

Table S1. Behavioral results in Experiment 2. Average proportion correct, misinformation selection, and confidence as a function of trial type (consistent, neutral, misleading) and warning group (no warning, pre-warning, post-warning) during the second memory test. Accuracy refers to proportion of correct recognition memory trials within each trial type. Selection refers to the proportion of responses for which a misleading detail was selected within each trial type. Confidence refers to participants' ratings of their accuracy in their memory decision (range 1-4).

	Consistent	Neutral	Misleading
No Warning			
Accuracy	.82 (0.05)	.63 (0.05)	.37 (0.05)
Selection	.04 (0.02)	.18 (0.03)	.52 (0.05)
Confidence	3.49 (0.10)	2.84 (0.10)	3.32 (0.10)
Reaction Time	6152.99 (47.90)	6283.21 (47.90)	6160.45 (47.90)
Pre-Warning			
Accuracy	.78 (0.05)	.64 (0.05)	.59 (0.05)
Selection	.09 (0.02)	.20 (0.03)	.29 (0.05)
Confidence	3.27 (0.10)	2.93 (0.10)	3.15 (0.10)
Reaction Time	6106.44 (48.75)	6175.41 (48.75)	6243.10 (48.75)
Post-Warning			
Accuracy	.78 (0.05)	.77 (0.05)	.56 (0.05)
Selection	.09 (0.02)	.12 (0.03)	.30 (0.05)
Confidence	3.14 (0.10)	3.10 (0.10)	3.05 (0.10)
Reaction Time	6255.82 (47.90)	6170.49 (47.90)	6238.29 (47.90)

Means and standard errors are reported.

Table S2. Whole-brain analysis of effect of warning on accurate memory decisions. Cortical regions demonstrating greater activity during accurate memory decisions (Hits > Misses) in participants who received a warning (pre-warning + post-warning) compared to participants who did not receive a warning (no warning).

<i>Region</i>	<i>BA</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>No v. Pre</i>	<i>No v. Post</i>	<i>Pre v. Post</i>	<i>r</i>
R Parahippocampal Cortex	36	18	-38	-12	4.23	.000	.000	.33	.23*
R Occipital Cortex	18	10	-84	20	4.15	.001	.001	.89	.18
L Occipital Cortex	18	-10	-86	18	4.95	.000	.001	.81	.31*

*Note: Whole-brain contrast was conducted at $p < .001$, $k = 8$; BA = Brodmann Area; Statistics for direct comparison of beta values in each warning group are reported in rightmost columns; No v. Pre = No Warning versus Pre-Warning; No v. Post = No Warning versus Post-Warning; Pre v. Post = Pre-Warning versus Post-Warning; r = Pearson's correlation coefficient for relationship between neural activity and memory performance on misleading trials. * $p < .05$.*

Table S3. Whole-brain analysis of effect of warning on memory retrieval during misleading trials. Cortical regions demonstrating reduced activity in participants who received a warning (pre-warning + post-warning) compared to participants who did not receive a warning (no warning) during misleading trials (Misleading > Baseline).

<i>Region</i>	<i>BA</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>No v. Pre</i>	<i>No v. Post</i>	<i>Pre v. Post</i>	<i>r</i>
R Anterior Prefrontal Cortex	10	40	52	-4	3.84	.001	.005	.72	-.27*
R DLPFC	9	0	48	46	4.25	.001	.000	.11	-.12
R DLPFC	46	40	44	2	3.92	.000	.001	.53	-.27*
R DLPFC	9	40	26	36	4.70	.000	.000	.41	-.10
R Premotor Cortex	6	52	2	38	3.92	.001	.006	.46	-.24*
R Inferior Temporal Gyrus	20	50	-6	-34	4.88	.000	.000	.93	-.16
R Somatosensory Cortex	1	48	-12	16	4.10	.001	.002	.46	-.33*
L Supramarginal Gyrus	40	-50	-18	20	3.22	.023	.000	.04*	-.34*
R Supramarginal Gyrus	40	52	-26	18	5.09	.001	.003	.78	-.45*
R Auditory Cortex	41	54	-30	10	4.07	.001	.000	.42	-.35*
L Posterior Cingulate Cortex	23	-6	-36	30	4.14	.002	.000	.88	-.23
R Fusiform Gyrus	37	44	-50.	-20	3.96	.001	.002	.86	-.24

*Note: Whole-brain contrast was conducted at $p < .001$, $k = 8$; BA = Brodmann Area; Statistics for direct comparison of beta values in each warning group are reported in rightmost columns; No v. Pre = No Warning versus Pre-Warning; No v. Post = No Warning versus Post-Warning; Pre v. Post = Pre-Warning versus Post-Warning. r = Pearson's correlation coefficient for relationship between neural activity and memory performance on misleading trials. * $p \leq .05$.*

Appendix A

No Warning:

- **Prior to the audio narrative:** "You will now hear an audio narrative of the video you just watched."
- **After the audio narrative:** "You will now answer a series of questions relating to the video you watched at the beginning of this experiment. Please answer each question to the best of your ability. If you do not know the answer, make your best guess. All questions must be answered."

Pre-warning

- **Prior to paying the audio narrative:** "You will have to answer questions regarding the video you previously watched for a second time. We will play a narrative of that video; however, we are uncertain as to the source of the narrative. Therefore, we were unable to verify the accuracy of the narrative. As such, base your answers only on what you saw in the video, and not on what you hear in the narrative."
- **After playing the audio narrative:** "You will now answer a series of questions relating to the video you watched at the beginning of this experiment. Please answer each question to the best of your ability. If you do not know the answer, make your best guess. All questions must be answered."

Post-warning

- **Prior to playing the audio narrative:** "You will now hear an audio narrative of the video you just watched."
- **After playing the audio narrative:** "You will have to answer questions regarding the video you previously watched for a second time. We just played a narrative of that video; however, we are uncertain as to the source of the narrative. Therefore, we were unable to verify the accuracy of the narrative. As such, base your answer only on what you saw in the video, and not on what you heard in the narrative. Please answer each question to the best of your ability. If you do not know the answer, make your best guess. All questions must be answered."