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 temporo-parietal 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).

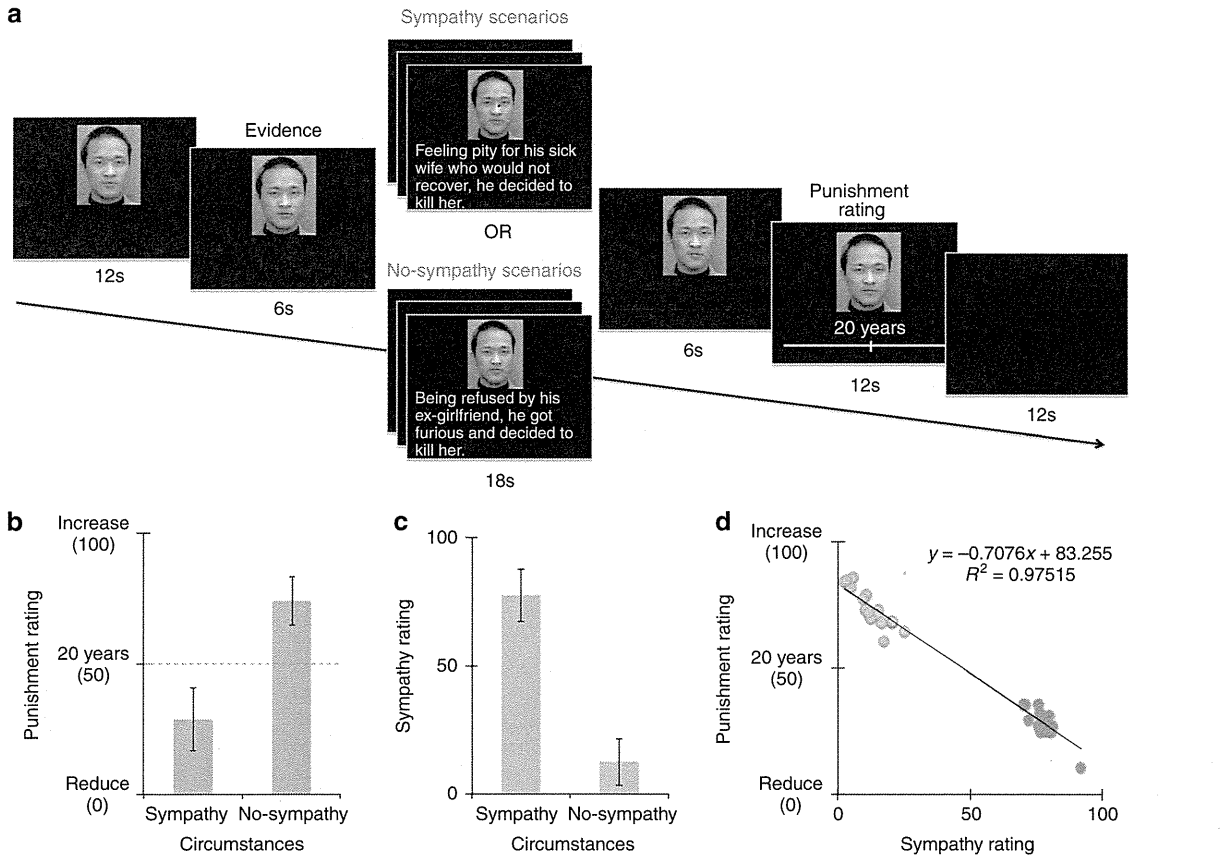

Figure 1 | Task design and behavioural performance. (**a**) Study paradigm. (**b**) Mean punishment ratings for sympathy and no-sympathy trials (n = 22, paired t-test, t_{21} = −18.94, P < 0.001). (**c**) Mean sympathy ratings for sympathy and no-sympathy trials (n = 22, paired t-test, t_{21} = 11.82, 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

Table 1 | Areas exhibiting a stronger response in sympathy scenarios.

Region	*Side*	*MNI coordinates*			Z	k	*SVC*
Posterior cingulate cortex	**R**	**2**	**−50**	**28**	**4.05**	**2946**	**0.001**
Precuneus	**L**	**−2**	**−68**	**42**	**3.85**		
Precuneus	**R**	**0**	**−18**	**28**	**3.73**		
Middle frontal gyrus	R	36	16	60	4.11	1501	
Superior frontal gyrus	R	20	32	58	4.08		
Superior frontal gyrus	L	−18	24	64	4.06		
Temporo-parietal junction (middle temporal gyrus)	**L**	**−50**	**−72**	**20**	**4.15**	**535**	**0.001**
Temporo-parietal junction (angular gyrus)	**L**	**−46**	**−64**	**34**	**3.58**		
Inferior frontal gyrus (p. orbitalis)	L	−40	34	−14	4.37	296	
Inferior frontal gyrus (p. triangularis)	L	−50	36	18	3.62		
Inferior frontal gyrus (p. triangularis)	L	−42	34	0	3.54		
Dorsomedial prefrontal cortex (superior medial gyrus)	**L**	**−4**	**44**	**22**	**3.50**	**217**	**0.008**
Dorsomedial prefrontal cortex (superior medial gyrus)	**L**	**−4**	**54**	**28**	**3.49**		
Superior frontal gyrus	**L**	**−12**	**52**	**34**	**3.33**		
Caudate nucleus	R	12	18	8	3.82	179	
Hippocampus	R	26	−32	−8	3.47	151	
Parahippocampal gyrus	R	28	−24	−22	3.43		
Caudate nucleus	L	−8	18	2	3.53	101	
Caudate nucleus	L	−12	14	10	3.37		
Middle temporal gyrus	L	−52	4	−32	3.72	81	
Lingual gyrus	R	16	−44	−8	3.50	59	
Parahippocampal gyrus	R	26	10	−30	3.68	48	
Dorsomedial prefrontal cortex (superior medial gyrus)	**R**	**10**	**44**	**50**	**3.37**	**23**	**0.011**
Superior frontal gyrus	**R**	**16**	**16**	**46**	**44**		
Inferior frontal gyrus (p. orbitalis)	L	−26	26	−26	3.56	22	
Temporal pole	L	−36	22	−32	3.19		
Medial temporal pole	R	44	16	−38	3.37	13	
Inferior frontal gyrus (p. orbitalis)	R	38	28	−12	3.22	13	
Temporal pole	L	−34	6	−20	3.17	12	

Abbreviation: SVC, small volume correction.
A priori regions of interest are in bold ($P<0.001$, $k>10$).

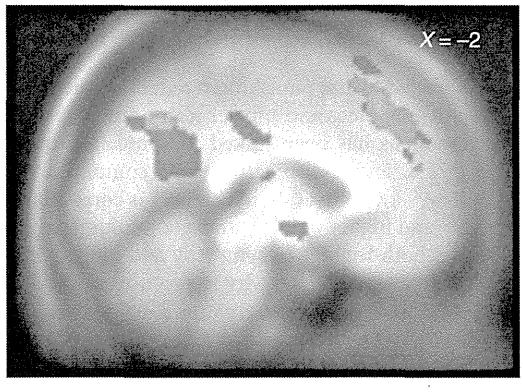

Figure 2 | Brain regions activated during trial-by-trial sympathy and punishment reduction. Regions in which activity correlated with parametric regressors of increasing sympathy (green) and reduced punishment (red). Common areas were found in precuneus (yellow). The image is shown at $P<0.001$ (uncorrected; $n = 22$, one-sample t-tests).

areas generally associated with moral and cognitive conflicts instead (DMPFC, precuneus and ACC).

The individual difference analysis showed that activity in the right mid-insula was related to the individual differences on the inclination to mitigate (mapping sympathy to sentences). Middle/posterior insula has been suggested to encode economic inequity[31], norm violation[32], somatosensory representation[33], monitoring internal states[34] and heart rate during a compassionate state[35]. These types of interoceptive processing in middle insula suggest that it is sensitive to emotions linked to sociality. Our study provides unusual evidence of this processing associated with a unique high-impact social judgment that affects others.

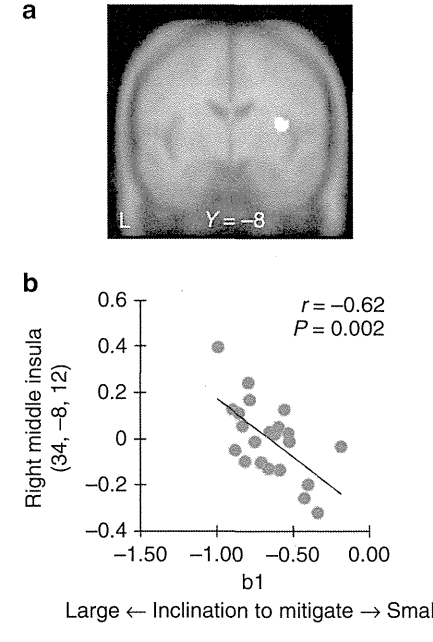

Figure 3 | Individual differences in inclination towards mitigation. Activation in right middle insula in the contrast sympathy minus no-sympathy trials (**a**, MNI 34/−8/12; $n = 22$, one-sample t-test, $P = 0.023$, small-volume-corrected) correlated with participant-wise coefficient of mitigation (**b**, $n = 22$, Pearson $r = -0.62$, $P = 0.002$).

The identified brain activity provides new insight into the capacity of the average brain to translate sympathetic feelings into appropriate legal action. A plausible neuro-legal standard is that the influence

of sympathetic reaction on difficult sentencing should recruit brain areas that process sympathy judgments in general, as well as areas that encode difficult decision conflict. Activity in these 'sympathy' regions is evident in our study when judging sympathy alone, and in choosing sentence mitigation. However, not every brain maps sympathy to prison sentences in the same numerical way (as reflected in differential mid-insula activity). Differences in these brain circuits between individuals, suggest that differential juror responses might need to be considered unequally. There is also mixed evidence about the normative basis of legal judgment, including a recent finding that judges' decisions are affected by timing of meals[36].

The current finding would also contribute to the attribution literature on situational correction, as there has not been any fMRI work on this. People often attribute behaviour to the corresponding personal disposition, which is, yet, corrected based on situational inducements too[37]. This is an apt psychological function behind mitigating circumstances that is to make someone less culpable, and the revealed brain activity might be associated with this type of situational correction.

Finally, we note that many legal principles treat emotional responses as likely to be prejudicial and prone to inflammatory manipulation (that is, an ideal juror would suppress them and legal rules limit their influence). Weighing mitigating circumstances during sentencing (*after* a verdict) represents an unusual case in which emotional sympathy judgment is actually required. Japanese criminal law, for example, requires that the decision among these sentences be based on mitigating circumstances[38].

Ironically, the fact that sympathy is clearly evident in brain activity, and influences sentence mitigation (as it should), raises interest in the opposite question: can people also suspend emotions when the law instructs them to? More generally, a deeper understanding of the brain could help figure out how highly evolved brain structures, which were sculpted to maintain order in small-scale ancestral societies, can be put to work under modern legal rules in much more challenging cases to create modern justice.

Methods

Participants. A total of 26 right-handed healthy subjects (12 males; mean age = 21.5 ± 1.8 (s.d.) years) participated in the study. All subjects had normal or corrected-to-normal vision, no history of neurological or psychiatric disorder and were not taking medications that interfere with the performance of fMRI. The study was approved by the Ethics and Radiation Safety Committee of the National Institute of Radiological Sciences (Chiba, Japan). Data from four participants were discarded owing to excessive movement (three subjects) and sleep (one subject). The data from 22 subjects (11 males; mean age = 21.5 ± 1.9 (s.d.) years) were included and analysed.

Stimuli and the task. A total of 32 faces of defendants and criminal acts were prepared in the study. Faces were drawn from a set of Asian face stimuli developed by CAS-PEAL face database[39]. Pilot participants ($n = 30$) rated attractiveness for each faces using an 11-point scale of −5 (not attractive at all) to +5 (very attractive). The average attractiveness for 32 faces was 0.15 ± 0.70.

All criminal acts were the pure murder cases, which were modified from real stories taken from a precedent search engine provided by Supreme Court of Japan (www.courts.go.jp) and from news articles on the web. We prepared 32 scenarios and consulted a judicial expert for advice on how externally valid the scenarios were. Each scenario composed of the fact (when, where and to whom) and the circumstance (why the defendant committed the murder). Half of the scenarios would induce sympathy for the defendants who suffer from domestic violence, death to disease or poverty (sympathy scenarios), and the other half that would not (no-sympathy scenarios). Responsibility, intentionality and severity of crime were matched between scenarios (see Supplementary Methods for all scenarios used in the study).

Subjects were instructed to judge crimes as if they were jurors (lay judges). Before entering the scanner, subjects read each criminal case and were asked to indicate if they have known them through media. The aim of this session was to ensure and ease comprehension in the scanner that constrain duration of presenting each scenario. During the fMRI task, subjects read each scenario and rated how much they would like to reduce or increment the sentence (all 20 years) given for the defendant using a VAS (with scores ranging from 0 = reduce as much as possible to 100 = increment as much as possible, and 50 = given sentence of 20 years; Fig. 1a). We chose a sentence of 20 years as a reference point, based on a sentencing guideline in Japan that gives 20 years in prison to a person who commits murder. Defendants' faces were presented together with the scenarios in order to make the experimental setting more realistic and to help participants to dedicate to the study. They completed two sessions, and each session contained 8 sympathy and 8 no-sympathy scenarios, 16 trials in total. The presentation of each scenario within a session was pseudorandomized. The order of sessions and the combination of faces and circumstance types were counterbalanced between subjects.

After the scan, subjects were shown the same scenarios again to rate how much they felt sympathy for the defendants using a VAS (with scores ranging from 0 = having no sympathy to 100 = having sympathy very much). They were also asked to fill out Interpersonal Reactivity Index (IRI), a self-report measure of dispositional empathy.

fMRI data acquisition. The functional imaging was conducted using a GE 3.0 Tesla Excite system to acquire gradient echo T2*-weighted echoplanar images with blood oxygenation level-dependent contrast. Each volume comprised 35 transaxial contiguous slices with a slice thickness of 3.8 mm to cover almost the whole brain (flip angle, 75; echo time, 25 ms; repetition time, 2000 ms; matrix, 64*64; and field of view, 24×24 cm^2).

fMRI data preprocessing. Image analysis was performed using SPM8 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK), except for the overlay image shown in Fig. 2 that was created using the MRIcron software (http://www.sph.sc.edu/comd/rorden/mricron/). All volumes were realigned to the first volume of each session to correct for subject motion and were spatially normalised to the standard space defined by the Montreal Neurological Institute (MNI) template. After normalisation, images were spatially smoothed using a Gaussian kernel with a full-width-at-half-maximum of 8 mm. Intensity normalisation and high-pass temporal filtering (using a filter width of 128 s) were also applied to the data.

Generalized linear model. For the primary whole-brain analyses, two modelling approaches were used. In the first approach, trials from each of the sympathy and no-sympathy circumstances were modelled as separate conditions. This allowed separate estimation of the evoked response for each of the conditions at each voxel. The model also included regressors for face, criminal evidence, punishment, and fixation, as well as six head-motion parameters as regressors of no interest. We first calculated a first-level single-subject contrast for reading sympathy minus no-sympathy circumstances. Then, we calculated a second-level group contrasts using a one-sample *t*-test. In the second approach, all trials from reading circumstances modelled using a single condition (that is, overall task-related activation), and two additional orthogonal parametric regressors were included representing (a) the amount of reducing punishment and (b) the level of sympathy rating. We first calculated first-level single-subject contrasts for reading circumstances modulated by reducing punishment and for reading circumstances modulated by sympathy rating. Then, we calculated second-level group contrasts using one-sample *t* tests on the single-subject contrasts to identify brain regions whose activation correlated with the amount of reducing punishment and the sympathy rating.

A statistical threshold of $P < 0.05$ corrected for multiple spatial comparisons across the whole brain at cluster level was used, except for a priori hypothesized regions that were voxel thresholded at $P < 0.001$ uncorrected (only clusters involving $k > 10$ or more contiguous voxels were reported[40]). Small volume correction (a 10-mm radius sphere) was used on a priori regions of interest: DMPFC/ACC, precuneus/posterior cingulate cortex, TPJ and the insula.

References

1. Whalen, D. H. & Blanchard, F. A. Effects of photographic evidence on mock juror judgment. *J. Appl. Soc. Psychol.* **12,** 30–41 (1982).
2. Salerno, J. M. & Bottoms, B. L. Emotional evidence and jurors' judgments: the promise of neuroscience for informing psychology and law. *Behav. Sci. Law* **27,** 273–296 (2009).
3. Haegerich, T. M. & Bottoms, B. L. Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. *Law Hum. Behav.* **24,** 421–448 (2000).
4. Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. An fMRI investigation of emotional engagement in moral judgment. *Science* **293,** 2105–2108 (2001).
5. Haidt, J. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* **108,** 814–834 (2001).
6. Buckholtz, J. W. *et al.* The neural correlates of third-party punishment. *Neuron* **60,** 930–940 (2008).
7. Kahneman, D., Schkade, D. & Sunstein, C. Shared outrage and erratic awards: the psychology of punitive damages. *J. Risk Uncertain.* **16,** 49–86 (1998).
8. Schkade, D., Sunstein, C. R. & Kahneman, D. Deliberating about dollars: the severity shift. *Columbia Law Rev.* **100,** 1139–1175 (2000).
9. Jones, O. D. Law, evolution and the brain: applications and open questions. *Philos. Trans. R Soc. Lond. B. Biol. Sci.* **359,** 1697–1707 (2004).
10. Hare, T. A., Camerer, C. F., Knoepfle, D. T. & Rangel, A. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* **30,** 583–590 (2010).

11. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7,** 268–277 (2006).
12. Lamm, C., Decety, J. & Singer, T. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* **54,** 2492–2502 (2011).
13. Decety, J. & Michalska, K. J. Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Dev. Sci.* **13,** 886–899 (2010).
14. Vogeley, K. *et al.* Neural correlates of first-person perspective as one constituent of human self-consciousness. *J. Cogn. Neurosci.* **16,** 817–827 (2004).
15. Ruby, P. & Decety, J. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* **4,** 546–550 (2001).
16. Farrer, C. & Frith, C. D. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* **15,** 596–603 (2002).
17. Greene, J. & Haidt, J. How (and where) does moral judgment work? *Trends Cogn. Sci.* **6,** 517–523 (2002).
18. Atique, B., Erb, M., Gharabaghi, A., Grodd, W. & Anders, S. Task-specific activity and connectivity within the mentalizing network during emotion and intention mentalizing. *Neuroimage* **55,** 1899–1911 (2011).
19. Young, L. & Saxe, R. Innocent intentions: a correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia* **47,** 2065–2072 (2009).
20. Corbetta, M., Patel, G. & Shulman, G. L. The reorienting system of the human brain: from environment to theory of mind. *Neuron* **58,** 306–324 (2008).
21. Immordino-Yang, M. H., McColl, A., Damasio, H. & Damasio, A. Neural correlates of admiration and compassion. *Proc. Natl Acad. Sci. USA* **106,** 8021–8026 (2009).
22. Farrow, T. F. *et al.* Investigating the functional anatomy of empathy and forgiveness. *Neuroreport* **12,** 2433–2438 (2001).
23. Kuo, W. J., Sjostrom, T., Chen, Y. P., Wang, Y. H. & Huang, C. Y. Intuition and deliberation: two systems for strategizing in the brain. *Science* **324,** 519–522 (2009).
24. Bhatt, M. & Camerer, C. F. Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games Econ. Behav.* **52,** 424–459 (2005).
25. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **15,** 85–93 (2011).
26. Shackman, A. J. *et al.* The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* **12,** 154–167 (2011).
27. Liu, X., Hairston, J., Schrier, M. & Fan, J. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* **35,** 1219–1236 (2011).
28. Botvinick, M. M., Cohen, J. D. & Carter, C. S. Conflict monitoring and anterior cingulate cortex: an update. *Trends. Cogn. Sci.* **8,** 539–546 (2004).
29. Harbaugh, W. T., Mayr, U. & Burghart, D. R. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* **316,** 1622–1625 (2007).
30. Fehr, E. & Camerer, C. F. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* **11,** 419–427 (2007).
31. Hsu, M., Anen, C. & Quartz, S. R. The right and the good: distributive justice and neural encoding of equity and efficiency. *Science* **320,** 1092–1095 (2008).
32. Montague, P. R. & Lohrenz, T. To detect and correct: norm violations and their enforcement. *Neuron* **56,** 14–18 (2007).
33. Danziger, N., Faillenot, I. & Peyron, R. Can we share a pain we never felt? Neural correlates of empathy in patients with congenital insensitivity to pain. *Neuron* **61,** 203–212 (2009).
34. Damasio, A. R. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999).
35. Lutz, A., Greischar, L. L., Perlman, D. M. & Davidson, R. J. BOLD signal in insula is differentially related to cardiac function during compassion meditation in experts vs. novices. *Neuroimage* **47,** 1038–1046 (2009).
36. Danziger, S., Levav, J. & Avnaim-Pesso, L. Extraneous factors in judicial decisions. *Proc. Natl Acad. Sci. USA* **108,** 6889–6892 (2011).
37. Gilbert, D. T., Pelham, B. W. & Krull, D. S. On cognitive busyness: when person perceivers meet persons perceived. *J. Pers. Soc. Psychol.* **54,** 733–740 (1988).
38. Egashira, K., Kobayakawa, K., Nishida, N., Takahashi, H. & Nomi, Y. *Roppo Zensho [Comdemium of Laws] (Japanese)* (Yuhikaku, 2011).
39. Gao, W. *et al.* The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Trans. Syst., Man and Cy- bernetics, Part A* **38,** 149–161 (2008).
40. Forman, S. D. *et al.* Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* **33,** 636–647 (1995).

Acknowledgements

We thank T. Kouchiyama for his advice in designing the task, and K. Suzuki and I. Izumida for their help as clinical research coordinators. This study was supported in part by JSPS KAKENHI 22791156, 23680045, MEXT Tamagawa University GCOE, MEXT SRPBS, MEXT KAKENHI 23011005, 23120009.

Author contributions

M.Y. and H.T. designed the experiment. M.Y. and F.S. prepared and conducted the experiment. M.Y. analysed the data and wrote the paper. C.F.C. edited the manuscript. H.T. coordinated subject recruitment. All authors discussed the results.

Additional information

Supplementary Information accompanies this paper on http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Yamada, M. *et al.* Neural circuits in the brain that are activated when mitigating criminal sentences. *Nat. Commun.* 3:759 doi: 10.1038/ncomms1757 (2012).

Psychiatry and Clinical Neurosciences 2012; **66**: 187–194 doi:10.1111/j.1440-1819.2011.02314.x

Regular Article

Relationships between exploratory eye movement dysfunction and clinical symptoms in schizophrenia

Masahiro Suzuki, MD, PhD,[1] Sakae Takahashi, MD, PhD,[1]* Eisuke Matsushima, MD, PhD,[2] Masahiko Tsunoda, MD, PhD,[4] Masayoshi Kurachi, MD, PhD,[4] Takashi Okada, MD, PhD,[5] Takuji Hayashi, MD, PhD,[6] Yohei Ishii, PhD,[7] Kiichiro Morita, MD, PhD,[7] Hisao Maeda, MD, PhD,[8] Seiji Katayama, MD, PhD,[9] Tatsui Otsuka, MD, PhD,[10] Yoshio Hirayasu, MD, PhD,[10] Mizuho Sekine, MD,[3] Yoshiro Okubo, MD, PhD,[3] Mai Motoshita, PhD,[2] Katsuya Ohta, MD, PhD,[2] Makoto Uchiyama, MD, PhD[1] and Takuya Kojima, MD, PhD[11]

[1]*Department of Psychiatry, Nihon University School of Medicine,* [2]*Section of Liaison Psychiatry and Palliative Medicine, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University,* [3]*Department of Neuropsychiatry, Nippon Medical School, Tokyo,* [4]*Department of Neuropsychiatry, University of Toyama Graduate School of Medicine and Pharmaceutical Sciences, Toyama,* [5]*Department of Neuropsychiatry, Graduate School of Medicine, Kyoto University, Kyoto,* [6]*Toyosato Hospital, Inukami,* [7]*Cognitive and Molecular Research Institute of Brain Diseases, Kurume University, Kurume,* [8]*Wakahisa Hospital, Fukuoka,* [9]*Yasugi Daiichi Hospital, Yasugi,* [10]*Department of Psychiatry, Yokohama City University School of Medicine, Yokohama and* [11]*Ohmiya-Kosei Hospital, Saitama, Japan*

Aim: Many psychophysiological tests have been widely researched in the search for a biological marker of schizophrenia. The exploratory eye movement (EEM) test involves the monitoring of eye movements while subjects freely view geometric figures. Suzuki *et al.* (2009) performed discriminant analysis between schizophrenia and non-schizophrenia subjects using EEM test data; consequently, clinically diagnosed schizophrenia patients were identified as having schizophrenia with high probability (73.3%). The aim of the present study was to investigate the characteristics of schizophrenia patients who were identified as having schizophrenia on EEM discriminant analysis (SPDSE) or schizophrenia patients who were identified as not having schizophrenia on EEM discriminant analysis (SPDNSE).

Methods: The data for the 251 schizophrenia subjects used in the previous discriminant-analytic study were analyzed, and the demographic or symptomatic characteristics of SPDSE and SPDNSE were investigated. As for the symptomatic features, a factor analysis of the Brief Psychiatric Rating Scale (BPRS) rating from the schizophrenia subjects was carried out.

Results: Five factors were found for schizophrenia symptoms: excitement/hostility; negative symptoms; depression/anxiety; positive symptoms; and disorganization. SPDSE had significantly higher factor scores for excitement/hostility, negative symptoms and disorganization than SPDNSE. Furthermore, the BPRS total score for the SPDSE was significantly higher than that for the SPDNSE.

Conclusion: SPDSE may be a disease subtype of schizophrenia with severe symptoms related to excitement/hostility, negative symptoms and disorganization, and EEM parameters may detect this subtype. Therefore, the EEM test may be one of the contributors to the simplification of the heterogeneity of schizophrenia.

Key words: biological marker, clinical symptoms of schizophrenia, exploratory eye movement, heterogeneity, schizophrenia.

Correspondence: Sakae Takahashi, MD, PhD, Department of Psychiatry, Nihon University School of Medicine, 30-1 Oyaguchi- Kamicho, Itabashi-ku, Tokyo 173-8610, Japan. Email: sakae@med.nihon-u.ac.jp
Received 4 September 2010; revised 29 August 2011; accepted 9 September 2011.

MANY PSYCHOPHYSIOLOGICAL TESTS have been performed in the search for a biological marker for schizophrenia.[1,2] Event-related potentials (ERP), P300,[3] P50[4] and mismatch negativity (MMN),[5,6] prepulse inhibition (PPI),[7,8] saccadic and smooth pursuit eye movements[9–12] and working memory tasks[13,14] have been widely researched. Moreover, many researchers have focused on abnormalities of working memory as an endophenotype for schizophrenia in molecular genetic studies.[15,16]

We have studied eye movements while subjects freely viewed geometric figures; this is called the exploratory eye movement (EEM) test. In most previous studies, only schizophrenia patients have consistently shown disturbances of EEM.[17–25] Moreover, the parents and siblings of schizophrenia patients had EEM dysfunctions.[26,27] In addition, EEM demonstrated a significant linkage to chromosome 22q11.[28] Chromosome 22q11 is one of the most interesting regions in the genetic etiology of schizophrenia. Microdeletions at chromosome 22q11 cause velocardio-facial syndrome (VCFS/DiGeorage syndrome: DGS), and patients with VCFS have a high risk of schizophrenia.[29,30] Furthermore, there is strong evidence that this deletion is a risk factor for schizophrenia in a genome-wide association study (GWAS) using copy number variants (CNV).[31] Therefore, we believe that EEM disturbance may be a biological marker of schizophrenia, in addition to the aforementioned physiological defects.

On the basis of these findings, we considered that the EEM test might be useful for the clinical diagnosis of schizophrenia as well. Suzuki *et al.* carried out a discriminant analysis between schizophrenia patients and non-schizophrenia subjects in a large sample using EEM test data.[32] EEM performance was recorded in 251 schizophrenia patients and 389 non-schizophrenia subjects (111 patients with mood disorder; 28 patients with neurotic disorder; 250 normal controls). As a result, 184 of the 251 clinically diagnosed schizophrenia patients were identified as having schizophrenia (sensitivity, 73.3%); and 308 of the 389 clinically diagnosed non-schizophrenia subjects were identified as non-schizophrenic (specificity, 79.2%). Based on this finding, we propose that the EEM test might be useful for the clinical diagnosis of schizophrenia.

In the discriminant-analytic study,[32] we were interested in characteristics of the schizophrenia patients who were identified as having schizophrenia on EEM discriminant analysis (SPDSE), or those who were identified as not having schizophrenia on EEM discriminant analysis (SPDNSE). Many researchers have indicated the potential heterogeneity of schizophrenia.[33–37] Hence, the EEM parameters may be able to detect different subtypes of schizophrenia. In the present study, to clarify the features of SPDSE and SPDNSE, we reanalyzed that data,[32] and focused on the demographic or symptomatic characteristics. If the characteristics of SPDSE and SPDNSE are clarified, further knowledge regarding the heterogeneity of schizophrenia may be yielded. Therefore, in the present study we discuss the features of SPDSE and SPDNSE and a further application of EEM for scientific research into schizophrenia.

METHODS

Subjects

Two hundred and fifty-one schizophrenia patients participated in the discriminant-analytic study (paranoid type, 65.3%; hebephrenic type, 15.9%; catatonic type, 1.2%; undifferentiated type, 5.2%; residual type, 9.6%; simple type, 1.6%; and unspecified type, 1.2%).[32] The patients were in/outpatients recruited from multiple centers, eight university hospitals and three affiliated hospitals. Diagnoses were made by one experienced psychiatrist according to the ICD-10 criteria for research at each university or hospital.[38] The demographic characteristics of the subjects were as follows: age, 37.9 ± 11.3 years; gender (M/F), 157/94; and duration of illness, 14.5 ± 13.1 years. The patients who had a history of alcohol abuse or illicit substance abuse, or head injury were excluded from the study; also excluded were those with convulsive, neurologic or ophthalmologic disorders.

The clinical symptoms of the schizophrenia patients were assessed using the Brief Psychiatric Rating Scale (BPRS),[39] which yielded an average total score of 41.5 ± 13.3. All BPRS ratings were done by one experienced psychiatrist in each university or hospital. Of the 251 patients with schizophrenia, 249 received neuroleptic medication. The average daily dosage is expressed as a haloperidol equivalent of 13.9 ± 10.7 mg.[40] This study was approved by the Ethics Committees of the eight universities. Written informed consent was obtained from all participants, after the procedures and possible risks of the study were fully explained.

Procedure

A standard test of the EEM using a digital eye-mark recording system (nac Image Technology, EMR-NS, Tokyo, Japan) was performed. An eye camera that detected corneal reflection of infrared light to identify eye movements, and a 15-in LCD monitor (1024 × 768 pixels) that displayed target figures for the EEM tasks (Fig. 1) were included in this system. According to the following method, three horizontal S-shaped figures (an original target figure and two figures slightly different from the original target figure) were individually displayed on the LCD monitor (Fig. 1). First, the retention task: the subject was shown the original S-shaped figure (Fig. 1a) for 15 s. Next, the comparison task: the subject was instructed to compare a figure with the original figure (Fig. 1a); they were then shown a figure slightly different from the original one, which had one bump in a different position (Fig. 1b), for 15 s. After 15 s had elapsed and with the figure still in view, the subject was asked whether it differed from the original figure and, if it did, how it differed. After the subject had replied and while the figure was still being shown, he/she was asked 'Are there any other differences?' The comparison task was then repeated with a figure without bumps (Fig. 1c).

In the digital eye-mark recording system, the detected eye movements were automatically analyzed by a digital computerized EEM analyzer. Consequently, four parameters emerged: number of eye fixations (NEF), total eye scanning length (TESL), mean eye scanning length (MESL) and responsive search score (RSS). The NEF, TESL and MESL were based on data of eye movements that occurred during 15 s of the retention task. In the comparison task, the RSS was based on data of eye movements that occurred for 5 s immediately after the question: 'Are there any other differences?' More detailed descriptions of the EEM test equipment and method are given in our previous studies.[17,20,32]

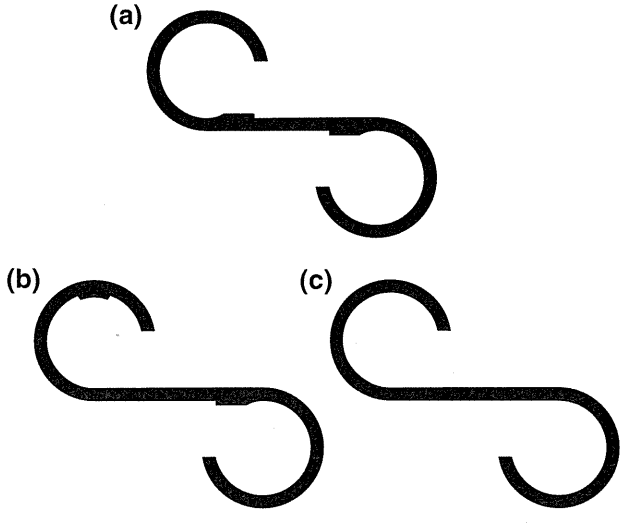

Figure 1. (a) Original target figure; (b,c) two figures slightly different from the target.

In our previous study, 184 of the 251 clinically diagnosed schizophrenia patients were identified as having schizophrenia on discriminant analysis using the EEM parameters (SPDSE).[32] The remaining 67 schizophrenia patients were identified as not having schizophrenia (SPDNSE). Table 1 lists the background data of the SPDSE and SPDNSE. In the present study we compared demographic and symptomatic characteristics of SPDSE with those of SPDNSE.

Statistical analysis

Group differences on the demographic and symptomatic data were assessed using the *t*-test or the χ^2 test. For group comparison of the symptomatic data, scores for factors extracted by factor analysis of BPRS ratings and BPRS total scores were used. In the factor analysis, we conducted a principal component analysis with orthogonal rotation (Varimax method) according to previous studies.[41–43] Moreover, based on prior studies, factors with eigenvalues >1.0 were considered to be meaningful.[41,43] All statistical analyses were performed using SPSS for Windows version 17.0. The statistical significance was set at $P < 0.05$ (two-tailed).

RESULTS

Group comparisons (SPDSE vs SPDNSE) of demographic characteristics

There were no significant differences for age, sex, duration of illness or drug dosage between SPDSE and SPDNSE.

Group comparisons (SPDSE vs SPDNSE) of subtypes and clinical symptoms

There were no significant differences for the subtypes between SPDSE and SPDNSE.

Table 1. Subject characteristics

	SPDSE ($n = 184$)	SPDNSE ($n = 67$)
Age (years), mean ± SD	38.0 ± 12.6	37.7 ± 12.0
Gender (M/F)	112/72	45/22
Duration of illness (years), mean ± SD	14.6 ± 13.9	14.3 ± 10.8
Equivalent dose of haloperidol (mg), mean ± SD†	14.4 ± 11.1	12.5 ± 9.7
Subtype, *n* (%)		
Paranoid	120 (65.3)	44 (65.6)
Hebephrenic	30 (16.3)	10 (14.9)
Catatonic	3 (1.6)	0 (0)
Undifferentiated	9 (4.9)	4 (6.0)
Residual	18 (9.8)	6 (9.0)
Simple	3 (1.6)	1 (1.5)
Unspecified	1 (0.5)	2 (3.0)

†In each group (SPDSE or SPDNSE), one patient did not receive neuroleptic medication, respectively.
SPDSE, schizophrenia patients identified as having schizophrenia on exploratory eye movement (EEM) discriminant analysis; SPDNSE, schizophrenia patients identified as not having schizophrenia on EEM discriminant analysis.

Factor analysis of BPRS items

Table 2 lists the factors and factor loadings derived using principal component analysis of BPRS rating. The principal component analysis extracted five factors that accounted for 70.0% of the variance. Based on previous studies, BPRS items with factor loadings >0.5 were considered to load on the

Table 2. Factors and factor loadings derived in BPRS principal component analysis

	Factor				
	1	2	3	4	5
BPRS items					
Somatic concern	0.033	0.080	0.615	0.505	−0.074
Anxiety	0.184	0.123	0.727	0.272	−0.126
Emotional withdrawal	0.070	0.879	0.139	0.043	0.140
Conceptual disorganization	0.401	0.298	0.113	0.356	0.629
Guilt feelings	0.091	−0.085	0.670	−0.157	0.487
Tension	0.416	0.404	0.543	0.106	−0.126
Mannerisms and posturing	0.383	0.457	0.178	0.339	0.393
Grandiosity	0.736	−0.115	0.133	0.124	0.158
Depressive mood	0.192	0.287	0.722	0.041	−0.058
Hostility	0.783	0.077	0.213	0.210	−0.118
Suspiciousness	0.477	0.126	0.273	0.546	−0.111
Hallucinatory behavior	0.246	0.171	0.045	0.805	0.067
Motor retardation	0.004	0.850	0.179	0.159	0.083
Uncooperativeness	0.677	0.432	−0.057	0.122	0.086
Unusual thought content	0.276	0.170	0.133	0.734	0.322
Blunted affect	0.021	0.857	0.083	0.168	0.160
Excitement	0.778	−0.023	0.195	0.218	0.153
Disorientation	−0.034	0.241	−0.241	0.056	0.659
Variance explained (total = 70.0%)†	17.5	17.5	14.1	12.6	8.4

†Cumulative or percentage of variance explained is rounded off; therefore, the cumulative percentage is not identical to the sum of each percentage. Underline, BPRS items with factor loadings >0.5.
BPRS, Brief Psychiatric Rating Scale.

Table 3. Mean factor scores and BPRS total score (mean ± SD)

	SPDSE ($n = 184$)	SPDNSE ($n = 67$)	t (d.f. = 249)	z
Factor				
1 Excitement/hostility	0.09 ± 1.07	−0.25 ± 0.74		−2.16*
2 Negative symptoms	0.10 ± 1.01	−0.27 ± 0.93	−2.57*	
3 Depression/anxiety	−0.03 ± 1.03	0.07 ± 0.92	0.70	
4 Positive symptoms	0.03 ± 1.03	−0.07 ± 0.92	−0.71	
5 Disorganization	0.08 ± 1.03	−0.21 ± 0.89	−2.06*	
BPRS total score (mean ± SD)	43.08 ± 13.48	37.51 ± 12.10	−2.98*	

*$P < 0.05$.

BPRS, Brief Psychiatric Rating Scale; SPDSE, schizophrenia patients identified as having schizophrenia on exploratory eye movement (EEM) discriminant analysis; SPDNSE, schizophrenia patients identified as not having schizophrenia on EEM discriminant analysis.

respective factor.[41,43] Consequently, we summarized the five factors as follows: factor 1 loaded heavily in grandiosity, hostility, uncooperativeness and excitement; factor 2 had heavy loadings in emotional withdrawal, motor retardation and blunted affect; factor 3 loaded heavily in somatic concern, anxiety, guilt feelings, tension and depressive mood; factor 4 had heavy loadings in somatic concern, suspiciousness, hallucinatory behavior and unusual thought content; factor 5 loaded heavily in conceptual disorganization and disorientation. Accordingly, we interpreted the five factors as having the following dimensions: factor 1, excitement/hostility (17.5% of total variance); factor 2, negative symptoms (17.5%); factor 3, depression/anxiety (14.1%); factor 4, positive symptoms (12.6%); and factor 5, disorganization (8.4%).

Group comparisons (SPDSE vs SPDNSE) of factor scores

Table 3 lists the mean factor scores of the five factors for SPDSE and SPDNSE. SPDSE had significantly higher scores of excitement/hostility ($P = 0.005$), negative symptoms ($P = 0.011$) and disorganization ($P = 0.040$) than SPDNSE. Furthermore, the BPRS total score of SPDSE was significantly higher than that of the SPDNSE ($P = 0.003$). For the excitement/hostility factor, the Levene test for equality of variance did not show homoskedasticity between the two groups. Therefore, the P-value for the excitement/hostility factor was based on an unequal-variance t-value. In order to confirm the result of the excitement/hostility factor, we also performed the non-parametric test, Mann–Whitney U-test. Consequently, SPDSE also demonstrated significantly higher scores of excitement/hostility than SPDNSE on non-parametric analysis ($P = 0.031$).

DISCUSSION

Suzuki *et al.* performed discriminant analysis between schizophrenia patients and non-schizophrenia subjects using the EEM test data.[32] As a result, 184 of the 251 clinically diagnosed schizophrenia patients were identified as having schizophrenia (sensitivity, 73.3%). In the present study, results of the factor analysis of BPRS ratings from the aforementioned 251 schizophrenia subjects produced five factors of symptoms (excitement/hostility; negative symptoms; depression/anxiety; positive symptoms; and disorganization). Excitement/hostility, negative symptoms and disorganization were more predominant in the 184 SPDSE subjects compared to the SPDNSE subjects. Furthermore, the BPRS total score of the SPDSE was significantly higher than that of the SPDNSE. Consequently, the SPDSE group may consist of patients with severe schizophrenia, and the severity of symptoms in SPDSE was found to be due mainly to excitement/hostility, negative symptoms and disorganization.

Evidence for five dimensions in schizophrenia symptoms was found in the present study. Many studies have proposed similar five-factor structures.[41–47] In these studies, the Positive and Negative Syndrome Scale (PANSS) has been used as the symptom rating scale. In contrast, the present data were based on the BPRS. All items of the BPRS, however, are included in the PANSS.[39,48] Therefore, it

is possible that the present findings reflect the past studies of the factor analysis using PANSS items. Consequently, although items included for each factor in previous studies and the present study were not identical, the present findings of the factor analysis are distinctly similar to previous factor-analytic study results. Thus, we consider that the present five-factor structure may be meaningful for the symptomatology of schizophrenia. The PANSS, however, is more informative than the BPRS, therefore the present study may be limited by this issue.

In the present study, demographic data, age, sex, duration of illness and drug dosage for SPDSE and SPDNSE were not significantly different. But there were significant differences for symptom, excitement/hostility, negative symptoms and disorganization between SPDSE and SPDNSE. In our previous study, EEM parameters were not influenced by the demographic data.[27,32] Moreover, one of the EEM parameter, RSS, which was principally used in the discriminant analysis of SPDSE, was associated with negative symptoms.[17] Altogether, we believe that differences between SPDSE and SPDNSE in the EEM may relate to symptoms of schizophrenia, but not demographic data, sex, age, course of illness or medication.

With regard to the ICD-10 subtypes, we also did not find significant differences between SPDSE and SPDNSE. This finding seems to conflict with the significant differences of the BPRS scores between the two groups. Lykouras *et al.* investigated relationships between the DSM-III-R schizophrenia subtypes and the PANSS scores.[49] As a result, paranoid type was associated with positive symptoms, and disorganized type linked to negative symptoms. In addition to disorganized type, however, catatonic type related to negative symptoms. Moreover, based on the DSM-IV-TR, the schizophrenia symptoms have been divided into three dimensions.[50] However, past reports and the present study propose that schizophrenia may be symptomatically more complex.[41–47] This has also been indicated by Wolthaus *et al.*[47] In this way, subtypes and dimensions of the diagnostic criteria are often not consistent with those of the symptomatic rating scales. There is, however, a possible limitation to the present study. Although we discussed diagnoses using the ICD-10 criteria and the BPRS scores in detail, inter-rater and intra-rater reliabilities for those were not formally assigned. Consequently, if they were formally assigned, the ICD-10 subtypes might coincide with the BPRS scores.

Based on the present findings, SPDSE may be associated with excitement/hostility, negative symptoms and disorganization in the present five symptomatic dimensions. Accordingly, SPDSE may have three different dimensions; but it can also be said that SPDSE may be a schizophrenia subtype characterized by these three dimensions. The present findings may indicate that there is a putative subtype of schizophrenia with severe symptoms related to excitement/hostility, negative symptoms and disorganization. Furthermore, the EEM abnormality may be a biological marker for this subtype of schizophrenia. There is another point worth making. As mentioned here, the EEM parameter, RSS was associated with negative symptoms.[17] Thus, negative symptoms may be the most specific of the three dimensions to the subtype.

In addition to the schizophrenia patients, their parents and siblings also had EEM dysfunction.[26,27] Therefore, we considered that the EEM abnormality may be an intermediate phenotype of schizophrenia, and may be useful for linkage studies of schizophrenia. Indeed, we found a significant linkage to chromosome 22q11.2–12.1 in our previous linkage study using EEM impairment as an endophenotype of schizophrenia.[28] Chromosome 22q11 is one of the most interesting regions for the etiology of schizophrenia. Moreover, in this area, there are several candidate genes for schizophrenia, for example COMT, PRODH and ZDHHC8, and so on.[29,30]

Many researchers have presented positive linkage and association findings with schizophrenia, but initial findings have often not been replicated.[30] One of the most significant causes of conflicting results in the present molecular genetic studies of schizophrenia may be the potential heterogeneity of schizophrenia. Several investigators have suggested that schizophrenia is not a single disease entity but may reflect common symptomatology caused by several distinct genetic abnormalities.[33–37] As mentioned here, the EEM deficits are linked to chromosome 22q11. If the EEM parameters are associated with a schizophrenia subtype with severe symptoms related to excitement/hostility, negative symptoms and disorganization, chromosome 22q11 and genes of 22q11 may relate to this subtype. In this manner, if we are able to find a new subtype using the EEM disturbances, and clarify the heterogeneity of schizophrenia, then linkage or association studies for schizophrenia using the subtype may yield further knowledge regarding the genetic influences on schizophrenia.

In conclusion, we have found evidence for the existence of five dimensions of schizophrenia symptoms: excitement/hostility; negative symptoms; depression/anxiety; positive symptoms; and disorganization. Schizophrenia patients with EEM abnormalities (SPDSE) may have severe symptoms related to excitement/hostility, negative symptoms and disorganization. In light of the heterogeneity of schizophrenia, SPDSE may be a disease subtype of schizophrenia with the aforementioned symptomatic features; and the EEM parameters may detect this subtype. Therefore, EEM may be one of the contributors to the simplification of the heterogeneity of schizophrenia. Consequently, we may apply EEM to other scientific studies as an endophenotype for schizophrenia.

ACKNOWLEDGMENT

The present study was supported by a Health and Labor Sciences Research Grant for Research on Psychiatric and Neurological Diseases and Mental Health (H16-KOKORO-003) from the Ministry of Health, Labor and Welfare, Japan.

REFERENCES

1. Braff DL. Information processing and attention dysfunctions in schizophrenia. *Schizophr. Bull.* 1993; **19**: 233–259.
2. Szymanski S, Kane JM, Lieberman JA. A selective review of biological markers in schizophrenia. *Schizophr. Bull.* 1991; **17**: 99–111.
3. Bharath S, Gangadhar BN, Janakiramaiah N. P300 in family studies of schizophrenia: Review and critique. *Int. J. Psychophysiol.* 2000; **38**: 43–54.
4. Potter D, Summerfelt A, Gold J, Buchanan RW. Review of clinical correlates of P50 sensory gating abnormalities in patients with schizophrenia. *Schizophr. Bull.* 2006; **32**: 692–700.
5. Devrim-Ucok M, Keskin-Ergen HY, Ucok A. Mismatch negativity at acute and post-acute phases of first-episode schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* 2008; **258**: 179–185.
6. Michie PT. What has MMN revealed about the auditory system in schizophrenia? *Int. J. Psychophysiol.* 2001; **42**: 177–194.
7. Parwani A, Duncan EJ, Bartlett E *et al.* Impaired prepulse inhibition of acoustic startle in schizophrenia. *Biol. Psychiatry* 2000; **47**: 662–669.
8. Weike AI, Bauer U, Hamm AO. Effective neuroleptic medication removes prepulse inhibition deficits in schizophrenia patients. *Biol. Psychiatry* 2000; **47**: 61–70.
9. Fukushima J, Fukushima K, Chiba T, Tanaka S, Yamashita I, Kato M. Disturbances of voluntary control of saccadic eye movements in schizophrenic patients. *Biol. Psychiatry* 1988; **23**: 670–677.
10. Haraldsson HM, Ettinger U, Magnusdottir BB, Sigmundsson T, Sigurdsson E, Petursson H. Eye movement deficits in schizophrenia: Investigation of a genetically homogenous Icelandic sample. *Eur. Arch. Psychiatry Clin. Neurosci.* 2008; **258**: 373–383.
11. Levy DL, Holzman PS, Matthysse S, Mendell NR. Eye tracking and schizophrenia: A selective review. *Schizophr. Bull.* 1994; **20**: 47–62.
12. Sereno AB, Holzman PS. Antisaccades and smooth pursuit eye movements in schizophrenia. *Biol. Psychiatry* 1995; **37**: 394–401.
13. Barrantes-Vidal N, Aguilera M, Campanera S *et al.* Working memory in siblings of schizophrenia patients. *Schizophr. Res.* 2007; **95**: 70–75.
14. Goldman-Rakic PS. The physiological approach: Functional architecture of working memory and disordered cognition in schizophrenia. *Biol. Psychiatry* 1999; **46**: 650–661.
15. Braff DL, Freedman R, Schork NJ, Gottesman II. Deconstructing schizophrenia: An overview of the use of endophenotypes in order to understand a complex disorder. *Schizophr. Bull.* 2007; **33**: 21–32.
16. Egan MF, Goldberg TE, Kolachana BS *et al.* Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl Acad. Sci. USA* 2001; **98**: 6917–6922.
17. Kojima T, Matsushima E, Ando K *et al.* Exploratory eye movements and neuropsychological tests in schizophrenic patients. *Schizophr. Bull.* 1992; **18**: 85–94.
18. Kojima T, Matsushima E, Iwama H *et al.* Visual perception process in amphetamine psychosis and schizophrenia. *Psychopharmacol. Bull.* 1986; **22**: 768–773.
19. Kojima T, Matsushima E, Nakajima K *et al.* Eye movements in acute, chronic, and remitted schizophrenics. *Biol. Psychiatry* 1990; **27**: 975–989.
20. Kojima T, Matsushima E, Ohta K *et al.* Stability of exploratory eye movements as a marker of schizophrenia: A WHO multi-center study. *Schizophr. Res.* 2001; **52**: 203–213.
21. Kojima T, Potkin SG, Kharazmi M, Matsushima E, Herrera J, Shimazono Y. Limited eye movement patterns in chronic schizophrenic patients. *Psychiatry Res.* 1989; **28**: 307–314.
22. Matsukawa Y, Takahashi S, Aoki M *et al.* Patients with systemic lupus erythematosus show a normal responsive search score in exploratory eye movement analysis: Comparison with schizophrenia. *Ann. Rheum. Dis.* 2002; **61**: 748–750.
23. Matsushima E, Kojima T, Ohbayashi S, Ando H, Ando K, Shimazono Y. Exploratory eye movements in schizophrenic patients and patients with frontal lobe lesions. *Eur. Arch. Psychiatry Clin. Neurosci.* 1992; **241**: 210–214.

24. Obayashi S, Matsushima E, Okubo Y, Ohkura T, Kojima T, Kakuma T. Relationship between exploratory eye movements and clinical course in schizophrenic patients. *Eur. Arch. Psychiatry Clin. Neurosci.* 2001; **251**: 211–216.
25. Tonoya Y, Matsui M, Kurachi M, Kurokawa K, Sumiyoshi T. Exploratory eye movements in schizophrenia: Effects of figure size and the instruction on visual search. *Eur. Arch. Psychiatry Clin. Neurosci.* 2002; **252**: 255–261.
26. Takahashi S, Matsushima E, Kojima T, Tanabe E, Yara K, Hagiwara M. Study of the families and twins of schizophrenics. **In**: Kojima T, Matsushima E, Ando K (eds). *Eyes and the Mind.* Karger, Basel, 2000; 75–79.
27. Takahashi S, Tanabe E, Yara K, Matsuura M, Matsushima E, Kojima T. Impairment of exploratory eye movement in schizophrenia patients and their siblings. *Psychiatry Clin. Neurosci.* 2008; **62**: 487–493.
28. Takahashi S, Ohtsuki T, Yu SY *et al.* Significant linkage to chromosome 22q for exploratory eye movement dysfunction in schizophrenia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 2003; **123**: 27–32.
29. Arinami T. Analyses of the associations between the genes of 22q11 deletion syndrome and schizophrenia. *J. Hum. Genet.* 2006; **51**: 1037–1045.
30. Kirov G, O'Donovan MC, Owen MJ. Finding schizophrenia genes. *J. Clin. Invest.* 2005; **115**: 1440–1448.
31. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **455**: 237–241.
32. Suzuki M, Takahashi S, Matsushima E *et al.* Exploratory eye movement dysfunction as a discriminator for schizophrenia: A large sample study using a newly developed digital computerized system. *Eur. Arch. Psychiatry Clin. Neurosci.* 2009; **259**: 186–194.
33. Sawa A, Snyder SH. Schizophrenia: Diverse approaches to a complex disease. *Science* 2002; **296**: 692–695.
34. Takahashi S, Faraone SV, Lasky-Su J, Tsuang MT. Genome-wide scan of homogeneous subtypes of NIMH genetics initiative schizophrenia families. *Psychiatry Res.* 2005; **133**: 111–122.
35. Tsuang M. Schizophrenia: Genes and environment. *Biol. Psychiatry* 2000; **47**: 210–220.
36. Tsuang MT, Faraone SV. The case for heterogeneity in the etiology of schizophrenia. *Schizophr. Res.* 1995; **17**: 161–175.
37. Tsuang MT, Lyons MJ, Faraone SV. Heterogeneity of schizophrenia. Conceptual models and analytic strategies. *Br. J. Psychiatry* 1990; **156**: 17–26.
38. World Health Organization. *The International Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research.* World Health Organization, Geneva, 1993.
39. Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol. Rep.* 1962; **10**: 799–812.
40. Inagaki A, Inada T. Dose equivalence of psychotropic drugs. Part 18: Dose equivalence of psychotropic drugs: 2006-version. *Jpn. J. Clin. Psychopharmacol.* 2006; **9**: 1443–1447 (in Japanese).
41. Klingberg S, Wittorf A, Wiedemann G. Disorganization and cognitive impairment in schizophrenia: Independent symptom dimensions? *Eur. Arch. Psychiatry Clin. Neurosci.* 2006; **256**: 532–540.
42. Lancon C, Auquier P, Nayt G, Reine G. Stability of the five-factor structure of the Positive and Negative Syndrome Scale (PANSS). *Schizophr. Res.* 2000; **42**: 231–239.
43. Mass R, Schoemig T, Hitschfeld K, Wall E, Haasen C. Psychopathological syndromes of schizophrenia: Evaluation of the dimensional structure of the positive and negative syndrome scale. *Schizophr. Bull.* 2000; **26**: 167–177.
44. Emsley R, Rabinowitz J, Torreman M, Group R-I-EPGW. The factor structure for the Positive and Negative Syndrome Scale (PANSS) in recent-onset psychosis. *Schizophr. Res.* 2003; **61**: 47–57.
45. Lindenmayer JP, Grochowski S, Hyman RB. Five factor model of schizophrenia: Replication across samples. *Schizophr. Res.* 1995; **14**: 229–234.
46. Marder SR, Davis JM, Chouinard G. The effects of risperidone on the five dimensions of schizophrenia derived by factor analysis: Combined results of the North American trials. *J. Clin. Psychiatry* 1997; **58**: 538–546.
47. Wolthaus JE, Dingemans PM, Schene AH *et al.* Component structure of the positive and negative syndrome scale (PANSS) in patients with recent-onset schizophrenia and spectrum disorders. *Psychopharmacology (Berl.)* 2000; **150**: 399–403.
48. Kay SR, Opler LA, Fiszbein A. *Positive and Negative Syndrome Scale (PANSS) Rating Manual.* Multi-Health Systems, Toronto, 1991.
49. Lykouras L, Oulis P, Daskalopoulou E *et al.* Clinical subtypes of schizophrenic disorders: A cluster analytic study. *Psychopathology* 2001; **34**: 23–28.
50. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision; DSM-IV-TR.* American Psychiatric Association, Washington DC, 2000.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 19 March 2013
doi: 10.3389/fpsyg.2013.00105

Cross-cultural differences in the processing of non-verbal affective vocalizations by Japanese and Canadian listeners

Michihiko Koeda[1,2]*, Pascal Belin[2], Tomoko Hama[3], Tadashi Masuda[4], Masato Matsuura[3] and Yoshiro Okubo[1]

[1] Department of Neuropsychiatry, Nippon Medical School, Tokyo, Japan
[2] Voice Neurocognition Laboratory, Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK
[3] Department of Biofunctional Informatics, Tokyo Medical and Dental University, Tokyo, Japan
[4] Division of Human Support System, Faculty of Symbiotic Systems Science, Fukushima University, Fukushima, Japan

Edited by:
Anjali Bhatara, Université Paris Descartes, France

Reviewed by:
Jan Van Den Stock, Katholieke Universiteit Leuven, Belgium
Keiko Ishii, Kobe University, Japan

***Correspondence:**
Michihiko Koeda, Department of Neuropsychiatry, Nippon Medical School, 1-1-5, Sendagi, Bunkyo-ku, Tokyo 113-8603, Japan.
e-mail: mkoeda@nms.ac.jp

The Montreal Affective Voices (MAVs) consist of a database of non-verbal affect bursts portrayed by Canadian actors, and high recognitions accuracies were observed in Canadian listeners. Whether listeners from other cultures would be as accurate is unclear. We tested for cross-cultural differences in perception of the MAVs: Japanese listeners were asked to rate the MAVs on several affective dimensions and ratings were compared to those obtained by Canadian listeners. Significant Group × Emotion interactions were observed for ratings of Intensity, Valence, and Arousal. Whereas Intensity and Valence ratings did not differ across cultural groups for sad and happy vocalizations, they were significantly less intense and less negative in Japanese listeners for angry, disgusted, and fearful vocalizations. Similarly, pleased vocalizations were rated as less intense and less positive by Japanese listeners. These results demonstrate important cross-cultural differences in affective perception not just of non-verbal vocalizations expressing positive affect (Sauter et al., 2010), but also of vocalizations expressing basic negative emotions.

Keywords: montreal affective voices, emotion, voice, cross-cultural differences, social cognition

INTRODUCTION

Vocal affective processing has an important role in ensuring smooth communication during human social interaction as well as facial affective processing. Facial expressions are generally recognized as the universal language of emotion (Ekman and Friesen, 1971; Ekman et al., 1987; Ekman, 1994; Izard, 1994; Jack et al., 2012): however, several studies have demonstrated cross-cultural differences in facial expression between Western and Eastern groups (Ekman and Friesen, 1971; Ekman et al., 1987; Matsumoto and Ekman, 1989; Izard, 1994; Yrizarry et al., 1998; Elfenbein and Ambady, 2002; Jack et al., 2009, 2012). Whether such cross-cultural differences also exist in the recognition of emotional vocalizations is not clear.

Most previous cross-cultural studies of auditory perception have investigated the processing of emotional Valence using word stimuli (Scherer and Wallbott, 1994; Kitayama and Ishii, 2002; Ishii et al., 2003; Min and Schirmer, 2011). One important study demonstrated cross-cultural differences in the rating of Intensity when subjects recognized meaning of the words with major emotions such as joy, fear, anger, sadness, and disgust (Scherer and Wallbott, 1994). Another previous study examined cross-cultural differences in the perception of emotional words (Kitayama and Ishii, 2002). This study indicated that native English speakers spontaneously pay more attention to verbal content than to vocal tone when they recognize emotional words, whereas native Japanese speakers spontaneously attend more to vocal tone than to verbal content. The other study has shown that Japanese are more sensitive to vocal tone compared to Dutch participants in the experiment of the multisensory perception of emotion (Tanaka et al., 2010). Further, one other study demonstrated cross-cultural differences in semantic processing of emotional words (Min and Schirmer, 2011), but found no difference in the processing of emotional prosody between native and non-native listeners. These studies suggest cross-cultural differences in auditory recognition of emotional words.

Studies of affective perception in speech prosody are made complex, in particular, by the potential interactions between the affective and the linguistic contents of speech (Scherer et al., 1984; Murray and Arnott, 1993; Banse and Scherer, 1996; Juslin and Laukka, 2003). To avoid this interaction, some studies have controlled the processing of semantic content using pseudo-words (Murray and Arnott, 1993; Schirmer et al., 2005) or pseudo-sentences (Ekman and Friesen, 1971; Pannekamp et al., 2005; Schirmer et al., 2005). The other previous study has employed a set of low-pass filtered vocal stimuli to select the final set of emotional utterances (Ishii et al., 2003), i.e., non-verbal vocalizations often accompanying strong emotional states such as laughs or screams of fear. Non-verbal affective vocalizations are ideally suited to investigations of cross-cultural differences in the perception of affective information in the voice since they eliminate the need to account for language differences between groups.

A recent study compared the perception of such non-verbal affective vocalizations by listeners from two highly different cultures: Westerners vs. inhabitants of remote Namibian villages. Non-verbal vocalizations expressing negative emotions could be recognized by the other culture much better than those expressing

positive emotions, which lead the authors to propose that a number of primarily negative emotions have vocalizations that can be recognized across cultures while most positive emotions are communicated with culture-specific signals (Sauter et al., 2010). However this difference could be specific to English vs. Namibian groups, reflecting for instance different amounts of exposure to vocalizations through media or social interactions, and might not generalize to other cultures.

In the present experiment we tested for cross-cultural differences in perception of affective vocalizations between two cultures much more comparable in socio-economic status and exposure to vocalizations: Canadian vs. Japanese participants. Stimuli consisted of the Montreal Affective Voices (MAVs; Belin et al., 2008), a set of 90 non-verbal affect bursts produced by 10 actors and corresponding to emotions of Anger, Disgust, Fear, Pain, Sadness, Surprise, Happiness, and Pleasure. The MAVs have been validated in a sample of Canadian listeners and showed high inter-reliability in judgments of emotional Intensity, Valence, and Arousal as well as hit rates in emotional recognition (Belin et al., 2008). Here, we collected affective ratings using similar procedures in Japanese listeners and compared those ratings to those obtained in the Canadian listeners. Before the experiment, we predicted that ratings of negative emotion are culturally universal although cross-cultural differences would exist in ratings of positive emotion.

MATERIALS AND METHODS

SUBJECTS

Thirty Japanese subjects (male 15, female 15) participated in this study. The average age was 22.3 ± 1.4 years. The educational years of Japanese subjects were 14.1 ± 0.3. The data of Japanese subjects were compared with 29 Canadian subjects (male 14, female 15); average age: 23.3 ± 1.5 years (Belin et al., 2008). Both Japanese and Canadian participants consisted exclusively of undergraduate students.

After a thorough explanation of the study, written informed consent was obtained from all subjects, and the study was approved by the Ethics Committee of Nippon Medical School.

VOICE MATERIALS

The MAVs: 10 French-Canadian actors expressed specific emotional vocalizations and non-emotional vocalizations (neutral sounds) using "ah" sounds. The eight emotional vocalizations were angry, disgusted, fearful, painful, sad, surprised, happy, and pleased. The simple "ah" sounds were used to control the influence of lexical-semantic processing. Since each of the eight emotional vocalizations and the neutral vocalization were spoken by 10 actors, the total number of MAVs sounds was 90. The MAVs are available at: http://vnl.psy.gla.ac.uk/

EVALUATION SCALE

Each emotional vocalization was evaluated using three criteria: perceived emotional Intensity in each of the eight Emotions, perceived Valence, and perceived Arousal. Each scale had a range from 0 to 100.

The Valence scale represented the extent of positive or negative emotion expressed by the vocalization: 0 was extremely negative, and 100 was extremely positive. The Arousal scale represented the extent of excitement expressed by the vocalization: 0 was extremely calm, and 100 was extremely excited. The Intensity scale represented the Intensity of a given emotion expressed by the vocalization: 0 was not at all intense, and 100 was extremely intense. The Intensity scale was used for eight emotions: Anger, Disgust, Fear, Pain, Sadness, Surprise, Happiness, and Pleasure.

METHODS OF EVALUATION BY PARTICIPANTS

The MAVs vocalizations were played on a computer in a pseudo-random order. The subjects listened with headphones at a comfortable hearing level, and they evaluated each emotional vocalization for perceived Intensity, Valence, and Arousal using a visual analog scale in English on a computer (10 ratings per vocalization: 8 Intensity ratings, 1 Valence rating, 1 Arousal rating). Simultaneously, participants were given a printed Japanese translation of the scale labels, and by referring to this Japanese sheet, the test was performed using exactly the same procedure as in the Canadian study (Belin et al., 2008). All Japanese participants performed the experiment using a translation sheet with emotional words translated from English to Japanese. Based on previous studies (Scherer and Wallbott, 1994), the Japanese translation of English emotional labels was independently assessed by three clinical psychologists. Through their discussion, the appropriate emotional labels were determined.

STATISTICAL ANALYSIS

Statistical calculations were made using SPSS (Statistical Package for Social Science) Version 19.0. The Japanese data and the Canadian published data, with permission to verify, were statistically analyzed. A previous study demonstrated gender effects in Canadian participants using the MAV (Belin et al., 2008). Using the same methods to reveal the gender effects, an ANOVA with Emotion, Actor gender, and Participant gender as factors was calculated for ratings by the Japanese listeners. Further, to clarify the cross-cultural effect between Japanese and Canadian participants, three mixed two-way ANOVAs were calculated on ratings of Intensity, Valence, and Arousal. For each mixed ANOVA, to verify the equality of the variance of the differences by Emotions, Mauchly's sphericity was calculated. If the sphericity could not be assumed using Mauchly's test, Greenhouse–Geisser's correction was calculated.

RELIABILITY AND ACCURACY

First, we analyzed the inter-subject reliability of the ratings using Cronbach's alpha. Next, we examined the Intensity ratings for their sensitivity (hit rate, by Emotion) and specificity (correct rejection rate, by rating scale). Based on the previous report (Belin et al., 2008), the accuracy of emotional recognition was investigated using measures of sensitivity (hit rate, by Emotion) and specificity (correct rejection rate, by rating scale). For each vocalization, participants rated the perceived emotional Intensity along each of eight different scales (Anger, Disgust, Fear, Pain, Sadness, Surprise, Happiness, and Pleasure). To calculate sensitivity, for a given portrayed emotion, a maximum Intensity rating in the corresponding scale (i.e., if Intensity rating of Anger was highest when the subject listened to angry vocalization) was taken as a hit; otherwise, as a miss. In other words, emotions with high hit rates are those that

are well recognized, i.e., that scored highest on the scale of the intended emotion. Conversely, specificity relates to the extent to which the rating scale measures what it is intended to measure. To calculate specificity for a given rating scale, if the maximum score was obtained for the corresponding portrayed emotion across the eight vocalizations from one actor (i.e., when the subject listened to disgusted vocalization by actor 1, if rating of Disgust was highest in the eight emotional items), it was taken as a correct rejection; otherwise, as a false alarm. A highly specific rating scale is one rating scale for which the corresponding vocalization obtains the highest score. In other words, it is a measure of how a rating scale is specific to an emotion.

RESULTS

AFFECTIVE RATING

Inter-participant (30 participants) reliability across the 90 items [10 ratings scales: (Valence, Arousal, eight emotional Intensities) × (9 Emotional sounds) was analyzed: Cronbach's alpha = Japanese: 0.941, $F(89, 299) = 230.6$, $p < 0.001$]. Since this reliability for 30 subjects is very high, the ratings of 10 actors' vocalizations were averaged with the ratings of all 30 Japanese participants. [Canadian participants had an inter-participant reliability rating of 0.978 (Belin et al., 2008)]. **Table 1** shows the averaged ratings of Intensity, Valence, and Arousal for the present sample of Japanese participants and the Canadian participants in the study of Belin et al. (2008). **Figure 1** shows the distribution (average ± 2 SD) of ratings of 1-1. Intensity, 1-2. Valence, and 1-3. Arousal in Japanese and Canadian participants.

INTENSITY

A mixed two-way ANOVA with listeners' Group (Japanese, Canadian) and Emotion ($n = 8$) as factors was calculated on Intensity scores. A significant main effect was revealed between listener's Groups [$F(1, 57) = 20.828$, $p < 0.001$] as well as among the Emotions [$F(5.5, 313.5) = 40.520$, $p < 0.001$; Greenhouse–Geisser's test]. Crucially, a significant interaction between Group and Emotion was observed, $F(5.5, 313.5) = 9.137$, $p < 0.001$, (**Figure 1A**) indicating that rating differences between the two groups varied with the specific Emotion considered. *Post hoc* tests showed that Intensity ratings from Japanese listeners were significantly lower than ratings from Caucasian listeners for Anger, Disgust, Fear, Surprise, and Pleasure (t-test, $p < 0.05/8$: Anger, $t = -4.358$; Disgust, $t = -4.756$; Fear, $t = -3.073$; Surprise, $t = -2.851$; Pleasure, $t = -6.737$: **Table 1**; **Figure 1A**).

VALENCE

A mixed two-way ANOVA with listeners' Group (Japanese, Canadian) and Emotion ($n = 9$) as factors was calculated on Valence scores. There was a significant main effect of listeners' Group: $F(1, 57) = 5.920$, $p < 0.018$, as well as a significant main effect of Emotion $F(4.3, 244.3) = 224.926$, $p < 0.001$ (Greenhouse–Geisser's test). Crucially, a significant interaction between Group and Emotion was observed: $F(4.3, 244.3) = 25.101$, $p < 0.001$ (**Figure 1B**) indicating that rating differences between the two groups varied with the specific Emotion considered. *Post hoc* tests showed that Valence ratings from Japanese listeners were significantly higher than ratings from Caucasian listeners for Anger, Disgust, Fear

Table 1 | The mean (*M*) ratings of 1. Intensity, 2. Valence, and 3. Arousal for 10 actors' voices (5 male actors, 5 female actors) by 30 Japanese and 30 Canadian participants.

	1-1: Intensity					1-2: Valence					1-3: Arousal				
Vocal expression	**Japan**		**Canada**		***t*-test**	**Vocal expression**	**Japan**		**Canada**		***t*-test**	**Vocal expression**	**Japan**		**Canada**		***t*-test**
	M	**SEM**	***M***	**SEM**	***P***		***M***	**SEM**	***M***	**SEM**	***P***		***M***	***SEM***	***M***	***SEM***	***P***
Neutral	–	–	–	–		Neutral	46.7	1.2	47.6	1.7	n.s.	Neutral	21.5	3.6	33.0	2.7	n.s.
Angry	55.0	3.5	74.7	2.9	<0.001*	Angry	39.2	2.9	16.6	1.7	<0.001*	Angry	78.3	1.6	73.0	2.1	n.s.
Disgusted	45.5	4.0	70.3	3.3	<0.001*	Disgusted	34.8	1.4	24.4	2.5	<0.001*	Disgusted	40.3	2.1	36.9	2.7	n.s.
Fearful	54.1	3.6	67.6	2.4	0.003*	Fearful	34.0	2.7	24.3	1.3	0.002*	Fearful	78.3	1.7	72.6	2.2	n.s.
Painful	52.4	3.7	57.8	3.0	n.s.	Painful	32.3	2.7	23.5	1.9	n.s.	Painful	69.3	2.0	66.4	2.1	n.s.
Sad	75.1	2.5	77.0	2.7	n.s.	Sad	21.7	2.2	19.3	1.6	n.s.	Sad	65.2	3.1	45.0	3.5	<0.001*
Surprised	65.4	3.4	77.2	2.4	0.006*	Surprised	45.4	2.3	39.1	1.1	n.s.	Surprised	70.5	2.3	71.2	1.9	n.s.
Happy	76.2	3.5	81.3	2.8	n.s.	Happy	84.9	2.0	85.2	2.1	n.s.	Happy	67.4	2.6	58.3	3.1	n.s.
Pleased	30.0	3.8	62.1	2.8	<0.001*	Pleased	43.5	2.6	70.8	2.2	<0.001*	Pleased	30.6	2.3	35.9	2.9	n.s.

*SEM is mean of standard deviation. A mixed two-way ANOVA demonstrated significant main effects of Group and Emotion in Intensity and Valence: $p < 0.001$; *represents significant difference of post hoc tests; Intensity: $p < 0.05/8$, Valence: $p < 0.05/9$, Arousal: $p < 0.05/9$.*

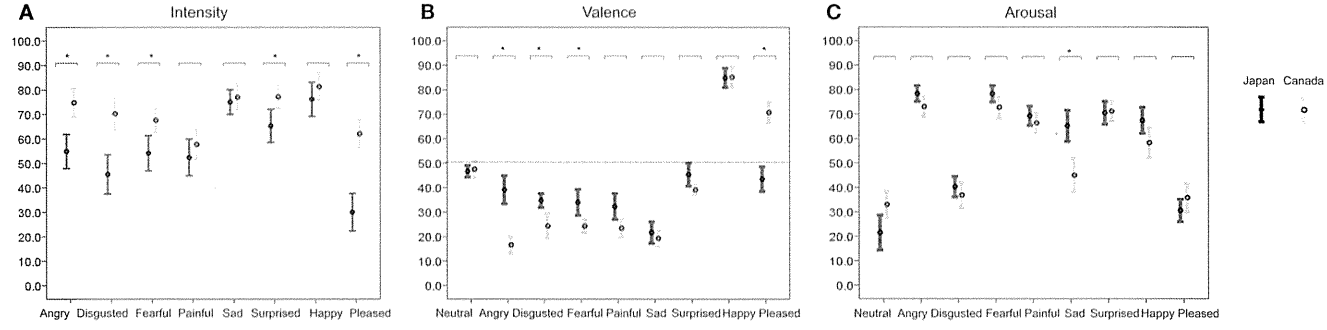

FIGURE 1 | Shows the distribution of ratings (error bar: mean ± SD) for each emotional sound judged by 30 Japanese and 30 Canadian participants for (A) Intensity, (B) Valence, and (C) Arousal. Each horizontal axis represents each rating score (0–100). Each vertical axis shows categories of emotional vocalizations. A mixed two-way ANOVA represents significant main effects of Subject group (Japanese and Canadian) and Emotion, respectively: $p < 0.001$; $^*p < 0.05/8$ (Intensity), $^*p < 0.05/9$ (Valence), $^*p < 0.05/9$ (Arousal); *post hoc* t-test.

(t-test, $p < 0.05/9$: t-test, $p < 0.05/9$: Anger, $t = 6.696$, Disgust, $t = 3.608$; Fear, $t = 3.232$: **Table 1**; **Figure 1B**), whereas the Valence rating from Japanese listeners was significantly lower than ratings from Caucasian listeners for Pleasure (t-test, $p < 0.05/9$; Pleasure, $t = -8.121$; **Table 1**, **Figure 1B**).

AROUSAL

A mixed two-way ANOVA with listeners' Group (Japanese, Canadian) and Emotion ($n = 9$) as factors was calculated on Arousal scores. There was no significant main effect of Group: $F(1, 57) = 2.099$, $p > 0.05$, whereas there was a significant main effect of Emotion $F(4.4, 250.5) = 158.524$, $p < 0.001$ (Greenhouse–Geisser's test). Crucially, a significant interaction between Group and Emotion was observed: $F(4.4, 250.5) = 8.955$, $p < 0.001$ (**Figure 1C**), indicating that rating differences between the two groups varied with the specific Emotion considered. *Post hoc* tests showed that the Arousal ratings from Japanese listeners were significantly higher than ratings from Caucasian listeners for sad vocalizations (t-test, $p < 0.05/9$: sad, $t = 4.334$: **Table 1**; **Figure 1C**), whereas the other Emotions were not significantly different between Japanese and Canadian participants (t-test, $p > 0.05/9$: **Table 1**; **Figure 1C**).

SENSITIVITY AND SPECIFICITY

We evaluated the Intensity ratings for their sensitivity (hit rate, by Emotion) and specificity (correct rejection rate, by rating scale). A maximum Intensity rating in the scale corresponding to the portrayed emotion was considered as a hit; otherwise, as a miss. **Table 2** shows the Intensity ratings of portrayed emotions for Japanese and Canadian participants: means of hit rates by participants and means of correct rejection rates by participants.

A Mixed two-way ANOVA with listener's Group and Emotion ($n = 8$) as factors were calculated on the score of sensitivity and specificity, respectively. In both sensitivity and specificity, a significant main effect of Group was observed [sensitivity: $F(1, 57) = 51.6$, $p < 0.001$; specificity: $F(1, 57) = 44.8$, $p < 0.001$] as well as main effects of Emotion [sensitivity: $F(5.4, 310) = 38.0$, $p < 0.001$; specificity: $F(5.6, 320) = 41.5$, $p < 0.001$, Greenhouse–Geisser's test]. Interaction effects (Group × Emotion) for sensitivity and specificity were also observed sensitivity: $F(5.4, 310) = 9.0$, $p < 0.001$; specificity: $F(5.6, 320) = 11.0$, $p < 0.001$, indicating that rating differences between the two Groups varied with the specific Emotion considered.

There were significant differences in hit rates between Japanese and Canadian participants for angry, disgusted, fearful, painful, and pleased actors' vocalizations ($p < 0.05/8$, t-test): hit rates for these emotions were all lower in Japanese participants. In correct rejection rate, there were significant differences between Japanese and Canadian participants for Disgust and Fear ratings scales, with lower correct rejection rates in Japanese listeners ($p < 0.05/8$).

In Japanese participants, hit rates for each Emotion varied greatly, from 25% for fearful to 79% for sad. Hit rates and correct rejection rate to happy, sad, and surprised vocalizations were relatively high (more than 50%), whereas hit rates and correct rejection rate to angry, disgusted, fearful, painful, and pleased vocalizations were lower (less than 50%).

In **Table 2**, the maximum Intensity rating for each portrayed emotion is shown in bold. For fearful vocalizations only, the Emotion with a maximum score by Japanese participants was different from the portrayed emotion. Japanese listeners on average gave higher Intensity rating in the Surprise scale (66%) than the Fear scale (54%) in response to fearful vocalizations. For all other Emotions, Japanese participants gave the maximum ratings in the scale corresponding to the portrayed emotion, as did the Canadian listeners.

GENDER DIFFERENCES OF ACTOR AND PARTICIPANT

We examined the effects of participant's and actor's gender on hit rates in Japanese participants (**Figure 2**). A three-way mixed ANOVA was calculated with the factors of actor's gender and participant's gender as well as Emotion in Japanese participants. In addition to a significant effect of the emotion [$F(1, 56) = 70.285$, $p < 0.001$], a significant effect of actor's gender [$F(1, 56) = 4.003$, $p \leq 0.05$] was observed, whereas no significant effect was revealed in participant's gender [$F(1, 56) = 3.727$, $p > 0.05$] or interaction effect: emotion × actor's

Table 2 | Intensity ratings (0–100) averaged across all actors for each portrayed emotion and Intensity ratings scale in Japanese and Canadian participants.

Intensity rating scale		Portrayed emotion																		Correct rejection rate (%)	
		Neutral		Angry		Disgusted		Fearful		Painful		Sad		Surprised		Happy		Pleased		Specificity	(Validity)
		M	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM	*M*	SEM
Anger	Japan	9	1.1	**55**[bd]	5.6	18	4.5	25	5.1	33	5.4	14	4.3	21	5.0	7	2.5	12	3.4	36	6.6
	Canada	9	0.5	**75**[ac]	2.4	14	1.1	19	2.1	33	3.8	9	0.7	17	0.8	3	0.3	7	0.8	77	4.4
Disgust	Japan	12	1.3	49	5.7	**45**[ac]	5.9	48	6.3	48	6.2	33	6.2	33	6.0	8	2.9	30	5.8	44*	4.6
	Canada	10	0.5	23	0.7	**70**[ac]	2.7	21	1.6	26	2.7	9	0.6	24	1.4	4	0.3	8	0.7	73*	4.5
Fear	Japan	7	0.8	30	3.7	15	4.0	54	5.9	25	5.5	21	5.1	34	5.9	5	1.7	14	3.2	18*	4.7
	Canada	9	0.5	16	2.0	11	0.6	**68**[bc]	2.5	21	2.0	10	0.7	45	2.6	3	0.2	6	1.0	69*	3.0
Pain	Japan	6	0.8	30	5.7	22	4.7	31	6.1	**52**[d]	5.7	30	6.0	23	5.1	5	1.6	13	3.9	32	8.0
	Canada	9	0.9	24	1.6	11	1.1	31	3.1	**58**[ac]	3.6	26	1.8	21	1.0	3	0.2	7	0.4	62	4.0
Sadness	Japan	10	1.1	15	4.1	23	4.8	21	4.8	27	5.0	**75**[ac]	4.7	13	3.8	7	2.3	26	5.4	75	5.2
	Canada	11	0.8	13	1.2	9	0.8	13	1.2	15	1.5	**77**[ac]	3.6	11	0.4	3	0.2	5	0.3	89	2.5
Surprise	Japan	7	0.8	46	6.2	20	4.5	**66**[a]	5.2	36	6.0	17	4.4	**65**[ac]	5.3	17	4.4	17	4.1	66	7.5
	Canada	9	0.6	26	1.8	26	1.8	57	3.0	35	3.0	11	2.7	**77**[ac]	2.0	18	1.1	25	2.2	64	2.7
Happiness	Japan	7	0.8	12	3.2	13	3.2	9	2.8	9	2.7	13	3.5	15	4.0	**76**[ac]	4.6	25	5.0	59	3.4
	Canada	14	0.5	6	0.4	9	0.8	7	0.4	10	1.1	11	2.4	15	1.3	**81**[c]	1.2	54	3.3	76	3.0
Pleasure	Japan	6	0.8	10	2.9	13	3.3	8	2.4	8	2.4	10	2.7	12	3.2	64	5.6	**32**	6.0	29	5.2
	Canada	13	0.3	6	0.4	9	0.9	6	0.4	11	1.9	10	2.4	12	1.0	76	1.1	**62**	3.8	39	4.0
Hit rate (%)	Japan			44*	7.3	51*	5.1	25*	3.4	35*	7.9	79	4.6	72	6.8	69	3.2	34*	5.7		
	Canada			78*	5.0	81*	3.7	56*	3.0	51*	3.0	86	2.0	75	2.9	60	4.5	59*	3.8		

Boldface indicates maximum average rating. Note the high hit rates for most affective categories.

[a]$p < 0.001$. [b]$p < 0.05$, strongest rating on the scale corresponding to the portrayed emotion (columns). [c]$p < 0.001$. [d]$p < 0.05$, strongest rating for the portrayed emotion corresponding to the rating scale (rows; Fisher's protected least significance test).

**$p < 0.05/8$, t-test.*

gender [$F(1, 56) < 1$, $p > 0.05$], emotion × participant's gender [$F(1, 56) = 2.496$, $p > 0.05$], and emotion × actor's gender × participant's gender [$F(1, 56) < 1$, $p > 0.05$]. Hit rates were higher for vocalizations portrayed by the female actors irrespective of participant's gender (**Figure 2**).

Further, we investigated cultural effect on hit rates including Japanese and Canadian participants. A three-way ANOVA was calculated with the factors of listener's group, actor's gender, and participant's gender. A significant main effect was observed in listener's Group: $F(1, 110) = 83.211$, $p < 0.001$, and actor's gender $F(1, 110) = 11.675$, $p < 0.001$, and participant's gender $F(1, 110) = 8.396$, $p = 0.005 < 0.05$. Interaction effect showed no significant effect of listener's group × participant's gender, $F(1, 110) = 0.054$, $p > 0.05$, listener's group × actor's gender, $F(1, 110) = 0.428$, $p > 0.05$, actor's gender × participant's gender $F(1, 110) = 0.804$, $p > 0.05$, and listener's group × actor's gender × participant's gender, $F(1, 110) = 0.071$, $p > 0.05$. These results indicate that in hit rates, the effect of actor's gender exists regardless of cultures.

Gender differences were analyzed on ratings of Intensity, Valence, Arousal, and correct rejection rates as well as hit rates. A significant effect of actor's gender was observed in Intensity: $F(1, 55) = 136.712$, $p < 0.001$; Valence: $F(1, 55) = 14.551$, $p < 0.001$; Arousal: $F(1, 55) = 182.899$, $p < 0.001$; correct rejection rates: $F(1, 55) = 23.131$, $p < 0.001$. There was no significant effect of participant's gender in Intensity: $F(1, 55) = 0.002$, $p > 0.05$; Valence: $F(1, 55) = 1.289$, $p > 0.05$; Arousal: $F(1, 55) = 0.655$, $p > 0.05$. In correct rejection rate, a significant effect of participant's gender was observed: $F(1, 55) = 6.343$, $p = 0.015$, <0.05. No interaction between actor's gender and participant's gender was observed [Intensity: $F(1, 55) = 1.459$, $p > 0.05$, Valence: $F(1, 55) = 0.316$, $p > 0.05$, Arousal: $F(1, 55) = 2.191$, $p > 0.05$, Correct rejection rate: $F(1, 55) = 0.797$, $p > 0.05$].

DISCUSSION

We investigated cross-cultural differences between Japanese and Canadian participants in their perception of non-verbal affective vocalization using MAVs. The most intriguing finding is that significant Group × Emotion interactions were observed for all emotional ratings (Intensity, Valence, and Arousal). Ratings of Intensity and Valence for happy and sad vocalizations were not significantly different between Japanese and Canadian participants, whereas ratings for angry and pleased vocalizations were significantly different. Especially, for the Valence ratings in angry vocalizations, Japanese subjects rated less negative than Canadian subjects. Further, in the Valence ratings for pleasure vocalizations, Japanese subjects rated less positive than Canadian subjects.

CROSS-CULTURAL EFFECT FOR POSITIVE EMOTION

Correct rejection rates (validity) of Happiness and Pleasure were not significantly different between Caucasian and Japanese subjects (**Table 2**: Happiness: Canadian 76% vs. Japanese 56%, Pleasure: Canadian 39% vs. Japanese 29%). These findings suggest that these two items are valid beyond the culture. In our study, there was a significant difference in the ratings (Intensity and Valence) for pleased vocalizations between Japanese and Canadian participants, whereas no significant difference was observed in the ratings for happy vocalizations. Although Happiness (laughter) was well recognized across cultures, there were apparent cultural differences in the perception of Pleasure.

A recent study between Western participants and Namibian participants demonstrated that the positive vocalizations of achievement, amusement, sensual pleasure, and relief were recognized as culture-specific signals although happy vocalizations

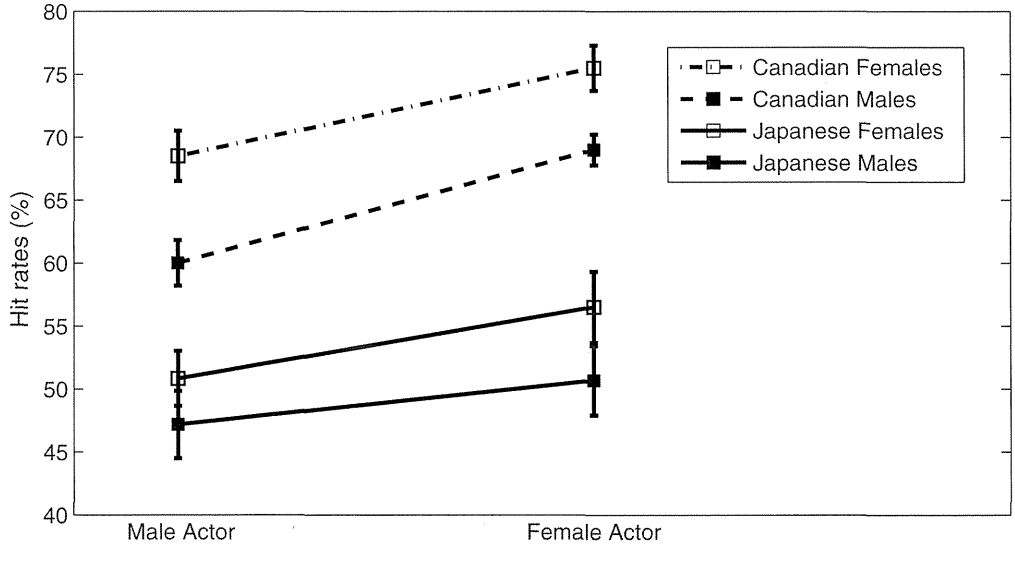

FIGURE 2 | Hit rates (percentage of test items with maximal rating on the scale corresponding to the portrayed emotion) split by actor's and participant's gender.