselin F (2010) Visual information extraction for static and dynamic facial expression of emotions: an eye-tracking experiment. Journal of Vision 10: 531.
44. Frijda NH (1953) The understanding of facial expression of emotion. Acta Psychologica 9: 294–362.
45. Berry DS (1990) What can a moving face tell us? Journal of Personality and Social Psychology 58: 1004–1014.
46. Bruce V, Valentine T (1988) When a nod's as good as a wink: The role of dynamic information in facial recognition. In Gruneberg MM, Morris PE, Sykes RN, eds. Practical aspects of memory: Current research and issues (Vol. 1, pp. 169–174). New York: John Wiley & Sons.
47. Harwood NK, Hall LJ, Shinkfield AJ (1999) Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. American Journal of Mental Retardation 104: 270–278.
48. Lander K, Christie F, Bruce V (1999) The role of movement in the recognition of famous faces. Memory and Cognition 27: 974–985.
49. Wehrle T, Kaiser S, Schmidt S, Scherer KR (2000) Studying the dynamics of emotional expression using synthesized facial muscle movements. Journal of Personality and Social Psychology 78: 105–119.
50. Sato W, Yoshikawa S (2004) The dynamic aspects of emotional facial expressions. Cognition and Emotion 18: 701–710.
51. Sato W, Kochiyama T, Yoshikawa S, Naito E, Matsumura M (2004) Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. Cognitive Brain Research 20: 81–91.
52. Schultz J, Pilz KS (2009) Natural facial motion enhances cortical responses to faces. Experimental Brain Research 194: 465–475.
53. Trautmann SA, Fehr T, Herrmann M (2009) Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain Research 1248: 100–115.
54. Hanawalt NG (1944) The role of the upper and the lower parts of the face as the basis for judging facial expressions: II. In posed expressions and "candid camera" pictures. The Journal of General Psychology 31: 23–36.
55. Ekman P, Friesen WV (1971) Constants across cultures in the face and emotion. Journal of Personality and Social Psychology 17: 124–129.
56. Sullivan LA, Kirkpatrick SW (1996) Facial interpretation and component consistency. Genetic, Social, and General Psychology Monographs 122: 389–404.
57. Rutherford MD, Towns AM (2008) Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. Journal of Autism & Developmental Disorders 38: 1371–1381.
58. Schyns PG, Bonnar L, Gosselin F (2002) Show me the features! Understanding recognition from the use of visual information. Psychological Science 13: 402–409.
59. Koda T, Ruttkay Z, Nakagawa Y, Tabuchi K (2010) Cross-cultural study on facial regions as cues to recognize emotions of virtual agents. Lecture Notes in Computer Science 6259: 16–27.
60. Liddell S (1980) American sign language syntax. The Hague: Mouton Publishers.
61. Nakamura K (2006) Deaf in Japan: Signing And the Politics of Identity. Ithaca: Cornell University Press.

nature COMMUNICATIONS

OPEN

ARTICLE

Received 24 Oct 2011 | Accepted 21 Feb 2012 | Published 27 Mar 2012

DOI: 10.1038/ncomms1757

Neural circuits in the brain that are activated when mitigating criminal sentences

Makiko Yamada[1,2], Colin F. Camerer[3], Saori Fujie[1], Motoichiro Kato[4], Tetsuya Matsuda[5], Harumasa Takano[1], Hiroshi Ito[1], Tetsuya Suhara[1] & Hidehiko Takahashi[1,2,5,6]

In sentencing guilty defendants, jurors and judges weigh 'mitigating circumstances', which create sympathy for a defendant. Here we use functional magnetic resonance imaging to measure neural activity in ordinary citizens who are potential jurors, as they decide on mitigation of punishment for murder. We found that sympathy activated regions associated with mentalising and moral conflict (dorsomedial prefrontal cortex, precuneus and temporo-parietal junction). Sentencing also activated precuneus and anterior cingulate cortex, suggesting that mitigation is based on negative affective responses to murder, sympathy for mitigating circumstances and cognitive control to choose numerical punishments. Individual differences on the inclination to mitigate, the sentence reduction per unit of judged sympathy, correlated with activity in the right middle insula, an area known to represent interoception of visceral states. These results could help the legal system understand how potential jurors actually decide, and contribute to growing knowledge about whether emotion and cognition are integrated sensibly in difficult judgments.

[1] Department of Molecular Neuroimaging, Molecular Imaging Center, National Institute of Radiological Sciences, 4-9-1 Anagawa, Inage-ku, Chiba 263-8555, Japan. [2] Decoding and Controlling Brain Information, Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. [3] Computation and Neural Systems and Humanities and Social Science Divisions, California Institute of Technology, Pasadena, California 91125-7700, USA. [4] Department of Neuropsychiatry, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. [5] Tamagawa University Brain Science Institute, 6-1-1 Tamagawa Gakuen, Machida, Tokyo 194-8610, Japan. [6] Department of Neuropsychiatry, Kyoto University School of Medicine, 54 Shogoin-kawara-cho, Sakyo-ku, Kyoto 606-8507, Japan. Correspondence and requests for materials should be addressed to M.Y. or H.T. (email: myamada@nirs.go.jp or hidehiko@nirs.go.jp).

Philosophers, psychologists and legal scholars have long debated whether mercy, sympathy and compassion should reduce moral culpability of legal defendants. People do have negative emotional responses to a wide range of situational factors that are not normatively justifiable because they are not considered evidence, or appeal to 'prejudices and sympathy', which jurors are typically instructed to ignore, for example, gruesome crime scene pictures lead to more mock jury convictions when they are in colour (compared with black and white[1]). Other studies show that evidence that provides negative emotions leads to more punitive judgments[2]. A mock trial study reported that jurors found a defendant less guilty when they heard a defense attorney urging sympathy for the defendant[3].

Studies of moral cognition in hypothetical non-legal scenarios have revealed the increased activity in emotion-related brain regions (that is, insula, amygdala and orbitofrontal cortex) and decreased activity in cognitive processes (that is, dorsolateral prefrontal and parietal areas) when participants contemplated morally charged actions, such as a trolley dilemma involving killing lives of some people to save others[4]. Other studies indicate that people do sometimes punish norm violators in non-legal or legal situations, driven by moral judgments in which negative emotional reactions have a critical role[5,6]. Prosocial emotions, such as sympathy, also influence decision making (for example, charitable giving)[3]; however, there is currently no direct cognitive and neural evidence for how sympathy is translated into legal outcomes.

All these studies raise an important question of whether moral judgments and accompanying emotions are used reasonably (as legal rules require) or not. Neuroscience is now entering this debate about defendants, but very little is known about whether mental activity of juries and judges conforms to legal principles. The case we explore here is 'mitigating circumstances', a rare case where emotions, such as sympathy, are allowed to influence juror judgment.

The legal domain is unusual because it may be especially challenging to map emotions into numerical legal outcomes. This difficulty is found in studies of hypothetical punitive damages in tort cases, which not only show that jurors typically agree on moral outrageousness of actions, but also show large disagreement about how outrage is mapped to punitive dollar awards[7,8].

Weighing mitigation puts an unusual burden on people (*qua* jurors) to have appropriate emotional sympathetic reactions, then encode their emotions into prison years. Uncovering the cognitive and neural mechanisms of sympathy that motivate mitigation will inform the role of emotion in jurors' decision process, and perhaps the ultimate policy issue of what role emotional evidence can and should have in trials. Understanding the neuroscientific basis of legal mitigation adds to a basic understanding of moral neuroscience. Neural evidence could also advance theory and practice of law, as so little is known about whether the mental activity of juries and judges conforms to normative legal principles[9].

In summary, our results revealed that sympathy activated brain regions associated with mentalising and moral conflict, including dorsomedial prefrontal cortex (DMPFC), precuneus and temporoparietal junction (TPJ). Sentence mitigation also recruited these sympathy regions, uncovering neural evidence for a close relationship between sympathy and mitigation. Furthermore, individual differences on the inclination to mitigate were reflected in differential middle insula activity. These findings do not just contribute to the field of neuroscience, but could help lawmakers to understand jurors' decision making and their individual differences in trials.

Results

Sympathy and no-sympathy scenarios. We measured brain activity using functional magnetic resonance imaging (fMRI) while subjects are making hypothetical sentence reduction decisions, in dramatic scenarios adapted from actual murder cases. Sympathy-related brain activity was collected during reading circumstances pertaining to defendants' crimes (Fig. 1a). Only actual Japanese murders were used, so the crime was serious, uniform across trials and lifelike. This simple design was chosen to generate engagement and limit nuisance brain activity due to subtle differences in crimes and plausibility of artificially created scenarios.

Mitigating circumstances were of two types: those that would induce sympathy and those that would not. The sympathy scenarios included desperate situations of defendants suffering from domestic violence, disease or poverty. Figure 1a gives one example of each type. The intentionality and severity of the murders were matched between conditions (see Supplementary Methods). After reading about the circumstances, subjects decided how much they would change the sentence given for the defendant (initially 20 years) if they were on a jury. After scanning, subjects were again presented with the same scenarios and asked to rate how much sympathy they felt for the defendant, using a visual analogue scale.

Sympathy and punishment ratings. The sympathy manipulation is internally valid because participants gave significantly higher sympathy ratings to those defendants with sympathy circumstances compared with those with no-sympathy circumstances (n = 22, paired t-test, $t_{21} = -18.94$, $P < 0.001$, Fig. 1b). They also reduced sentences much more in sympathy circumstances (n = 22, paired t-test, $t_{21} = 11.82$, $P < 0.001$, Fig. 1c). Unsurprisingly, sympathy and punishment (sentence length) were highly negatively correlated (n = 32 stories, linear regression analysis, $P < 0.001$, $R^2 = 0.97$, Fig. 1d).

fMRI results were analysed using standard generalized linear model regression techniques (see Methods). A short block design was used where regressors were included for the various events of the trials (Fig. 1a). Interaction terms corresponding to punishment and sympathy ratings interacted with trial onsets that were added as parametric regressors.

Brain regions associated with sympathy scenarios. We first analysed brain areas exhibiting a stronger response in sympathy scenarios than no-sympathy scenarios. Precuneus, left TPJ and DMPFC showed larger activities for sympathy than for no-sympathy scenarios (n = 22, one-sample t-test, Table 1). These regions are related to mentalization and sympathy as discussed below in detail, which confirms that the experimental manipulation of sympathy produced results consistent with the sympathy ratings data.

Brain regions associated with punishment and sympathy. We then searched for brain regions that responded, during the description, to the subjects' trial-by-trial ratings of sympathy and their amounts of punishment reduction. Activity in precuneus, DMPFC and left TPJ were correlated with sympathy ($P < 0.05$, small-volume-corrected, Fig. 2, Supplementary Table S1). Signal increase in precuneus and DMPFC were also associated with the reduction of punishment ($P < 0.05$, small-volume-corrected, Fig. 2: note that a small TPJ region was also activated in sentence reduction, but only with $k = 6$ voxels). Thus, precuneus and DMPFC were commonly activated by both sympathy and reduction of punishment. Sentence reduction was also associated with activity in anterior cingulate cortex (ACC). All regions showing a whole-brain correlation at $P < 0.001$ are listed in Supplementary Tables S1 and S2.

Brain regions associated with an inclination to mitigate. Next, we constructed an individual-specific measure of an inclination to mitigate, by reducing sentences, as a function of sympathy. This measure comes from a simple linear regression on each individual's decisions: punishment = $b_0 + b_1$*sympathy + error[10]. A measure of an inclination to mitigate, the reduction in sentence per unit of sympathy, was given by the b_1 coefficient of the regression. This number represents a complex mapping from an emotional response to a number representing prison time for a defendant (a years-per-emotion coefficient).

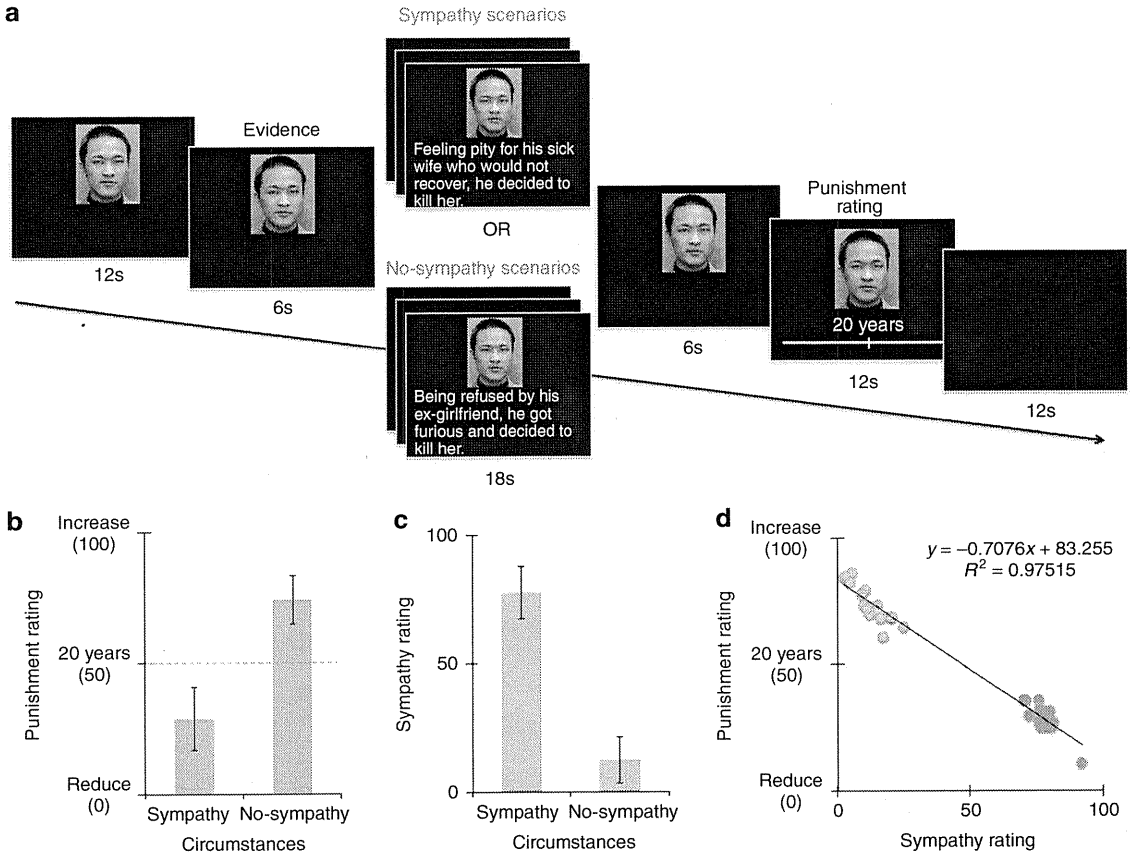

Figure 1 | Task design and behavioural performance. (**a**) Study paradigm. (**b**) Mean punishment ratings for sympathy and no-sympathy trials (n = 22, paired t-test, t_{21} = −18.94, P<0.001). (**c**) Mean sympathy ratings for sympathy and no-sympathy trials (n = 22, paired t-test, t_{21} = 11.82, P<0.001). (**d**) Correlation between sympathy and punishment ratings for sympathy stories (red circles) and no-sympathy stories (green circles). Error bars indicate s.d.

These b_1 coefficients were negative for all subjects (mean = −6.5±0.2, linear regression analysis, all P<0.001), indicating that the feeling of sympathy did correlate with reduction of punishment, but to different degrees across subjects. A negative linear regression between the individual-specific b_1 coefficient and blood oxygenation level-dependent responses in sympathy minus no-sympathy trials found activity in the right middle insula (P<0.05, small-volume-corrected, Fig. 3). Individuals who had larger activities in the insula when reading circumstances showed higher tendencies to mitigate, reducing sentencing years more as their sympathy increased.

Discussion

Comparison of activity during judgments of sympathy, and sentence reduction, suggest that activity in DMPFC (also known as paracingulate), precuneus (also known as posterior cingulate) and TPJ reflect a judgment-action circuit, which is illustrated in Fig. 2. Strength of sympathy judgments is associated with activity in DMPFC, precuneus and TPJ. DMPFC is involved in general mentalising[11] and is active when empathizing[12] or sympathizing[13] with others in pain. Precuneus has been linked to subjective perspective taking[14–16]. TPJ is also commonly identified as a part of a theory-of-mind circuit[17], including mentalising about intentions[18], and was activated in one study on judging innocence of intentions of people who caused harm[19]. This suggests that the sympathy judgment is an engagement with a reasoned simulation of what the defendant was thinking when committing the crime or how most people would judge the normative basis for mitigation. Note that although precuneus and TPJ can each be activated by non-social demands as well (for example, attention reorienting[20]), they are rarely co-activated as a group unless social cognitive demands are present.

Regions activated by the punishment reduction judgment include a large region of precuneus and smaller regions in DMPFC and ACC. As noted, precuneus has been linked to simulation on the self to understand others, and is also active when compassionating[21] and forgiving others[22]. The activation in this area is correlated with more iterated steps[23] and higher-value[24] strategic thinking in game theory tasks. Note that both feelings of sympathy and judged mitigation of punishment were encoded in activity in precuneus. This overlapping activity suggests that the precuneus may be a region that accepts emotional judgment input and maps it into concrete punishment actions.

ACC is a region now thought to be activated by negative affect[25,26], positive reward[27] and cognitive control[26,28]. In our context, mapping emotional sympathy into numerical sentencing requires high-level executive function by weighing negative affective reaction to murder, positive sympathy for the murder's mitigating circumstances and exerting cognitive control to choose numerical punishments that weigh these emotions consistently.

Smaller regions in caudate also showed differential activation in response to punishment reduction (Supplementary Table S2). Caudate activity is consistent with the hypothesis that sentence reduction is encoded as a special kind of prosocial 'charity' (as mitigation is like giving to charity[29]), as other prosocial choices activate the caudate too[30].

One fMRI study elicited punishment judgments in artificial scenarios varying offender culpability[6] and reports right DLPFC activity associated with responsibility judgments. We speculate that the absence of right DLPFC activation in our study is because there is no doubt about the defendants' guilt, so the most morally burdensome question of guilt versus innocence is resolved (right DLPFC is discharged from jury duty, so to speak). Activity then shifts to