

# SCIENTIFIC REPORTS



OPEN

## Inhibition-excitation balance in the parietal cortex modulates volitional control for auditory and visual multistability

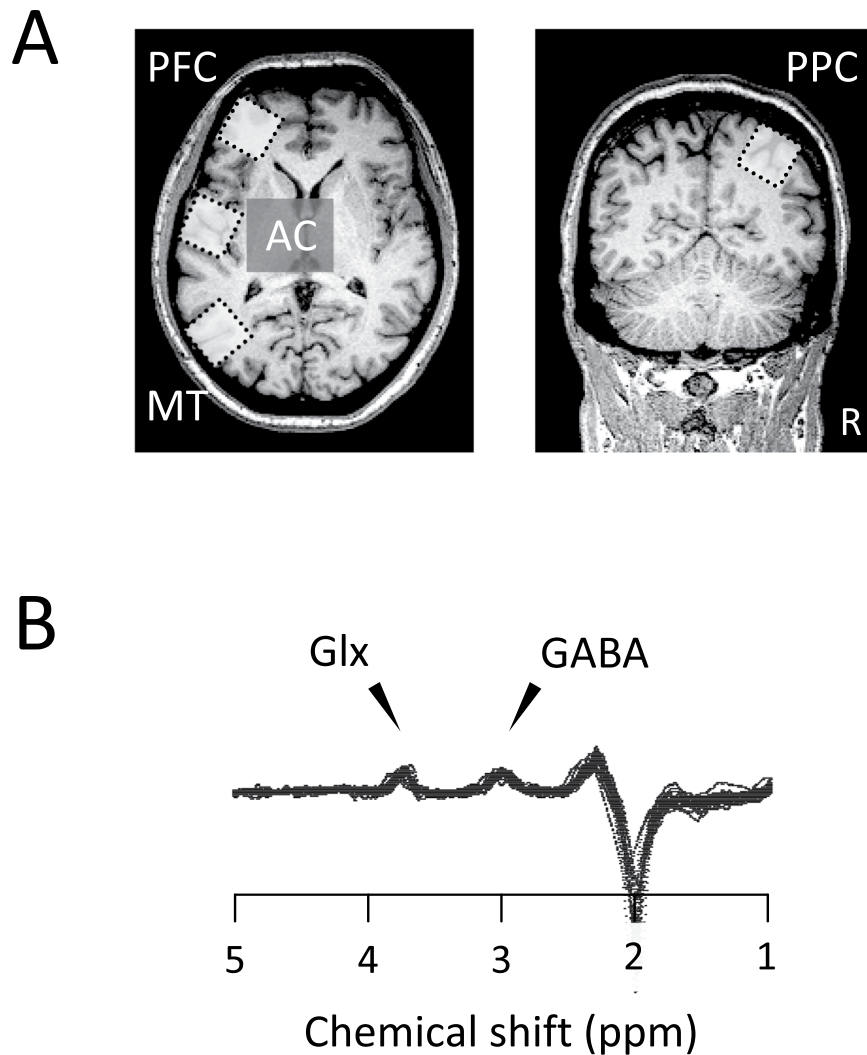
Hirohito M. Kondo<sup>1,2</sup>, Daniel Pressnitzer<sup>3,4</sup>, Yasuhiro Shimada<sup>5</sup>, Takanori Kochiyama<sup>5,6</sup> & Makio Kashino<sup>7,8</sup>

Perceptual organisation must select one interpretation from several alternatives to guide behaviour. Computational models suggest that this could be achieved through an interplay between inhibition and excitation across competing types of neural population coding for each interpretation. Here, to test for such models, we used magnetic resonance spectroscopy to measure non-invasively the concentrations of inhibitory  $\gamma$ -aminobutyric acid (GABA) and excitatory glutamate-glutamine (Glx) in several brain regions. Human participants first performed auditory and visual multistability tasks that produced spontaneous switching between percepts. Then, we observed that longer percept durations during behaviour were associated with higher GABA/Glx ratios in the sensory area coding for each modality. When participants were asked to voluntarily modulate their perception, a common factor across modalities emerged: the GABA/Glx ratio in the posterior parietal cortex tended to be positively correlated with the amount of effective volitional control. Our results provide direct evidence implicating that the balance between neural inhibition and excitation within sensory regions resolves perceptual competition. This powerful computational principle appears to be leveraged by both audition and vision, implemented independently across modalities, but modulated by an integrated control process.

Perceptual multistability describes an intriguing situation, whereby an observer reports random changes in conscious perception for a physically unchanging stimulus<sup>1,2</sup>. Multistability is a powerful tool with which to probe perceptual organisation, as it highlights perhaps the most fundamental issue faced by perception for any reasonably complex natural scene. And because the information encoded by sensory receptors is never sufficient to fully specify the state of the outside world<sup>3</sup>, at each instant perception must always choose between a number of competing alternatives. In realistic situations, the process produces a stable and useful representation of the world. In situations with intrinsically ambiguous information, the same process is revealed as multistable perception.

A number of theoretical models have converged to pinpoint the generic computational principles likely to be required to explain multistability, and hence perceptual organisation<sup>4-9</sup>. All of these models consider three core ingredients: inhibition between competing neural populations, adaptation within these populations, and neuronal noise. The precise role of each ingredient and their respective importance is still being debated. Noise is introduced to induce fluctuations in each population and initiate the stochastic perceptual switching in some models<sup>7-9</sup>, whereas switching dynamics are solely determined by inhibition in others<sup>5,6</sup>. Functional brain imaging in humans has provided results qualitatively compatible with those computational principles at several levels of the visual processing hierarchy<sup>10</sup>. But, for most functional imaging techniques in humans such as fMRI or MEG/EEG, changes

<sup>1</sup>School of Psychology, Chukyo University, Nagoya, Aichi, Japan. <sup>2</sup>Human Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, Atsugi, Kanagawa, Japan. <sup>3</sup>Laboratoire des Systèmes Perceptifs, CNRS UMR 8248, Paris, France. <sup>4</sup>Département d'Études Cognitive, École Normale Supérieure, Paris, France. <sup>5</sup>Brain Activity Imaging Center, ATR-Promotions, Seika-cho, Kyoto, Japan. <sup>6</sup>Department of Cognitive Neuroscience, Advanced Telecommunications Research Institute International, Seika-cho, Kyoto, Japan. <sup>7</sup>Sports Brain Science Project, NTT Communication Science Laboratories, NTT Corporation, Atsugi, Kanagawa, Japan. <sup>8</sup>School of Engineering, Tokyo Institute of Technology, Yokohama, Kanagawa, Japan. Correspondence and requests for materials should be addressed to H.M.K. (email: [kondo@lets.chukyo-u.ac.jp](mailto:kondo@lets.chukyo-u.ac.jp))



**Figure 1.** The size and location of voxels and the edited GABA spectra ( $N = 36$ ). (A) The voxel in the auditory cortex (AC) included the Heschl gyrus and the anterior part of the temporal plane. The voxel in the motion-sensitive area (MT) was placed at the ventrolateral occipital cortex. The voxel in the prefrontal cortex (PFC) was located at the anterior part of the middle frontal gyrus. The voxel in the posterior parietal cortex (PPC) was centred on the intraparietal sulcus. (B) The GABA and Glx peaks were calculated using the differences between AC spectra obtained by editing radio frequency on/off pulses. GABA,  $\gamma$ -aminobutyric acid; Glx, glutamate-glutamine; R, right.

in activation are difficult to trace back to neural mechanisms. For instance, an increase in activation in a given brain region could be the result of more excitation from upstream parts of the recruited network, or, equivalently, of less local inhibition. In addition, such an increase in activation could be associated with either more excitatory or more inhibitory projections to downstream parts of the network<sup>11</sup>. Therefore, direct experimental evidence of the implication of inhibition, excitation, or both, as theorized by the models, has been difficult to obtain.

We used magnetic resonance spectroscopy (MRS) to obtain non-invasive, *in vivo* measurements of  $\gamma$ -aminobutyric acid (GABA) and glutamate-glutamine (Glx) concentrations in the human brain<sup>12</sup> (Fig. 1). Using the same MRS technique, van Loon *et al.* (2013) investigated the link between GABA and Glx concentrations and visual multistability tasks for three different types of stimuli: structure from motion, binocular rivalry, and motion-induced blindness<sup>13</sup>. They showed that in all cases a higher GABA concentration was associated with longer percept durations. Intriguingly, a link with Glx was only found in the stimulus of structure from motion, so the inhibition-excitation balance model was not fully supported by their results. A more recent study using similar auditory streaming has demonstrated that the Glx concentration in the auditory cortex (AC) was related to the dominance of “segregated” percepts, where a complex scene is broken into separate streams<sup>14</sup>. In another study investigating the priming of ambiguous motion by non-ambiguous motion, Glx concentration in the prefrontal cortex (PFC) was correlated with the dominance of visual motion assimilation between context and test<sup>15</sup>. These two later sets of results point to the usefulness of studying the neurotransmitter concentration with MRS in relation to multistable perception tasks, but this is not tested directly in the inhibition-excitation balance model. Thus, even if there are now results strongly suggesting that neural inhibition or excitation is an integral part of

perceptual selection, the implication of a more specific model of the inhibition-excitation balance remains to be demonstrated.

Here, we examined the interplay between inhibition and excitation for multistable stimuli, comparing two sensory modalities using auditory streaming<sup>16,17</sup> and plaid motion<sup>18,19</sup> (Fig. 2A). The use of two modalities has several advantages. First, given the spatial resolution of MRS, it is possible to image brain regions involved in the processing of auditory or visual stimuli separately. Auditory streaming has neural correlates in AC<sup>17,20</sup>, whereas moving plaids induce activity in the motion-sensitive area (MT) of the visual cortex<sup>21,22</sup>. On the basis of previous findings and in relation to the theoretical model, we hypothesized that higher GABA/Glx ratios (more inhibition) in AC and MT might slow alternations and thus lead to longer percept durations in auditory streaming and moving plaids. Second, in spite of their superficial differences, auditory streaming and moving plaids both involve a basic competition between a one-object (Grouped) and a two-object (Split) interpretation. Perhaps as a consequence, deep similarities in the behavioural characteristics of multistability have been found for these stimuli<sup>16,23</sup>. Comparing the two modalities is thus a useful test of the generic principles of perceptual organisation, beyond specific stimulus characteristics. Finally, the comparison across modalities allows to address one of the longest-standing questions regarding multistability<sup>10,24,25</sup>, namely is the multistable selection process distributed throughout the brain, including in some low-level sensory areas, or is there a dominant role for higher-level regions through cognitive control? In particular, a causal role for the posterior parietal cortex (PPC) and PFC has long been argued as regards perceptual switching in auditory and visual multistability<sup>26–29</sup>, but this has recently been disputed<sup>30</sup>. Accordingly, we tested the further hypothesis that the GABA/Glx ratios in non-sensory brain regions, here PPC and PFC, could be involved in the behavioural characteristics of auditory and visual multistability. The PPC in particular has not been examined in previous studies using MRS.

The role of higher brain regions could be to initiate the switches, but it could also be to modulate perception when the participant tries to exert some control over it. To test for this last hypothesis, we manipulated the volition of participants during the task (also sometimes referred to as selective attention or willful intent) in addition to collecting spontaneous reports. It is known that participants are to some extent able to influence their percepts in multistable situations, either when asked to slow down or speed up the rate of reversals<sup>31–34</sup> or when asked to favour one of the two possible interpretations<sup>16,35</sup>. Interestingly, if a top-down volitional mechanism originates from PPC or PFC, it should reasonably be expected to be shared across perceptual modalities and thus correlate with behaviour irrespective of the modality of entry.

In brief, we acquired spectroscopy measurements in four brain regions. This provided one estimate of neurotransmitter balance per region, per participant. We then explored the correlations, over participants, between these measures of inhibition-excitation balance and inter-individual variations in perception - as measured in a separate behavioural session.

## Results

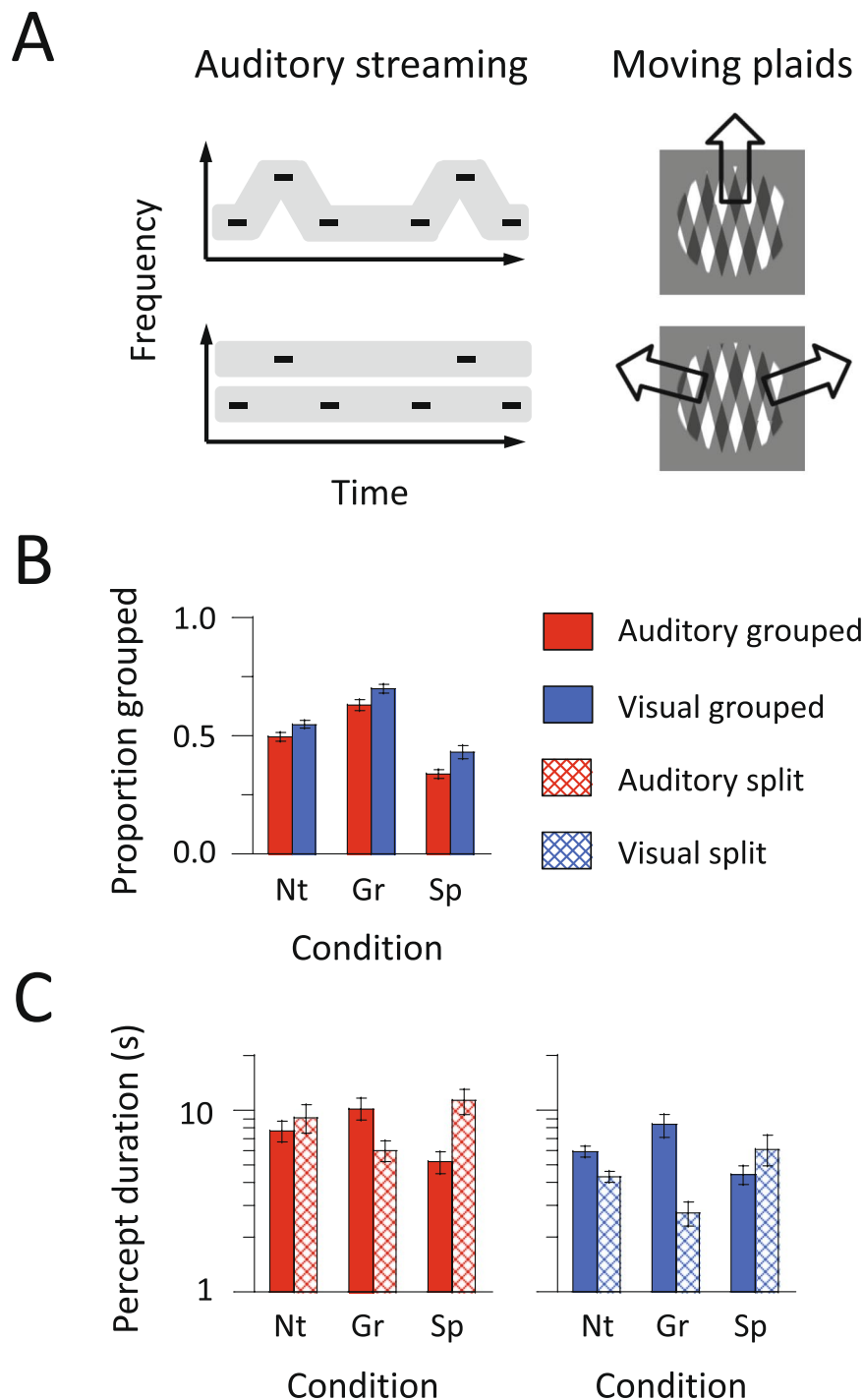
**Behavioural percept durations.** We first investigated the pattern of reported percepts under different volitional conditions, namely Neutral for spontaneous reports and Grouped or Split for the two different instruction cues (Fig. 2B). A repeated-measures analysis of variance (ANOVA) showed that the proportion (mean  $\pm$  standard error) of reported grouped percepts was larger overall for moving plaids ( $55.9 \pm 1.2\%$ ) than for auditory streaming ( $48.8 \pm 1.3\%$ ):  $F_{(1, 35)} = 22.98$ ,  $\eta_p^2 = 0.40$ ,  $p < 0.001$ . For both modalities, as expected, the proportion of grouped percepts increased under the Grouped condition ( $66.4 \pm 1.8\%$ ) compared with the Neutral condition ( $52.3 \pm 1.4\%$ ). Also, as expected, the proportion of grouped percepts decreased under the Split condition ( $38.4 \pm 2.1\%$ ):  $F_{(2, 70)} = 60.67$ ,  $\eta_p^2 = 0.63$ ,  $p < 0.001$ . The interaction between modality and condition was not significant:  $F_{(2, 70)} = 0.76$ ,  $\eta_p^2 = 0.02$ ,  $p > 0.47$ . This lack of interaction indicates that the amount of volitional control did not differ between auditory streaming and moving plaids.

We then characterized the finer effects of volitional intent on percept durations (Fig. 2C). For auditory streaming, a  $2 \times 3$  ANOVA revealed an interaction between percept type and volitional condition:  $F_{(2, 70)} = 33.03$ ,  $\eta_p^2 = 0.49$ ,  $p < 0.001$ . The duration of grouped percepts according to the volitional condition displayed the following pattern: Grouped ( $10.2 \pm 1.4$  s)  $>$  Neutral ( $7.7 \pm 1.0$  s)  $>$  Split ( $5.2 \pm 0.7$  s). For split percepts, the pattern was reversed: Split ( $11.3 \pm 1.8$  s)  $>$  Neutral ( $9.1 \pm 1.6$  s)  $>$  Grouped ( $6.0 \pm 0.8$  s). Similar results were observed for moving plaids, with an interaction between percept type and condition:  $F_{(2, 70)} = 46.44$ ,  $\eta_p^2 = 0.57$ ,  $p < 0.001$ . The duration of grouped percepts according to the volitional condition displayed the following pattern: Grouped ( $8.3 \pm 1.2$  s)  $>$  Neutral ( $5.9 \pm 0.4$  s)  $>$  Split ( $4.4 \pm 0.5$  s). For split percepts, the pattern was reversed: Split ( $6.1 \pm 1.2$  s)  $>$  Neutral ( $4.3 \pm 0.3$  s)  $>$  Grouped ( $2.7 \pm 0.4$  s).

In summary, we observed the same effect of volitional control for both modalities, with volition resulting in increased durations for the target percepts and shortened durations for the non-target percepts.

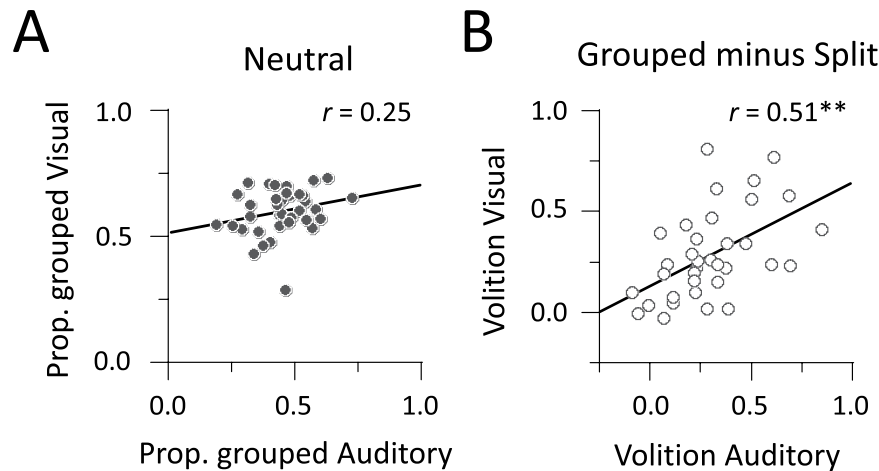
**Behavioural correlations across modalities.** Next, we looked for correlations across modalities under the different conditions of volitional control. Under the Neutral condition, the correlation between the proportion of grouped percepts across modalities did not reach statistical significance:  $r = 0.25$  (95% confidence interval,  $-0.08$  to  $0.53$ ),  $p > 0.14$  (Fig. 3A). This suggests a functional independence of modalities in the Neutral condition: participants who tended to perceive the grouped interpretation in audition were not the same as those who tended to perceive the grouped interpretation in vision.

We extended the analysis to situations where volitional control was applied. We computed the amount of effective volitional control by using the proportion of grouped percepts under the Grouped versus Split conditions. The amount of effective volitional control as a whole did not differ between modalities: auditory streaming ( $29.1 \pm 3.6\%$ ) versus moving plaids ( $26.9 \pm 3.5\%$ ),  $t = 0.59$ ,  $\eta^2 = 0.01$ ,  $p > 0.50$ . However, on a participant-by-participant basis, the amount of effective volitional control was positively correlated across



**Figure 2.** Stimuli and behavioural results for auditory and visual multistability. (A) Schematic representation of stimuli. A sound sequence of triplet tones was presented in auditory streaming. The auditory stimulus produces perceptual switches between one and two streams. Two rectangular gratings were moving in plaids. The visual stimulus induces perceptual switches between upward grouped and sideward split motion. (B,C) The average proportion of time spent in grouped percepts is shown as a function of condition (Neutral, Grouped, and Split) and input modality (auditory and visual presentation). Red and blue bars indicate the results for auditory streaming and moving plaids, respectively. The duration of alternating percepts is converted to a logarithmic scale for display purposes. Error bars represent the standard error of means.

modalities:  $r=0.51$  (95% CI, 0.22 to 0.72),  $p < 0.001$  (Fig. 3B). This means that in contrast to the Neutral case, individuals who were better at controlling their auditory percepts were also better at controlling their visual percepts.



**Figure 3.** Relationship between perceptual multistability and volitional control. **(A)** Correlation of the proportion of grouped percepts between auditory streaming and moving plaids. Symbols indicate individual data under the Neutral conditions. **(B)** Correlation of the amount of effective volitional control between auditory streaming and moving plaids. The amount of effective volitional control is computed by subtracting the proportion of grouped percepts between the Grouped and Split conditions.  $**p < 0.01$ .

In summary, we found no evidence for correlation across modalities in the Neutral condition. However, a correlation appeared under conditions of volitional control. This suggests an additional common factor across modalities that shapes perceptual alternations under conditions of volitional control only.

**GABA/Glx ratio compared with percept durations.** Figure 4 shows individual results comparing the percept duration with the GABA/Glx ratio. Auditory percept durations were positively correlated with the GABA/Glx ratios in the AC voxel:  $r = 0.59$  (95% CI, 0.32 to 0.77),  $p < 0.001$ . Importantly, the correlation in AC with auditory behaviour was larger than that in MT, PPC or PFC:  $t > 2.08$ ,  $p < 0.05$ . Mirroring this observation, visual percept durations were positively correlated with the GABA/Glx ratios in the MT voxel:  $r = 0.57$  (95% CI, 0.29 to 0.76),  $p < 0.001$ . The correlation in MT was larger than that in AC, PPC or PFC:  $t > 2.41$ ,  $p < 0.05$  (Fig. 5). We found a similar pattern of correlations in terms of GABA- and Glx-alone measures (Supplementary Table S1). In sensory areas, the correlations of the GABA/Glx ratios with percept durations were generally greater than those of the GABA and Glx concentrations, but the overall pattern of result was identical.

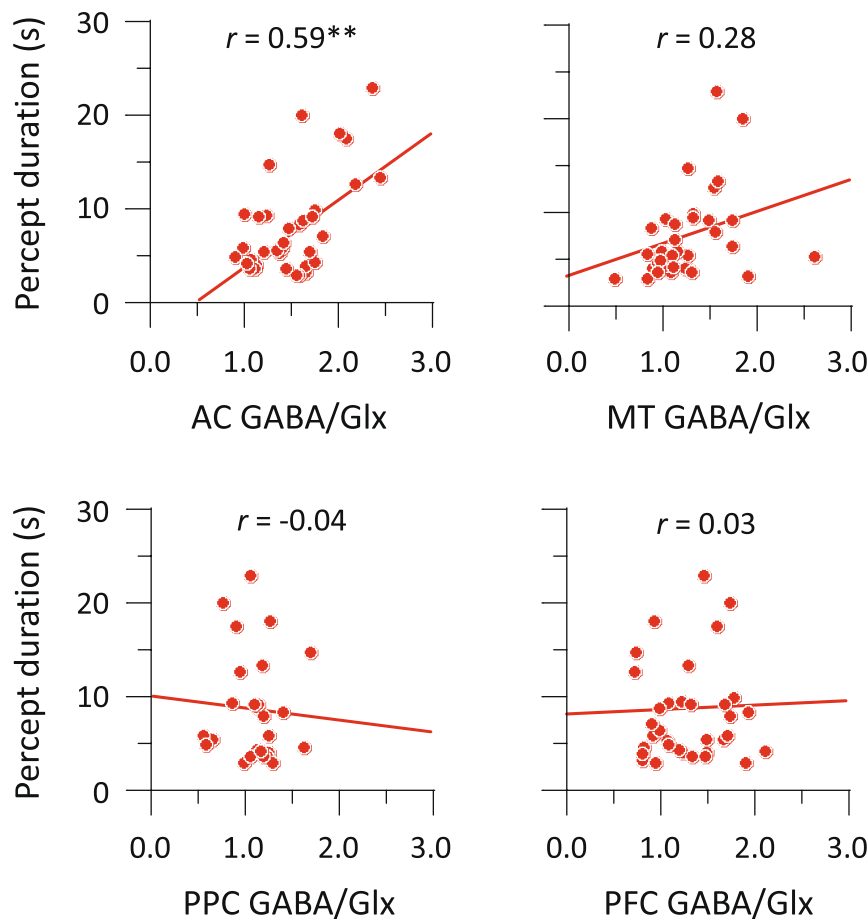
In summary, we observed strikingly similar patterns for the auditory and visual cases: the GABA/Glx ratio within the sensory area coding for the stimulus was reliably correlated with percept durations. In both modalities, participants who had higher GABA/Glx ratios displayed longer average percept durations. The GABA and Glx concentrations accounted for a sizeable proportion of the variance, which is remarkable given the many possible sources of experimental noise potentially involved in each measure. Finally, these correlations were, highly specific: they were only observed in the sensory area coding for the modality of interest, and not in other sensory areas nor prefrontal or parietal areas.

**GABA/Glx ratio compared with volitional control.** We examined whether the GABA/Glx ratio could account for the common factor between audition and vision under conditions of volitional control suggested by the behavioural analysis. This was accomplished by correlating the amount of behavioural volitional control (Fig. 3B) with the GABA/Glx ratio in the AC, MT, PFC, and PPC voxels.

For auditory streaming, the GABA/Glx ratio in PPC showed a significant correlation with the amount of effective volitional control:  $r = 0.39$  (95% CI, 0.00 to 0.67),  $p < 0.05$  (Fig. 6). There was no trend towards correlation between the amount of effective volitional control and the GABA/Glx ratio in any of the other voxels:  $r < 0.23$  in all cases. The correlation in PPC with volitional control was larger than that in PFC:  $t = 2.18$ ,  $p < 0.05$ . For moving plaids, there was a marginal correlation between the amount of effective volitional control and the GABA/Glx ratio in PPC:  $r = 0.35$  (95% CI,  $-0.04$  to 0.64),  $p = 0.082$ , in the same direction as that observed for audition and the PPC voxel (Fig. 7). Again, we found no significant correlation between the amount of effective volitional control in vision and GABA/Glx ratio in any of the other voxels:  $r < 0.27$ . The correlation in PPC with volitional control was larger than that in PFC:  $t = 2.03$ ,  $p < 0.05$ . Both sets of results were thus overall similar for audition and vision, especially considering that the difference between the auditory and visual correlation coefficients in PPC was not significant ( $t < 1$ ,  $p > 0.80$ ). When the same analyses of volitional control were repeated using GABA and Glx measures separately, correlations did not reach statistical significance (see Supplementary Table S1), suggesting that the ratio GABA/Glx is a more accurate measure of volitional control of perception.

Therefore, even if some of the analyses just fall short of significance, the results suggest that the best candidate to explain the correlation between auditory and visual volitional control is the inhibition-excitation balance in PPC, but not in PFC.

## Auditory streaming



**Figure 4.** Correlations between MRS and behavioural results in auditory streaming. Scatter plots with linear regression fit for the relationship between GABA/Glx ratios and median percept durations. Circles indicate individual data under the Neutral conditions. **\*\*** $p < 0.01$ .

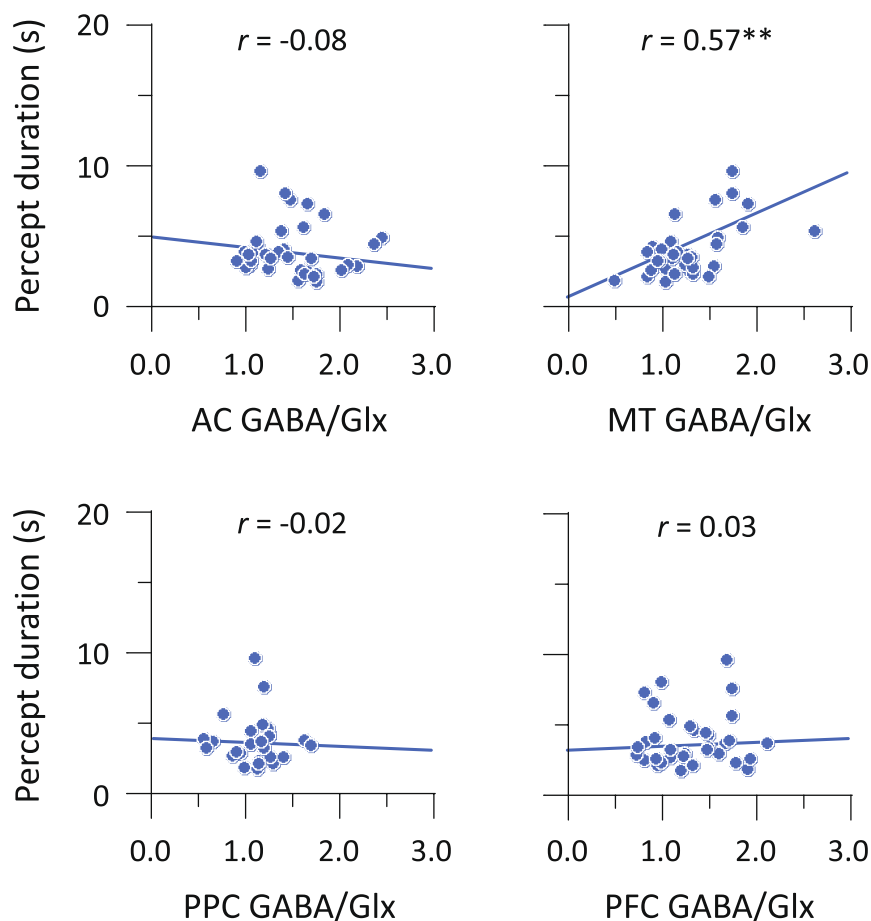
### Discussion

Using a comparison between perceptual multistability in audition and vision, we showed that higher GABA/Glx ratios within sensory regions correlated with longer percept durations. This is precisely predicted by computational models of perceptual organisation based on a balance of excitation and inhibition<sup>13,25</sup>. Importantly, the correlation was only found within sensory areas specialised for the stimulus modality: AC for audition and MT for vision. Under conditions of spontaneous reports, such a correlation did not extend to higher brain areas such as PPC and PFC. However, when volitional control was applied, a behavioural correlation across modality appeared and the GABA/Glx ratio in PPC tended to be linked to the effectiveness of volitional control.

**Direct evidence in humans of distributed competition during perceptual organisation.** The first set of novel findings in the present study relates to the modality-specific nature of the correlation between GABA and Glx concentrations and behaviour, in the neutral-instructions case. On a methodological level, this specificity precludes an interpretation of the MRS results as a generic effect of GABA/Glx ratios in the brain, because for instance percept durations in the auditory task were not correlated with GABA/Glx ratios in MT. Rather, the modality-specific but qualitatively similar correlations can be interpreted as a common computational principle, namely the balance between inhibition and excitation, implemented independently across each sensory modality<sup>36</sup>.

The nature of the correlation (more inhibition leading to longer percepts) is consistent with previous MRS results obtained with some visual stimuli<sup>13</sup>, extending them to yet another visual motion stimulus. To the best of our knowledge, this is the first direct experimental evidence for such processes in the auditory modality. This novel finding has theoretical importance because the ‘features’ in competition for visual stimuli, such as the dominant percept in binocular rivalry, are better understood than those at play in auditory streaming, which presumably involve a strong timing element. The traditional cortical accounts of auditory streaming do not in fact require an inhibition-excitation mechanism<sup>37,38</sup>. Our data show that framing the streaming problem in terms of

## Moving plaids



**Figure 5.** Correlations between MRS and behavioural results in moving plaids. Scatter plots with linear regression fit for the relationship between GABA/Glx ratios and median percept durations. Circles indicate individual data under the Neutral conditions.  $**p < 0.01$ .

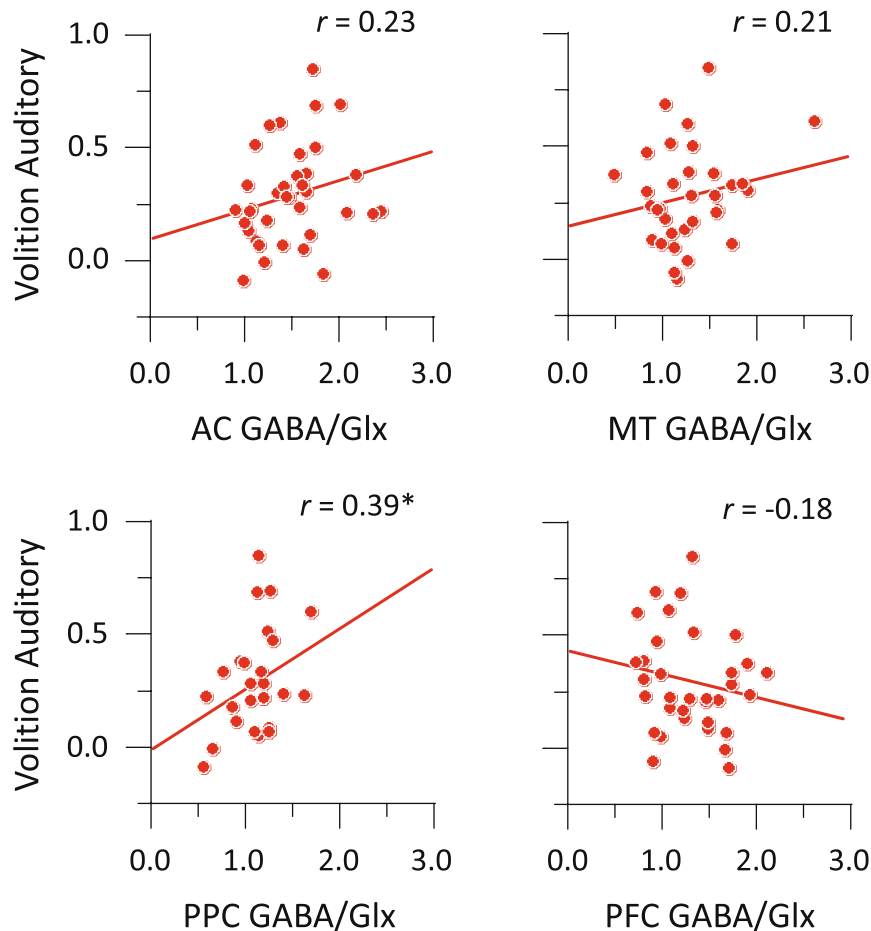
competing neural populations<sup>39,40</sup> or the accumulation of evidence with an inhibitory process to effect classification<sup>41</sup> may be promising avenues for guiding future neurophysiological investigations<sup>25</sup>.

**Idiosyncratic variability in multistable perception.** Our results add another cause of the puzzling but well-documented idiosyncratic variability in the behavioural characteristics of perceptual multistability<sup>42</sup>. Genotype variations impacting the dopamine system have been found to correlate with individual differences in auditory streaming and verbal transformations<sup>23</sup>. Structural correlates have been uncovered for visual bistable dynamics, for cortical volume in frontal and visual areas<sup>43</sup>, for cortical volume in parietal regions<sup>29</sup>, as well as for inter-region connectivity<sup>44</sup>. Here, we demonstrate that the relative balance of the GABA and Glx concentrations is also an idiosyncratic factor affecting percept durations for both visual and auditory multistability. This has potential clinical implications, as developmental disorders that occur during the maturation of the GABAergic system can lead to the long-term disruption of cortical circuits<sup>45</sup>, with decreases in GABA production<sup>46</sup>. In addition, alterations in glutamate signalling in the cortex can induce sensory memory impairments<sup>47</sup> and the recurrence of auditory hallucinations<sup>48</sup>. Thus, our findings have implications pertaining to the impact of disorders of the GABAergic and glutamatergic systems on mid-level functions such as perceptual organisation.

**Possible mechanisms of volitional control.** Another set of findings relates to the effect of volition. The behavioural data showed that volition could both lengthen a target percept and shorten an unwanted percept. If we equate volitional intent with a strengthening of the target percept, this seems at odd with Levelt's second proposal for binocular rivalry<sup>49</sup> and with our own previous reports<sup>16</sup>. However, there are other such examples of the lengthening of the strengthened percept in vision that depend on details of the experimental situation<sup>50</sup>. It would seem that the larger cohort tested here allowed the measurement of such an effect for auditory streaming.

We demonstrated that a higher GABA/Glx ratio in PPC tended to be associated with a larger magnitude of effective volitional control for perceptual organisation. The correlation just failed to reach our statistical significance criterion, perhaps because the sample size for the PPC voxel analysis was smaller than for other voxels, or

## Auditory streaming



**Figure 6.** Correlations between MRS measure and volitional control for auditory streaming. Amounts of effective volitional control as a function of GABA/Glx ratio. \* $p < 0.05$ .

because the imaging and behavioural data were collected further apart in time than for other voxels. In any case, the correlation with the GABA/Glx ratio in PPC was stronger than in the PFC region tested here.

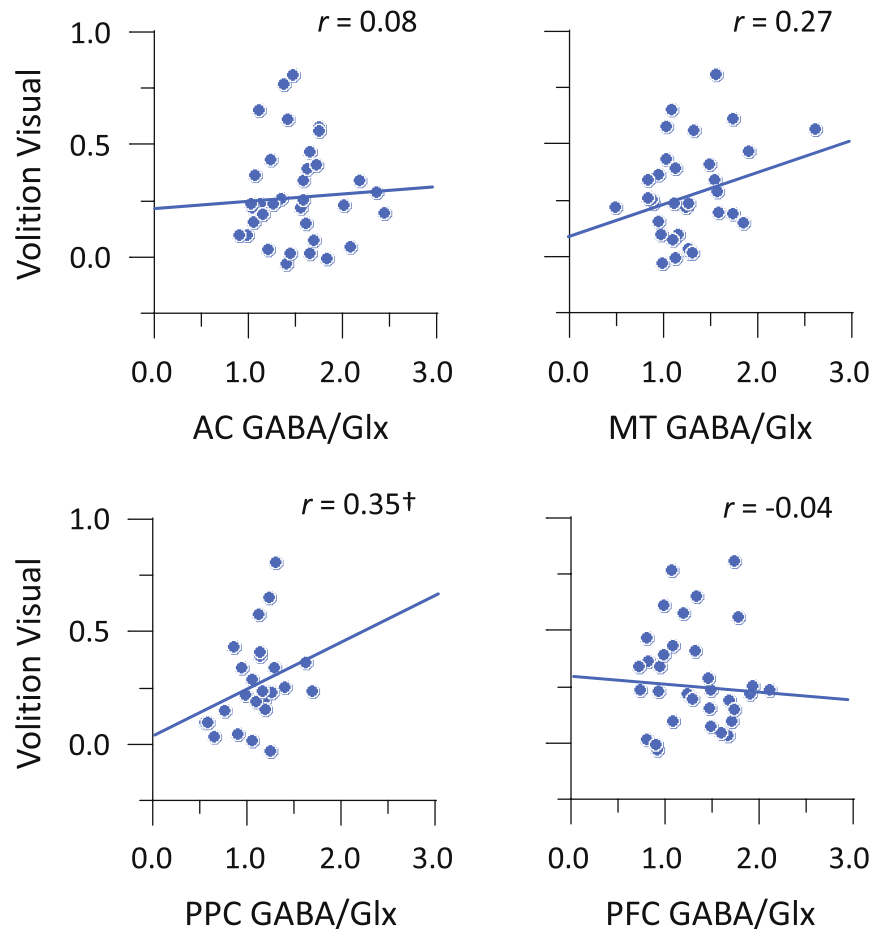
A recent study using an attentional blink paradigm, where two stimuli must be attended in rapid succession, showed that GABA concentrations in PPC were linked to an inhibition of attentional orienting from the first to the second target<sup>51</sup>. One interpretation of our own findings could thus be that inhibitory mechanisms in the parietal cortex play a role in the control of selective attention, which in turn is a part of volitional control. This interpretation would be consistent with previous findings using fMRI. For auditory streaming, the activity in the right intraparietal sulcus (mostly coincident with our PPC voxel) was correlated with attention and awareness of the competing streams, the two being confounded in this study<sup>28</sup>. In a visual paradigm, the activity of the intraparietal sulcus has been found to be modulated by instruction-guided attention, whereas the activity of the visual areas was positively correlated with the visibility of visual targets, regardless of attentional levels<sup>52</sup>. In EEG studies, activity in the right inferior parietal cortex was elicited before perceptual switches of the Necker cube<sup>53</sup> and binocular rivalry<sup>54</sup>, suggesting a causal link. Hence, it is plausible that the parietal cortex contains subregions for manipulating perceptual content.

Interestingly, unlike previous suggestions<sup>33,34</sup>, we could not relate volitional control to the PFC function. Note that we manipulated volitional control with instructions to select one perceptual interpretation<sup>35</sup>, and not with instructions to speed up or slow down the alternations<sup>32–34</sup>. This was motivated by the hypothesis that instructions to speed up may be mediated by general attention or arousal effects, whereas percept selection is more specifically related to bistable competition<sup>35</sup>. When this distinction is made, it does not seem that PFC is implicated in the initiation of perceptual switches themselves<sup>55</sup>.

**Predictive coding.** Current formulations of hierarchical perception usually appeal to some form of hierarchical predictive processing or predictive coding<sup>56–59</sup>. This perspective ties together several themes we have referred to above. In brief, in predictive coding, perception is cast as Bayesian belief updating through the use of prediction errors. In other words, expectations about the causes of sensory input are used to generate top-down predictions.



## Moving plaids



**Figure 7.** Correlations between MRS measure and volitional control for moving plaids. Amounts of effective volitional control as a function of GABA/Glx ratio. † $p < 0.10$ .

These predictions are then compared with the actual input (or expectations at lower levels in cortical hierarchies) to form a prediction error. This prediction error is then returned to revise expectations - produce better predictions and eliminate prediction error. Crucially, the influence of prediction errors is controlled by optimising their precision (i.e., reliability or inverse variance). Technically, predictive coding can be regarded as a form of Kalman filtering and the precision corresponds to the Kalman gain applied to prediction errors. This later formulation is relevant to the present findings from several perspectives; first, it links cortical gain and the inhibition-excitation balance to the precision of ascending prediction errors in cortical hierarchies. Physiologically, this rests upon the post-synaptic gain and the sort of synaptic balance found in our spectroscopy measurements. Psychologically, the modulation of precision is generally thought of in terms of attentional gain or selection<sup>60</sup>, which provides a nice link between the inhibition-excitation balance and the attentional mediation of volitional control for perceptual organisation. Furthermore, it is simple to show that the precision corresponds to the rate of evidence accumulation in predictive coding, which fits neatly with our results pertaining to the duration of perceptual switches<sup>61</sup>. Multistable perception in the context of predictive coding has already been discussed<sup>62</sup>. When the precision of lower level prediction errors is attenuated in a modality-specific sensory cortex, the rate of evidence accumulation is suppressed and the duration of any given multistable percept increases. In contrast, during volitional control, the precision of higher level prediction errors is attenuated (in the parietal cortex), thereby rendering parietal expectations more sensitive to the sensory features of interest<sup>63</sup>. This interpretation is potentially important because many neuropsychiatric symptoms can now be understood in terms of aberrant precision control, particularly conditions such as autism and schizophrenia<sup>64</sup>. It may be no coincidence that a loss of inhibition-excitation balance is now considered one of the primary pathophysiologicals of schizophrenia<sup>65,66</sup>.

**Integrated-distributed models of perceptual organisation.** The account emerging from the current findings is of inhibitory-excitatory processes<sup>13</sup> mediating neural competition within sensory regions, additionally modulated by higher-order mechanisms such as volitional control. This is highly consistent with long-standing and influential hierarchical models of perceptual organisation, where basic computational principles are

implemented in a distributed manner<sup>4</sup>, with an additional modulation by higher-level processes such as attention or volition<sup>67</sup>. Using perceptual multistability and MRS imaging, we thus observed further experimental evidence for such models directly in human participants<sup>68</sup>.

## Materials and Methods

**Ethics statement.** The study was carried out in accordance with the Declaration of Helsinki. All procedures reported in this study were approved by the Ethics and Safety Committees of NTT Communication Science Laboratories and ATR-Promotions (protocol nos: H24-004 and AN14-001). All participants gave written informed consent after the procedures had been fully explained to them.

**Participants.** Thirty-six participants were recruited for the experiment (22 males and 14 females;  $M_{\text{age}} = 36.9$ ,  $SD_{\text{age}} = 10.9$ ). All participants were right-handed Japanese people with normal hearing and with normal or corrected-to-normal vision. None had any history of neurological or psychiatric disorders. All participants experienced an MRS session with AC, MT and PFC voxels. Due to the limited time available for an experimental session, another MRS session for the PPC voxel was run on a subsequent day, on a subset of 26 participants (20 males and 6 females,  $M_{\text{age}} = 34.1$ ,  $SD_{\text{age}} = 8.3$ ). Behavioural data were obtained immediately after the first MRS session.

**MRS data acquisition.** To minimize confounding factors affecting the GABA and Glx concentrations, we conducted the acquisition of MR spectra at a fixed time for all participants, namely between 1:00 p.m. and 4:00 p.m. Data were acquired with a 3 T MRI scanner with a 12-channel receive-only head coil (MAGNETOM Trio, Siemens). Head motion was minimized by providing comfortable padding around the participant's head. For an assessment of cortical thickness and volume, anatomical images were obtained with a T1-weighted pulse sequence (isotropic voxel size of 1 mm<sup>3</sup>).

MR spectra were acquired from four  $3 \times 3 \times 3$  cm<sup>3</sup> voxels of interest, positioned in AC, MT, PPC, and PFC (Fig. 1). Voxels were positioned by using internal landmarks to achieve a consistent position between participants. The AC voxel was aligned with the first transverse sulcus. It contained the Heschl gyrus (Brodmann area: BA 41) and included the anterior part of the temporal plane (BA 42). The MT voxel (BA 19) was centred at the junction of the ascending limb of the inferior temporal sulcus and the lateral occipital sulcus. The PPC voxel was centred on the intraparietal sulcus (BA 7) of the right hemisphere. The voxel included a part of the superior and inferior parietal lobules. The PFC voxel was located at the anterior part of the middle frontal gyrus (BA 46). All the voxels except the PPC voxel were angled parallel to the surface of the left hemisphere.

Four runs (for each of the four regions of interest) were acquired for each participant from the different voxels. Before each run, we carefully carried out manual shimming (approximately 5 min) of the magnetic field in the voxel. We used the Mescher-Garwood proton resolved spectroscopy (MEGA-PRESS) technique<sup>69</sup>. For each spectrum, 64 spectral averages of 1024 data points were acquired with a repetition time of 1500 ms and an echo time of 68 ms, resulting in a scan duration of 3 min 18 sec (Supplementary Figure S1). An editing pulse with a bandwidth of 44 Hz was applied at 1.9 ppm (on) and 7.5 ppm (off) in interleaved scans. An unsuppressed water signal was also acquired from the same voxel. The difference in the edited spectra yielded the GABA and Glx peaks. As with all metabolic imaging studies using MRS, the GABA and Glx measures reflect both the intracellular fluid and the synaptic pool.

**Behavioural tasks.** After the MRS data acquisition, the participants performed the multistability tasks outside the scanner (Fig. 2A). The behavioural experiment lasted approximately 1 hr. Stimulus presentation and response collection were managed using MATLAB with the Psychophysics Toolbox<sup>70</sup>. Two tasks were conducted separately: auditory streaming and moving plaids<sup>23</sup>.

The auditory stimuli were made from repetitions of an ABA- tone pattern, where A and B represent two different pure tones and the hyphen represents a silent interval. The A and B tones were centred on 1 kHz with a four-semitone frequency difference between them (frequency for A = 891 Hz; frequency for B = 1122 Hz). The duration of each tone was 40 ms, including 10-ms rising and falling cosine ramps. The stimulus onset asynchrony between successive tones was 100 ms. The presentation level was set at 70 dB SPL. Stimuli were delivered through Sennheiser HDA 200 headphones.

The visual stimuli consisted of two rectangular-wave gratings (velocity = 1.25 deg/sec; spatial frequency = 0.5 cycle/deg; duty cycle = 0.5). The gratings were moving in directions 120 deg apart. The stimuli were presented through a circular aperture on a grey background, at a viewing distance of 57 cm and covering a visual angle of 5 deg. A small point for fixation was added to the centre of the visual stimuli.

We first explained the two tasks using a visual illustration of the stimuli. For auditory streaming, participants were asked to report whether they heard one stream (ABA-ABA-...) with a galloping rhythm, or two streams (A-A-... and -B-B-...) with an isochronous rhythm for each stream. For plaids, participants were asked to report whether they saw a single plaid moving upward, or two superimposed gratings moving sideways in opposite directions. In the first part of the experiment, participants were simply instructed to pay attention to the stimulus (Neutral instructions). They reported their perception continuously during a 5-min presentation of each stimulus. Their responses were collected via two buttons on a computer keyboard. A response indicated by a button press was held until a subsequent button press.

Following the Neutral condition, two further conditions were run with volitional control manipulations<sup>16</sup>. In the Grouped condition, the participants were instructed to try as far as possible to maintain perception of one stream or one plaid. In the Split condition, the participants were instructed to try to maintain perception of two streams or two gratings. The order of the two volitional conditions was randomized across the participants.

**Data analyses.** The MRS data were analysed using TARQUIN (version 4.2.10)<sup>71</sup>. The data were Fourier-transformed to a spectrum of 2048 data points, the signal was smoothed by a 3 Hz Lorentzian filter,

phased and referenced to a water signal at 4.7 ppm. A basis set in the software was fitted to the average spectrum allowing peak amplitudes, widths, and frequencies to be optimised (Voigt function). The final results were expressed as GABA and Glx signals (peaks at 3.00 and 3.76 ppm, respectively) relative to the unsuppressed water signals. The GABA and Glx concentrations were quantified in arbitrary units (a.u). The GABA and Glx concentrations (mean  $\pm$  standard error) were  $1.78 \pm 0.07$  a.u. and  $1.22 \pm 0.04$  a.u. for AC;  $1.47 \pm 0.05$  a.u. and  $1.31 \pm 0.05$  a.u. for MT;  $1.44 \pm 0.06$  a.u. and  $1.36 \pm 0.08$  a.u. for PPC;  $1.53 \pm 0.09$  a.u. and  $1.13 \pm 0.05$  a.u. for PFC. GABA/Glx ratios were computed to assess the inhibition-excitation balance in the voxels.

For the behavioural data analysis, time-series of percept durations were analysed separately for each auditory streaming and plaid motion block. All durations were longer than 300 ms. It has been previously reported that the duration of the first percept is longer than the duration of subsequent percepts for auditory streaming and moving plaids<sup>16,72</sup>. Thus, we excluded the duration of the first percept from the analysis. From the remaining durations, we evaluated the proportion of grouped percepts relative to the total duration. The raw distributions of the percept durations in all cases are provided as Supplementary Figure 2. Volitional control, as expected, changed the percept duration distribution. The amount of effective volitional control was estimated for each participant by subtracting the proportion of grouped percepts between blocks with grouped and split instructions<sup>35</sup>. Next, we computed the median of percept durations for each condition and participant. The pooled durations did not follow a normal distribution (Kolmogorov-Smirnov tests,  $p < 0.01$ ), so we transformed the data to a logarithmic scale to enable us to perform a repeated-measures ANOVA. The Sidák correction was used for post hoc comparisons ( $\alpha$ -level = 0.05). Pearson correlation coefficients and 95% CIs were computed between the behavioural and MRS measures. Statistical analyses were carried out with IBM SPSS Statistics (version 22) and R (version 3.1.2).

## Data Availability

Data will be made available on request.

## References

1. Leopold, D. A. & Logothetis, N. K. Multistable phenomena: changing views in perception. *Trends Cogn. Sci.* **3**, 254–264 (1999).
2. Schwartz, J. L., Grimault, N., Hupé, J. M., Moore, B. C. J. & Pressnitzer, D. Multistability in perception: binding sensory modalities, an overview. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367**, 896–905, <https://doi.org/10.1098/rstb.2011.0254> (2012).
3. Helmholtz, H. V. *Helmholtz's Treatise on Physiological Optics. Volume III. (Translated from the third German edition by J. P. C. Southall)*. (The Optical Society of America (Original work published in 1867), 1924).
4. Wilson, H. R. Computational evidence for a rivalry hierarchy in vision. *Proc. Natl. Acad. Sci. USA* **100**, 14499–14503, <https://doi.org/10.1073/pnas.2333622100> (2003).
5. Lankheet, M. J. M. Unraveling adaptation and mutual inhibition in perceptual rivalry. *J. Vis.* **6**, 304–310, <https://doi.org/10.1167/6.4.1> (2006).
6. Noest, A. J., van Ee, R., Nijs, M. M. & van Wezel, R. J. A. Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J. Vis.* **7**, 1–14, <https://doi.org/10.1167/7.8.10> (2007).
7. Klink, P. C., van Ee, R. & van Wezel, R. J. A. General validity of Levelt's propositions reveals common computational mechanisms for moral rivalry. *PLoS ONE* **3**, e3473, <https://doi.org/10.1371/journal.pone.0003473> (2008).
8. Moreno-Bote, R., Rinzel, J. & Rubin, N. Noise-induced alternations in an attractor network model of perceptual bistability. *J. Neurophysiol.* **98**, 1125–1139, <https://doi.org/10.1152/jn.00116.2007> (2007).
9. Hugué, G., Rinzel, J. & Hupé, J. M. Noise and adaptation in multistable perception: noise drives when to switch, adaptation determines percept choice. *J. Vis.* **14**, 19, <https://doi.org/10.1167/14.3.19> (2014).
10. Tong, F., Meng, M. & Blake, R. Neural bases of binocular rivalry. *Trends Cogn. Sci.* **10**, 502–511, <https://doi.org/10.1016/j.tics.2006.09.003> (2006).
11. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878, <https://doi.org/10.1038/nature06976> (2008).
12. Puts, N. A. J. & Edden, R. A. E. *In vivo magnetic resonance spectroscopy of GABA: a methodological review. Prog. Nucl. Magn. Reson. Spectrosc.* **60**, 29–41, <https://doi.org/10.1016/j.pnmrs.2011.06.001> (2012).
13. van Loon, A. M. *et al.* GABA shapes the dynamics of bistable perception. *Curr. Biol.* **23**, 823–827, <https://doi.org/10.1016/j.cub.2013.03.067> (2013).
14. Kondo, H. M., Farkas, D., Denham, S. L., Asai, T. & Winkler, I. Auditory multistability and neurotransmitter concentrations in the human brain. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, 20160110, <https://doi.org/10.1098/rstb.2016.0110> (2017).
15. Takeuchi, T., Yoshimoto, S., Shimada, Y., Kochiyama, T. & Kondo, H. M. Individual differences in visual motion perception and neurotransmitter concentrations in the human brain. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, 20160111, <https://doi.org/10.1098/rstb.2016.0111> (2017).
16. Pressnitzer, D. & Hupé, J. M. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr. Biol.* **16**, 1351–1357, <https://doi.org/10.1016/j.cub.2006.05.054> (2006).
17. Kondo, H. M. & Kashino, M. Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. *J. Neurosci.* **29**, 12695–12701, <https://doi.org/10.1523/JNEUROSCI.1549-09.2009> (2009).
18. Hupé, J. M. & Rubin, N. The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. *Vision Res.* **43**, 531–548 (2003).
19. Adelson, E. H. & Movshon, J. A. Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525 (1982).
20. Micheyl, C. *et al.* The role of auditory cortex in the formation of auditory streams. *Hear. Res.* **229**, 116–131, <https://doi.org/10.1016/j.heares.2007.01.007> (2007).
21. Huk, A. C. & Heeger, D. J. Pattern-motion responses in human visual cortex. *Nat. Neurosci.* **5**, 72–75, <https://doi.org/10.1038/nn774> (2001).
22. Castelo-Branco, M. *et al.* Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc. Natl. Acad. Sci. USA* **99**, 13914–13919, <https://doi.org/10.1073/pnas.202049999> (2002).
23. Kondo, H. M. *et al.* Separability and commonality of auditory and visual bistable perception. *Cereb. Cortex* **22**, 1915–1922, <https://doi.org/10.1093/cercor/bhr266> (2012).
24. Sterzer, P., Kleinschmidt, A. & Rees, G. The neural bases of multistable perception. *Trends Cogn. Sci.* **13**, 310–317, <https://doi.org/10.1016/j.tics.2009.04.006> (2009).
25. Szabo, B. T., Denham, S. L. & Winkler, I. Computational models of auditory scene analysis: a review. *Front. Neurosci.* **10**, 524, <https://doi.org/10.3389/fnins.2016.00524> (2016).
26. Lumer, E. D., Friston, K. J. & Rees, G. Neural correlates of perceptual rivalry in the human brain. *Science* **280**, 1930–1934 (1998).
27. Sterzer, P. & Kleinschmidt, A. A neural basis for inference in perceptual ambiguity. *Proc. Natl. Acad. Sci. USA* **104**, 323–328, <https://doi.org/10.1073/pnas.0609006104> (2007).

28. Cusack, R. The intraparietal sulcus and perceptual organization. *J. Cogn. Neurosci.* **17**, 641–651, <https://doi.org/10.1162/0898929053467541> (2005).
29. Kanai, R., Bahrami, B. & Rees, G. Human parietal cortex structure predicts individual differences in perceptual rivalry. *Curr. Biol.* **20**, 1626–1630, <https://doi.org/10.1016/j.cub.2010.07.027> (2010).
30. Brascamp, J., Blake, R. & Knapen, T. Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nat. Neurosci.* **18**, 1672–1678, <https://doi.org/10.1038/nn.4130> (2015).
31. Meredith, G. M. & Meredith, C. G. W. Effect of instructional conditions on rate of binocular rivalry. *Percept. Mot. Skills* **15**, 655–664 (1962).
32. van Ee, R., van Dam, L. C. I. & Brouwer, G. J. Voluntary control and the dynamics of perceptual bi-stability. *Vision Res.* **45**, 41–55, <https://doi.org/10.1016/j.visres.2004.07.030> (2005).
33. Windmann, S., Wehrmann, M., Calabrese, P. & Güntürkün, O. Role of the prefrontal cortex in attentional control over bistable vision. *J. Cogn. Neurosci.* **18**, 456–471, <https://doi.org/10.1162/089892906775990570> (2006).
34. de Graaf, T. A., de Jong, M. C., Goebel, R., van Ee, R. & Sack, A. T. On the functional relevance of frontal cortex for passive and voluntarily controlled bistable vision. *Cereb. Cortex* **21**, 2322–2331, <https://doi.org/10.1093/cercor/bhr015> (2011).
35. Meng, M. & Tong, M. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* **4**, 539–551, <https://doi.org/10.1167/4.7.2> (2004).
36. Hupé, J. M., Joffo, L. M. & Pressnitzer, D. Bistability for audiovisual stimuli: perceptual decision is modality specific. *J. Vis.* **8**, 1–15, <https://doi.org/10.1167/8.7.1> (2008).
37. Fishman, Y. I., Reser, D. H., Arezzo, J. C. & Steinschneider, M. Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear. Res.* **151**, 167–187 (2001).
38. Micheyl, C., Tian, B., Carlyon, R. P. & Rauschecker, J. P. Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* **48**, 139–148, <https://doi.org/10.1016/j.neuron.2005.08.039> (2005).
39. Mill, R. W., Böhm, T. M., Bendixen, A., Winkler, I. & Denham, S. L. Modelling the emergence and dynamics of perceptual organisation in auditory streaming. *PLoS Comput. Biol.* **9**, e1002925, <https://doi.org/10.1371/journal.pcbi.1002925> (2013).
40. Rankin, J., Sussman, E. & Rinzel, J. Neuromechanistic model of auditory bistability. *PLoS Comput. Biol.* **11**, e1004555, <https://doi.org/10.1371/journal.pcbi.1004555> (2015).
41. Barniv, D. & Nelken, I. Auditory streaming as an online classification process with evidence accumulation. *PLoS ONE* **10**, e0144788, <https://doi.org/10.1371/journal.pone.0144788> (2015).
42. Kleinschmidt, A., Sterzer, P. & Rees, G. Variability of perceptual multistability: from brain state to individual trait. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367**, 988–1000, <https://doi.org/10.1098/rstb.2011.0367> (2012).
43. Watanabe, T., Masuda, N., Megumi, F., Kanai, R. & Rees, G. Energy landscape and dynamics of brain activity during human bistable perception. *Nat. Commun.* **5**, 4765, <https://doi.org/10.1038/ncomms5765> (2014).
44. Genç, E., Bergmann, J., Singer, W. & Kohler, A. Interhemispheric connections shape subjective experience of bistable motion. *Curr. Biol.* **21**, 1494–1499, <https://doi.org/10.1016/j.cub.2011.08.003> (2011).
45. Powell, E. M. *et al.* Genetic disruption of cortical interneuron development causes region- and GABA cell type-specific deficits, epilepsy, and behavioral dysfunction. *J. Neurosci.* **23**, 622–631 (2003).
46. Schuler, V. *et al.* Epilepsy, hyperalgesia, impaired memory, and loss of pre- and postsynaptic GABA<sub>B</sub> responses in mice lacking GABA<sub>B(1)</sub>. *Neuron* **31**, 47–58 (2001).
47. Javitt, D. C., Steinschneider, M., Schroeder, C. E. & Arezzo, J. C. Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proc. Natl. Acad. Sci. USA* **93**, 11962–11967 (1996).
48. MacDonald, M. L. *et al.* Altered glutamate protein co-expression network topology linked to spine loss in the auditory cortex of schizophrenia. *Biol. Psychiatry* **77**, 959–968, <https://doi.org/10.1016/j.biopsych.2014.09.006> (2015).
49. Levelt, W. J. M. *On Binocular Rivalry*. (Institute for Perception RVO-TNO, 1965).
50. Brascamp, J. W., Klink, P. C. & Levelt, W. J. M. The 'laws' of binocular rivalry: 50 years of Levelt's propositions. *Vision Res.* **109**, 20–37, <https://doi.org/10.1016/j.visres.2015.02.019> (2015).
51. Kihara, K., Kondo, H. M. & Kawahara, J. I. Differential contributions of GABA concentration in frontal and parietal regions to individual differences in attentional blink. *J. Neurosci.* **36**, 8895–8901, <https://doi.org/10.1523/JNEUROSCI.0764-16.2016> (2016).
52. Tsubomi, H. *et al.* Dissociable neural activations of conscious visibility and attention. *J. Cogn. Neurosci.* **24**, 496–506, [https://doi.org/10.1162/jocn\\_a\\_00122](https://doi.org/10.1162/jocn_a_00122) (2012).
53. Britz, J., Landis, T. & Michel, C. M. Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb. Cortex* **19**, 55–65, <https://doi.org/10.1093/cercor/bhn056> (2009).
54. Britz, J., Pitts, M. A. & Michel, C. M. Right parietal brain activity precedes perceptual alternation during binocular rivalry. *Hum. Brain Mapp.* **32**, 1432–1442, <https://doi.org/10.1002/hbm.21117> (2011).
55. Paffen, C. L. E. & Alais, D. Attentional modulation of binocular rivalry. *Front. Hum. Neurosci.* **5**, 105, <https://doi.org/10.3389/fnhum.2011.00105> (2011).
56. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
57. Friston, K. & Kiebel, S. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 1211–1221, <https://doi.org/10.1098/rstb.2008.0300> (2009).
58. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711, <https://doi.org/10.1016/j.neuron.2012.10.038> (2012).
59. Kashino, M. & Kondo, H. M. Functional brain networks underlying perceptual switching: auditory streaming and verbal transformations. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367**, 977–987, <https://doi.org/10.1098/rstb.2011.0370> (2012).
60. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215, <https://doi.org/10.3389/fnhum.2010.00215> (2010).
61. FitzGerald, T. H., Moran, R. J., Friston, K. J. & Dolan, R. J. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. *Neuroimage* **107**, 219–228, <https://doi.org/10.1016/j.neuroimage.2014.12.015> (2015).
62. Hohwy, J. *The Predictive Mind*. (Oxford University Press, 2013).
63. Kanai, R., Komura, Y., Shipp, S. & Friston, K. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, <https://doi.org/10.1098/rstb.2014.0169> (2015).
64. Powers, A. R. 3rd, Gancsos, M. G., Finn, E. S., Morgan, P. T. & Corlett, P. R. Ketamine-induced hallucinations. *Psychopathology* **48**, 376–385, <https://doi.org/10.1159/000438675> (2015).
65. Lisman, J. Excitation, inhibition, local oscillations, or large-scale loops: what causes the symptoms of schizophrenia? *Curr. Opin. Neurobiol.* **22**, 537–544, <https://doi.org/10.1016/j.conb.2011.10.018> (2012).
66. Anticevic, A. *et al.* Early-course unmedicated schizophrenia patients exhibit elevated prefrontal connectivity associated with longitudinal change. *J. Neurosci.* **35**, 267–286, <https://doi.org/10.1523/JNEUROSCI.2310-14.2015> (2015).
67. Dieter, K. C. & Tadin, D. Understanding attentional modulation of binocular rivalry: a framework based on biased competition. *Front. Hum. Neurosci.* **5**, 155, <https://doi.org/10.3389/fnhum.2011.00155> (2011).
68. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
69. Mescher, M., Merkle, H., Kirsch, J., Garwood, M. & Gruetter, R. Simultaneous *in vivo* spectral editing and water suppression. *NMR Biomed.* **11**, 266–272 (1998).

70. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 443–446 (1997).
71. Wilson, M., Reynolds, G., Kauppinen, R. A., Arvanitis, T. N. & Peet, A. C. A constrained least-squares approach to the automated quantitation of *in vivo* <sup>1</sup>H magnetic resonance spectroscopy data. *Magn. Reson. Med.* **65**, 1–12, <https://doi.org/10.1002/mrm.22579> (2011).
72. Denham, S. L. & Winkler, I. The role of predictive models in the formation of auditory streams. *J. Physiol. Paris* **100**, 154–170, <https://doi.org/10.1016/j.jphysparis.2006.09.012> (2006).

### Acknowledgements

We thank an anonymous reviewer for suggesting the Discussion section about predictive coding, and another anonymous reviewer for suggesting the analysis in Supplementary Table S1. We also thank Keith Heberlein, Mark A. Brown and Heiko Meyer for providing a GABA spectral editing sequence (a work-in-progress version) and the Institute for Advanced Collaborative Researches at Chukyo University for generous support. This study was partly funded by JSPS KAKENHI grant (nos. 17K04494 to H.M.K. and 15K01857 to T.K.) and ERC grant ADAM (no. 295603 to D.P.). D.P. was supported by grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

### Author Contributions

H.M.K., D.P. and M.K. conceived of and designed the study. H.M.K. and Y.S. collected the data. H.M.K. and T.K. analysed the data. H.M.K. and D.P. wrote the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-32892-3>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018