

# A cortical network that marks the moment when conscious representations are updated



Elisabeth Stöttinger<sup>a,\*</sup>, Alex Filipowicz<sup>a</sup>, Derick Valadao<sup>a</sup>, Jody C. Culham<sup>b</sup>,  
Melvyn A. Goodale<sup>b</sup>, Britt Anderson<sup>a,c,1</sup>, James Danckert<sup>a,1</sup>

<sup>a</sup> University of Waterloo, Department of Psychology, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1

<sup>b</sup> The Brain and Mind Institute, The Natural Sciences Centre, Western University, London, Ontario, Canada, N6A 5B7

<sup>c</sup> Centre for Theoretical Neuroscience, University of Waterloo, Canada

## ARTICLE INFO

### Article history:

Received 10 June 2015

Received in revised form

15 October 2015

Accepted 30 October 2015

Available online 31 October 2015

## ABSTRACT

In order to survive in a complex, noisy and constantly changing environment we need to categorize the world (e.g., Is this food edible or poisonous?) and we need to update our interpretations when things change. How does our brain update when object categories change from one to the next? We investigated the neural correlates associated with this updating process. We used event-related fMRI while people viewed a sequence of images that morphed from one object (e.g., a plane) to another (e.g., a shark). All participants were naïve as to the identity of the second object. The point at which participants ‘saw’ the second object was unpredictable and uncontaminated by any dramatic or salient change to the images themselves. The moment when subjective perceptual representations changed activated a circumscribed network including the anterior insula, medial and inferior frontal regions and inferior parietal cortex. In a setting where neither the timing nor nature of the visual transition was predictable, this restricted cortical network signals the time of updating a perceptual representation. The anterior insula and mid-frontal regions (including the ACC) were activated not only at the actual time when change was reported, but also immediately before, suggesting that these areas are also involved in processing alternative options after a mismatch has been detected.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Updating mental representations is a fundamental human skill that enables optimal decision making in a noisy and constantly changing environment. Updating can happen in response to surprising, unexpected events (e.g., a bolt of lightning indicates one should head quickly inside), or as a result of accumulating evidence in a noisy, uncertain environment (e.g., a gradual accumulation of dark clouds suggests that it is time to head indoors; McGuire et al., 2014). There is a considerable amount of research exploring the brain networks involved in updating in response to abrupt, transient sensory events (Craig, 2009; Menon and Uddin, 2010), as well as updating of multi stable perceptions that occur without any physical changes (e.g., ambiguous figures; binocular

rivalry; Long and Toppino, 2004; Lumer et al., 1998; Tong et al., 1998). However updating in response to gradually changing events also occurs commonly in everyday life but has been little studied. For example, one might monitor the colour of a steak on a barbeque to determine when it is cooked to perfection or the facial expression of a partner to determine if the mood has shifted. These decisions may be categorical (Liberman et al., 1957; Lieberman et al., 2005; Hartendorp et al., 2010), and can be critical for social interaction and even survival (Grinband et al., 2006; Calder et al., 1996; Etcoff and Magee 1992; McCullough and Emmorey, 2009). But while the answers can be categorical, the stimuli they depend on vary continuously (e.g., the colour of the meat; slight shifts in the curve of a lip). How is it that small changes in perceptual stimuli can lead to abrupt, categorical shifts in experience? How does our brain update when object categories change from one to the next?

Understanding how conscious interpretations of our environment change from moment to moment has been a long-standing preoccupation of psychologists, philosophers, and neuroscientists. Since the 1800s, reversible, ambiguous or bistable figures like the Rubin's face/vase picture or Necker cube have been used to study how people spontaneously alternate conscious percepts between

\* Correspondence to: Department of Psychology, 200 University Avenue West, Waterloo, ON, Canada, N2L 3G1.

E-mail addresses: [elisabeth.stoettinger@gmail.com](mailto:elisabeth.stoettinger@gmail.com) (E. Stöttinger), [alsfilip@uwaterloo.ca](mailto:alsfilip@uwaterloo.ca) (A. Filipowicz), [dvaladao@uwaterloo.ca](mailto:dvaladao@uwaterloo.ca) (D. Valadao), [jculham@uwo.ca](mailto:jculham@uwo.ca) (J.C. Culham), [mgoodale@uwo.ca](mailto:mgoodale@uwo.ca) (M.A. Goodale), [britt@uwaterloo.ca](mailto:britt@uwaterloo.ca) (B. Anderson), [jdanker@uwaterloo.ca](mailto:jdanker@uwaterloo.ca) (J. Danckert).

<sup>1</sup> These two authors contributed equally to the work.

two mutually exclusive interpretations of objectively stable pictures (Long and Toppino, 2004). Similarly, in the phenomenon known as binocular rivalry, people spontaneously alternate their conscious interpretations between two monocular views presented separately to each eye instead of reporting the two images as being superimposed (Lumer et al., 1998). Studies using fMRI, TMS and EEG have demonstrated that frontal and parietal areas—especially within the right hemisphere—are involved when participants switch between the two conflicting percepts (Britz et al., 2009; Kleinschmidt et al., 1998; Lumer et al., 1998; Zaretskaya et al., 2010). Research on brain networks involved in updating in response to abrupt, transient sensory events have demonstrated activation in the anterior insula and anterior cingulate cortex (ACC; Craig, 2009; Menon and Uddin, 2010 for a review), as well as activation in the temporal parietal junction (TPJ). Although these studies provide a hint as to which brain areas are involved in updating conscious interpretations of the world, processes in those paradigms are not typical of the stimuli that provoke updating when things in our environment change gradually. Our world is in constant flux. A shift in perception in this context is usually induced by an accumulation of mismatching evidence.

In a paradigm known as perceptual decision making participants have to gather sensory information in a noisy, impoverished or perceptually degraded context (e.g., noise in an image is slowly reduced, thereby gradually revealing the object). These studies typically report increased activity in the anterior insula and anterior cingulate cortex at the moment of recognition (Ploran et al., 2007; Heekeren et al., 2008 for a review). Although perceptual decision making does require participants to accumulate changing sensory information, participants in this task do not have to abandon an old representation in favour of a new one. Instead, they gradually build a representation from scratch. Relatively few studies have used slowly evolving and ambiguous sensory signals between two distinct interpretations (e.g., Thielscher and Pessoa, 2007; Grinbald et al., 2006), and those studies that have used such stimuli often used a small set of stimuli that allowed participants to anticipate the timing and nature of the perceptual alternations. To our knowledge no study has examined the neural correlates of perceptual updating in circumstances that most closely approximate its natural occurrence; that is when participants have to (1) abandon an established representation, (2) do so when the change is not announced by marked or abrupt changes in the stimulus, and (3) when participants do not know what new state to expect.

We have developed a quick and easy task to measure updating conscious interpretations in response to continuously changing stimuli: in different picture sets a commonly known object (e.g., rabbit) morphed over 15 iterations into a completely different object (e.g., duck) (Stöttinger et al., 2014, 2015). Similar to the gradual changes in real world environments, the images in this picture morphing task change gradually. Despite equivalent changes between each transition, at one point observers suddenly shift from perceiving one item to perceiving another. Using this task we demonstrated that damage to the right hemisphere selectively and differentially impairs updating: compared to patients with left brain damage and healthy controls, patients with right brain damage (RBD) needed to view significantly more pictures in the morphing sequence before successfully identifying the second image. Lesion overlay analysis implicated that the anterior insula is involved in updating (see Danckert et al., 2012 for the same finding, using a different updating task). This impairment in RBD patients did not represent an inability to detect change *per se*. All participants typically reported perceived differences between the pictures. RBD patients however tended to interpret these changes in favor of the first object (e.g., it is a rabbit, but the ears have changed; now the rabbit is looking up, etc.) before they eventually

shifted their conscious interpretation towards a new object (e.g., oh, it is a duck). In a similar vein, Vocat et al. (2013) showed that patients with anosognosia for hemiplegia (a syndrome in which patients deny their left-sided paralysis after right hemisphere brain damage), tend to persist reporting their initial beliefs even in the face of highly incongruent cues.

We hypothesised that focal brain injury to the right hemisphere—presumably the right anterior insula—interfered with a network important for representational updating; a notion supported by findings from fMRI studies on perceptual decision making, updating in response to salient events, ambiguous figures and binocular rivalry (Britz et al., 2009; Craig, 2009; Heekeren et al., 2008; Kleinschmidt et al., 1998; Lumer et al., 1998; Menon and Uddin, 2010; Ploran et al., 2007; Zaretskaya et al., 2010). Lesions in stroke patients however reflect vascular anatomy and not functional networks. The common involvement of the insula in brain damaged patients might partially reflect the fact that the insula is commonly injured in strokes due to its immediate vicinity to the middle cerebral artery (Fink et al., 2005).

Hence we employed functional neuroimaging (fMRI) in healthy individuals to evaluate whether the insula and/or other structures are involved in updating. In the present study, participants viewed sets of images (Stöttinger et al., 2015) that morphed from one unambiguous object to another (Fig. 1). Participants were asked to indicate via button press whether they saw the first or another object. Using the participant's own report allowed us to pinpoint the moment of a shift in conscious interpretation for each individual and each picture set. We could do this despite all participants being naïve to the identity of the first and second objects, and where subtle changes of visual features provoked categorical shifts in perception. As participants could not predict *when* they would 'see' a shift, or the nature of the shift, our activations cannot reflect anticipations. Our data show that a network of structures including the dorsal-medial frontal cortex, inferior parietal lobes, and anterior insula is activated when participants indicate a change in their subjective perception of the object. Further, activity in the anterior insula and mid-frontal regions including the ACC, medial (BA 9) and superior frontal (BA 6, BA8) areas was seen seconds *before* participants reported this categorical shift. The network others have found to be involved in shifts in conscious interpretations might be best understood more generally as a network for updating mental representations, signalled either by salient events, accumulation of bottom-up sensory events, or internal signals.

## 2. Methods

### 2.1. Participants

A total of seventeen (6 male) neurologically healthy participants with normal or corrected to normal vision took part in this study (mean age 27.65 years, SD=8.01). Two participants reported being left handed; the remaining fifteen were right-hand dominant. The individual activation patterns of the two left handed participants did not reveal any deviant pattern of activation compared to the main group. Given that left-handed people represent a portion of the population we decided not to exclude those individuals from the sample (see Willems et al., 2014 for that argument). None of the participants had a history of brain injury. Participants received \$30 for their participation. All participants provided informed consent prior to participation. The research protocol was approved by the Office of Research Ethics at the University of Waterloo and the Tri-Hospital Research Ethics Board of the Region of Waterloo in Ontario, Canada.

a

**Stimuli**

Animate - Inanimate



Inanimate - Animate



Animate - Animate



Inanimate - Inanimate



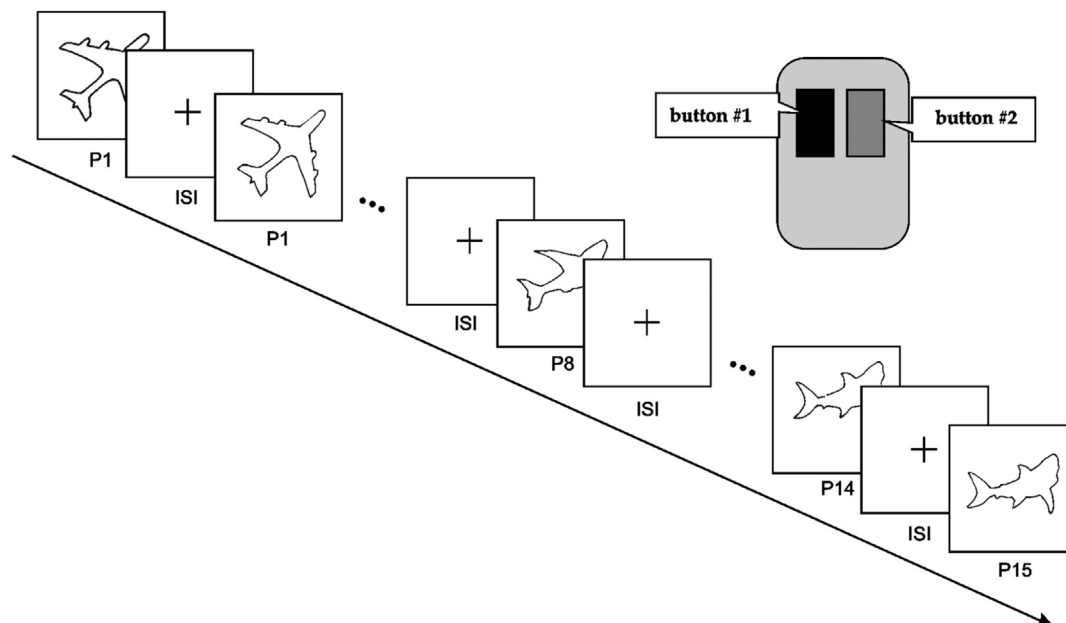
Early abrupt shift



Late abrupt shift



b

**Task**

**Fig. 1.** Objects morphed over 15 iterations (a), preceded and followed by inter stimulus intervals (ISI) of 2, 4, 6 or 8 s (b). Participants pressed one button for 'seeing object 1' and another for 'seeing object 2'. Two additional control sets were used, that either induced a late or an early abrupt shift (a).

## 2.2. The task

### 2.2.1. Stimuli

Twenty picture sets were used in which line drawings of common objects morphed over 15 iterations into a different object (Fig. 1a). Picture sets were selected from a larger set validated in an earlier study (Stöttinger et al., 2015). Participants reported verbally what they saw by typing in the name of the object. This reporting method guaranteed that participants could freely indicate what they saw without being restricted by categories (e.g., forced choice). Picture sets for the current study were selected with the criterion that all sets were perceived categorically in the normative study (i.e., that the majority of participants in the normative study categorized each picture they saw as either the first object or second object in the set; Stöttinger et al., 2015). The average proportion of first object reports for all twenty selected picture sets can be found in Fig. 3a (left panel). Selecting only the categorically perceived image sets guaranteed that pictures in the middle of the morphing sequence were not simply 'noisier' than pictures at the beginning or end. In other words, the ambiguous images were still easily categorised by participants as either object 1 or object 2. All images were a standard size (315–316 × 316 pixels) and were displayed on a white background. In ten picture sets, objects morphed within the same category (5 morphed from one animate object to another animate object, 5 morphed from one inanimate object to another inanimate object). In ten picture sets, objects changed class (from animate to inanimate or vice versa). In addition, participants were presented with two kinds of control picture sets to ensure that they were responding to changes in the pictures in the set rather than simply to the position in the set (e.g., always switching after the 8th picture). In these control picture sets, a salient deviating picture was presented either after three pictures or after thirteen pictures resulting in an early or late abrupt shift. Those sets served as controls and were not analysed further (Fig. 1a).

### 2.2.2. Behavioural paradigm

The picture morphing task consisted of five experimental runs. Each run included four picture sets of each kind of morphing sequence (animate–animate, animate–inanimate, inanimate–animate, inanimate–inanimate) together with one control picture set inducing either an early or late abrupt shift between the first and second object (Fig. 1a). We randomized the order in which the picture sets were presented in each run and kept this randomized order consistent across participants. Runs were counterbalanced between participants using a Latin square. Picture morphing in each picture set occurred over fifteen discrete steps, each corresponding with the acquisition of a whole-brain image. Each picture within a set was presented for two seconds. Pictures were randomly intermixed with eight inter-stimulus-interval periods (2, 4, 6 or 8 s) during which participants saw a fixation cross (Fig. 1b). Participants provided their responses in the scanner using two buttons on a four button Cedrus fibre optic system. In a two-alternative forced-choice task participants were asked to press the first button when they 'saw the first object' and the second button when they 'saw the second object'. All participants were ignorant as to the identity of the second object in each picture set. Although speed was not emphasised, participants were encouraged to make a button press within the 2-s time window during which the picture was presented on the screen. At the end of each set of 15 images the word END was presented for 2 s to indicate that the next picture set would begin shortly. Infrequently, a participant failed to press any button (i.e., in 0.8% of individual presentations of an image). Most of the time omissions were preceded and followed by the same button press, suggesting no change in the participant's conscious percept. In three cases the omission

occurred between a switch in button presses. In those cases the moment of change was assigned to the first occasion when the second button was pressed.

To familiarise participants with the procedure and timing of the task, they were trained with five different picture sets (none of which were used in the actual scanning period) a few days prior to scanning. Instructions were repeated before each run of the actual experiment.

## 2.3. fMRI data collection

Functional data were acquired using gradient echo-planar T2\*-weighted images collected on a 1.5 T Phillips scanner located at Grand River Hospital in Waterloo, Ontario (TR=2000 ms; TE=40 ms; slice thickness=5 mm with no gap; 26 slices/volume; FOV=220 × 220 mm<sup>2</sup>; voxel size=2.75 × 2.75 × 5 mm<sup>3</sup>; flip angle=90°). Each experimental run consisted of 26 slices per volume and 285 volumes. At the beginning of each session, a whole brain T1-weighted anatomical image was collected for each participant (TR=7.4 ms; TE=3.4 ms; voxel size=1 × 1 × 1 mm<sup>3</sup>; FOV=240 × 240 mm<sup>2</sup>; 150 slices with no gap; flip angle=8°). The experimental protocol was programmed using E-Prime experimental presentation software (v1.1 SP3; Psychology Software Tools, Pittsburgh, PA). Stimuli were presented on an Avotec Silent Vision™ fibre-optic presentation system using binocular projection glasses (Model SV-7021). The onset of each trial was synchronized with the onset of data collection for the appropriate functional volume using trigger pulses from the scanner.

## 2.4. fMRI data analysis

Functional data were analysed using Brain Voyager QX software (v2.8, Brain Innovation B.V., Maastricht, Netherlands). Preprocessing included slice scan-time correction, 3D motion correction, linear trend removal and temporal high-pass filtering. No spatial smoothing was applied to the data. All functional runs were corrected for motion artefacts using a sinc interpolation. All low-resolution functional data for each subject and each run were aligned to their respective high-resolution anatomical image and transformed into Talairach space (Talairach and Tournoux, 1988). Predictors were created for each of the 15 images and were labelled as 'stable' for the first and last three images (i.e., periods in which the reported object seen never changed), 'change' (i.e., the image at which the subject indicated for the first time that they now saw object 2), 'immediately pre-change' (i.e., the image just prior to indicating a perceived change) and 'post-change' (i.e., the image just after a change was indicated). All predictors were convolved using the two gamma hemodynamic response function. Whole-brain, random-effects group analyses for the picture morphing task was conducted with contrast *t* maps thresholded at a false discovery rate (FDR)=0.05. Brain Voyager's cluster-level threshold plug-in was used to define the minimum significant cluster size threshold (Monte Carlo simulation=1000 times). Data were analysed in two steps using a random effects general linear model (GLM). In the first step a balanced contrast between two different predictors was used: change (defined as the picture when button presses changed from 1 to 2 and the immediately preceding and succeeding pictures) vs. stable (the three pictures at the start and the end of each set [−P1 −P2 −P3 +P(change-1) +P(change) +P(change+1) −P13 −P14 −P15; Fig. 3b]). Note that stable pictures at the beginning and end were collapsed. Due to counterbalancing, the picture sets both stable periods contained the exact same amount of animate and inanimate objects. A contrast between the three pictures at the beginning (P1-3) and the three pictures at the end (P13-15) of each series did not reveal any significant activations at FDR=0.05.



In the second step, each of the change periods (the picture immediately before the change [−P1 −P2 −P3 +P(change−1)−P13 −P14 −P15], the picture at the change [−P1 −P2 −P3 +P(change)−P13 −P14 −P15], as well as the picture immediately after the change [−P1 −P2 −P3 +P(change+1)−P13 −P14 −P15]) were compared separately against the stable pictures in a balanced contrast (Fig. 3b).

### 3. Results

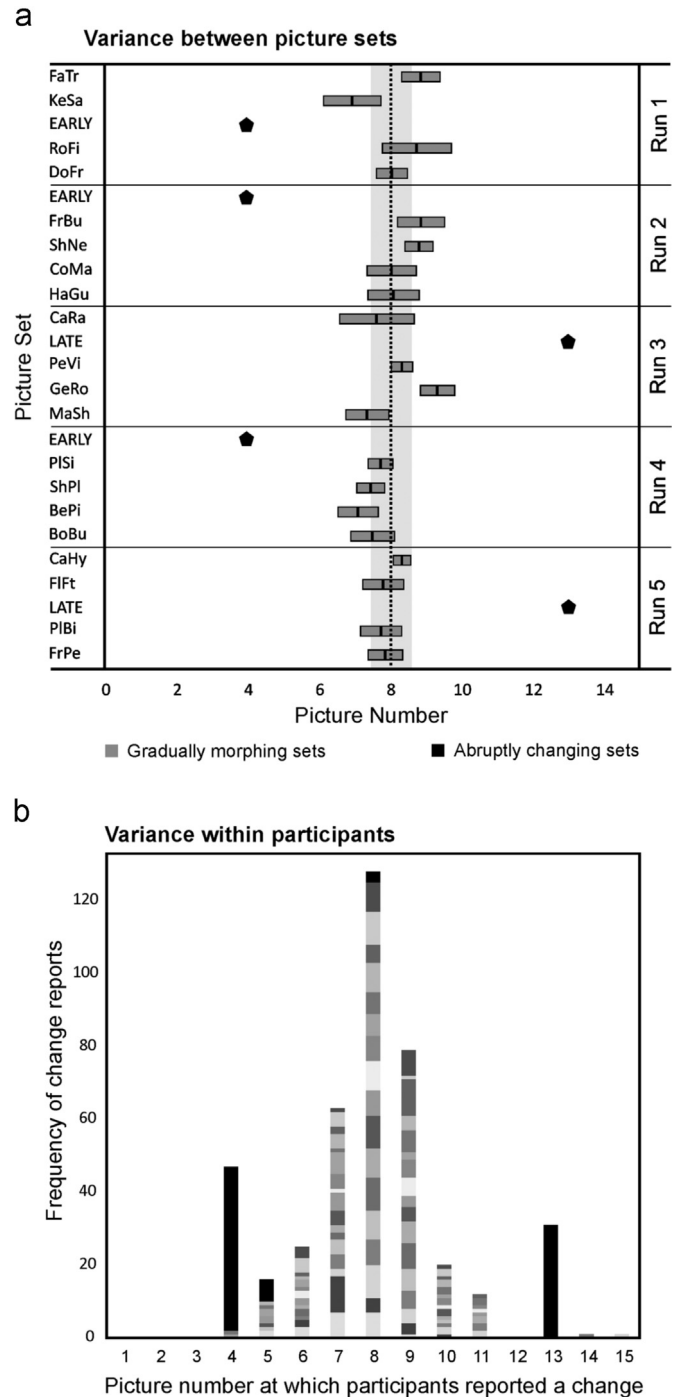
#### 3.1. Behavioural results

Although participants on average reported a change from one object to the next between the 7<sup>th</sup> and 9<sup>th</sup> pictures in the sequence (Mean=8.09, SD=1.50) the number at which participants reported a change varied widely between picture sets and within each participant (Fig. 2; Average range=5.41; SD=1.77; Min=5.53, SD=0.80; Max=10.94, SD=1.75). Overall percentage of first object reports were collapsed for all 20 picture sets and all 17 participants (Fig. 3a, right panel, black line) and compared against the normative data (Fig. 3a, left panel, black line; Stöttinger et al., 2015). A repeated measure analysis for the between subject factor experiment (normative study vs. current study) and the within subject factor picture number (1–15) did not find a significant difference between the proportion of first object reports in our study and the normative data. There was neither a significant main effect for experiment [ $F(1,38)=1.81$ ,  $p=0.19$ ,  $\eta^2=.05$ ] nor a significant interaction between image number and experiment [ $F(14,532)=.73$ ,  $p=0.74$ ,  $\eta^2=0.02$ ]. In other words, the difference in reporting method (verbal vs. forced choice button press) did not influence response patterns.

#### 3.2. Imaging results

We contrasted brain activity when perceptual reports changed (defined as the picture when button presses changed from button 1 to button 2 and the immediately preceding and succeeding pictures), with activations for stable periods (three pictures at the start and end of each set). When perceptual reports changed, higher activations were observed in a network of brain regions including the anterior insula, dorsal medial frontal cortex (including the anterior cingulate (ACC), medial (BA9) and superior frontal (BA6, BA8) areas), inferior frontal, and inferior parietal cortex (Table 1). This contrast did not simply reflect the fact that participants were searching more intensely for a change in the middle images compared to the beginning of each series: A contrast between the stable pictures (first and last three pictures) vs. stable pictures closer to the middle of the sequence (i.e., for which participants did not report a change) did not reveal any significant activations at an FDR=0.05. Furthermore, the contrast change vs. stable images (middle pictures) revealed similar activations as seen in our original contrast of change vs. stable images (first and last three pictures). Higher activations for change images compared to stable pictures were observed in the anterior insula, inferior frontal, and inferior parietal cortex bilaterally, as well as the dorsal medial frontal cortex including the anterior cingulate (ACC).

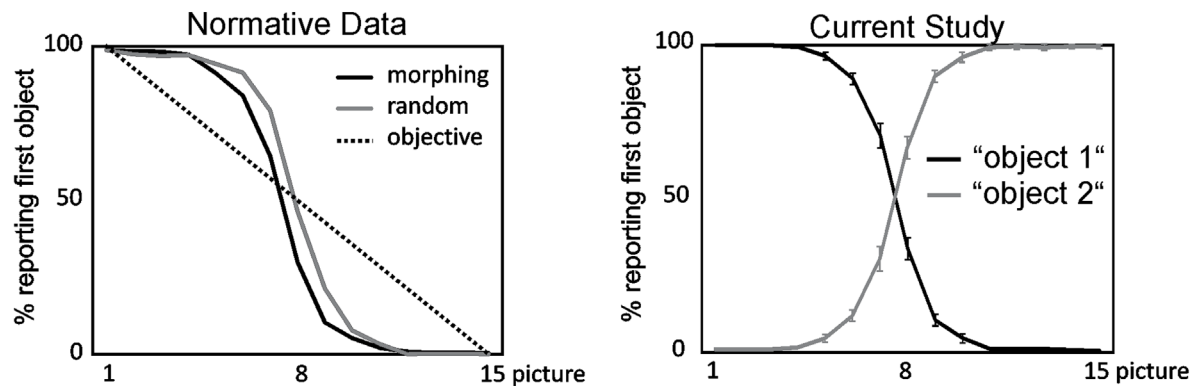
Immediately before participants reported a change (on average 5 s prior to button 2 presses) increased activation was observed in the anterior insula and mid-frontal cortex including the ACC, medial (BA 9) and superior frontal (BA 6, BA8) areas (Fig. 3c, panel (a), Table 2a)—indicating that those areas were not only active at the time when participants reported a change (Ploran et al., 2007, Thielscher and Pessoa, 2007) but were also active several seconds



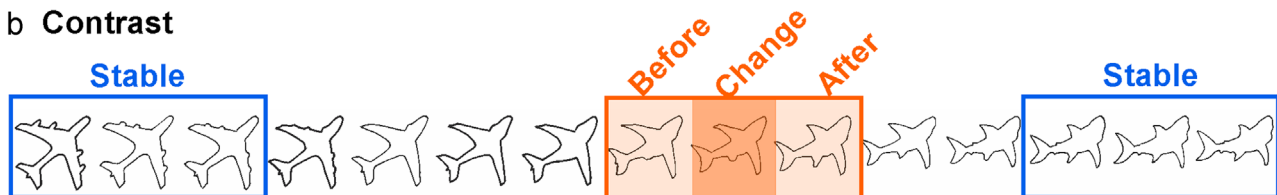
**Fig. 2.** Variance between picture-sets (a) and participants (b). (a) displays the mean ( $\pm$  95% confidence interval) for each of the picture sets and each of the five runs as well as the mean (dotted line) and 95% confidence interval (vertical grey bar) for all picture sets and all participants. Picture sets: FaTr=Face-Tree; KeSa=Key-Saw. RoFi=Rocket-Fish; DoFr=Dog-Frog; FrBu=Frog-Buddha; Sh-Ne=Shark-Syringe; CoMa=Coffee Maker-Man; HaGu=Hair dryer-Gun; CaRa=Cat-Rabbit; Pe-Vi=Pear-Violin; GeRo=Gecko-Rose; MaSh=Mallet-Shark; PISI=Pliers-Scissors; ShPi=Shark-Plane; BePi=Bear-Pig; BoBu=Bowtie-Butterfly; CaHy=Candelabra-Hay fork; FIFt=Fly-Fir tree; PIBi=Plane-Bird; FrPe=Frog-Person. The black polygons indicate the picture number at which either an early or late shift was induced. The histogram in (b) plots the frequency of change reports for each picture within the sequence. Data from each participant are rendered in a different grey scale and stacked. In black are reports for early and late control sets. In total, changes were most frequently reported for the eighth picture. Note, however, that each participant reported changes across a range of picture numbers and the modal picture differed between participants.

## Results

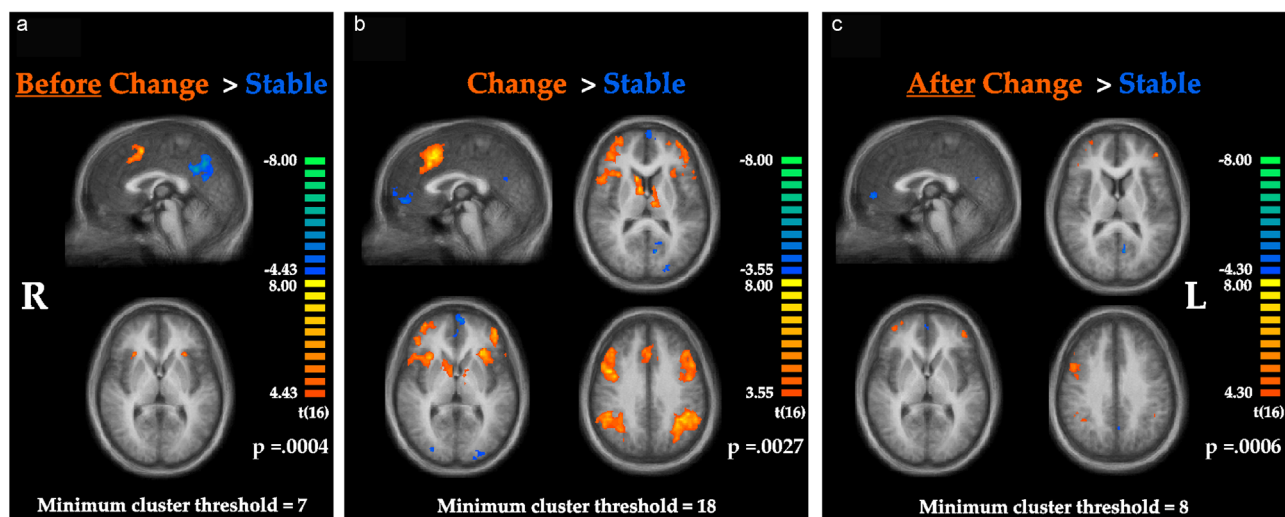
### a Behavioral Results



### b Contrast



### Imaging result



**Fig. 3.** The picture sets for the current experiment were selected from a normative sample (Stöttinger et al., 2015) in which picture sets were either presented in a sequential order as in our fMRI study (morphing gradually from object one to object two) or individually—essentially random non-morphing sequences. (a) (left panel) displays the average proportion of first object reports (y-axis) for morphing (black line) and random (grey line) presentation orders for each of the 15 pictures (x-axis) of the 20 selected picture sets (which represent a subset of those tested in our normative study). The dotted line displays the actual proportion for each picture ranging from 100% object 1/ 0% object 2 for the first image, to a middle (50% object 1/ 50% object 2) and final image (0% object 1; 100% object 2). Regardless of whether pictures were presented within the morphing context or randomly, participants only reported seeing either the first or second object. Hence, the intermediate images were not “noisy” or unclassifiable. Results (Fig a; right panel) show mean ( $\pm$  SE) percentage of responses across all picture sets and participants for the current study; initial-object reports=black line, second-object reports=grey line. (b) Each of the change periods (the picture immediately before the change, the picture at the change, as well as the picture immediately after the change) were compared separately against the stable pictures in a balanced contrast. (c) Whole-brain, random-effects group analyses conducted with contrast  $t$  maps thresholded at an FDR of=0.05. Stable pictures (the first and last three pictures of each series) were contrasted with the picture (a) before (b) at, and (c) after the change (b). Brain Voyager’s cluster-level threshold plug-in was used to define the minimum significant cluster size threshold.

prior to that decision (Fig. 3c, panel (b); Table 2b). After a change was reported, activity in the anterior insula and mid-frontal regions was not different between the change and stable pictures (Fig. 3c, panel (c); Table 2c).

Beta weights were extracted using a region-of-interest general

linear model (ROI-GLM) for the anterior insula bilaterally and the mid-frontal area identified in the omnibus contrast: change vs. stable (Table 1). For this analysis 13 different time points were used: the time point of change (T0) plus 6 pictures before (–T6, –T5, –T4, –T3, –T2, –T1) and after participants reported a change (T1,

**Table 1**  
Neural correlates of the change period.

	Size	Talairach coordinates			Peak value	
Regions: Change > Stable	# of voxels	x	y	z	t	p
left caudate	915	−10	1	15	6.034112	0.000017
left cerebellum	2416	−22	−68	−30	6.432366	0.000008
left fusiform	1045	−49	−50	−18	6.225165	0.000012
left inferior parietal	8874	−31	−53	33	7.663129	0.000001
left insula	2304	−28	22	3	10.51637	< 0.000001
left frontal (middle frontal gyrus)	1222	−28	49	18	6.23856	0.000012
left frontal (precentral gyrus)	6690	−46	−2	33	8.654748	< 0.000001
medial frontal	6638	2	19	48	11.19495	< 0.000001
right caudate	959	5	1	6	6.149346	0.000014
right inferior frontal operculum/ insula	18078	50	4	21	12.32694	< 0.000001
right fusiform	1323	50	−50	−15	5.431866	0.000055
right inferior parietal	9368	41	−44	42	9.181044	< 0.000001
right superior frontal	581	14	10	63	6.157094	0.000014
Regions: Stable > Change						
Left superior temporal gyrus	790	−49	−32	18	−5.5557	0.000043
Left mid temporal gyrus	1279	−52	−74	21	−5.67861	0.000034
Left occipital	858	−19	−86	9	−6.96545	0.000003
Precuneus	13406	8	−50	3	−7.12144	0.000002
Prefrontal	1795	−4	61	9	−6.08971	0.000016
Right superior temporal	908	50	−29	12	−5.01019	0.000128

**Table 1:** Brain activity when perceptual reports changed (defined as the picture when button presses changed from 1 to 2 and the immediately preceding and succeeding periods), was contrasted with activations for stable periods (the three pictures at the start and end of each set). The opposing contrast (Stable > Change) is presented in the lower portion of the table.

T2, T3, T4, T5, T6). Fig. 4 displays the beta weights (after correction of serial correlations using AR(2) model) for the three regions. Activations peak at the actual change point. Activations, however, start to ramp up before a change has been reported. This finding most likely indicates that those areas are involved in the processing of alternative options once a change had been detected.

#### 4. Discussion

Our external environment is complex, noisy and constantly changing. By necessity we have to categorize events in the world. Just as importantly, we need to update our representations when things change. This mechanism is crucial for survival (e.g., deciding whether a piece of meat is safe to eat) and for social interactions (e.g., interpreting facial expressions; Grinband et al., 2006; Calder et al., 1996; Etcoff and Magee 1992; McCullough and Emmorey, 2009). Although many studies have investigated the neural correlates involved in updating in response to salient, transient events (Craig, 2009; Menon and Uddin, 2010), less is known about how our brain updates in response to gradually changing events. Here we used our gradually morphing picture task (Stöttinger et al., 2015) to investigate such a mechanism.

Similar to gradually changing scenarios in the real world, in our picture morphing task each image is only slightly altered from the one before and the interval changes are equivalent in magnitude across the series. Despite equivalent changes at each transition, at some point people categorically shift from perceiving one item to perceiving another. We looked at the neural correlates of these categorical shifts in perceptual experience. The neural correlates associated with such shifts in conscious interpretations occur when neither the timing of the shift nor the outcome of the shift can be anticipated.

The cortical areas active at the time of these categorical shifts in perception were most pronounced in the anterior insula, the medial and inferior frontal regions and the inferior parietal cortex. The anterior insula and a mid-frontal area were not only activated

at the actual time when change was reported, but also immediately before (~ 5 s; Fig. 3c).

Although we cannot rule out the possibility that the activations observed here reflect an unconscious mental process initiated long before a deliberate, conscious decision has been made (Falk et al., 2010; Soon et al., 2008), a more plausible assumption is that the observed network is involved in processing alternative options after a mismatch has been detected. In order to update a mental representation in an uncertain, ambiguous, slowly-changing, and open-ended environment, one needs to detect mismatches between one's current model (e.g., it is a plane) and new incoming information (i.e., the next morphed picture in the sequence). One has to flexibly decide either to stay with the current model or to explore new, alternative options (e.g., maybe it is a bird, maybe a fish, maybe a seal, etc.; Collins and Koechlin, 2012). The assumption that the anterior insula is involved in processing alternative options fits well with the results of (Donoso et al., 2014) our previous patient studies showing that damage to the anterior insula—especially on the right<sup>2</sup>—can selectively and differentially disrupt updating. Patients in those studies were able to detect changes but were impaired when they had to explore alternative options once a change had been detected (Dancikert et al., 2012; Filipowicz et al., 2015; Sepahvand et al., 2014; Stöttinger, et al., 2014).

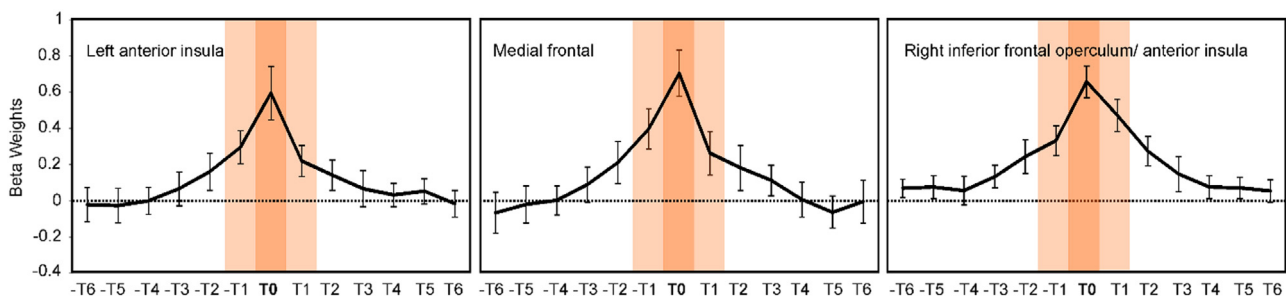
Activations in our study were not due to greater amounts of image change when switches were reported. The objective change in images—measured as a percentage of pixels changed from one image to the next—was constant across all series (~4% with no significant difference between the individual picture positions). Also, classifications for all pictures were consistent and categorical (Stöttinger et al., 2015; Fig. 3a). Hence, intermediate images were

<sup>2</sup> Using a separate subjects GLM the number of voxels (activated above a  $p < 0.001$  threshold) for the left (LH) and right (RH) insula were extracted and compared against each other. Although the anterior insula was activated bilaterally in our study, the number of activated voxels was slightly higher on the right than on the left [ $t(15) = 2.07, p = 0.056$ ]. Please note that due to technical difficulties the individual activations for one participant could not be extracted.

**Table 2**

Each of the change periods (the picture immediately before the change, the picture at the change, as well as the picture immediately after the change) were contrasted separately against the stable pictures.

A Neural correlates preceding the change image.						
Regions: Change > Stable	Size	Talairach coordinates			Peak value	
	# of voxels	x	y	z	T	p
left insula	208	−28	22	3	8.008277	0.000001
medial frontal	1847	−1	10	51	9.755181	< .0000001
right insula	460	32	16	9	8.470898	< .0000001
Regions: Stable > Change						
left mid temporal	277	−46	−77	15	−5.78503	0.000028
left post central	221	−61	−29	15	−6.44969	0.000008
right mid frontal	515	−22	25	45	−6.84217	0.000004
right mid temporal	317	41	−74	27	−6.22546	0.000012
precuneus	5920	−4	−44	39	−9.05185	< 0.0000001
b Neural correlates at the moment of change.						
left cerebellum	782	−43	−53	−39	6.188355	0.000013
left fusiform	492	−49	−50	−15	7.43193	0.000001
left inferior parietal	11632	−49	−38	39	9.286848	< 0.0000001
left insula	3631	−28	19	−3	7.872611	0.000001
left mid frontal gyrus	13574	−43	49	6	7.811516	0.000001
left thalamus and caudate	1457	−10	1	15	6.196401	0.000013
medial frontal	7655	2	19	51	9.118768	< .0000001
right caudate	2562	11	13	12	7.899613	0.000001
right frontal	2028	14	4	60	8.113521	< 0.0000001
right inferior parietal	11620	41	−47	42	10.14605	< 0.0000001
right frontal/insula	25692	44	7	36	9.290281	< 0.0000001
Regions: Stable > Change						
left cerebellum	487	−31	−65	−12	−6.20276	0.000013
left mid frontal BA10	1829	−1	52	−3	−6.24231	0.000012
left occipital	1596	−34	−80	−6	−6.00077	0.000018
left posterior cingulate	1061	−10	−68	12	−6.12292	0.000015
right occipital	1912	29	−83	3	−6.88266	0.000004
c Neural correlates at the picture immediately after the change.						
left cerebellum	824	−22	−68	−30	6.900126	0.000004
left cerebellum (medial)	681	−13	−77	−30	7.456213	0.000001
left fusiform	325	−49	−50	−18	7.774301	0.000001
left inferior parietal	290	−49	−41	36	5.574762	0.000042
left mid frontal gyrus	576	−43	43	12	6.354239	0.00001
right fusiform	369	44	−56	−12	6.414	0.000009
right inferior parietal	1926	41	−44	42	8.218021	0
right mid frontal gyrus	3021	44	13	39	7.293252	0.000002
right mid frontal gyrus (anterior)	415	29	58	6	7.171705	0.000002
Regions: Stable > Change						
precuneus	1461	−4	−65	30	−6.89842	0.000004
prefrontal	327	2	55	3	−6.8635	0.000004



**Fig. 4.** Beta weights (y-axis) at the time point of change (T0) plus the six pictures before and after the change displayed for the anterior insula (bilaterally) and the mid frontal area. Error bars represent 95-confidence interval. The vertical orange bars highlight the picture at which participants changed their reports (T0) plus the picture before (−T1) and after (T1) the change.

not simply more “noisy” or unclassifiable. In a similar vein, our results cannot be due to participants anticipating the identity of the second object. Participants were ignorant to the identity of all objects, and the second object was not predictable from the first. The results are also not likely to be due to anticipating the timing of the switch. Although participants on average switched between the 7th and 9th pictures in the sequence, the number at which

participants reported a change varied widely within each participant (typically ranging from reports at picture #6 to picture #11). In addition, each participant was exposed to one control series per run, inducing either an early (picture #4) or late shift (picture #13). In other words, there was no consistent ‘switch image’ built into the sequences and the images when people reported the switches varied widely within participant and across picture sets



(Fig. 2). Thus, the increased activations did not simply reflect a predictable time when switches occurred.

The pattern of activity we found in our study is similar to that seen during perceptual decision making. Ploran et al. (2007) used a task where pictures obscured by white noise were gradually revealed via gradual reductions in the amount of noise. They found a circumscribed network including the medial frontal cortex, anterior insula/frontal operculum, and thalamus associated with the perceptual decision at the moment of recognition. Thielscher and Pessoa. (2007) used a different task for assessing perceptual decision making in which faces morphed from a fearful to a disgusted expression (or vice versa). Despite the procedural differences they reported a similar network activated at the moment the perceptual decision was being made (i.e., anterior cingulate cortex, middle frontal gyrus, and inferior frontal gyrus/insula).

One interpretation of this prior work is that those areas reflect the engagement of additional attentional resources to maintain effective decision making under perceptually uncertain or impoverished conditions (Heekeren, 2008 for a review). Similarly, it has been argued that the anterior insula (together with the ACC) is part of a salience network (Craig, 2009; Menon and Uddin, 2010; Uddin, 2015). This model of the anterior insula suggests that the region processes salient events in the service of bottom-up error detection to initiate attentional control. This results in a switch between large-scale networks by activating the central executive network (dorsolateral prefrontal cortex and posterior parietal cortex) and deactivating the default network (ventromedial prefrontal cortex and posterior cingulate cortex). The coupling with the ACC then facilitates rapid access to the motor system.

A salience account, however cannot explain our results. Our morphing images changed gradually, without local surprise; updating from one image to another was not driven by past expectations, but rather required abandoning them. Our results show that the same network that others have identified as a “salience” network (Craig, 2009; Menon and Uddin, 2010; Uddin, 2015) might be best understood more generally as a network for updating mental representations, signalled either by bottom-up salience, or internal signals.

The ability to update conscious representations as things change is crucial to survival, and may be compromised in neurological disorders including stroke, but also developmental disorders such as Autism Spectrum Disorder (ASD). ASD has also been associated with abnormal brain activity in the ACC and insula (Di Martino et al., 2009; Uddin, 2015 for a review). In addition, ASD patients are impaired in changing their conscious percept of objects in a gradually morphing picture task (Burnett and Jellema, 2013) similar to ours. Impairments typically associated with ASD, including a difficulty interpreting socially relevant input (Gervais et al., 2004; Kleinhans et al., 2008) and local perceptual biases (Happé, 1999) may reflect damage to networks critical for updating conscious representations.

Thus, the network (anterior insula–mid-frontal area) others have found to be involved in salience detection and perceptual decision making might be best understood as more generically involved in updating mental representations. A failure to update may explain some of the deficits seen in ASD and stroke patients. Being unable to cope with a noisy and ever changing environment represents a substantial barrier to successful rehabilitation. Consequently, understanding the neural networks involved represents a vital step towards developing improved rehabilitation protocols.

## Acknowledgements

Natural Sciences and Engineering Research Council ([http://www.nserc-crsng.gc.ca/index\\_eng.asp](http://www.nserc-crsng.gc.ca/index_eng.asp)) of Canada Discovery

(#261628-07); Canada Research Chair grants; Heart and Stroke Foundation of Ontario (<http://www.heartandstroke.on.ca>; #NA 6999 to J.D.); Canadian Institutes of Health Research (<http://www.cihr-irsc.gc.ca/e/193.html>; #219972 operating grant to J.D. and B. A.). The above-mentioned funding agencies had no role in the study design, data collection and analysis, decision to publish, or the preparation of the manuscript.

## References

- Britz, J., Landis, T., Michel, C.M., 2009. Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb. Cortex* 19 (1), 55–65.
- Burnett, H.G., Jellema, T., 2013. (Re-) conceptualisation in Asperger's syndrome and typical individuals with varying degrees of autistic-like traits. *J. Autism Dev. Disord.* 43, 211–223.
- Calder, A.J., Young, A.W., Perrett, D.I., Etcoff, N.L., Rowland, D., 1996. Categorical perception of morphed facial expressions. *Vis. Cogn.* 3, 81–118.
- Collins, A., Koechlin, E., 2012. Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLOS Biol.* 10 (3), e1001293.
- Craig, A.B., 2009. How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.
- Danckert, J., Stöttinger, E., Quehl, N., Anderson, B., 2012. Right hemisphere brain damage impairs strategy updating. *Cereb. Cortex* 22, 2745–2760.
- Di Martino, A., Ross, K., Uddin, L.Q., Sklar, A.B., Castellanos, F.X., Milham, M.P., 2009. Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. *Biol. Psychiatry* 65, 63–74.
- Donoso, M., Collins, A.G., Koechlin, E., 2014. Foundations of human reasoning in the prefrontal cortex. *Science* 344 (6191), 1481–1486.
- Etcoff, N.L., Magee, J.J., 1992. Categorical perception of facial expressions. *Cognition* 44, 227–240.
- Falk, E.B., Berkman, E.T., Mann, T., Harrison, B., Lieberman, M.D., 2010. Predicting persuasion-induced behavior change from the brain. *J. Neurosci.* 30, 8421–8424.
- Filipowicz, A., Stöttinger, E., Willms, A., Anderson, B., Danckert, J., 2015. Updating impairments following right brain damage: A failure of exploration? Poster presented at: 5th Annual Symposium on the Biology of Decision Making, Paris, France.
- Fink, J.N., Selim, M.H., Kumar, S., Voetsch, B., Fong, W.C., Caplan, L.R., 2005. Insular cortex infarction in acute middle cerebral artery territory stroke: predictor of stroke severity and vascular lesion. *Arch. Neurol.* 62 (7), 1081.
- Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthélémy, C., Brunelle, B., Samson, Y., Zilbovicius, M., 2004. Abnormal cortical voice processing in autism. *Nat. Neurosci.* 7, 801–802.
- Grinband, J., Hirsch, J., Ferrera, V.P., 2006. A neural representation of categorization uncertainty in the human brain. *Neuron* 49, 757–763.
- Happé, F., 1999. Autism: cognitive deficit or cognitive style? *Trends Cogn. Sci.* 3, 216–222.
- Hartendorp, M.O., Van der Stigchel, S., Burnett, H.G., Jellema, T., Eilers, P.H., Postma, A., 2010. Categorical perception of morphed objects using a free-naming experiment. *Vis. Cogn.* 18, 1320–1347.
- Heekeren, H.R., Marrett, S., Ungerleider, L.G., 2008. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479.
- Kleinhans, N.M., Richards, T., Sterling, L.M., Stegbauer, K.C., Mahurin, R., Johnson, L.C., Greenon, J., Dawson, G., Aylward, E., 2008. Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain* 131, 1000–1012.
- Kleinschmidt, A., Büchel, C., Zeki, S., Frackowiak, R.S.J., 1998. Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc. R. Soc. Lond. Ser. B: Biol. Sci.* 265 (1413), 2427–2433.
- Liberman, A.M., Harris, K.S., Hoffman, H.S., Griffith, B.C., 1957. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54, 358–368.
- Liebsenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T., Medler, D.A., 2005. Neural substrates of phonemic perception. *Cereb. Cortex* 15, 1621–1631.
- Long, G.M., Toppino, T.C., 2004. Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol. Bull.* 130 (5), 748.
- Lumer, E.D., Friston, K.J., Rees, G., 1998. Neural correlates of perceptual rivalry in the human brain. *Science* 280 (5371), 1930–1934.
- McCullough, S., Emmorey, K., 2009. Categorical perception of affective and linguistic facial expressions. *Cognition* 110, 208–221.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667.
- Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., Wheeler, M.E., 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J. Neurosci.* 27, 11912–11924.
- Sepahvand, N., Stöttinger, E., Danckert, J., Anderson, B., 2014. Sequential decisions: a computational comparison of observational and reinforcement accounts. *PLOS One* 9, e94308.
- Soon, C.S., Brass, M., Heinze, H.J., Haynes, J.D., 2008. Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.
- Stöttinger, E., Filipowicz, A., Marandi, E., Quehl, N., Danckert, J., Anderson, B., 2014.

- Statistical and perceptual updating: correlated impairments in right brain injury. *Exp. Brain Res.* 232, 1971–1987.
- Stöttinger, E., Sepahvand, N.M., Danckert, J., Anderson, B., 2015. Assessing perceptual change with an ambiguous figures task: normative data for 40 standard picture sets. *Behav. Res. Methods*, 1–22.
- Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging.
- Thielscher, A., Pessoa, L., 2007. Neural correlates of perceptual choice and decision making during fear–disgust discrimination. *J. Neurosci.* 27, 2908–2917.
- Tong, F., Nakayama, K., Vaughan, J.T., Kanwisher, N., 1998. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21, 753–759.
- Uddin, L.Q., 2015. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* 16, 55–61.
- Vocat, R., Saj, A., Vuilleumier, P., 2013. The riddle of anosognosia: does unawareness of hemiplegia involve a failure to update beliefs? *Cortex* 49, 1771–1781.
- Willems, R.M., Van der Haegen, L., Fisher, S.E., Francks, C., 2014. On the other hand: including left-handers in cognitive neuroscience and neurogenetics. *Nat. Rev. Neurosci.* 15, 193–201.
- Zaretskaya, N., Thielscher, A., Logothetis, N.K., Bartels, A., 2010. Disrupting parietal function prolongs dominance durations in binocular rivalry. *Curr. Biol.* 20 (23), 2106–2111.