# NORMAL AGING SLOWS SPONTANEOUS SWITCHING IN AUDITORY AND VISUAL BISTABILITY

**HIROHITO M. KONDO** [a,b*] **AND TAKANORI KOCHIYAMA** [c,d]

[a] *School of Psychology, Chukyo University, Nagoya, Aichi 466-8666, Japan*

[b] *Human Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, Atsugi, Kanagawa 243-0198, Japan*

[c] *Department of Cognitive Neuroscience, Advanced Telecommunications Research Institute International*

[d] *Brain Activity Imaging Center, ATR-Promotions, Seika-cho, Kyoto 619-0288, Japan*

**Abstract**—Age-related changes in auditory and visual perception have an impact on the quality of life. It has been debated how perceptual organization is influenced by advancing age. From the neurochemical perspective, we investigated age effects on auditory and visual bistability. In perceptual bistability, a sequence of sensory inputs induces spontaneous switching between different perceptual objects. We used different modality tasks of auditory streaming and visual plaids. Young and middle-aged participants (20–60 years) were instructed to indicate by a button press whenever their perception changed from one stable state to the other. The number of perceptual switches decreased with participants' ages. We employed magnetic resonance spectroscopy to measure non-invasively concentrations of the inhibitory neurotransmitter ($\gamma$-aminobutyric acid, GABA) in the brain regions of interest. When participants were asked to voluntarily modulate their perception, the amount of effective volitional control was positively correlated with the GABA concentration in the auditory and motion-sensitive areas corresponding to each sensory modality. However, no correlation was found in the prefrontal cortex and anterior cingulate cortex. In addition, effective volitional control was reduced with advancing age. Our results suggest that sequential scene analysis in auditory and visual domains is influenced by both age-related and neurochemical factors.

*This article is part of a Special Issue entitled: Sequence Processing.* © 2017 The Author(s). Published by Elsevier Ltd on behalf of IBRO. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Correspondence to: H. M. Kondo, School of Psychology, Chukyo University, 101-2 Yagoto Honmachi, Showa, Nagoya, Aichi 466-8666, Japan.
E-mail address: kondo@lets.chukyo-u.ac.jp (H. M. Kondo).
*Abbreviations:* AC, auditory cortex; ACC, anterior cingulate cortex; ANOVA, analysis of variance; GABA, $\gamma$-aminobutyric acid; MRS, magnetic resonance spectroscopy; MT, motion-sensitive area; PFC, prefrontal cortex.

**Key words:** perceptual organization, age, attention, awareness, scene analysis, MRS.

# INTRODUCTION

Normal aging impairs speech communication in noisy environments or multitalker situations. This reflects changes in peripheral and central auditory processing, as well as more general changes in cognitive and attentional processing (Working Group on Speech Understanding and Aging, 1988). The difficulty in speech comprehension is probably derived from age-related declines in listener's abilities of perceptual organization (see Alain et al., 2006 for a review). The purpose of perceptual organization is to structure a mixture of sensory inputs into meaningful perceptual objects. For hearing, this task is termed auditory scene analysis (Bregman, 1990). The principles of auditory scene analysis are based on sequential cues, such as similarity in acoustic features over time, and concurrent cues, such as harmonicity. In the laboratory, it has been found that concurrent scene analysis declines with age (Alain et al., 2001; Snyder and Alain, 2005), whereas sequential scene analysis is preserved in the elderly (Trainor and Trehub, 1989; Alain et al., 1996; Snyder and Alain, 2007). In the context of auditory streaming, a recent study has demonstrated that older listeners do not benefit more from background (i.e., unattended) predictable cues than younger listeners do (Rimmele et al., 2012a). Therefore, there is the possibility that advancing age impairs the ability to exploit sequential information outside the focus of attention.

A parallel exists in vision. In the absence of disease or injury, we can suffer from age-related impairments of spatial vision, motion perception, and object recognition (Ball and Sekuler, 1986; Spear, 1993; Sekuler and Sekuler, 2000). Surprisingly, there are rather few studies investigating age effects on visual perceptual organization. Most previous studies have used binocular rivalry, where monocular images are presented to different eyes, and shown that the rate of rivalry alternations slows down with age (Jalavisto, 1964; Ukai et al., 2003; Tarita-Nistor et al., 2006). The present study investigated whether visual motion perception is influenced by advancing age, because visual motion provides sequential information to segregate objects from the background.

We focused on bistable stimuli to clarify age effects on perceptual organization. In perceptual bistability, a physically unchanging stimulus induces spontaneous switching of conscious perception (Schwartz et al.,

2012). To compare two sensory modalities, we employed auditory streaming as an auditory stimulus (Pressnitzer and Hupé, 2006; Kondo and Kashino, 2009) and moving plaids as a visual stimulus (Adelson and Movshon, 1982; Hupé and Rubin, 2003) (Fig. 1A). The two types of stimuli have an advantage for probing perceptual organization, because they involve a basic competition between a one-object (grouped) and a two-object (split) interpretation. In daily life, we frequently have difficulty taking signals from sensory inputs due to spatial and temporal occluders. From the perspective of adaptive behavior, it is thus important for the brain to create some possible percepts from insufficient information and fluctuate its perceptual interpretations (Kondo et al., 2017a).

We not only collected spontaneous bistable reports, but also manipulated the volition (or sometimes called selective attention) of participants during the task. Perception and attention interact dynamically to facilitate information processing. In general, if the relevant
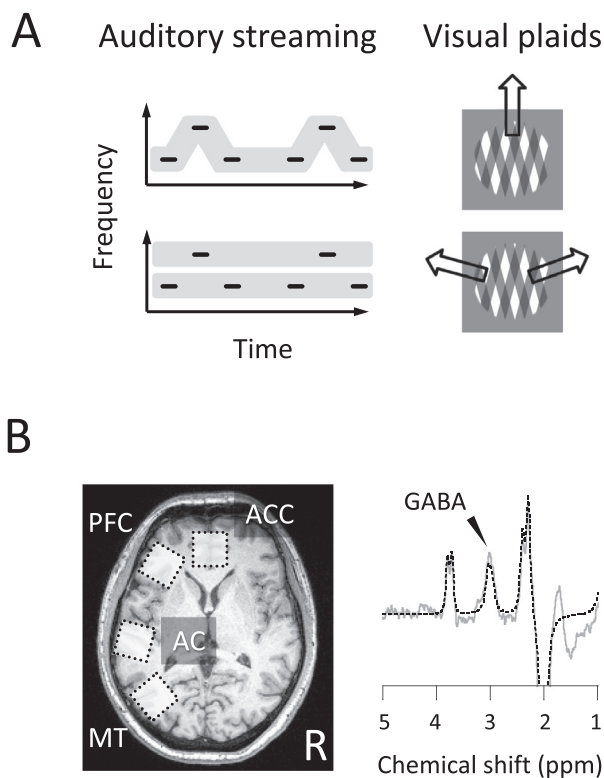


**Fig. 1.** Perceptual tasks and MRS data acquisition. (A) A sound sequence of triplet tones was presented in auditory streaming. The auditory stimulus produced perceptual switches between one and two streams. Two rectangular gratings were moving in visual plaids. The visual stimulus led to perceptual switches between upward grouped and sideward split motion. (B) The size and location of voxels (left) and the edited GABA spectrum of one representative participant (right). 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 anterior cingulated cortex (ACC) was centered on the interhemispheric fissure. The γ-aminobutyric acid (GABA) peak was obtained from the differences in spectra by editing radio frequency on/off pulses and fitted by a Gaussian function (dashed line).

information is perceptually salient, attentional demand would be low (Desimone and Duncan, 1995). Even when observers select what to attend based on low-level features, attention operates on perceptual objects (Shinn-Cunningham, 2008). In a bistability paradigm, the predictability of sequential information gives no priority to either of the percepts (Bendixen, 2014). Thus, we can expect that bottom-up attention is essentially required for the formation and selection of perceptual objects and that spontaneous switching in perceptual bistability is sensitive to attentional abilities for each individual.

We hypothesized that participants' ages would affect perceptual organization and volitional control. What neural mechanisms contribute to the age effects? One possibility is that the efficacy of cortical inhibition based on γ-aminobutyric acid (GABA) mediates neurophysiological changes in perception and attention (Leventhal et al., 2003), because the GABAergic system plays a critical role in all sorts of brain functions, such as sharpening tuning curves of sensory neurons and orienting the focus of attention (Tadin and Blake, 2005). We used magnetic resonance spectroscopy (MRS) to obtain a non-invasive, *in vivo* measurement of GABA in the human brain (Puts and Edden, 2012). Previous studies have shown that the GABA concentration in different cortical areas predicts individual differences in orientation discrimination (Edden et al., 2009), visual awareness (van Loon et al., 2013), visual attention (Kihara et al., 2016), and auditory multistability (Kondo et al., 2017a). The interest in the MRS technique has grown in recent years, but few studies have investigated the relationship between GABA levels and age-related changes in brain functions (see Duncan et al., 2014 for a review). The present study specified GABA levels in a broad range of young and middle-aged people.

In MRS studies, *a priori* assumptions are important for regional specificity (Duncan et al., 2014). From the perspective of neurochemical measures, we here explore the outstanding issue of whether the bistable selection process is supported by low-level sensory areas or high-level cognitive ones. Auditory streaming has neural correlates in the auditory cortex (AC) (Gutschalk et al., 2005; Micheyl et al., 2007; Kondo and Kashino, 2009), whereas moving plaids induce activity in the motion-sensitive area (MT) of the visual cortex (Huk and Heeger, 2001; Castelo-Branco et al., 2002). In addition, it has long been argued that the prefrontal cortex (PFC) and anterior cingulate cortex (ACC) are involved in perceptual switching in visual bistability (Lumer et al., 1998; Sterzer and Kleinschmidt, 2007), but this has been recently disputed (Brascamp et al., 2015). We clarified whether GABA concentrations in the four cortical regions are linked with individual differences in auditory and visual bistability.

## EXPERIMENTAL PROCEDURES

### Participants

Thirty-eight young and middle-aged participants were recruited for this study (22 males and 16 females; $M_{age}$ = 36.6, $SD_{age}$ = 10.8, range 20–60 years). The pool of the participants consisted of the following age

groups: the twenties ($n$ = 11), thirties ($n$ = 12), forties ($n$ = 9), and fifties ($n$ = 6). The score of the Edinburgh Handedness Inventory was 94.5 ± 5.9, indicating that they were strongly right-handed. All participants were Japanese people with normal hearing and with normal or corrected-to-normal vision. None had any history of neurological or psychiatric disorders. All procedures reported in this study were approved by the Ethics and Safety Committees of NTT Communication Science Laboratories and ATR-Promotions (protocol numbers: H24-004 and AN14-001). The study was carried out in accordance with the Declaration of Helsinki. All participants gave written informed consent after the procedures had been fully explained to them.

## MRS data acquisition

To minimize confounding factors affecting neurotransmitter concentrations, we conducted the acquisition of MR spectra at a fixed time for all participants, which was between 1:00 p.m. to 2:30 p.m. For the reduction of noise or bias, participants were requested to refrain from alcohol for 24 h before the experiment and caffeine on the day of the experiment. Data were acquired with a 3T MRI scanner with a 12-channel receive-only head coil (Magnetom Trio, Siemens). Head motion was minimized with comfortable padding around the participant's head. For assessment of cortical thickness and volume, anatomical images were obtained with a T1-weighted pulse sequence (isotropic voxel size of 1 mm$^3$).

MR spectra were acquired from four $3 \times 3 \times 3$ cm$^3$ voxels of interest, positioned in the AC, MT, PFC, and ACC (Fig. 1B). Voxels were positioned using internal landmarks in order to achieve a consistent position between participants. The AC voxel was aligned to 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 centered at the junction of the ascending limb of the inferior temporal sulcus and the lateral occipital sulcus. The PFC voxel was located at the anterior part of the middle frontal gyrus (BA 46). The ACC voxel (including portions of BAs 32 and 9) was located superior to the genu of the corpus callosum and centered on the interhemispheric fissure. All the voxels except the ACC one were angled parallel to the brain surface of the left hemisphere.

Four consecutive runs were acquired from the different voxels for each participant. Before each run, we carefully carried out manual shimming (approximately 5 min) of the magnetic field in the voxel to avoid line broadening. We used the Mescher-Garwood proton resolved spectroscopy (MEGA-PRESS) technique (Mescher et al., 1998), in which editing pulses were applied at 1.9 ppm (on) and 7.5 ppm (off). Thus, the subtraction of edit-off from edit-on spectra allows separation of the GABA signal at 3.00 ppm from the overlying creatine signal. 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 scan duration of 3 min 18 s. The unsuppressed water signal was also acquired from the same voxel. As with all metabolic imaging studies using MRS, the neurotransmitter concentration measures reflect both the intracellular concentration and synaptic pool.

## Perceptual bistability tasks

Immediately after the MRS data acquisition, participants performed the perceptual tasks outside the scanner. The whole psychophysical experiment lasted approximately 1 h. Stimulus presentation and response collection were managed using MATLAB (Mathworks Inc.) with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The auditory streaming and visual plaids tasks were conducted separately (Kondo et al., 2012).

Auditory streaming 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 centered 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.

Visual plaids stimuli consisted of two rectangular-wave gratings (velocity = 1.25 deg/s; 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 gray background, at a viewing distance of 57 cm and covering a visual angle of 5 deg. A small point for fixation was added to center of the visual stimuli.

We first explained the two tasks by means of a visual illustration of the stimuli (Fig. 1A). 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 visual 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 behavioral 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. Responses were collected via two buttons on a computer keyboard. A response indicated by a button press was held until a subsequent button press. The order of auditory or visual blocks was assigned randomly.

Following the Neutral condition, two further conditions with volitional control manipulations were run (Pressnitzer and Hupé, 2006). In the Grouped condition, participants were instructed to try and maintain perception of one-stream or one-plaid as much as possible. In the Split condition, participants were instructed to try and maintain perception of two-stream or two-grating as much as possible. The order of the two volitional conditions was randomized across participants.

**Data analyses**

GABA quantification was performed using the Gannet toolbox for MATLAB (Edden et al., 2014). MRS data analysis comprised the following steps: (1) alignment of edit-on and -off spectra; (2) subtraction of edited spectra, to yield GABA spectra, and averaging across acquisitions; and (3) fitting a Gaussian function to the GABA peak at 3.00 ppm and quantifying GABA based on the area under the curve. The water level was obtained from a Gaussian–Lorentzian fit to the unsuppressed water signal. The GABA concentrations were quantified in institutional units (i.u.) (Mullins et al., 2014). The GABA concentrations (mean ± standard error) were 1.72 ± 0.05 i.u. for AC; 1.49 ± 0.04 i.u. for MT; 1.51 ± 0.08 i.u. for PFC; 1.29 ± 0.05 i.u. for ACC. The fit errors related to the GABA values were 8.6 ± 0.4% for AC; 8.2 ± 0.4% for MT; 11.2 ± 0.5% for PFC; 9.8 ± 0.4% for ACC. Thus, the reliability of neurotransmitter levels generally reached a satisfactory level because the fit errors were around 10%. A correlation analysis was conducted to investigate the relationship between participants' ages and neurotransmitter measures. However, it did not reveal any significant correlation: $|r| < 0.28$, $p > 0.10$ (Table 1). Thus, advancing age is relatively independent of the GABA levels identified in the present study.

For the behavioral data analysis, the time series of percept durations was analyzed separately for each auditory streaming and visual plaids blocks. All durations were longer than 300 ms. It has been reported that the duration of the first percept is longer than the duration of subsequent ones for auditory streaming and visual plaids (Denham and Winkler, 2006; Pressnitzer and Hupé, 2006). Thus, we excluded the first-percept duration from the analysis. From the remaining durations, for each participant we evaluated the proportion of grouped percepts relative to the total duration. 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 (Meng and Tong, 2004). Statistical analyses were carried out with IBM SPSS Statistics (version 22). In an analysis of variance (ANOVA), the Greenhouse-Geisser correction ($\varepsilon$) was used to adjust degrees of freedom of a within-subject test and correct for violations of the assumption of sphericity. In such cases, adjusted $p$-values were acquired. Tukey's HSD test was used for post hoc comparisons ($\alpha$-level = 0.05). The Jonckheere-Terpstra test was used as a trend test (two-tailed).

## RESULTS

**Perceptual Switching, Age, and GABA**

We found negative correlations between participants' ages and switch numbers. Under the Neutral condition, the number of perceptual switches decreased with advancing age: $r = -0.332$, $p = 0.041$ for auditory streaming; $r = -0.373$, $p = 0.021$ for visual plaids (Fig. 2A). We classified all participants into the four groups to examine the relationship between participants' ages and switch numbers. A mixed-design ANOVA was conducted to analyze effects of the input modality (a within-subject factor), age group (a between-subject factor), and volitional condition (a within-subject factor). First, the number of perceptual switches (mean ± standard error) was greater for visual plaids (60.2 ± 4.5) than for auditory streaming (35.0 ± 3.0): $F_{(1, 34)} = 33.97$, $\eta_p^2 = 0.500$, $p < 0.001$. Second, a main effect of age group was marginally significant: $F_{(3,34)} = 2.81$, $\eta_p^2 = 0.199$, $p = 0.054$. Post-hoc tests revealed that although the number of perceptual switches did not differ between the twenties (61.2 ± 5.7) and thirties (50.3 ± 5.4) groups, it was greater for the twenties group than for the forties (42.9 ± 6.3) and fifties (36.1 ± 7.7) groups. We performed a trend test to examine age effects on the number of perceptual switches. There was a significant trend of smaller switch numbers with older age groups: $T_{JT} = 175.50$, $z = -2.36$, $p = 0.018$ for auditory streaming; $T_{JT} = 187.00$, $z = -2.09$, $p = 0.040$ for visual plaids (Fig. 2B). Third, a main effect of volitional condition did not reach statistical significance: the Neutral (48.9 ± 3.0), Grouped (45.2 ± 3.4), and Split (48.7 ± 4.0) conditions; $F_{(2, 68)} = 1.36$, $\eta_p^2 = 0.039$, $p = 0.263$. There was no interaction of modality × group, $F_{(3, 34)} = 0.93$, $\eta_p^2 = 0.076$, $p = 0.436$; modality × condition, $F_{(2, 68)} = 1.65$, $\eta_p^2 = 0.046$, $p = 0.204$; group × condition, $F_{(6, 68)} = 1.14$, $\eta_p^2 = 0.091$, $p = 0.350$. A three-way interaction was not significant: $F_{(6, 68)} = 0.68$, $\eta_p^2 = 0.056$, $p = 0.643$. Our results indicate that advancing age influences the number of perceptual switches for both input modalities, but the degree of volitional control does not.

**Table 1.** Correlations between participants' ages, perceptual measures and GABA concentrations

| Measure | Age | | AC GABA | | MT GABA | | PFC GABA | | ACC GABA | |
|---|---|---|---|---|---|---|---|---|---|---|
| | r | p | r | p | r | p | r | p | r | p |
| Age | n/a | n/a | -0.100 | 0.551 | 0.226 | 0.184 | −0.273 | 0.108 | 0.010 | 0.955 |
| Switch numbers | | | | | | | | | | |
|   Auditory | **−0.332** | **0.041** | −0.235 | 0.168 | −0.004 | 0.980 | 0.152 | 0.376 | −0.077 | 0.660 |
|   Visual | **−0.373** | **0.021** | −0.073 | 0.664 | **−0.346** | **0.034** | 0.161 | 0.348 | −0.221 | 0.203 |
| Volitional control | | | | | | | | | | |
|   Auditory | **−0.394** | **0.014** | **0.346** | **0.033** | −0.142 | 0.408 | −0.176 | 0.304 | −0.140 | 0.423 |
|   Visual | −0.244 | 0.140 | 0.228 | 0.168 | 0.308 | 0.068 | −0.158 | 0.358 | −0.001 | 0.994 |

*Note:* A value indicated in bold is significant ($p < 0.05$, $N = 38$). Auditory, auditory streaming; Visual, visual plaids. r, Pearson's correlation coefficient.
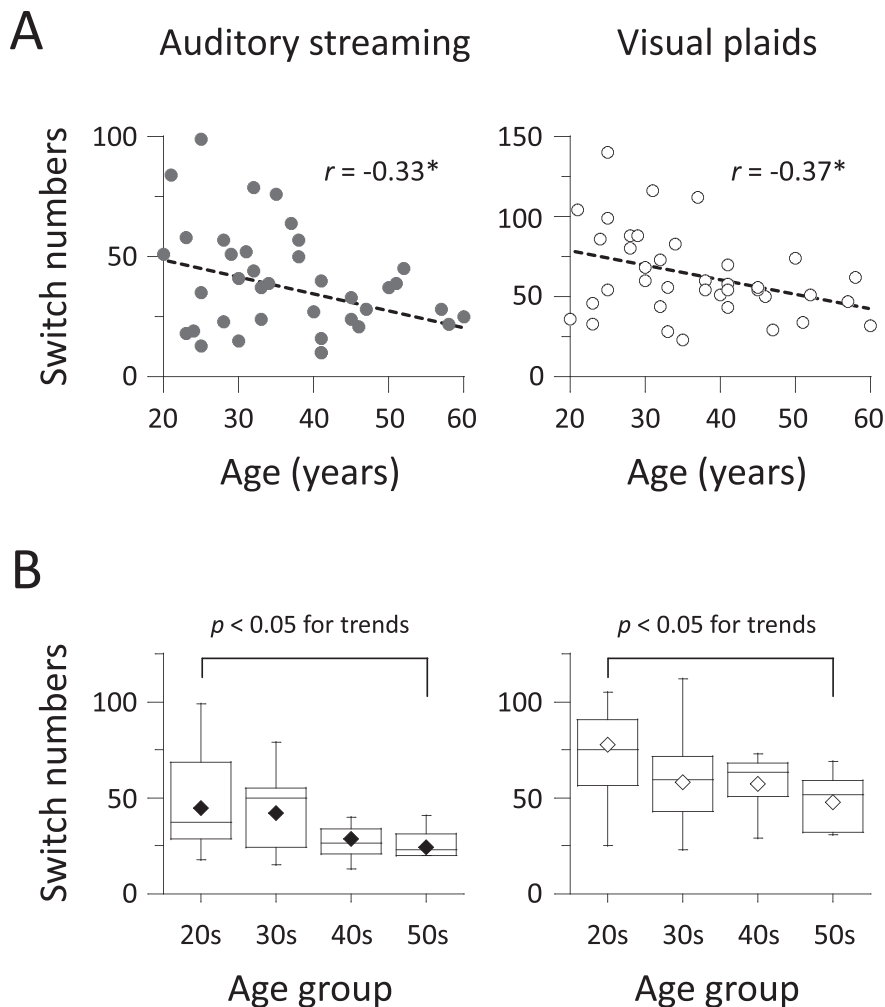
## A



## B



**Fig. 2.** Relationship between participants' ages and the number of perceptual switches. (A) Scatter plots with linear regression fit for auditory streaming and visual plaids. Circles indicate individual data ($N = 38$) under the Neutral condition. Switch numbers were obtained from a 5-min experimental block for each task. $^*p < 0.05$. (B) Box plots for comparison of switch numbers between the four age groups: the twenties ($n = 11$), thirties ($n = 12$), forties ($n = 9$), and fifties ($n = 6$). Diamonds indicate the mean of grouped data. The Jonckheere–Terpstra test was used as a trend test.

We performed a correlation analysis to investigate whether GABA concentrations are associated with the number of perceptual switches under the Neutral condition (Table 1). We found only a significant correlation between GABA concentrations in MT and switch numbers in visual plaids: $r = -0.346$, $p = 0.034$. For auditory streaming, there was a negative but low correlation between AC GABA concentrations and switch numbers: $r = -0.235$, $p = 0.168$. A partial correlation analysis was used to specify the relationship between switch numbers, participants' ages, and GABA levels. When age effects were partialed out, partial correlations between switch numbers and GABA concentrations did not reach statistical significance: $r = -0.171$, $p = 0.327$ for auditory streaming; $r = -0.286$, $p = 0.095$ for visual plaids. In addition, GABA concentrations in PFC and ACC did not show any significant correlation with the number of perceptual switches: $|r| < 0.22$, $p > 0.20$. These results suggest

that GABA concentrations have a limited effect on the number of perceptual switches.

### Volitional Control, Age, and GABA

We next focused on volitional control for perceptual organization. For each participant, we computed the amount of effective volitional control by subtracting the proportion of grouped percepts under the Split condition from that under the Grouped condition. For auditory streaming, the amount of effective volitional control was negatively correlated with participants' ages: $r = -0.394$, $p = 0.014$ (Fig. 3A). In addition, for visual plaids, there was a low but negative correlation between the two variables: $r = -0.236$, $p = 0.142$. The results suggest that normal aging impairs attentional abilities of top-down control for auditory bistability. We performed a 2 (input modality) × 4 (age group) ANOVA on the amount of effective volitional control. The measure did not differ between auditory streaming (0.270 ± 0.037) and visual plaids (0.247 ± 0.039): $F_{(1,34)} = 0.34$, $\eta_p^2 = 0.010$, $p = 0.563$. A main effect of age group was not significant: 0.301 ± 0.059 for the twenties; 0.324 ± 0.057 for the thirties; 0.263 ± 0.066 for the forties; 0.145 ± 0.080 for the fifties; $F_{(3, 34)} = 1.19$, $\eta_p^2 = 0.095$, $p = 0.328$. In addition, we did not find a monotonic trend between the age groups and effective volitional control: $T_{JT} = 196.50$, $z = -1.80$, $p = 0.071$ for auditory streaming; $T_{JT} = 235.00$, $z = -0.80$, $p = 0.424$ for visual plaids (Fig. 3B). An interaction of input modality × age group was also not significant: $F_{(3, 34)} = 0.44$, $\eta_p^2 = 0.037$, $p = 0.726$. Thus, it appears that volitional control abilities for visual perceptual organization do not deteriorate with advancing age.

We performed a correlation analysis to investigate whether GABA concentrations are related to the amount of effective volitional control (Table 1). For auditory streaming, we found a positive correlation between AC GABA levels and effective volitional control: $r = 0.346$, $p = 0.033$ (Fig. 4). For visual plaids, there was a marginal significant correlation between MT GABA levels and effective volitional control: $r = 0.308$, $p = 0.068$. When age effects were removed from the zero-order correlations, partial correlations between GABA concentrations and effective volitional control were significant: $r = 0.388$, $p = 0.021$ for auditory streaming; $r = 0.385$, $p = 0.022$ for visual plaids. Thus,
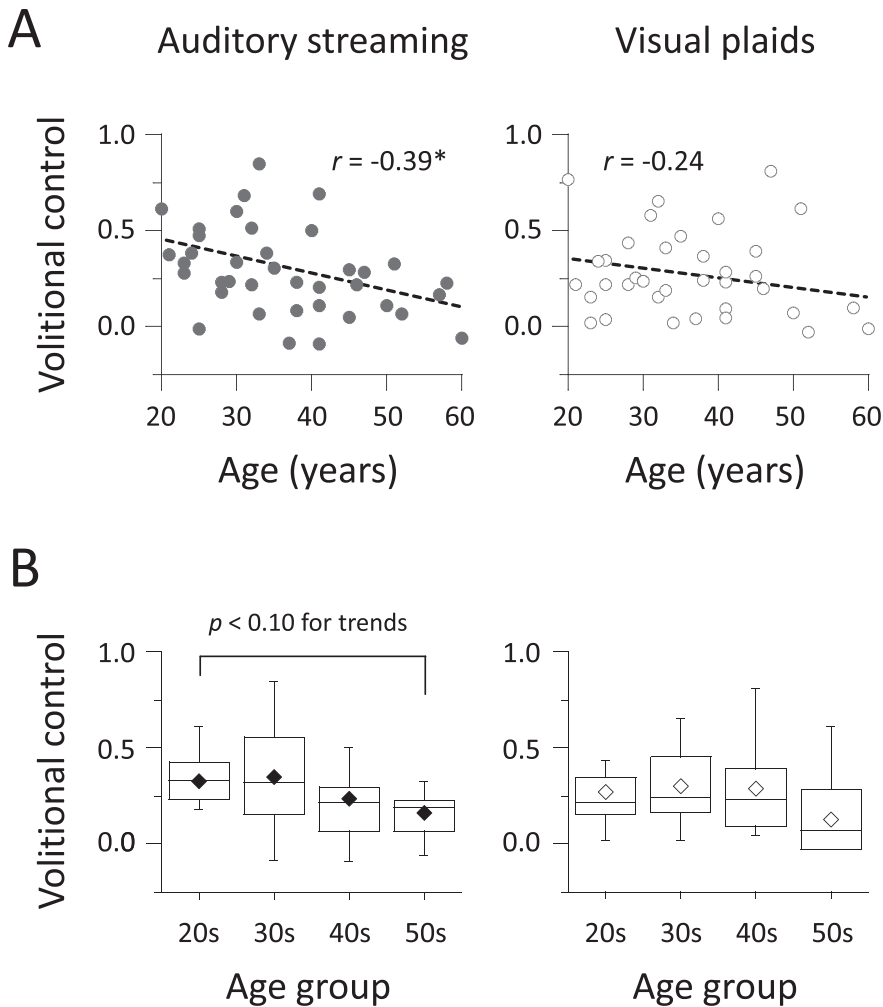
# A

### Auditory streaming
### Visual plaids



**Fig. 3.** Relationship between participants' ages and volitional control for perceptual organization. (A) Scatter plots with linear regression fit for auditory streaming and visual plaids. The computation for the amount of effective volitional control is described in the main text (see Data Analyses section). (B) Box plots for comparison of effective volitional control between the four age groups.
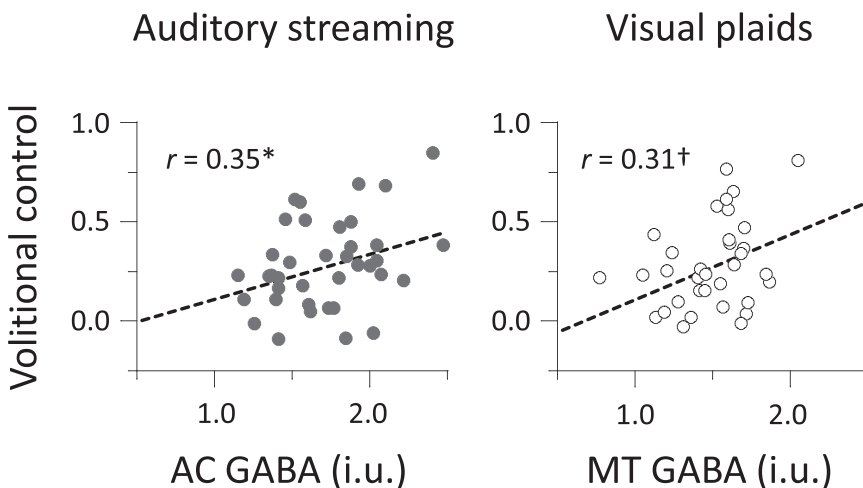


**Fig. 4.** Correlations between GABA concentrations and volitional control for perceptual organization. Circles represent individual data. The horizontal axis indicates the institutional unit (i.u.). $^*p < 0.05$, $^†p < 0.10$.

for each participant, the amount of effective volitional control could be explained by GABA concentrations within modality-specific sensory areas.

## DISCUSSION

We showed that normal aging reduced the number of perceptual switches, even when participants performed simple auditory streaming or visual plaids task. For both modalities, the amount of effective volitional control decreased with age when participants had to favor one perceptual interpretation over the other. In addition, the perceptual measure was positively correlated with GABA concentrations within sensory areas corresponding to the stimulus modality – the AC for audition and MT for vision. No correlation was found in the PFC and ACC. Our results suggest that sequential scene analysis in auditory and visual domains is influenced by both age-related and neurochemical factors.

For audition, it has been found that age-related declines in concurrent stream segregation are robust, but that age effects on sequential stream segregation are subtle (see Alain et al., 2006 for a review). In the latter case, older listeners as well as younger ones have abilities to exploit cues of sequential information by means of stimulus features, such as differences in frequency (Snyder and Alain, 2007) and presentation rate (Trainor and Trehub, 1989). Snyder and Alain (2007) have demonstrated that the proportion of a grouped percept during the build-up (approximately 10 s) of auditory streaming does not differ between the young ($M_{age} = 29.5$), middle-aged ($M_{age} = 46.6$), and elderly ($M_{age} = 72.7$) when a sequence of triplet-tone sounds induced ambiguous perception. In the present study, the proportion of a grouped percept relative to a 5-min stimulus duration was similar in both young and middle-aged groups: 49.8 ± 4.1% (mean ± standard error) for the twenties, 49.3 ± 4.0% for the thirties, 48.5 ± 2.1% for the forties, and 52.0 ± 5.3% for the fifties. However, the number of percep-

tual switches decreased with advancing age. Thus, there is the possibility that a factor to initiate perceptual switches is functionally independent of a factor to stabilize perceptual states.

The same pattern of auditory streaming results was found in visual plaids: the number of perceptual switches was negatively correlated with participants' ages, but the proportion of a grouped percept was not: $r = 0.116$, $p = 0.486$. Spontaneous switching in visual plaids, as well as in auditory streaming, is probably susceptible to age-related impairments in using unattended background cues. In a binocular rivalry paradigm, both younger ($M_{age} = 22.4$) and older ($M_{age} = 66.4$) observers found it more difficult to detect a probe when it was presented to the non-dominant eye, but the magnitude of accuracy deterioration was greater for older observers (Norman et al., 2007). Roudaia et al. (2010) used a two-frame apparent motion paradigm to investigate the properties of low motion detectors. The spatial and temporal integration of motion information was reduced with advancing age. These results have been taken as evidence that older people cannot use unattended cues of visual perceptual organization.

We found that the number of perceptual switches did not differ between the different volitional conditions, although volitional control modulated the dominance of grouped or split percepts. It is known that participants are able to influence their percepts under bistable situations when they are instructed to attend to a specific target (Meng and Tong, 2004; Pressnitzer and Hupé, 2006). However, participants cannot fully prevent perceptual switching from occurring (Paffen et al., 2006). The inability of volitional control for perceptual switching may be explained by a two-level neural network model of binocular rivalry (Wilson, 2003), in which neural adaptation among the dominant neurons plays an important role in evoking rivalry alternations. This model suggests that although directing attention to a specific target enhances the contrast between different stable percepts, the strongly driven neurons also strongly adapt. Given that that is the case, one could expect to observe the pattern of our results described above.

Another set of findings relates to the amount of effective volitional control. The measure of auditory streaming decreased with advancing age. This may be caused by the deterioration of sensory memory. Several researchers have argued that regular patterns embedded within auditory stimuli are stored in auditory sensory memory and used to create a predictive model of the auditory scene (Winkler et al., 2009; Bendixen et al., 2010). Incoming input that matches the prediction is integrated into the auditory object, whereas deviations from the prediction are detected in auditory sensory memory and processed as changes in auditory objects or borders of auditory objects (Shinn-Cunningham, 2008). Older adults' deficits in regularity processing in auditory sensory memory may affect the ability to use regular patterns for stream segregation (Alain and Woods, 1999; Rimmele et al., 2012b).

The amount of effective volitional control was positively correlated with GABA concentrations in the AC and MT corresponding to the input modalities, but not with those in the PFC and ACC. The MRS results provide evidence that the GABAergic system within sensory areas is implicated in attentional modulation for perceptual organization. Neuroimaging findings have been mixed in bistable paradigms. Indeed, perceptual switching induces activations of sensory and subcortical areas in auditory (Gutschalk et al., 2005; Kondo and Kashino, 2009; Schadwinkel and Gutschalk, 2011) and visual pathways (Tong et al., 1998; Polonsky et al., 2000; Wunderlich et al., 2005). However, several researchers have argued that the frontal areas are somehow involved in perceptual switching during binocular rivalry and apparent motion, even without the participant's volitional intent (e.g., Lumer et al., 1998; Sterzer and Kleinschmidt, 2007). A transcranial magnetic stimulation study has shown that the PFC is causally relevant for voluntary control over perceptual switches of a structure-from-motion stimulus (de Graaf et al., 2011). In an electrophysiological study, elderly people have reduced frontal activities related to auditory attention (Chao and Knight, 1997). However, the PFC may not be related to the initiation of perceptual switching itself (Paffen and Alais, 2011; Brascamp et al., 2015). Another explanation is that the GABAergic system of the frontal areas is not associated with task-specific attentional modulation, but rather a general mental set: GABA concentrations in PFC play an important role in identifying stimulus features in a top-down manner (Kihara et al., 2016). Therefore, future study should clarify the relationship between neural and neurochemical measures in terms of brain networks.

The interest in individual differences in conscious perception has been growing in recent years (Kanai and Rees, 2011; Kondo et al., 2012; Deike et al., 2014; Kondo et al., 2017b). However, it is unclear whether sequential scene analysis is influenced by advancing age. The present study used a bistable paradigm to investigate age effects on auditory and visual perceptual organization. Normal aging reduced the number of perceptual switches in auditory streaming and visual plaids. Intriguingly, the amount of effective volitional control was positively correlated with GABA concentrations within sensory areas coding for each modality. Further, our findings have clinical implications in that dysfunctions of the GABAergic system impact perceptual organization.

## AUTHOR CONTRIBUTIONS

H.M.K. designed the research; H.M.K. performed the research; H.M.K. and T.K. analyzed and interpreted data; and H.M.K. wrote the paper.

# REFERENCES

Adelson EH, Movshon JA (1982) Phenomenal coherence of moving visual patterns. Nature 300:523–525.

Alain C, Dyson BJ, Snyder JS (2006) Aging and the perceptual organization of sounds: a change of scene? In: Conn PM, editor. Handbook of models for human aging. New York: Academic Press. p. 759–770.

Alain C, McDonald KL, Ostroff JM, Schneider B (2001) Age-related changes in detecting a mistuned harmonic. J. Acoust. Soc. Am. 109:2211–2216.

Alain C, Ogawa KH, Woods DL (1996) Aging and the segregation of auditory stimulus sequences. J Gerontol B Psychol Sci Soc Sci 51:91–93.

Alain C, Woods DL (1999) Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. Psychol Aging 14:507–519.

Ball K, Sekuler R (1986) Improving visual perception in older observers. J Gerontol 41:176–182.

Bendixen A (2014) Predictability effects in auditory scene analysis: a review. Front Neurosci 8:60.

Bendixen A, Denham SL, Gyimesi K, Winkler I (2010) Regular patterns stabilize auditory streams. J Acoust Soc Am 128:3658–3666.

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:443–446.

Brascamp J, Blake R, Knapen T (2015) Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. Nat Neurosci 18:1672–1678.

Bregman AS (1990) Auditory Scene Analysis: the Perceptual Organization of Sound. Cambridge, MA: MIT Press.

Castelo-Branco M, Formisano E, Backes W, Zanella F, Neuenschwander S, Singer W, Goebel R (2002) Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. Proc Natl Acad Sci USA 99:13914–13919.

Chao LL, Knight RT (1997) Prefrontal deficits in attention and inhibitory control with aging. Cereb Cortex 7:63–69.

de Graaf TA, de Jong MC, Goebel R, van Ee R, Sack AT (2011) On the functional relevance of frontal cortex for passive and voluntarily controlled bistable vision. Cereb Cortex 21:2322–2331.

Deike S, Denham SL, Sussman E (2014) Probing auditory scene analysis. Front Neurosci 8:293.

Denham SL, Winkler I (2006) The role of predictive models in the formation of auditory streams. J Physiol Paris 100:154–170.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Duncan NW, Wiebking C, Northoff G (2014) Associations of regional GABA and glutamate with intrinsic and extrinsic neural activity in humans—a review of multimodal imaging studies. Neurosci Biobehav Rev 47:36–52.

Edden RAE, Muthukumaraswamy SD, Freeman TCA, Singh KD (2009) Orientation discrimination performance is predicted by GABA concentration and gamma oscillation frequency in human primary visual cortex. J Neurosci 29:15721–15726.

Edden RAE, Puts NAJ, Harris AD, Barker PB, Evans CJ (2014) Gannet: a batchprocessing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. J Magn Reson Imaging 40:1445–1452.

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25:5382–5388.

Huk AC, Heeger DJ (2001) Pattern-motion responses in human visual cortex. Nat Neurosci 5:72–75.

Hupé JM, Rubin N (2003) The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. Vision Res 43:531–548.

Jalavisto E (1964) The phenomenon of retinal rivalry in the aged. Gerontologia 9:1–8.

Kanai R, Rees G (2011) The structural basis of inter-individual differences in human behaviour and cognition. Nat Rev Neurosci 12:231–242.

Kihara K, Kondo HM, Kawahara JI (2016) Differential contributions of GABA concentration in frontal and parietal regions to individual differences in attentional blink. J Neurosci 36:8895–8901.

Kondo HM, Farkas D, Denham SL, Asai T, Winkler I (2017a) Auditory multistability and neurotransmitter concentrations in the human brain. Philos Trans R Soc Lond B Biol Sci 372:20160110.

Kondo HM, Kashino M (2009) Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. J Neurosci 29:12695–12701.

Kondo HM, Kitagawa N, Kitamura MS, Koizumi A, Nomura M, Kashino M (2012) Separability and commonality of auditory and visual bistable perception. Cereb Cortex 22:1915–1922.

Kondo HM, van Loon AM, Kawahara JI, Moore BCJ (2017b) Auditory and visual scene analysis: an overview. Philos Trans R Soc Lond B Biol Sci 372:20160099.

Leventhal AG, Wang Y, Pu M, Zhou Y, Ma Y (2003) GABA and its agonists improved visual cortical function in senescent monkeys. Science 300:812–815.

Lumer ED, Friston KJ, Rees G (1998) Neural correlates of perceptual rivalry in the human brain. Science 280:1930–1934.

Meng M, Tong M (2004) Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. J Vis 4:539–551.

Mescher M, Merkle H, Kirsch J, Garwood M, Gruetter R (1998) Simultaneous in vivo spectral editing and water suppression. NMR Biomed 11:266–272.

Micheyl C, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Wilson EC (2007) The role of auditory cortex in the formation of auditory streams. Hear Res 229:116–131.

Mullins PG, McGonigle DJ, O'Gorman RL, Puts NAJ, Vidyasagar R, Evans CJ, Cardiff Symposium on MRS of GABA, Edden RAE (2014) Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. Neuroimage 86:43–52.

Norman JF, Norman HF, Pattison K, Taylor MJ, Goforth KE (2007) Aging and the depth of binocular rivalry suppression. Psychol Aging 22:625–631.

Paffen CLE, Alais D (2011) Attentional modulation of binocular rivalry. Front Hum Neurosci 5:105.

Paffen CLE, Alais D, Verstraten FAJ (2006) Attention speeds binocular rivalry. Psychol Sci 17:752–756.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442.

Polonsky A, Blake R, Braun J, Heeger DJ (2000) Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. Nat Neurosci 3:1153–1159.

Pressnitzer D, Hupé JM (2006) Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. Curr Biol 16:1351–1357.

Puts NAJ, Edden RAE (2012) *In vivo magnetic resonance spectroscopy* of GABA: a methodological review. Prog Nucl Magn Reson Spectrosc 60:29–41.

Rimmele J, Schröger E, Bendixen A (2012a) Age-related changes in the use of regular patterns for auditory scene analysis. Hear Res 289:98–107.

Rimmele J, Sussman E, Keitel C, Jacobsen T, Schröger E (2012b) Electrophysiological evidence for age effects on sensory memory processing of tonal patterns. Psychol Aging 27:384–398.

Schadwinkel S, Gutschalk A (2011) Transient bold activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. J Neurophysiol 105:1977–1983.

Schwartz JL, Grimault N, Hupé JM, Moore BCJ, Pressnitzer D (2012) Multistability in perception: binding sensory modalities, an overview. Philos Trans R Soc Lond B Biol Sci 367:896–905.

Sekuler R, Sekuler AB (2000) Visual perception and cognition. In: Evans JG, Williams TF, Beattie BL, Michel J-P, Wilcock GK, editors. Oxford textbook of geriatric medicine. New York: Oxford University Press.

Shinn-Cunningham BG (2008) Object-based auditory and visual attention. Trends Cogn Sci 12:182–186.

Snyder JS, Alain C (2005) Age-related changes in neural activity associated with concurrent vowel segregation. Cogn Brain Res 24:492–499.

Snyder JS, Alain C (2007) Sequential auditory scene analysis is preserved in normal aging adults. Cereb Cortex 17:501–512.

Spear PD (1993) Neural bases of visual deficits during aging. Vision Res 33:2589–2609.

Sterzer P, Kleinschmidt A (2007) A neural basis for inference in perceptual ambiguity. Proc Natl Acad Sci USA 104:323–328.

Tadin D, Blake R (2005) Motion perception getting better with age? Neuron 45:325–327.

Tarita-Nistor L, González EG, Markowitz SN, Steinbach MJ (2006) Binocular interactions in patients with age-related macular degeneration: acuity summation and rivalry. Vision Res 46:2487–2498.

Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) Binocular rivalry and visual awareness in human extrastriate cortex. Neuron 21:753–759.

Trainor LJ, Trehub SE (1989) Aging and auditory temporal sequencing: ordering the elements of repeating tone patterns. Percept Psychophys 45:417–426.

Ukai K, Ando H, Kuze J (2003) Binocular rivalry alternation rate declines with age. Percept Mot Skills 97:393–397.

van Loon AM, Knapen T, Scholte HS, St. John-Saaltink E, Donner TH, Lamme VAF (2013) GABA shapes the dynamics of bistable perception. Curr Biol 23:823–827.

Wilson HR (2003) Computational evidence for a rivalry hierarchy in vision. Proc Natl Acad Sci USA 100:14499–14503.

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 13:532–540.

Working Group on Speech Understanding and Aging (1988) Speech Understanding and Aging. J Acoust Soc Am 83:859–895.

Wunderlich K, Schneider KA, Kastner S (2005) Neural correlates of binocular rivalry in the human lateral geniculate nucleus. Nat Neurosci 8:1595–1602.