

*****IN PRESS AT NEUROPSYCHOLOGIA*****

Exploring the Neural Substrates of Misinformation Processing

Andrew Gordon¹, Jonathan C.W. Brooks^{1,2}, Susanne Quadflieg¹,

Ullrich K. H. Ecker³ & Stephan Lewandowsky^{1,3}

Authors' Affiliations:

¹School of Experimental Psychology, University of Bristol, Bristol, UK.

²Clinical Research and Imaging Centre, University of Bristol, Bristol, UK

³School of Psychological Science, University of Western Australia, Perth, Australia.

Corresponding author:

Andrew Gordon (andrew.gordon@bristol.ac.uk)

School of Experimental Psychology, University of Bristol
5 Priory Road, Bristol BS8 1TU
Telephone: 01179546621

Word count abstract: 168

Word count manuscript (excluding references, tables, and figure captions): 5350

1: Abstract

It is well known that information that is initially thought to be correct but then revealed to be false, often continues to influence human judgement and decision making despite people being aware of the retraction. Yet little research has examined the underlying neural substrates of this phenomenon, which is known as the ‘continued influence effect of misinformation’ (CIEM). It remains unclear how the human brain processes critical information that retracts prior claims. To address this question in further detail, 26 healthy adults underwent functional magnetic resonance imaging (fMRI) while listening to brief narratives which either involved a retraction of prior information or not. Following each narrative, subjects’ comprehension of the narrative, including their inclination to rely on retracted information, was probed. As expected, it was found that retracted information continued to affect participants’ narrative-related reasoning. In addition, the fMRI data indicated that the continued influence of retracted information may be due to a breakdown of narrative-level integration and coherence-building mechanisms implemented by the precuneus and posterior cingulate gyrus.

Keywords: continued influence effect; misinformation; memory updating

2: Introduction

The persistence of erroneous beliefs appears ubiquitous in modern societies (Lewandowsky et al., 2012). One need look no further than the purported link between a common childhood vaccine and autism which has no scientific basis but is widely accepted by the public (Horne et al., 2015), or the persistent belief in the existence of Weapons of Mass Destruction in Iraq before the invasion of 2003 even though none were ever found (Jacobson, 2010; Kull et al., 2003). Much misinformation now spreads online, and accordingly, the World Economic Forum recently listed digital misinformation as one of the main challenges to societies globally (WEF, 2013).

Decades of behavioral research indicate that discredited information can continue to influence people's beliefs and reasoning even after that information has been retracted—a phenomenon known as the continued influence effect of misinformation (CIEM) (Ecker et al., 2011; Ecker et al., 2010; Johnson & Seifert, 1994; Wilkes & Leatherbarrow, 1988; for reviews see Lewandowsky et al., 2012; Schwarz et al., 2016). The CIEM can occur even when people (a) remember the retraction (Johnson & Seifert, 1994; Marsh et al., 2003), and/or (b) receive prior warning about their exposure to false information (Ecker et al., 2010).

Attempts to retract misinformation in real world settings therefore often struggle to prove effective (Miton & Mercier, 2015; Nyhan & Reifler, 2015). Difficulties in creating effective interventions against misinformation may also partially reflect the fact that the cognitive mechanisms underlying the CIEM remain elusive. To date, two main explanations have been discussed in the literature. According to the selective-retrieval account, both the incorrect (i.e. the misinformation) and the correct (i.e., the retraction) information get frequently stored

in memory concurrently, and the CIEM arises when the former gets activated but is insufficiently suppressed (Catarino et al., 2015; Ecker et al., 2011; Jacoby & Whitehouse, 1989; Lewandowsky et al., 2012). By contrast, according to the mental-model-updating account, all information regarding a particular event is usually integrated in a unified mental model that requires updating whenever new relevant information becomes available. Retractions, however, necessarily threaten a model's internal coherence by invalidating some of its original content and may, therefore, be particularly poorly encoded and/or retrieved (Ecker et al., 2010; Johnson-Laird, 2012; Verschueren et al., 2005). In a nutshell, whereas the selective-retrieval account links the CIEM to a failure in memory selection and suppression, the model-updating account holds a failure in memory updating responsible for the effect.

Though both models postulate competing mental mechanisms, studying these mechanisms and their relative contribution to the CIEM via behavioural experiments has proven difficult in the past. A more promising avenue to differentiate between competing explanations of the effect may involve the identification of their underlying neural substrates. If there was indeed a systematic difference between the encoding of retracting and non-retracting information, this difference should be associated with observable changes in brain activity. Above all, the encoding of retracting information should violate a person's prior assumptions more strongly than the encoding of non-retracting information. Violation of assumption and detection of error has been shown to elicit stronger activity in the anterior cingulate cortex (ACC; Braver et al., 2001; Bush et al., 2000; Carter et al., 1998; Carter & van Veen, 2007). In consequence, higher activity in this region when processing information that contradicts prior claims (by retracting them) would be expected.

In addition, according to the model-updating account, the encoding of retracting information and its integration with existing material should be shallower than the encoding of non-retracting information. The presentation of information that invalidates earlier assertions of the narrative results in a 'gap' in the mental representation. The representation will therefore no longer be coherent unless the false assertion, the misinformation, is maintained (Lewandowsky et al., 2012). One way to achieve this coherence is to limit the processing of the retracting information. Shallower encoding of information has previously been linked to reduced activity in the left hippocampus (HC) and in the left inferior frontal gyrus (IFG; Otten et al., 2001; Wagner, 1998). Hence, traces of reduced activity in these regions during the encoding of retractions compared to non-retractions could provide further empirical evidence for a model-updating account of the CIEM.

Aside from these neural changes during the encoding of retracting and non-retracting information, differences in brain activity can also be expected during the retrieval of both types of information. Initial evidence indicates that the right dorsolateral prefrontal cortex, including the IFG and middle frontal gyrus (MFG), can exert systematic inhibition of hippocampal retrieval processes when subjects actively suppress memories (such as retracted information; Anderson et al., 2004; Anderson & Hanslmayr, 2014; Benoit & Anderson, 2012; Depue, 2012; Levy & Anderson, 2012). In addition, research on memory substitution has demonstrated that recalling a correct memory (such as a retraction) in the context of unwanted memories (such as initial misinformation) engages regions of the left ventrolateral prefrontal cortex that approximate to the pars opercularis and triangularis substructures of the left IFG (Benoit & Anderson, 2012; Depue, 2012; Nee & Jonides, 2008). In light of these findings, differences in activity in the right IFG, the right MFG, and/or the left IFG

during the retrieval of narratives with or without retractions promise to capture the role of memory suppression and/or substitution in the CIEM.

In summary, based on the above tentative explanations, tracking prefrontal as well as hippocampal activity during the encoding of retracting and non-retracting information as well as during the retrieval of retracted versus non-retracted information promises to enhance our understanding of the mechanisms that contribute to the CIEM. Following precedents, the current study asked participants to read a series of brief news reports and to complete a comprehension test about each report, while undergoing fMRI scanning. Importantly, across reports, we manipulated whether the same sentence towards the end of a report either retracted information that was provided earlier or not. We predicted this manipulation to elicit differences in neural activity during the encoding and/or the subsequent retrieval of the key sentence in brain regions previously linked to expectancy violations, memory suppression, and memory substitution (i.e., the ACC, HC, IFG, and MFG).

3: Method

3.1: Participants

We recruited 26 native English speakers, aged between 18-34 years ($M = 22.88$; $SD = 4.17$; 15 females, 11 males) through an advertisement on the University of Bristol's Psychology department website. Participants received a £20 cash stipend. Three further participants were tested but were excluded from the final analysis, one because of repeated and excessive head motion during scanning ($>3\text{mm}$), one because of non-completion due to illness, and a third due to a technical malfunction with the presentation software. All participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and reported

normal or corrected-to-normal vision and hearing. None had a history of neurological or neuropsychiatric disorder or was currently taking psychoactive medication. Written informed consent was obtained from all individuals. The study protocol was approved by the Human Research Ethics Committee of the University of Bristol's Faculty of Science.

3.2: Stimuli

Table 1.

Example stimulus materials: Across retraction and control conditions, news reports differed solely in sentence 2 (in bold italics). Depending on the content in sentence 2, the key sentence 5 (in bold), which was identical across conditions, did or did not act as a retraction.

1A: Example News Reports			
Event	Sentence	Control	Retraction
A	Sentence 1	A new office building lies in rubble today after it suddenly collapsed.	
	Sentence 2	<i>Onlookers reported that the collapse occurred around 11am this morning.</i>	<i>Onlookers reported an out of control fire on the lower floors was responsible.</i>
	Sentence 3	The building was only finished two months ago and opened shortly after.	
	Sentence 4	The surrounding area has been closed off to the public while the rubble is cleared.	
	Sentence 5	Fire crews at the scene stated that there was no evidence of a fire.	
	Sentence 6	The company that owned the building has been contacted for comment.	
B	Sentence 1	Flight AA9035 from Chicago to LA was evacuated before take-off yesterday.	
	Sentence 2	<i>Passengers said that the evacuation took place shortly after the plane had begun to taxi.</i>	<i>Passengers said that the evacuation was due to a fire in the engine bay.</i>
	Sentence 3	Passengers were asked to return to the terminal and await further news.	
	Sentence 4	The airline responsible for the flight were unavailable for interview.	
	Sentence 5	Fire crews at the scene stated that there was no evidence of a fire.	
	Sentence 6	The flight was eventually cancelled and passengers were recompensed.	
1B: Example Comprehension Probes			
A/B	Question 1	Based on this incident, measures designed to prevent fires should be reviewed.	
	Question 2	It was necessary for the fire services to be at the scene.	
	Question 3	It is likely that fire alarms would have been activated.	

Brief, fictional news reports similar to those previously used in behavioural studies on the CIEM served as stimuli (cf. Johnson & Seifert, 1994). All stimuli have been made available on the Open Science Framework (OSF - <https://osf.io/346pn/>). Each news report was six sentences long and began by introducing a news-worthy event (e.g., the collapse of a building or the evacuation of a plane). As shown in Table 1A, for each event two different report versions were created that differed only in the content of sentence 2. For example, in one version sentence 2 may have contained the (mis-)information that the building collapsed (or the plane was evacuated) due to a fire; whereas in another version sentence 2 may have contained only arbitrary contextual information of little relevance. Thus, depending on the content of sentence 2, the key sentence 5 - which was identical in both versions of the report - either did or did not act as a retraction of previously presented information. In this manner, retraction and control reports were created for 22 different events, resulting in 44 unique reports used in the study. Retraction and control reports were equivalent in terms of their mean lexical frequency [control reports: $M=2.07$, $SD=0.17$; retraction reports: $M=2.08$, $SD=0.19$; $t(21)=0.75$, $p=.461$], script concreteness [control reports: $M=410.84$, $SD=26.69$; retraction reports: $M=412.33$, $SD=28.87$; $t(21)=0.62$, $p=.545$], and Flesch reading ease [control reports: $M=60.89$, $SD=8.62$; retraction reports: $M=60.44$, $SD=9.76$; $t(21)=.60$, $p=.553$].

In order to measure reliance on misinformation, each report was followed by a set of three comprehension probes (see Table 1B). Participants were asked to rate each probe, using their memory of the report, on a 9-point scale from “strongly disagree” to “strongly agree”. Importantly, all memory-based comprehension probes referred to critical information in sentence 5 (i.e., to information that was present in both control and retraction reports). The probes were phrased in such a manner that

higher endorsement ratings could capture the degree to which participants incorporated misinformation from sentence 2 (if applicable) in their understanding of the news report. Additionally, reports for two separate events were paired such that they shared an identical sentence 5 and identical probe statements (see Table 1). This allowed us to counterbalance reports such that each participant received the identical sentence 5 and the identical probe statements in both retraction and control conditions across two different event scenarios. For example, half the participants were presented with Event A in the control condition and with Event B in the retraction condition; for the other half of participants, this assignment was reversed. Each participant received the same number of retraction and control reports, and report order was randomised across participants, with the constraint that two paired event reports were never presented in immediate succession.

3.3: fMRI Task and Procedure

To assess whether participants could hear the auditory presentation over the background noise of the scanner, each subject was presented with an example report during a short (30 s) functional imaging scan, and volume levels were adjusted if necessary. Presentation of stimuli was controlled and responses were recorded using Matlab (v 2012a) in conjunction with Psychtoolbox (v 3.0.8; Brainard, 1997); stimuli were presented via MRI-compatible headphones (Sensimetrics model S14). Participants listened to the 22 scripted news reports distributed across two functional runs of approximately 22 min each, with a short break in between. Participants were instructed to listen carefully to each news report and to subsequently respond to the comprehension statements as accurately as possible.

The beginning of each trial was signalled by an exclamation mark (!) on the screen for 1000 ms. The exclamation mark was then replaced by a fixation cross that

remained visible for the duration of the report, which was presented auditorily one sentence at a time. Pseudorandom uniform temporal jitter (ranging from 1 to 6 sec, with a mean of 3.5 sec) between sentences ensured that the haemodynamic response to the critical sentence 5 could be modelled separately from the adjacent stimuli (see Ollinger et al., 2001). Temporal jitter (ranging from 15 to 18 sec, with a mean of 16.5 sec) also separated the end of the report from the three corresponding comprehension probes.

Each set of comprehension probes was preceded by a question mark (?) presented on screen for 2000 ms. Probes were presented auditorily and were preceded by a warning tone of 1000 ms. After each probe, a response scale was shown on screen, which remained visible until participants entered a response. Participants logged their response on an MRI-compatible button response box (LU400, Cedrus Lumina) by moving a centrally-placed cursor from left to right on a 9-point rating scale (from “strongly agree” to “strongly disagree”) with the index and middle fingers of their right hand. A response was registered and the scale removed from the screen when no further cursor movement was detected for 1500 ms. The temporal offset of each endorsement statement and the onset of its corresponding scale were again separated by temporal jitter (ranging from 2 to 5 sec, with a mean of 3.5 sec), as was the offset of the scale and the onset of the next statement (jitter ranging from 4 to 8 sec, with a mean of 6 sec). Once a participant had replied to the third and final statement for a given report, there was a fixed inter-trial interval of 10 sec.

3.4: Image Acquisition

Images were acquired on a 3 tesla Siemens Skyra MRI scanner with a 32 channel receive-only head coil. Memory foam was used to minimize head

movement. Functional images were acquired using a whole-brain T2*-weighted gradient echo sequence [echo planar imaging (EPI), TR = 2500 ms, TE = 30 ms, flip angle = 90°, 3 x 3 mm in-plane resolution; field of view (FOV) = 192 mm, phase encoding anterior to posterior, parallel acceleration factor 2]. Each volume consisted of 36 axial slices aligned parallel to AC-PC line (anterior commissure – posterior commissure) with 3 mm slice thickness and 0 mm gap. For each participant, approximately 528 volumes were collected during each run of the comprehension task. For each subject, a high resolution (0.9 x 0.9 x 0.9 mm) T1-weighted 3D volume scan was acquired with the MP-RAGE sequence [slice thickness = 0.9mm; TR = 1800 ms; TE = 2.25 ms; FOV = 240 mm; flip angle = 90°]. Between the two functional runs, field-maps were acquired for each subject [slice thickness = 3.0 mm; resolution = 3 x 3 x 3mm; TR = 520 ms; TE1/TE2 = 4.92/7.38 ms; flip angle = 60°; FOV = 192 mm].

3.5: Data Analysis

Behavioural data were analysed using Matlab (v 2012b) for Windows. Image processing and statistical inference was performed with FSL software (v 5.0.8; Oxford Centre for Functional MRI of the BRAIN; FMRIB; Smith et al., 2004). Prior to model estimation, functional images were distortion and motion corrected (Jenkinson et al., 2002) by pre-processing with FEAT (FMRIB's Expert Analysis Tool), which also included spatial smoothing (FWHM = 5 mm), high-pass temporal filtering (cut-off 90 seconds). To facilitate group analysis the spatial transformation between the pre-processed EPI data and each subject's T1-weighted structural scan was determined using the boundary-based registration algorithm (BBR; Greve & Fischl, 2009) and FLIRT (FMRIB's Linear Image Registration Tool; Jenkinson et al., 2002; Jenkinson & Smith, 2001). To improve registration, brain tissue was segmented from structural

scans by using an in-house brain extraction tool (“VBM8BET”), based on the output from VBM8 (“VBM at Structural Brain Mapping Group”, n.d., <http://www.neuro.uni-jena.de/vbm/>). The final registration step included spatial normalisation of each subject’s brain extracted T1-weighted structural scan to a “standard space” template brain (Montreal Neurological Institute [MNI] averaged-152 subject 2 mm template), achieved using an initial 12-parameter affine registration with FLIRT, followed by non-linear registration using FNIRT (FMRIB’s non-linear image registration tool).

Parameter estimates for each explanatory variable were calculated using a general linear model (GLM) implemented in FEAT, which used pre-whitening with FILM (Woolrich et al., 2001) to address issues related to temporal auto-correlation. Participant-specific motion parameters were included as regressors at the first-level. Following subject-level modelling, parameter-estimate maps and associated variance images were transformed to standard space and input to a group-level mixed effects model, estimated using FLAME 1 + 2 (FMRIB's Local Analysis of Mixed Effects). Two types of subject-level analyses were conducted. First, standard univariate whole-brain contrasts examined the effects of information type (retraction versus control) during encoding (i.e., during sentence 5) and retrieval (i.e., during the comprehension probes). Second, additional parametric whole-brain analyses examined whether participants’ behavioural responses to comprehension probes were linked to their brain activity at encoding or retrieval.

For the first set of analyses, an event-related design was modelled using a canonical hemodynamic response function (HRF) with four regressors of interest (sentence 5 at encoding in both the control and retraction conditions, and all three comprehension probes at retrieval in both conditions) and eighteen nuisance regressors (all non-critical sentences at encoding, presentation of warning tone,

presentation of scale and subsequent response, and six motion parameters). All regressors (excluding confound regressors for motion) coded the onset and duration of each event. Based on the regressors of interest, contrast maps (retraction versus control) were computed separately for encoding and the retrieval phase. Within each subject, data from the two separate fMRI runs were combined using a fixed-effects model. These maps were then entered into a group-level paired *t*-test, treating participant as a random effect. Statistical inference was performed using Gaussian random field theory (Worsley et al., 1992), with a cluster-forming threshold of $z > 2.3$, and a (corrected) cluster-significance threshold of $p < 0.05$ (FWE corrected). Results were interrogated using AUTOAQ for automated anatomical labelling of activated clusters (Winkler, 2012). An arbitrary threshold of $> 10\%$ probability was used to determine regions contained within each cluster.

For the second set of analyses, participants' responses (at retrieval) were included as a parametric modulator of the two regressors modelling activity at (i) encoding (averaged across the three probes for each report) and (ii) retrieval. Corresponding statistical parametric maps were computed for each participant and entered into a second-level one-sample *t*-test for each condition, treating participant as a random effect. In order to identify regions for which correlations differed significantly between conditions for the regressors of interest (sentence 5 at encoding in both the control and retraction conditions, and comprehension probes at retrieval) a second-level paired *t*-test was conducted. The approach to determining significance levels for the parametric analyses was the same as for the whole-brain analyses.

4: Results

4.1: Behavioural data

Participants' endorsement of the comprehension probes, on a 9 point scale, was significantly higher following the presentation and subsequent retraction of misinformation (retraction condition: $M = 4.42$, $SD = 1.06$) compared to when no misinformation had been presented at all (control condition: $M = 2.89$, $SD = 0.89$), $t(25) = 11.31$, $p < .001$. This result shows that a retraction of misinformation was not sufficient to bring endorsement ratings back to a baseline level, demonstrating the expected continued influence effect of misinformation (CIEM).¹

4.2: Main effects of experimental condition

To examine the effects of condition on participants' brain activity during encoding and retrieval, we first computed a set of whole-brain contrasts that directly compared participants' neural response during control and retraction reports. During the encoding of critical information in sentence 5, the contrast retraction > control failed to return any supra-threshold activations. The reverse contrast, however, yielded enhanced activity in one brain region (347 voxels; peak voxel: $x = 18$, $y = -38$, $z = 46$; max z -value = 3.79). The region spanned the right precuneus (PrC) extending into the right posterior cingulate cortex (PCC) and postcentral gyrus (PCG). To illustrate the nature of neural activity in this region we extracted its mean parameter estimates for both experimental conditions (Fig. 1). No significant differences in brain activity emerged across the two experimental conditions at retrieval.

¹ It was not necessary to include a condition in which misinformation is presented and never retracted, as is common in most research on the CIEM, as the control and retraction conditions alone are sufficient to demonstrate the effect, i.e. significantly higher endorsement in the retraction condition.

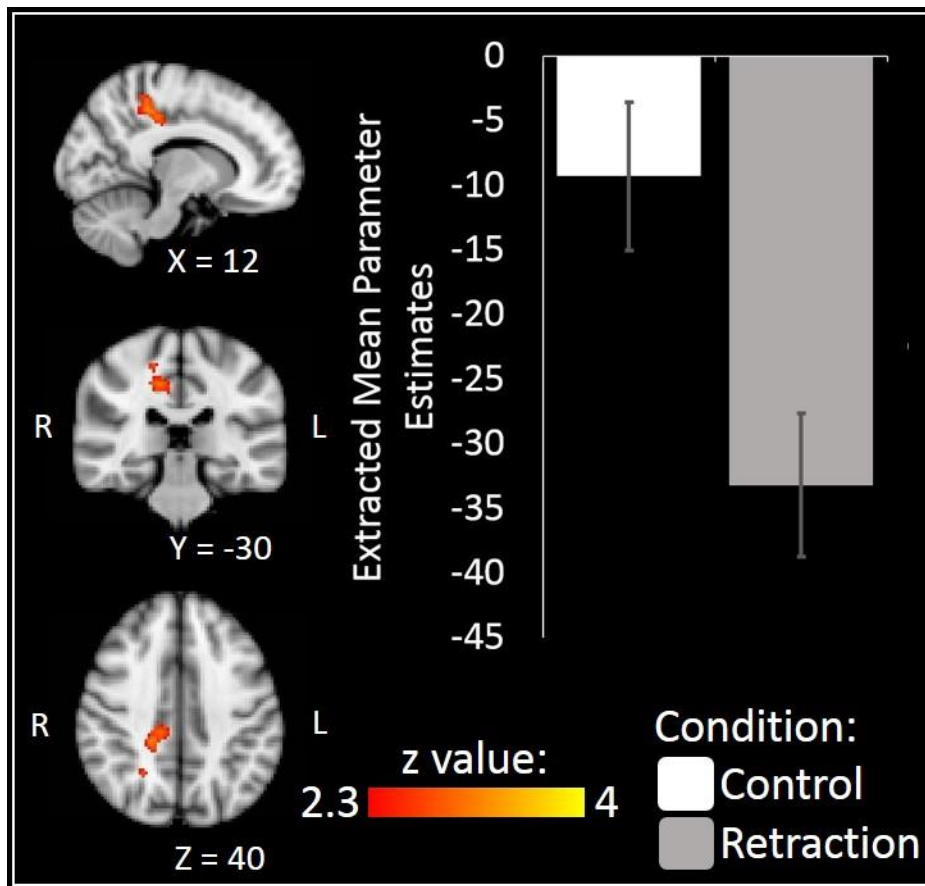


Figure 1. Whole-brain contrasts revealed a main effect of condition (control > retraction) during the encoding of critical information in sentence 5 in the right precuneus (extending into the posterior cingulate gyrus and the postcentral gyrus). To visually illustrate the nature of this difference we extracted and plotted the mean parameter estimates from the area in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on coronal, sagittal, and axial sections. Data were obtained from a mixed effects model (paired t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

4.3: Parametric effects of endorsement ratings at encoding

Our next set of analyses related changes in brain activity during the encoding of critical information in sentence 5 to participants' endorsement ratings on the reports' corresponding comprehension probes (the average ratings given on the three comprehension probes for each report were included as parametric modulators of the regressors modelling activity in each condition during sentence 5). In the control condition, positive correlations between neural activity during encoding of

sentence 5 and subsequent endorsement ratings were observed in two regions of the right parietal lobe (see Table 1A and Fig. 2A). Thus, enhanced activity in these two regions predicted less accurate responding to comprehension probes (recall that higher endorsement ratings reflect a continued influence of misinformation). More importantly, several brain regions, including the left and right MFG, showed a negative correlation (see Table 1B and Fig. 2B). Thus, enhanced activity in these regions predicted more accurate responding to comprehension probes. In the retraction condition, no supra-threshold activations emerged, regardless whether negative or positive correlations were examined.

Table 1.

Peak voxel in MNI coordinates and number of voxels for brain regions that showed changes in activity during encoding of critical information (as presented in sentence 5) which correlated with participants' subsequent endorsement ratings in the control condition as identified by an exploratory whole brain analysis at a cluster forming threshold of $z > 2.3$ and $p < 0.05$ (FWE-corrected).

Region	Voxels	Hemisphere	Max z-value	x	y	z
<i>A) Brain regions with positive correlations</i>						
Superior Parietal Lobule (extending into the Postcentral Gyrus)	335	R	3.58	28	-46	66
Supramarginal Gyrus (extending into the Postcentral Gyrus and the Parietal Operculum)	443	R	3.45	62	-20	32
<i>B) Brain regions with negative correlations</i>						
Cerebellum (Crus I extending into Crus II)	332	L	3.17	-10	-88	-24
Frontal Pole	591	L	3.52	-26	60	26
Middle Frontal Gyrus	377	R	3.39	54	22	32
Middle Frontal Gyrus	318	L	3.17	-52	22	34

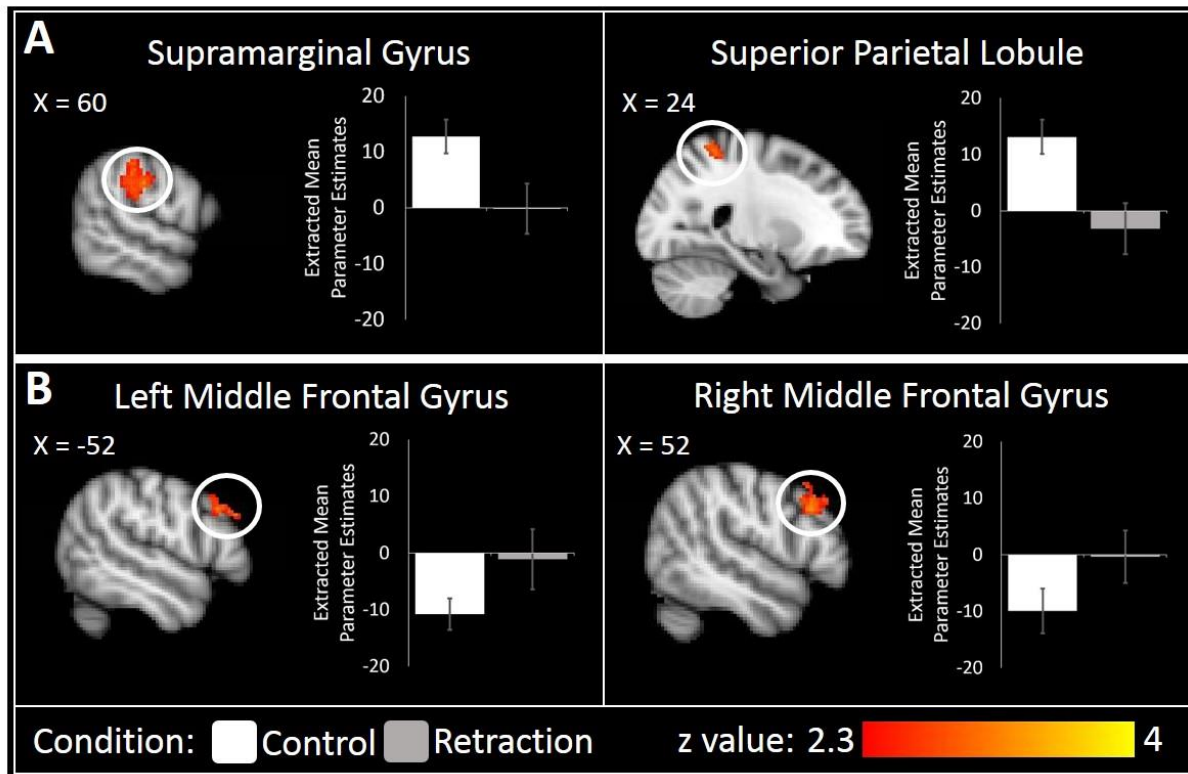


Figure 2. During the encoding of critical information (i.e., sentence 5) in control reports, neural activity correlated with participants' subsequent endorsement ratings in several brain regions. Panel A shows brain regions with positive correlations, panel B shows selected brain regions with negative correlations. For illustrative reasons, corresponding bar graphs portray each region's parametric response in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on sagittal sections. Data were obtained from a mixed effects model (one sample t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

Besides exploring the presence of parametric brain activity at encoding for each experimental condition separately, we also examined whether there were significant differences in this activity across conditions. This analysis returned only a single brain region that spanned the left PCG, supramarginal gyrus, and the parietal operculum (430 voxels; peak voxel: $x = -56$, $y = -18$, $z = 32$; z -value = 3.32). Extracting the region's mean parameter estimates for each participant per experimental condition illustrates that changes in its neural activity were more

predictive of participants' endorsement ratings in the control condition than in the retraction condition (see Fig. 3).

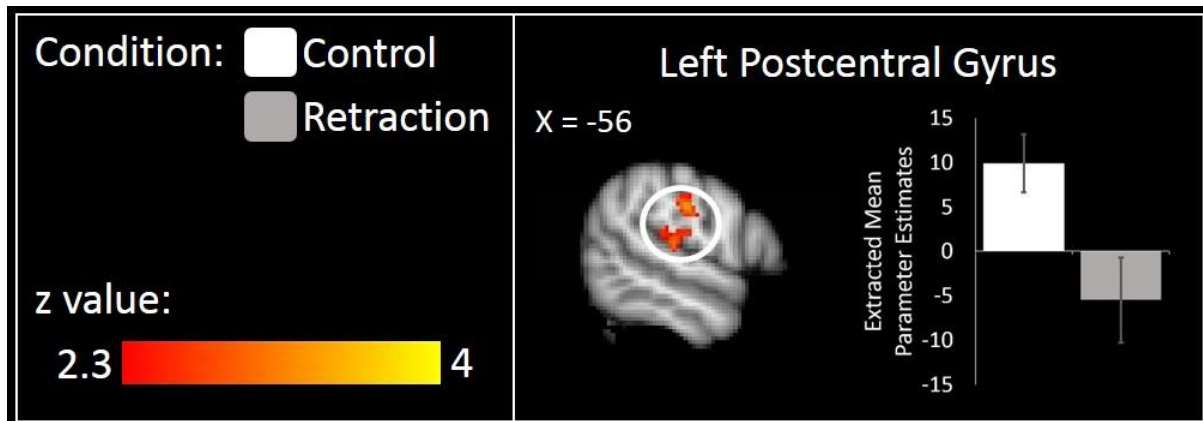


Figure 3. During the encoding of critical information (i.e., sentence 5), neural activity correlated more strongly with participants' subsequent endorsement ratings in the control condition than in the retraction condition in the left postcentral gyrus. The corresponding bar graph portrays the region's parametric response in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on a sagittal section (see also Table 3). Data were obtained from a mixed effects model (paired samples t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

4.4: Parametric effects of endorsement ratings at retrieval

An additional set of parametric analyses examined whether changes in brain activity during the processing of the comprehension probes (i.e., at retrieval) were correlated with participants' endorsement ratings in response to these probes (the ratings given on the three comprehension probes for each report were included as parametric modulators of the regressors modelling activity in each condition during retrieval). In the control condition (see Table 2), no supra-threshold activations indicative of positive correlations emerged, but several brain regions with negative correlations were identified. Among them were the right MFG, the left PCG (extending into the left precentral gyrus), and the PCC (see Fig. 4).

Table 2.

Peak voxel in MNI coordinates and number of voxels for brain regions that showed changes in activity during processing of the comprehension probes (i.e., at retrieval) which correlated with participants' subsequent endorsement ratings in the control condition as identified by an exploratory whole brain analysis at a cluster forming threshold of $z > 2.3$ and $p < 0.05$ (FWE-corrected).

Region	Voxels	Hemisphere	Max z-value	x	y	z
<i>A) Brain regions with positive correlations</i>						
No supra-threshold activation						
<i>B) Brain regions with negative correlations</i>						
Angular Gyrus (extending into Supramarginal Gyrus and the Occipital Cortex)	588	L	3.33	-42	-50	42
Middle Frontal Gyrus	578	R	3.34	44	8	42
Middle Temporal Gyrus (extending into the Angular Gyrus)	1216	R	3.64	66	-38	-8
Occipital Pole (extending into the Inferior Occipital Gyrus)	565	L	3.33	-34	-94	4
Postcentral Gyrus (extending into the Precentral Gyrus)	1164	L	4.11	-46	-20	50
Posterior Cingulate Gyrus	1075	L	3.33	-8	-48	24
Superior Frontal Gyrus	539	L/R	3.21	0	32	44

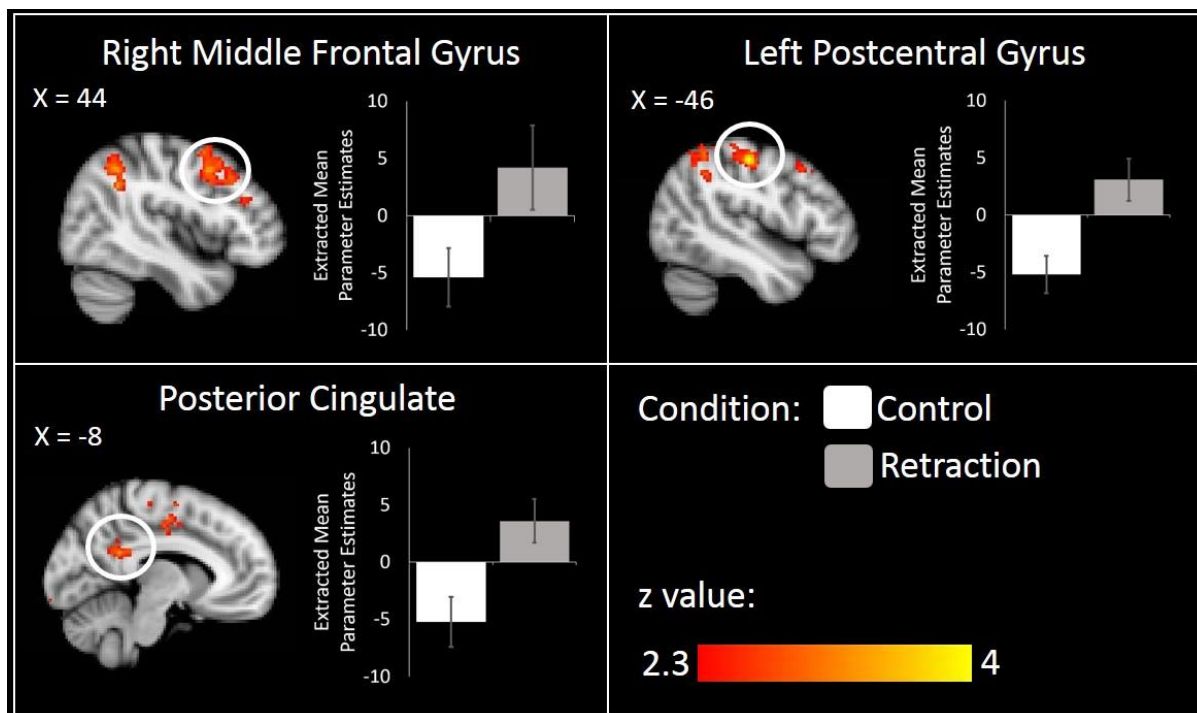


Figure 4. During comprehension probes (i.e., at retrieval) following control reports, neural activity correlated negatively with participants' subsequent endorsement ratings in several brain regions, including the right middle frontal gyrus (MFG), the left postcentral gyrus (PCG), and the posterior cingulate cortex (PCC). For illustrative reasons, corresponding bar graphs portray each region's parametric response in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on sagittal sections. Data were obtained from a mixed effects model (one sample t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

In the retraction condition (Table 3), only brain regions displaying positive correlations emerged (see Fig. 5), but no brain regions with negative correlations. The two regions with positive correlations were located in the left and the right PCG.

Table 3.

Peak voxel in MNI coordinates and number of voxels for brain regions that showed changes in activity during processing of the comprehension probes (i.e., at retrieval) which correlated with participants' subsequent endorsement ratings in the retraction condition as identified by an exploratory whole brain analysis at a cluster forming threshold of $z > 2.3$ and $p < 0.05$ (FWE-corrected).

Region	Voxels	Hemisphere	Max z-value	x	y	z
<i>A) Brain regions with positive correlations</i>						
Postcentral Gyrus (extending into the Precentral Gyrus)	687	R	3.16	46	-4	60
Postcentral Gyrus	567	L	3.35	-34	-32	62
<i>B) Brain regions with negative correlations</i>						
No supra-threshold activation						

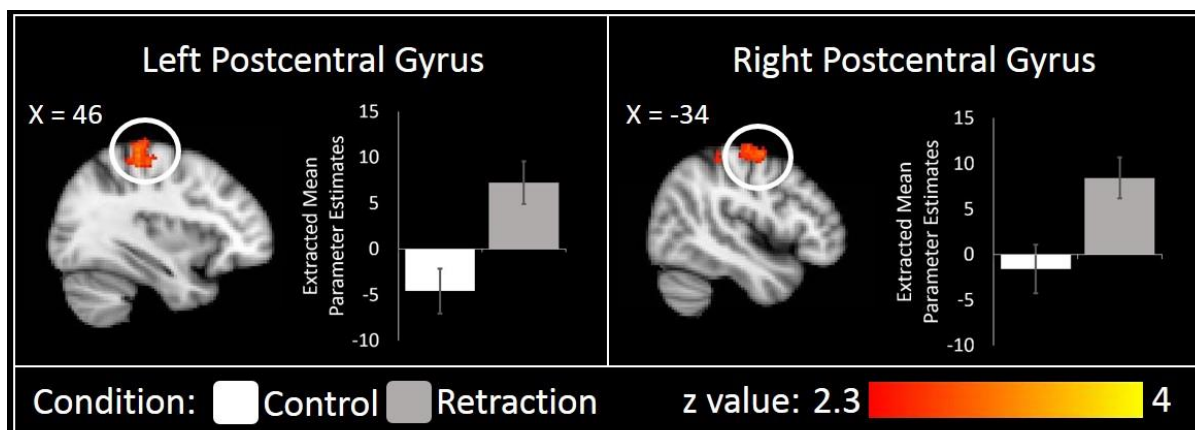


Figure 5. During comprehension probes (i.e., at retrieval) following retraction reports, neural activity correlated positively with participants' subsequent endorsement ratings in the left and right postcentral gyrus (PCG). For illustrative reasons, corresponding bar graphs portray each region's parametric response in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on sagittal sections. Data were obtained from a mixed effects model (one sample t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

Finally, six brain regions showed significant differences in their parametric activity at retrieval across conditions (see Table 4), including the left PCG (see Fig. 6). Importantly, the latter observation implied that the left PCG showed a statistically significant negative correlation with endorsement ratings in the control condition, a statistically significant positive correlation with endorsement ratings in the retraction condition, and a significant difference in correlations across conditions (Fig. 7).

Table 4. *Peak voxels in MNI coordinates and number of voxels for brain regions showing significant differences in correlations between participants' neural response to comprehension probes (i.e., at retrieval) and their endorsement ratings across experimental conditions (control vs. retraction) as identified by an exploratory whole brain analysis at a cluster forming threshold of $z > 2.3$ and $p < 0.05$ (FWE-corrected).*

Region	Voxels	Hemisphere	Max z-value	x	y	z
<i>Control Condition > Retraction Condition</i>						
No supra-threshold activation						
<i>Retraction Condition > Control Condition</i>						
Frontal Pole	543	L	3.42	-30	62	6
Fusiform Gyrus	829	L	3.51	-40	-48	-22
Inferior Occipital Cortex	870	R	3.36	52	-70	-12
Insular Cortex	418	L	3.26	-32	-22	12
Planum Temporale	1361	R	3.58	32	-26	4
Postcentral Gyrus	5557	L	4.34	-34	-34	60

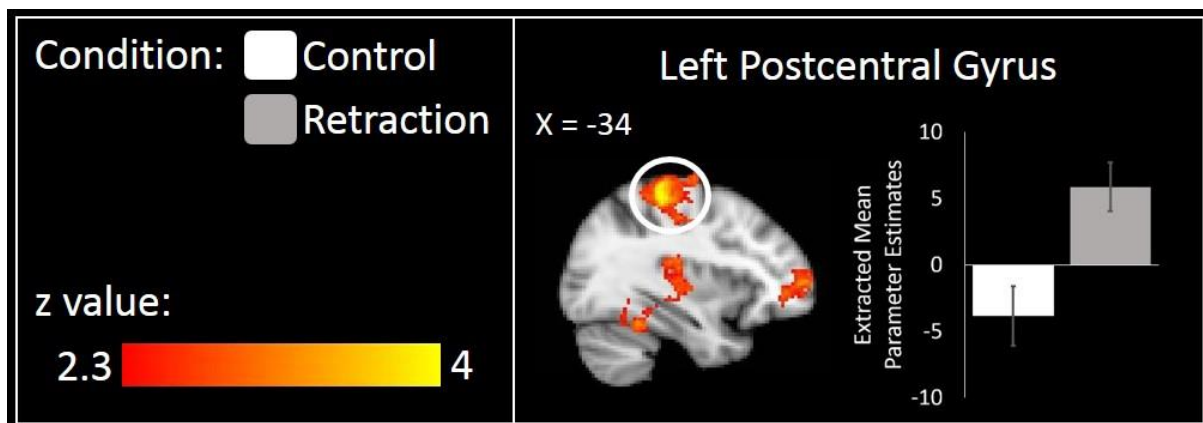


Figure 6. During comprehension probes (i.e., at retrieval), neural activity correlated more strongly with participants' subsequent endorsement ratings in the retraction condition than in the control condition in the left postcentral gyrus (PCG). The corresponding bar graph portrays the region's parametric response in both experimental conditions (mean \pm standard error bars). The figure shows group data from 26 participants displayed on the MNI 2mm brain template in radiological format on a sagittal section (see also Table 3). Data were obtained from a mixed effects model (paired samples t-test), with cluster forming threshold of $z > 2.3$, and $p < 0.05$ (FWE-corrected).

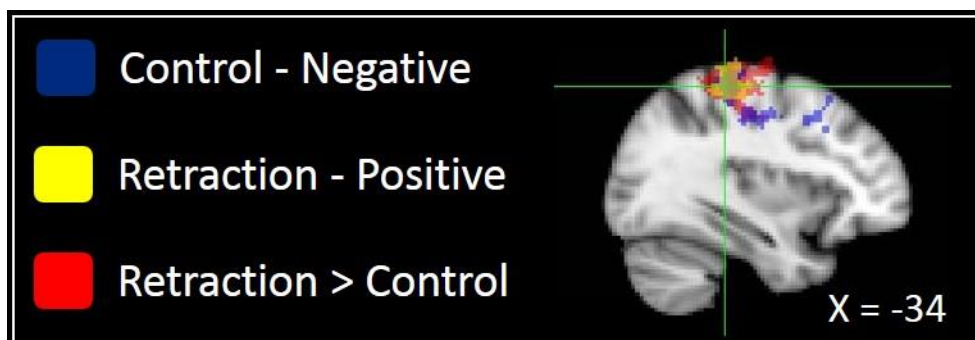


Figure 7: During comprehension probes (i.e. at retrieval), the left postcentral gyrus (PCG) showed a statistically significant negative correlation with endorsement ratings in the control condition, a statistically significant positive correlation with endorsement ratings in the retraction condition, and a significant difference in correlations across conditions. The figure shows the overlap between masks of functional results for these three contrasts at the location of the peak voxel for the correlation contrast (retraction > control) analysis, displayed on the MNI 2mm brain template in radiological format on a sagittal section.

5: Discussion

In line with a growing body of behavioural work, the current study demonstrated that simple retractions fail to cancel out the lasting effect of previously acquired misinformation (Ecker et al., 2011; Ecker et al., 2010; Ecker et al., 2017; Gilbert et al., 1990; Johnson & Seifert, 1994; Schul & Mazursky, 1990; Seifert, 2002; Wilkes & Leatherbarrow, 1988; for reviews see Lewandowsky et al., 2012; Schwarz et al., 2016). Going beyond existing work, however, the current study also investigated the neural substrates of this continued influence effect of misinformation (CIEM).

As predicted by the model-updating account, the encoding of identical pieces of information elicited differences in brain activity depending on whether the information was processed as a retraction or not. Specifically, retractions elicited less activity in the right PrC (extending into the PCC and PCG) than non-retractions. This observation is particularly notable in the context of prior reports on the role of the PrC/PCC. The region has repeatedly been shown to support the integration of continuous pieces of verbal information in a coherent mental model (Silbert et al., 2014; Whitney et al., 2009; Wilson et al., 2008; Xu et al., 2005). In addition, reductions in PrC activity have recently been linked to integration difficulties (e.g., Lahnakoski et al., 2017). In light of these findings, our data suggest that basic integrative and coherence-building mechanisms are less engaged whenever new pieces of information that challenge a person's prior understanding of an event require encoding.

It shall be acknowledged, however, that the PrC also forms a pivotal component of the so-called default mode network (DMN). While the DMN is typically associated with inward-focus and rest, brain activity in this network is also known to

be affected by task difficulty (Greicius & Menon, 2004; Mckiernan et al., 2003). Fortunately, our study design excludes the possibility that retractions were simply more difficult than non-retractions in terms of their linguistic or semantic properties given that the relevant statements were identical. Consequently, any potential differences in processing difficulty across conditions could only be related to the overall narrative in which the statements were encoded.

Our surprising finding that the encoding of retractions (compared to non-retractions) failed to elicit any increase in neural activity throughout the brain is deserving of future investigation. For example, retraction statements did not elicit stronger activity in the ACC, as is usually found when participants experience a violation of their expectancies (Braver et al., 2001; Bush et al., 2000; Carter et al., 1998; Carter & van Veen, 2007). Given that our sample of participants was comparable to many prior fMRI studies that reported ACC activity, this absence of an effect may not simply reflect an issue of statistical power. Instead, it may indicate that humans have a tendency to encode conflicting pieces of information about an event that they did not personally witness as alternative, rather than as mutually exclusive, accounts of an event (Ayers & Reder, 1998; Seifert, 2002). This possibility deserves particular attention in future research as it could partially explain why simple retraction messages tend to be rather ineffective at overcoming misinformation.

The current study also entailed several parametric analyses that examined whether brain activity during the encoding of critical new information in sentence 5 (e.g., “Firecrews [...] stated that there was no evidence of a fire.”) was correlated with participants’ subsequent responses to comprehension probes that re-visited this information (e.g., “Based on this incident, measures [...] to prevent fires should be

reviewed”). In the control condition, participants’ responses to these comprehension probes were more accurate (i.e., elicited *lower* endorsement ratings), the more they engaged the left and right MFG during the encoding of the critical (albeit non-retracting) information. These findings align well with prior reports that the MFG region supports the successful integration of new knowledge in an individual’s existing understanding of an event (Moss et al., 2011). Though a comparable effect appeared to be absent in the retraction condition (see Fig. 2), a direct comparison of parametric MFG activity across conditions failed to reach statistical significance.

In fact, at encoding, only the left PCG displayed a statistically significant difference in parametric activity across conditions. Specifically, it was found that activity increases in the left PCG tentatively predicted less accurate responses towards comprehension probes (i.e., higher endorsement ratings) in the control condition, but not in the retraction condition (see Fig. 3). Our result meshes well with other findings that increased activity in the left PCG during the encoding of information can signal insufficient processing, which in turn predicts subsequent forgetting (Rizio & Dennis, 2013). In light of this observation, our data could indicate that participants occasionally failed to refute misleading comprehension probes in the control condition because they failed to successfully memorize the critical information in the first place. In comparison, a similar correlational effect should be harder to detect in the retraction condition, as sufficient encoding would be more consistently and/or more severely compromised in this condition.

While our neuroimaging results so far lend empirical support to the view that the encoding of information differs depending on whether it acts as a retraction or not (as postulated by the model updating account; e.g. Ecker et al., 2010; Kendeou et al., 2014; Verschueren et al., 2005), we found less evidence that both types of

information engage different neural mechanisms during retrieval. First, no systematic differences in brain activity were observed when participants processed comprehension statements that invited them to re-visit critical information which contained retracting (and retracted) information or not. Second, subsequent parametric analyses that linked participants' brain activity during the processing of the comprehension statements to their replies to these statements revealed largely unexpected results.

Only in the control condition (but not in the retraction condition), did we detect increases in neural activity during the processing of comprehension probes that predicted more accurate replies in response to these probes. Among the brain regions displaying this pattern of activity were again the right MFG and left PCG. The right MFG overlapped considerably with the cluster identified in our previous encoding-based parametric analysis. In other words, activity increases predictive of accurate replies to comprehension probes were found in a similar portion of the right MFG during both encoding and retrieval. This convergence suggests that the region may not only play a major role in constituting an individual's original understanding of a (retraction-free) narrative, but also in maintaining it.

In contrast to the right MFG, localization of parametric activity in the left PCG differed across encoding and retrieval. Specifically, parametric activity in the left PCG was located more dorsally at retrieval than at encoding. In addition, it was found that at retrieval, increased activity in the dorsal portion of the left PCG predicted not only more accurate responding towards comprehension probes in the control condition, but also *less* accurate responding towards them in the retraction condition. This difference across conditions was statistically significant, emphasizing that retrieval-related parametric activity in the left dorsal PCG was directly affected

by the presence or absence of retractions in the original news reports. One ad-hoc interpretation of this finding draws from the observation that activity in the left dorsal PCG can be systematically enhanced whenever participants encounter statements that they find implausible (e.g., when Non-Christians read that 'Jesus Christ really performed the miracles attributed to him in the Bible'; Harris et al., 2009). Thus, activity in the left dorsal PCG in the current study may have signalled how implausible participants found each comprehension probe. In case of particularly implausible probes, additional memory checks may have been performed in both processing conditions, but with markedly different results. While enhanced scrutiny in the control condition should have exclusively activated memories that directly contradicted the seemingly implausible probe (thereby helping participants to refute it), such scrutiny in the retraction condition may have elicited some (actually retracted) memories validating the probe (thereby undermining a person's willingness to simply refute it). Although speculative, these results may provide tentative support for the notion that people encode both the incorrect (misinformation) and correct (retraction) information concurrently in memory (e.g. Ayers & Reder, 1998; Catarino et al., 2015; Ecker et al., 2011; Jacoby & Whitehouse, 1989).

In summary, this study aimed to enhance our understanding of the cognitive mechanisms that contribute to the CIEM. To do so, it investigated neural processing differences during the encoding and retrieval of retracting and non-retracting information using both exploratory and theory-driven data analysis methods. In partial support of the model-updating account of the CIEM (e.g., Ecker et al., 2010; Kendeou et al., 2014; Verschueren et al., 2005), we found that the same piece of information was processed differently in the right PrC/PCC depending on whether it

retracted prior information or not. Although previous theorizing strongly suggests that the CIEM is also due to inadequate memory suppression and substitution mechanisms (Catarino et al., 2015; Ecker et al., 2011), the current study provided little evidence in favour of these mechanisms. Future research is therefore needed to clarify whether this absence of evidence indicates a minor role of these mechanisms in the CIEM or whether fMRI designs need to be further optimized in order to adequately capture the impact of these mechanisms.

6: References

- Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, 18(6), 279–282. <http://doi.org/10.1016/j.tics.2014.03.002>
- Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., ... Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, 303(5655), 232–235. <http://doi.org/10.1126/science.1089504>
- Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, 5(1), 1-21. <http://doi.org/10.3758/BF03209454>
- Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron*, 76(2), 450–460. <http://doi.org/10.1016/j.neuron.2012.07.025>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.
- Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. *Cerebral Cortex*, 11(9), 825–836. <http://doi.org/10.1093/cercor/11.9.825>
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215–222. [http://doi.org/10.1016/S1364-6613\(00\)01483-2](http://doi.org/10.1016/S1364-6613(00)01483-2)
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280(5364), 747–749. <http://doi.org/10.1126/science.280.5364.747>
- Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective & Behavioral Neuroscience*, 7(4), 367–379. <http://doi.org/10.3758/cabn.7.4.367>

- Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to forget: inhibitory-control deficits compromise memory suppression in posttraumatic stress disorder. *Psychological Science*, 26(5), 604–616. <http://doi.org/10.1177/0956797615569889>
- Depue, B. E. (2012). A neuroanatomical model of prefrontal inhibitory modulation of memory retrieval. *Neuroscience and Biobehavioral Reviews*, 36(5), 1382–1399. <http://doi.org/10.1016/j.neubiorev.2012.02.012>
- Ecker, U. K. H., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition*, 1–8. <http://doi.org/10.1016/j.jarmac.2017.01.014>
- Ecker, U. K. H., Lewandowsky, S., & Apai, J. (2011). Terrorists brought down the plane!--No, actually it was a technical fault: processing corrections of emotive information. *Quarterly Journal of Experimental Psychology*, 64(2), 283–310. <http://doi.org/10.1080/17470218.2010.497927>
- Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, 18(3), 570–578. <http://doi.org/http://dx.doi.org/10.3758/s13423-011-0065-1>
- Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. W. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8), 1087–1100. <http://doi.org/10.3758/mc.38.8.1087>
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59(4), 601–613.
- Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, 16(9), 1484–1492. <http://doi.org/10.1162/0898929042568532>
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72. <http://doi.org/10.1016/j.neuroimage.2009.06.060>
- Harris, S., Kaplan, J. T., Curiel, A., Bookheimer, S. Y., Iacoboni, M., & Cohen, M. S. (2009). The neural correlates of religious and nonreligious belief. *PLoS one*, 4(10), e7272. <https://doi.org/10.1371/journal.pone.0007272>
- Horne, Z., Powell, D., Hummel, J. E., & Holyoak, K. J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), 10321–10324. <http://doi.org/10.1073/pnas.1504019112>
- Jacobson, G. C. (2010). Perception, memory, and partisan polarization on the Iraq war. *Political Science Quarterly*, 125(1), 31–56. <http://doi.org/10.1002/j.1538-165X.2010.tb00667.x>

- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology*, 118(2), 126–135. <http://doi.org/10.1037/0096-3445.118.2.126>
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841. [http://doi.org/10.1016/S1053-8119\(02\)91132-8](http://doi.org/10.1016/S1053-8119(02)91132-8)
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2), 143–156. [http://doi.org/10.1016/S1361-8415\(01\)00036-6](http://doi.org/10.1016/S1361-8415(01)00036-6)
- Johnson-Laird, P. N. (2012). Mental models and consistency. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 225–243). New York, NY: Guilford Press.
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1420–1436. <http://doi.org/10.1037/0278-7393.20.6.1420>
- Kendeou, P., Walsh, E. K., Emily, R., & Brien, E. J. O. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, 51(5-6), 374-397 <http://doi.org/10.1080/0163853X.2014.913961>
- Kull, S., Ramsay, C., & Lewis, E. (2003). Misperceptions, the media, and the Iraq war. *Political Science Quarterly*, 118(4), 569–598. <http://doi.org/10.2307/30035697>
- Lahnakoski, J. M., Jääskeläinen, I. P., Sams, M., & Nummenmaa, L. (2017). Neural mechanisms for integrating consecutive and interleaved natural events. *Human Brain Mapping*, 38(7), 3360-3376. <http://doi.org/10.1002/hbm.23591>
- Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(47), 16785–16794. <http://doi.org/10.1523/JNEUROSCI.2640-12.2012>
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. <http://doi.org/10.1177/1529100612451018>
- Marsh, E. J., Meade, M. L., & Roediger, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language*, 49(4), 519–536. [http://doi.org/10.1016/S0749-596X\(03\)00092-5](http://doi.org/10.1016/S0749-596X(03)00092-5)
- Mckiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., & Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience*, 15(3), 394–408. <http://doi.org/10.1162/089892903321593117>

- Miton, H., & Mercier, H. (2015). Cognitive obstacles to pro-vaccination beliefs. *Trends in Cognitive Sciences*, 19(11), 633–636. <http://doi.org/10.1016/j.tics.2015.08.007>
- Moss, J., Schunn, C. D., Schneider, W., McNamara, D. S., & VanLehn, K. (2011). The neural correlates of strategic reading comprehension: Cognitive control and discourse comprehension. *NeuroImage*, 58(2), 675–686. <http://doi.org/10.1016/j.neuroimage.2011.06.034>
- Nee, D. E., & Jonides, J. (2008). Dissociable interference-control processes in perception and memory. *Psychological Science*, 19(5), 490–500. <http://doi.org/10.1111/j.1467-9280.2008.02114.x>
- Nyhan, B., & Reifler, J. (2015). Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine*, 33(3), 459–464. <http://doi.org/10.1016/j.vaccine.2014.11.017>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Ollinger, J. M., Corbetta, M., & Shulman, G. L. (2001). Separating processes within a trial in event-related functional MRI (II. Analysis). *Neuroimage*, 13(1), 218–229. <http://doi.org/10.1006/nimg.2000.0711>
- Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain*, 124(2), 399–412. <http://doi.org/10.1093/brain/124.2.399>
- Rizio, A. A., & Dennis, N. A. (2013). The neural correlates of cognitive control: successful remembering and intentional forgetting. *Journal of Cognitive Neuroscience*, 25(2), 297–312. http://doi.org/10.1162/jocn_a_00310
- Schul, Y., & Mazursky, D. (1990). Conditions facilitating successful discounting in consumer decision making. *Journal of Consumer Research*, 16(4), 442–451. <https://doi.org/10.1086/209229>
- Schwarz, N., Newman, E., & Leach, W. (2016). Making the truth stick and the myths fade: Lessons from cognitive psychology. *Behavioural Science and Policy*, 2(1), 85–95.
- Seifert, C. M. (2002). The continued influence of misinformation in memory: What makes a correction effective? *Psychology of Learning and Motivation*, 41, 265–292. [http://doi.org/10.1016/S0079-7421\(02\)80009-3](http://doi.org/10.1016/S0079-7421(02)80009-3)
- Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences of the United States of America*, 111(43), E4687–E4696. <http://doi.org/10.1073/pnas.1323812111>
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(SUPPL. 1), 208–219. <http://doi.org/10.1016/j.neuroimage.2004.07.051>

- VBM at Structural Brain Mapping Group. (n.d.). Retrieved from <http://dbm.neuro.uni-jena.de/vbm/> (accessed 10.08.16)
- Verschueren, N., Schaeken, W., & D'Ydewalle, G. (2005). Everyday conditional reasoning: a working memory-dependent tradeoff between counterexample and likelihood use. *Memory & Cognition*, 33(1), 107–119. <http://doi.org/10.3758/BF03195301>
- Wagner, A. D. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281(5380), 1188–1191. <http://doi.org/10.1126/science.281.5380.1188>
- Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., & Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. *NeuroImage*, 47(1), 360–366. <http://doi.org/10.1016/j.neuroimage.2009.04.037>
- Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology Section A*, 40(2), 361–387. <http://doi.org/10.1080/02724988843000168>
- Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, 18(1), 230–242. <http://doi.org/10.1093/cercor/bhm049>
- Winkler, A.M. (2012). AutoAQ: Automatic atlas queries in FSL [Automated labelling of clusters of activations]. Retrieved from <http://brainder.org/tag/autoaq/> (accessed 20.08.16)
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage*, 14(6), 1370–1386. <http://doi.org/10.1006/nimg.2001.0931>
- World Economic Forum (2013). Digital wildfires in a hyperconnected world. Retrieved from <http://reports.weforum.org/global-risks-2013/risk-case-1/digital-wildfires-in-a-hyperconnected-world/> (accessed 15.02.16)
- Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 12(6), 900–918. <http://doi.org/10.1038/jcbfm.1992.127>
- Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, 25(3), 1002–1015. <http://doi.org/10.1016/j.neuroimage.2004.12.013>