A causal role of the NMDA receptor in recurrent processing during perceptual inference: NMDA receptor antagonist memantine selectively improves EEG decoding of the Kanizsa illusion

Samuel Noorman[1,2*], Timo Stein[1,2], Jasper Zantvoord[3,4],

Johannes J. Fahrenfort[1,2,5**], Simon van Gaal[1,2**]

[1]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

[2]Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands

[3]Amsterdam UMC location University of Amsterdam, Department of Psychiatry, Amsterdam, The Netherlands

[4]Amsterdam Neuroscience, Amsterdam, The Netherlands

[5]Department of Applied and Experimental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[*]Correspondence: sgnoorman@gmail.com

[**]Shared senior authorship

This research was supported by a grant from the H2020 European Research Council (ERC STG 715605, SVG).

Abstract

Perceptual inference requires the integration of visual features through recurrent processing, the dynamic exchange of information between higher and lower level cortical regions. While animal research has demonstrated a crucial role of NMDA receptors in recurrent processing, establishing a causal link between NMDA receptors and recurrent processing in humans has remained challenging. Here, we report two pharmacological studies with randomized, double-blind, crossover designs in which we administered the NMDA antagonist memantine, while collecting human electroencephalography (EEG). We trained and tested EEG classifiers to reflect the processing of specific stimulus features with increasing levels of complexity, namely differences in stimulus contrast, collinearity between local line elements, and illusory surfaces of a Kanizsa triangle. In two experiments involving different participants and visual tasks, we found that memantine selectively improved decoding of the Kanizsa illusion, known to depend on recurrent processing, while leaving decoding of contrast and collinearity largely unaffected. Interestingly, the results from an attentional blink (experiment 1) and task-relevance manipulation (experiment 2) showed that memantine was only effective when the stimulus was attended and consciously accessed. These findings suggest that NMDA inhibition through memantine enhances recurrent processing, especially for attended objects, and thereby provide a crucial step toward bridging animal and human research, shedding light on the neural mechanisms underpinning perceptual inference and conscious perception.

Introduction

Different theories of consciousness converge on a central principle: consciousness entails integration of information through recurrent neural processes that exceeds the capabilities of mere feedforward processing (Mashour et al., 2020; Seth & Bayne, 2022). Recurrent neural mechanisms encompass feedback connections between higher- and lower-order cortical areas as well as lateral connections within sensory areas. Studies in mice and monkeys have revealed a causal role of feedback from higher cortical areas to sensory areas for tactile perception (Manita et al., 2015; Sachidhanandam et al., 2013; Takahashi et al., 2016), detection of odors (Allen et al., 2017), and visual stimulus attributes (Ciceri et al., 2023; Lamme et al., 1998; van Vugt et al., 2018; Zatka-Haas et al., 2021). In macaques, general anesthetics targeted at reducing the overall level of consciousness of the animal reduce global cortical recurrent interactions (Mashour & Hudetz, 2018) and block feedback connections to V1, while sparing feedforward processing (Lamme et al., 1998). Although in humans indirect evidence for a role of recurrent processing in consciousness and perceptual integration has been provided, for example by focusing on neural markers of figure-ground segregation or backward masking (Fahrenfort et al., 2007, 2008, 2012, 2017; Kok et al., 2016; Kok & de Lange, 2014; Noorman et al., 2023; Scholte et al., 2008; Wyatte et al., 2014), the underlying biochemistry has remained so far elusive. In this study, we aimed to provide direct evidence for a causal role of NMDA receptors in recurrent processing (through lateral and feedback connections) and their role in perceptual inference and conscious access, factoring in potential modulating influences of attention (Flohr et al., 1998; Lamme, 2010; Sachidhanandam et al., 2013; Self et al., 2012; Thiele, 2012; Thiele & Bellgrove, 2018; Van Kerkoerle et al., 2014; Wang, 2008).

NMDA has a slower time constant than most other receptors (decay time constant ~100 ms, voltage-gated) and is known to be densely populated in recurrent connections in the cortex (Wang, 2008). In macaque monkeys, blocking NMDA receptors selectively disrupted late onset modulations of neural activity in V1 (i.e., recurrent processing) during figure-ground detection, whereas blocking AMPA receptors did not, but instead reduced the strength of the first feedforward sweep (Self et al., 2012; Van Kerkoerle et al., 2014). This suggests that feedback connections target neuronal synapses that are rich in NMDA receptors (among other receptors), whereas feedforward connections target synapses especially rich in AMPA receptors (Thiele, 2012; Thiele & Bellgrove, 2018). In mice, blocking NMDA receptors strongly reduced hit rates for the detection of whisker movements (Sachidhanandam et al., 2013). In humans, few studies have investigated the role of NMDA receptors in recurrent processing and (conscious) perception. Blocking NMDA receptors via administration of ketamine

resulted in impaired feature integration in a texture discrimination task (Meuwese et al., 2013) and a distortion of the neural representations of objects in early visual cortex (van Loon et al., 2016). However, as these human studies did not clearly distinguish between markers of feedforward and recurrent processing, they could not determine whether these effects were specific to recurrent processing or reflected other non-specific effects. Ketamine has rather broad adverse side effects, such as sedation and dissociation, which could potentially explain the observed behavioral deficits and neural distortions. Also, ketamine is known for its complex pharmacological dynamics, having effects on several neurotransmitter systems, including catecholaminergic, cholinergic, nicotinic, muscarinic and GABAergic receptors (Kocsis et al., 2013; Mathew et al., 2012; Müller et al., 2011).

In this study, we used a novel Kanizsa stimulus whose different visual features were designed to target distinct processing stages with increasing levels of complexity, namely local contrast differences, differences in collinearity of local line elements and the presence/absence of an illusory triangle (**Fig. 1A**) (see also Noorman et al., 2023). The processing of collinearity of line elements predominantly relies on lateral connections (Bosking et al., 1997; Gilbert & Wiesel, 1979; Li, 1998; Liang et al., 2017; Schmidt et al., 1997; Stettler et al., 2002), while processing of the illusory surface of the Kanizsa illusion relies on feedback connections (Halgren et al., 2003; Kok et al., 2016; Kok & de Lange, 2014; Lee & Nguyen, 2001; Pak et al., 2020; Wokke et al., 2013). Therefore, multivariate EEG decoding of the presence/absence of the illusionary Kanizsa triangle served as a marker of feedback processing (Fahrenfort et al., 2017; Noorman et al., 2023). The ancillary features of the stimulus, the two-legged white circles, could be aligned to form a non-illusionary triangle, enabling us to use collinearity decoding as a marker of lateral processing. By rotating the Kanizsa by 180 degrees, thus either having one black Pacman on top of the screen or two, we created differences in contrast between the upper and lower visual field. Because local contrast differences are processed early in the visual system, this allowed us to employ local contrast decoding as a marker of feedforward processing. In recent work we have validated that the decoding profiles of these features of different complexities at different points in time, in combination with the associated topography, can indeed serve as EEG markers of feedforward, lateral and feedback processes (Fahrenfort et al., 2017; Noorman et al., 2023).

We tested how these neural markers were affected by memantine, a low-affinity, voltage-dependent, uncompetitive NMDA antagonist. Memantine is well tolerated and has minimal side effects compared to ketamine.

We performed two separate experiments. In experiment 1, we perturbed conscious access through masking and inattention using a modified attentional-blink (AB) paradigm (**Fig. 1B**). In experiment 2, we manipulated whether the illusion, collinearity, or local contrast were task-relevant or not. We hypothesized that disrupting the reentrant glutamate signal via blocking NMDA receptors by memantine would impair illusion and possibly collinearity decoding, as putative markers of feedback and lateral processing, but would spare the decoding of local contrast differences, our marker of feedforward processing. To foreshadow our results, memantine indeed specifically affected illusion decoding, but enhancing rather than impairing it. In the Discussion, we offer explanations for this surprising finding, including the effect of memantine on extrasynaptic NMDA receptors in GABAergic cells, which may have resulted in boosted visual activity.

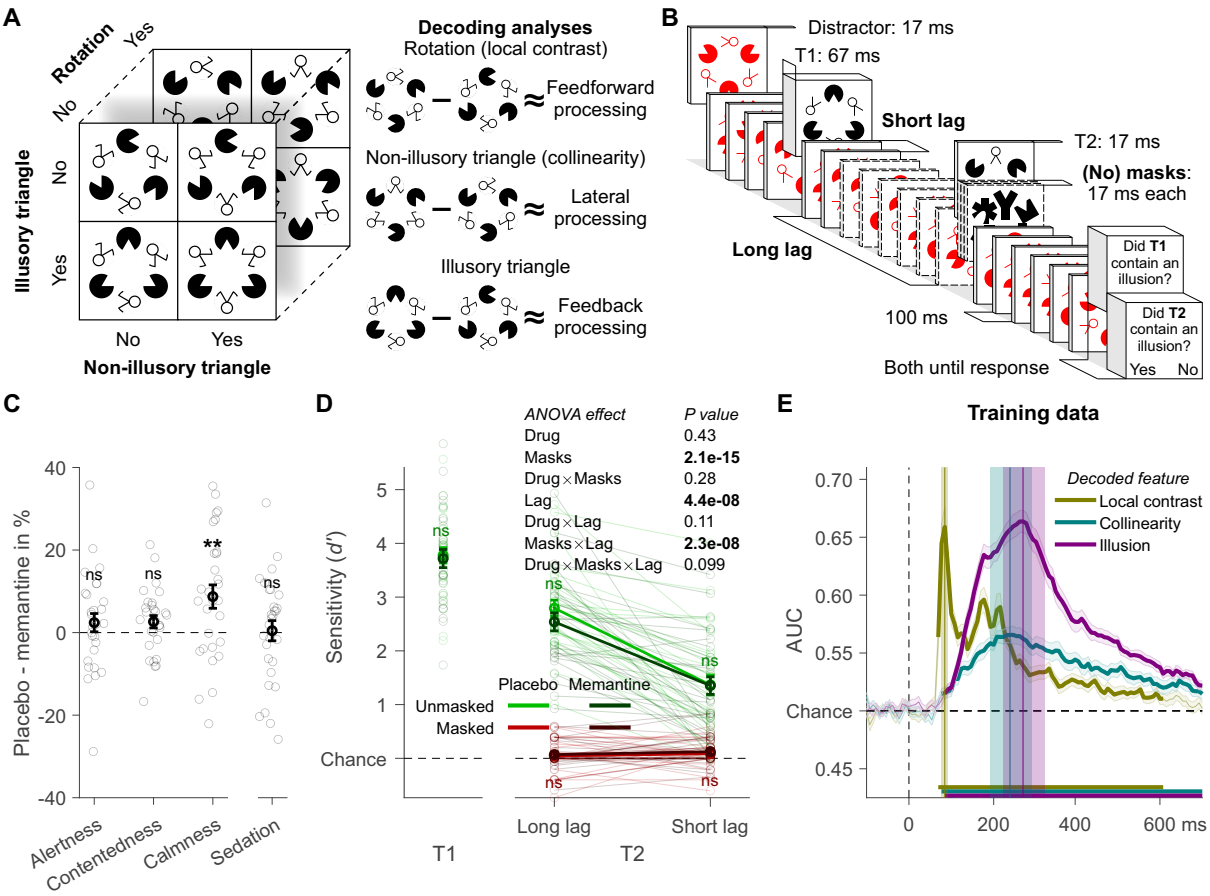


Figure 1. Experiment 1: design, behavior, and training data. (A) Target stimulus set and schematic of the markers for the different types of processing (showing example stimuli). (B) Trial design. (C) Effects of memantine on subjective state assessed with the visual analogue scale. The categories alertness, contentedness, and calmness are based on Bond and Lader (1974), sedation on Danion et al. (1989). For details see Methods. (D) Participants' ability to detect the Kanizsa illusion in the second target stimulus (T2). (E) Tenfold cross-validation decoding scheme applied to an independent training dataset. Mean decoding performance, area under the receiver operating characteristic curve (AUC), over time ± standard error of the mean (SEM) is shown. Thick lines differ from chance: P<0.05, cluster-based permutation test. Local contrast, collinearity, and illusion decoding accuracy peaked at 86, 242, and 273 ms after stimulus onset and these peaks were encompassed by a 78-94, 192-292, and 223-323 ms time window, respectively. Error bars are mean ± SEM. Individual data points are plotted using low contrast. Ns is not significant (*P*>0.05). ***P*=0.005.

Results

Memantine caused minimal side effects

The study had a randomized, double-blind, crossover design in which a placebo was administered during one session and memantine (20 mg) during another. Drug order was counterbalanced across the 29 healthy participants. Potential side effects of memantine on different aspects of the participant's subjective state were assessed with the visual analog scale (VAS) (Bond & Lader, 1974; Danion et al., 1989). The only significant effect was a moderate decrease in the category "calmness" for memantine compared to placebo (t_{28}=-23.07, *P*=0.005, BF_{10}=8.71; **Fig. 1C**, for details see Methods). Note that in experiment 2 (**Fig. 4**) none of these subjective measures differed between memantine and placebo, indicating that overall subjective side effects of memantine were minimal.

Memantine selectively improves decoding of the Kanizsa illusion

We recorded EEG signals from participants who detected illusory surfaces in two black targets (T1 and T2) amongst red distractors in a rapid serial visual presentation (RSVP) task (**Fig. 1B**). Visibility of T2 was manipulated in two ways, resulting in a 2×2 factorial task design. First, T2 was either masked or unmasked. Second, shortening the interval (lag) between the two targets from 900 to 200/300 ms induced an attentional blink (AB), which reduced T2 visibility as well. At the end of a trial, participants indicated for each target whether they had perceived an illusory surface or not. For all main analyses of behavior and EEG decoding we conducted repeated-measures ANOVAs with the factors drug (memantine vs. placebo), T1-T2 lag (short, inducing an AB, vs. long, no AB), and masks (masked vs. unmasked). The complete results of these ANOVAs are presented in the corresponding results figures. In the text, we only highlight the key results and report selectively additional statistics such as post-hoc tests.

As expected, both masking and the AB impaired participant's ability to detect the Kanizsa illusion in the second target (T2). Our masking procedure was highly effective, leading to poor T2 performance on masked trials (red lines in **Fig. 1D**). The AB was also effective, leading to a sharp drop in T2 performance on short lag versus long lag trials (green lines in **Fig. 1D**). Memantine had however no effect on perceptual sensitivity (**Fig. 1D**).

To establish our markers for the different neural processes of interest, the processing of different visual features (local contrast, collinearity, and the Kanizsa illusion) were decoded (**Fig. 1A**). For these decoding analyses, classifiers were trained using an independent training set obtained from a separate experimental task where the same stimuli were presented, but outside the context of the RSVP task (see Methods for details). No drugs were administered during this separate "localizer" task. In brief, during the localizer, participants had to match the target stimulus to different example stimuli ("templates"). For successful performance in this template-matching task, all three features of the stimulus had to be attended, and all features were thus task-relevant. We also used the independent training set from this template-matching task to determine the time windows for local contrast, collinearity, and illusion decoding. First, we decoded each visual feature from the localizer using a tenfold cross-validation scheme by training linear discriminant classifiers to categorize the feature as either present or absent based on the raw EEG activity across all electrodes. After testing the classifiers, each decoded visual feature showed a clear peak in classifier performance (area under the receiver operating characteristic curve [AUC]) (Hand & Till, 2001) over time: local contrast at 86 ms after stimulus onset, collinearity at 242 ms, and the illusion at 273 ms. In the following, we refer to "time window-based analyses" when we focus on the averages of the time windows encompassing these peaks (for the exact time windows used see **Fig. 1E**). The timing and mostly occipital nature (see the covariance/class separability maps of **Fig. 2A**) (Haufe et al., 2014) of these windows were consistent with the neural processes they are thought to reflect and in line with our previous EEG decoding studies (Fahrenfort et al., 2017; Noorman et al., 2023). Because of this topology, subsequent decoding analyses were restricted to occipitoparietal electrodes. To increase power, we averaged each time window in the localizer task's raw EEG data. Classifiers were trained on these averages and tested on all time samples of each condition of the main RSVP task's data (see Methods) (King & Dehaene, 2014). Note that this approach therefore reflects "off-diagonal decoding" in the generalization across time approach. We show all "on-diagonal" analyses in the **supplementary figures S2-S4**, which revealed very similar results.

We first examined the effect of memantine on the decoding of the different features of the first target (T1) from the RSVP task. For each time sample, we tested whether the memantine conditions differed in classifier performance from the placebo conditions. Cluster-based permutation testing was used to correct for multiple comparisons (Maris & Oostenveld, 2007). Crucially, only illusion decoding showed a significant cluster ranging from 211 to 352 ms after stimulus onset (**Fig. 2A**, blue lines). A similar result was obtained from the time window-based analysis (**Fig. 2B**, ANOVA results in the figure). Memantine significantly improved overall T1 decoding ($F_{1,28}$=8.33, P=0.007) and

drug interacted with visual feature ($F_{1,28}$=5.16, P=0.011). Post hoc t-tests confirmed that memantine selectively affected illusion decoding accuracy (t_{28}=-4.14, P<0.001, BF_{10}=101.92), thus leaving local contrast (t_{28}=-0.60, P=0.560, BF_{10}=0.23) and collinearity decoding unaffected (t_{28}=-1.04, P=0.306, BF_{10}=0.32). As memantine is an NMDA blocker, we had initially hypothesized that it would impair recurrent processing, and therefore decrease illusion decoding accuracy. Surprisingly however, although memantine specifically affected illusion decoding, it *increased* instead of decreased illusion decoding. Recent studies using memantine have however mostly observed increases in task performance and/or neural markers of sensory processing (Korostenskaja et al., 2007; Light et al., 2017; Swerdlow et al., 2016, 2020), and are therefore mostly in line with our observations here. To foreshadow our further results, this effect was robust as we obtained the same *increase* in Kanizsa decoding for T2s in this study (**Fig. 3**), and the effect was replicated in study 2 (**Fig. 5**). We address this issue further in the discussion.

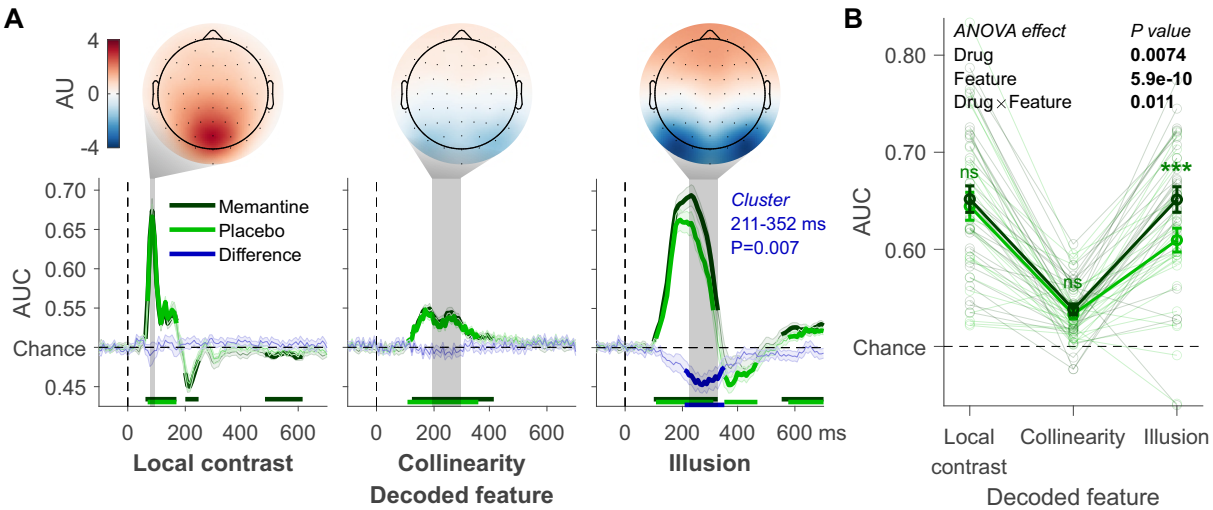
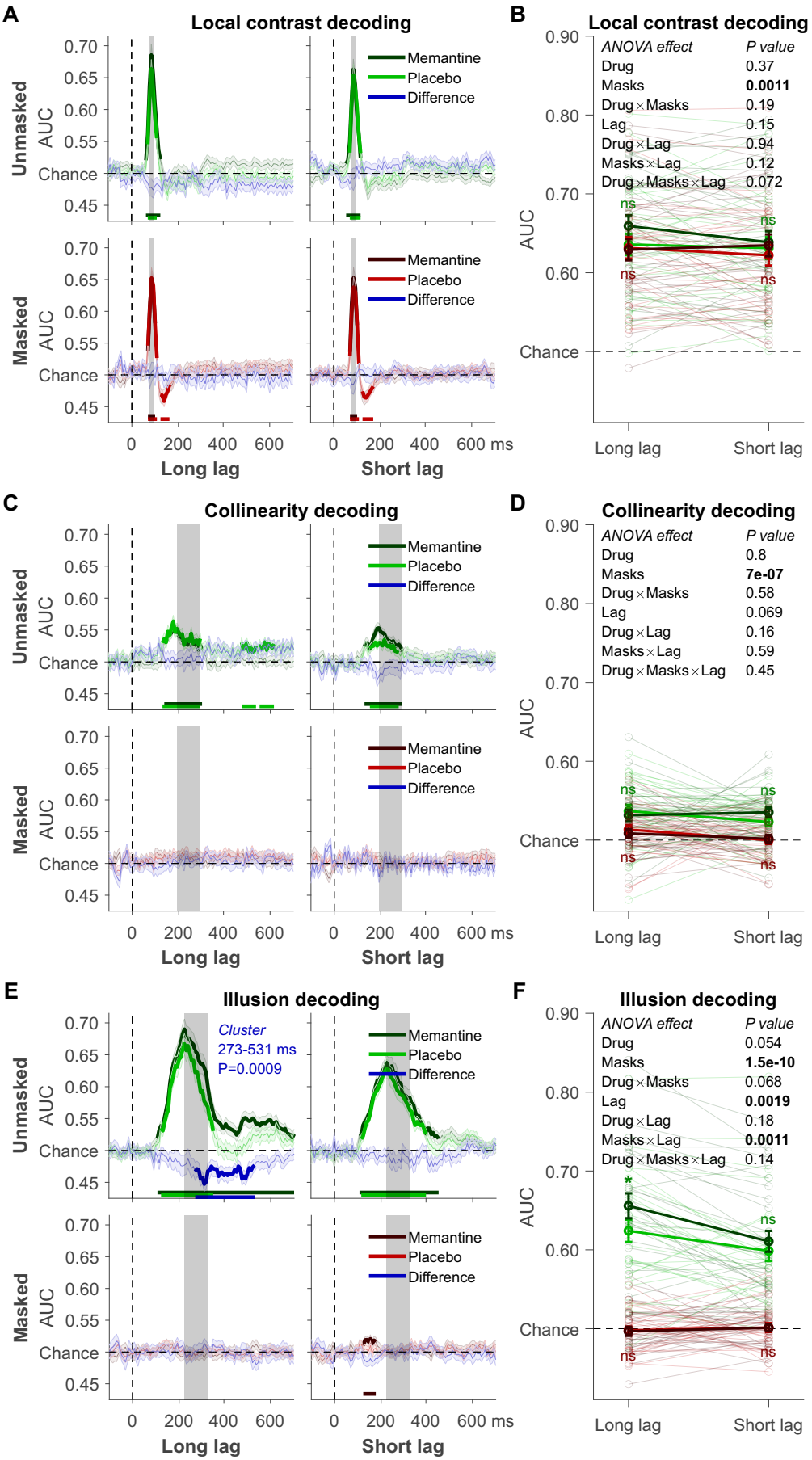


Figure 2. Decoding applied to the T1 data. (A) Local contrast, collinearity, and Kanizsa illusion decoding. These three visual features' time windows: 78-94, 192-292, and 223-323 ms, respectively. Covariance/class separability maps with arbitrary units (AU) reflecting underlying neural sources are shown. Below these maps: mean decoding performance, area under the receiver operating characteristic curve (AUC), over time ± standard error of the mean (SEM). Thick lines differ from chance: P<0.05, cluster-based permutation test. (B) Mean AUC for every time window. Error bars are mean ± SEM. Individual data points are plotted using low contrast. Ns is not significant (P>0.05). ***P<0.001.

The role of consciousness and attention in memantine's effect on illusion decoding

To determine how memantine interacted with our manipulations of consciousness (masking) and attention (T1-T2 lag, i.e., attentional blink [AB]), we examined the decoding of the different features of the second target (T2) from the RSVP task (**Fig. 3A**, **C**, and **E**). First, we tested whether masking and the AB affected local contrast, collinearity and illusion decoding in the way we expected. Indeed, local contrast decoding was unaffected by the AB (**Fig. 3B**), and although it was impaired by masking, the drop in decoding accuracy was numerically small (0.012 AUC on average) and considerably smaller than the effect of masking on collinearity (0.026 AUC) and illusion (0.123 AUC) decoding. These results are in line with theoretical proposals and empirical findings that suggest limited effects of masking and attentional manipulations on feedforward processes (Dehaene et al., 2006; Fahrenfort et al., 2007, 2017; Lamme, 2010; Lamme et al., 2002). As would be expected for markers of lateral and feedback processing, both collinearity and illusion decoding were abolished by masking (**Fig. 3C-F**). The AB exclusively impaired illusion decoding, but decoding accuracy for the illusion remained well above chance during the AB (**Fig. 3E-F**).

Similar to T1, also for T2 decoding, memantine increased illusion decoding, but only when T2 was presented outside the AB and when it was unmasked (**Fig. 3E**, top left, note that masking fully abolished illusion decoding). Also our time window-based analyses showed a significant effect of memantine only when the illusion was both unmasked and presented outside the AB (t_{28}=-2.76, P=0.010, BF_{10}=4.53; **Fig. 3F**). Note, however, that although these post-hoc tests of the effect of memantine on illusion decoding were significant, for our time window-based analyses we did not obtain a statistically significant interaction between the AB and memantine. While the interaction between masking and memantine only approached significance (P=0.068), the absence of an effect of memantine in the masked condition could reflect a floor effect, given that illusion decoding in the masked condition was not significantly better than chance. Thus, although these memantine effects were slightly less robust than for T1, probably due to reduced trial counts, these results point to (but do not conclusively demonstrate) a selective effect of memantine on illusion-related feedback processing that depends on the availability of attention. In addition to the lack of the interaction effect, another potential concern with this interpretation is that in the RSVP task, the illusion was always task-relevant and hence attended, whereas local contrast and collinearity were never task-relevant. It is therefore an open question whether local contrast and collinearity decoding were unaffected by memantine, because these visual features were unattended (not task-relevant) and/or because memantine exclusively targets feedback processes. To alleviate this potential concern, and to replicate the surprising result of increased illusion decoding under memantine, we carried out a second experiment. A new group of participants was recruited and we systematically varied task-relevance of local contrast, collinearity, and the illusion in this task. This allowed us to determine whether memantine indeed specifically affected feedback processing.


A
Local contrast decoding
Unmasked
Masked
AUC
Chance
Memantine
Placebo
Difference
Long lag
Short lag
B
Local contrast decoding
ANOVA effect P value
Drug 0.37
Masks 0.0011
Drug×Masks 0.19
Lag 0.15
Drug×Lag 0.94
Masks×Lag 0.12
Drug×Masks×Lag 0.072
ns
C
Collinearity decoding
D
Collinearity decoding
ANOVA effect P value
Drug 0.8
Masks 7e-07
Drug×Masks 0.58
Lag 0.069
Drug×Lag 0.16
Masks×Lag 0.59
Drug×Masks×Lag 0.45
E
Illusion decoding
Cluster
273-531 ms
P=0.0009
F
Illusion decoding
ANOVA effect P value
Drug 0.054
Masks 1.5e-10
Drug×Masks 0.068
Lag 0.0019
Drug×Lag 0.18
Masks×Lag 0.0011
Drug×Masks×Lag 0.14
*

242

Figure 3. Decoding of the T2 data. (A) Local contrast decoding: mean decoding performance, area under the receiver operating characteristic curve (AUC), over time ± standard error of the mean (SEM). Thick lines differ from chance: P<0.05, cluster-based permutation test. (B) Mean AUC for local contrast (rotation) decoding's time window: 78-94 ms. Error bars are mean ± SEM. Individual data points are plotted using low contrast. Ns is not significant (*P*>0.05). (C) Collinearity decoding: mean AUC over time. (D) Mean AUC for collinearity decoding's time window: 192-292 ms. (E) Kanizsa illusion decoding: mean AUC over time. (F) Mean AUC for Kanizsa illusion's time window: 223-323 ms. **P*=0.010.

Experiment 2 replicates memantine's selective effect on illusion decoding

In experiment 2, the pharmacological design was the same as before and VAS measurements now revealed no side effects from memantine (**Fig. 4B**). Twenty-seven new participants performed a discrimination task in which each visual feature, in separate blocks, was task-relevant (**Fig. 4A**). Approximately every 900-1100 ms an individual target stimulus (no RSVP) was presented for 83 ms when it was unmasked, or for 33 ms when it was masked. The timing of the masked condition was increased compared to experiment 1 to circumvent floor or ceiling effects. Performance on the Kanizsa illusion was staircased to 75% correct in a practice session by varying mask contrast, and these mask settings were used for all features (see Methods). For each target, participants indicated whether the current task-relevant feature (e.g., the Kanizsa illusion) was present or absent. In experiment 2, each visual feature was task-relevant and thus attended in some blocks, while it was task-irrelevant and thus less attended in other blocks. This allowed us to directly examine whether memantine's effects were specific to feedback processes or dependent on task-relevance of the stimulus features. For all main analyses of behavior and EEG decoding we conducted repeated-measures ANOVAs with the factors drug (memantine vs. placebo), masking (masked vs. unmasked), and task-relevant feature (local contrast, collinearity, illusion). The complete results of these ANOVAs are presented in the corresponding results figures. We found that masking impaired perceptual sensitivity, most prominently for the detection of the Kanizsa illusion, and less so for collinearity and rotation. Note that we employed less strong masking in this experiment than in experiment 1, leading to overall above chance perceptual sensitivity for the masked trials. For unmasked trials performance was at ceiling for all visual features (no differences in performance). Perceptual sensitivity was unaffected by memantine (**Fig. 4C**).

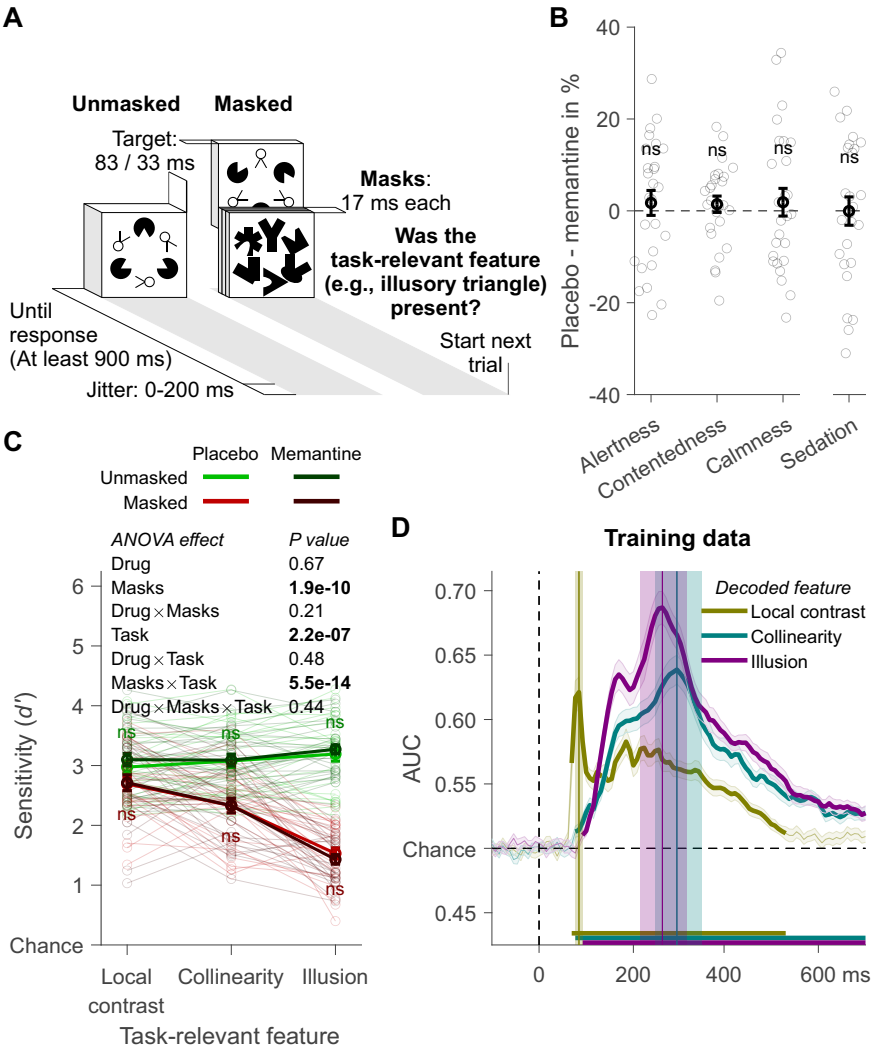
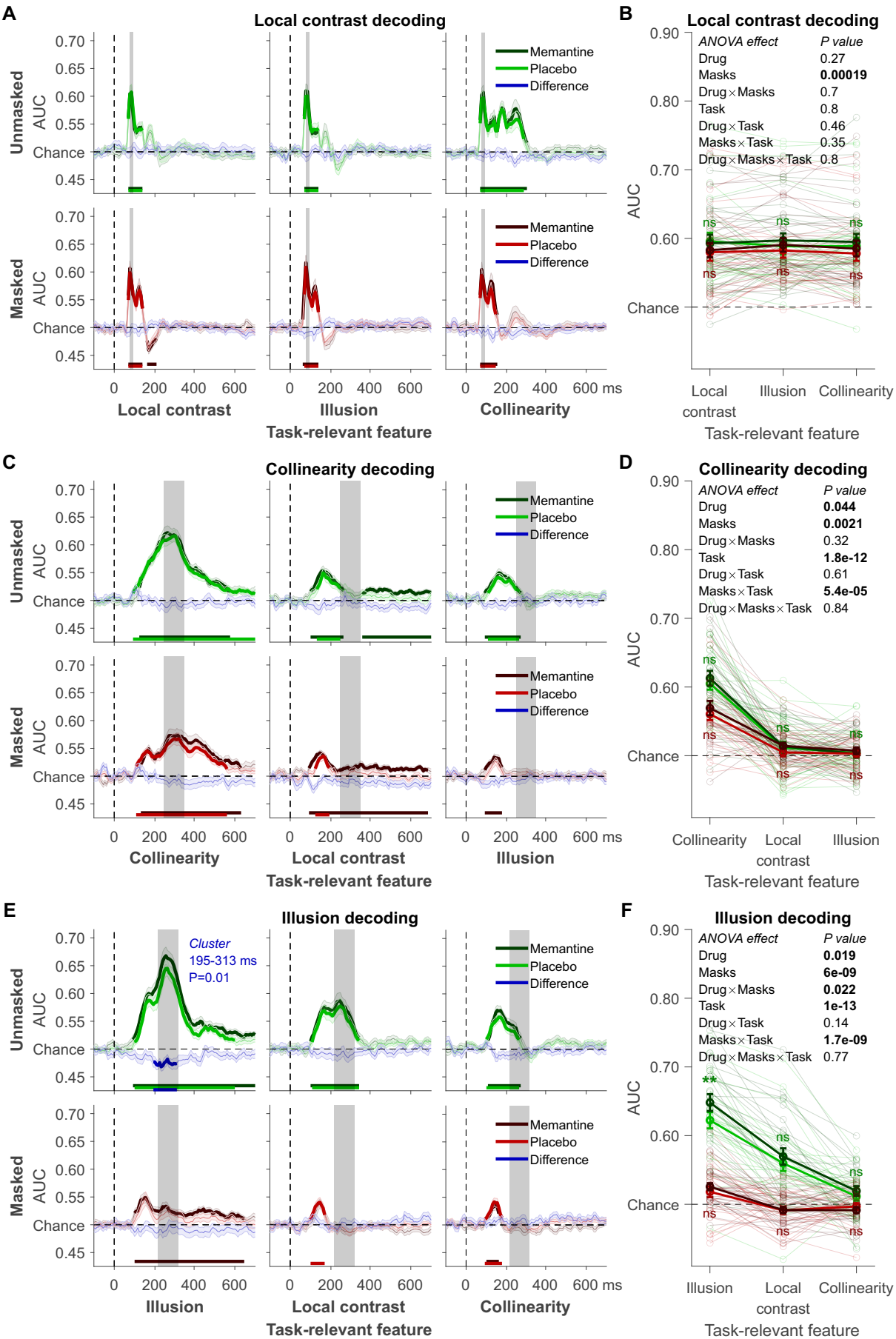


Figure 4. Experiment 2: design, behavior, and training data. (A) Trial design. (B) Effects of memantine on subjective state assessed with the visual analogue scale. The categories alertness, contentedness, and calmness are based on Bond and Lader (1974), sedation on Danion et al. (1989). For details see Methods. (C) Participants' ability to detect the task-relevant feature. (D) Tenfold cross-validation decoding scheme applied to an independent training dataset. Mean decoding performance, area under the receiver operating characteristic curve (AUC), over time ± standard error of the mean (SEM) is shown. Thick lines differ from chance: P<0.05, cluster-based permutation test. Local contrast, collinearity, and illusion decoding accuracy peaked at 86, 297, and 266 ms after stimulus onset and these peaks were encompassed by a 78-94, 247-347, and 217-317 ms time window, respectively. Error bars are mean ± SEM. Individual data points are plotted using low contrast. Ns is not significant (*P*>0.05).

The EEG decoding schemes were the same as for experiment 1 and similar time windows were obtained from the independent training dataset (**Fig. 4D**). The localizer task was similar to the main experiment, but no masking was used. This localizer task was performed in all three sessions (intake, and two main session involving the pharmacological manipulation, 1728 trials per session). In the intake session, no drugs were administered during the localizer task. In the other two sessions, the localizer task was performed right after pill intake and this was long before memantine reached peak plasma levels (after ~4 hrs.). Overall, the timing of the decoding of the three different features in the main experiment replicated our findings from experiment 1 (**Fig. 5**). For local contrast, task-relevance had no effect and the significant drop in decoding accuracy caused by masking was again numerically

small (0.010 AUC on average); smaller than for collinearity (0.015 AUC) and illusion decoding (0.069 AUC). Collinearity (**Fig. 5C-D**) and illusion (**Fig. 5E-F**) decoding were impaired by masking and boosted when the corresponding feature was task-relevant. Crucially, our cluster-based analyses again revealed that memantine significantly improved decoding only for the illusion and only when the illusion was unmasked and task-relevant (**Fig. 5E**, top left), thus replicating the results from experiment 1. Also the time window-based analysis revealed an overall main effect of drug for illusion decoding, and post-hoc test revealed that the drug effect was only significant when the illusion was unmasked and task-relevant (post hoc *t*-tests comparing memantine vs. placebo, t_{28}=-3.36, P=0.002, BF_{10}=15.66; **Fig. 5F**). For our time window-based analyses of illusion decoding, the specificity of the memantine effect to the unmasked condition was supported by a significant interaction between drug and masking (note, however, given overall much lower decoding accuracy in the masked condition, the lack of a memantine effect could reflect a floor effect). However, we did not obtain a statistically significant interaction between memantine and task-relevance. Thus, although the memantine effect was significant only when the illusion was unmasked and task-relevant, just like for the effect of temporal attention in experiment 1, these results do not conclusively demonstrate a selective effect of memantine that depends attention (task-relevance). Our time window-based ANOVAs now also revealed a main effect of memantine on collinearity decoding (**Fig. 5D**), in line with the hypothesized role of NMDA receptors in recurrent (feedback and lateral) processing.


A
Local contrast decoding
Unmasked
Masked
AUC
Chance
Memantine
Placebo
Difference
Local contrast
Illusion
Collinearity
Task-relevant feature
B
Local contrast decoding
ANOVA effect P value
Drug 0.27
Masks 0.00019
Drug×Masks 0.7
Task 0.8
Drug×Task 0.46
Masks×Task 0.35
Drug×Masks×Task 0.8
ns
Local contrast
Illusion
Collinearity
Task-relevant feature
C
Collinearity decoding
D
Collinearity decoding
ANOVA effect P value
Drug 0.044
Masks 0.0021
Drug×Masks 0.32
Task 1.8e-12
Drug×Task 0.61
Masks×Task 5.4e-05
Drug×Masks×Task 0.84
E
Illusion decoding
Cluster
195-313 ms
P=0.01
F
Illusion decoding
ANOVA effect P value
Drug 0.019
Masks 6e-09
Drug×Masks 0.022
Task 1e-13
Drug×Task 0.14
Masks×Task 1.7e-09
Drug×Masks×Task 0.77
**

312

Figure 5. Decoding results of experiment 2. (A) Local contrast decoding: mean decoding performance, area under the receiver operating characteristic curve (AUC), over time ± standard error of the mean (SEM). Thick lines differ from chance: P<0.05, cluster-based permutation test. (B) Mean AUC for local contrast decoding's time window: 78-94 ms. Error bars are mean ± SEM. Individual data points are plotted using low contrast. Ns is not significant (*P*>0.05). (C) Collinearity decoding: mean AUC over time. (D) Mean AUC for collinearity decoding's time window: 217-317 ms. (E) Kanizsa illusion decoding: mean AUC over time. (F) Mean AUC for Kanizsa illusion's time window: 223-323 ms. ***P*=0.002.

Discussion

In this study, we investigated the impact of NMDA receptor blockade on multivariate EEG decoding of different features of a visual stimulus (Kanizsa illusion, collinearity, and local contrast). We used the NMDA receptor antagonist memantine to selectively target recurrent processing while minimizing non-specific side effects, a notable improvement over previous studies blocking NMDA receptors using ketamine. Two separate experiments demonstrated selective effects of memantine on decoding accuracy of the Kanizsa illusion, our marker of feedback processing, while leaving collinearity and contrast decoding relatively unaffected, our markers for lateral and feedforward processing. Interestingly, in both experiments the effect of memantine on illusion decoding tended to be specific to attended, task-relevant stimuli presented without a backward mask. In experiment 1, the illusion was task-relevant and the effect was only found for stimuli presented unmasked and outside the attentional blink (AB), when full attention was available. In experiment 2, we obtained the effect only when the stimulus was unmasked and the illusion was task-relevant. Notably, rather than impairing illusion decoding, in both experiments memantine improved illusion decoding. These findings provide new insight into the role of NMDA-receptor mediated recurrent processing in human visual perception and will be further discussed below.

Our approach relied on using EEG decoding of different stimulus features at different points in time, together with their topography, as markers of distinct neural processes. Although such non-invasive, indirect measures of neural activity cannot provide direct evidence for feedforward vs. recurrent processes, the timing, topography, and susceptibility to masking of the decoding signatures obtained in the present study are consistent with neurophysiology (e.g., Bosking et al., 1997; Kandel et al., 2000; Lamme & Roelfsema, 2000; Lee & Nguyen, 2001; Liang et al., 2017; Pak et al., 2020), as well as with our previous work (Fahrenfort et al., 2017; Noorman et al., 2023). The specificity of the effect of memantine our EEG marker of illusion-related feedback processing is broadly consistent with previous studies indicating that ketamine, another NMDA receptor antagonist, affects perception and neural coding in visual tasks thought to rely on recurrent mechanisms (Meuwese et al., 2013; van Loon et al., 2016). The effect of ketamine in these studies likely affected intra-area lateral connections in early visual cortex (Meuwese et al., 2013) and inter-area feedback mechanisms (van Loon et al., 2016), respectively. The present effect of memantine was largely specific to illusion decoding, our marker of feedback processing, while collinearity decoding, our marker of lateral processing, was not (experiment 1) or only weakly (experiment 2) affected by memantine. We have no straightforward explanation for why NMDA receptor blockade would impact inter-areal feedback connections more strongly than intra-areal lateral connections, considering their strong functional interdependency and interaction in grouping and segmentation processes (Liang et al., 2017). One possibility is that this finding reflects properties of our EEG decoding markers for feedback vs. lateral processing: for example, decoding of the Kanizsa illusion may have been more sensitive to the relatively subtle effect of our pharmacological manipulation, either because overall decoding was better than for collinearity or because NMDA receptor dependent recurrent processes more strongly contribute to illusion decoding than to collinearity decoding.

While the effect of memantine in our study was largely specific to illusion decoding, the direction of this effect was in the opposite direction of what we had initially predicted and of the effect of ketamine in previous studies on perceptual representation: whereas ketamine impaired perceptual

organization and visual cortex representations (Meuwese et al., 2013; van Loon et al., 2016), memantine improved illusion decoding. Although both drugs are known to inhibit NMDA receptors by occupying the receptor's ion channel and are thereby blocking current flow (Glasgow et al., 2017; Molina et al., 2020), the drugs have different actions at receptors other than NMDA, with ketamine acting on dopamine D2 and serotonin 5-HT2 receptors, and memantine inhibiting several subtypes of the acetylcholine (ACh) receptor as well as serotonin 5-HT3 receptors. Memantine and ketamine are also known to target different NMDA receptor subpopulations, with their inhibitory action displaying different time courses and intensity (Glasgow et al., 2017; Johnson et al., 2015). Blockade of different NMDA receptor subpopulations can result in markedly different and even opposite results. For example, Self and colleagues (2012) found overall reduced or elevated visual activity after microinfusion of two different selective NMDA receptor antagonists (2-amino-5-phosphonovalerate and ifendprodil) in macaque primary visual cortex. Although both drugs impaired the feedback-related response to figure vs. ground, similar to the effects of ketamine (Meuwese et al., 2013; van Loon et al., 2016) such opposite effects on overall activity demonstrate that the effects of NMDA antagonism strongly depend on the targeted receptor subpopulation, each with distinct functional properties.

Indeed, previous studies have also found opposite effects of memantine and ketamine. For example, while ketamine diminishes the amplitude of the so-called mismatch negativity (MMN) EEG response to (auditory) oddball stimuli (Rosburg & Kreitschmann-Andermahr, 2016), a single dose of memantine enhances MMN amplitude (Korostenskaja et al., 2007; Light et al., 2017; Swerdlow et al., 2016). Although memantine was found to impair learning as well as markers of perceptual and neural processing in some tasks (e.g., Rammsayer, 2001, 2006; Schugens et al., 1997), most studies found memantine to improve learning and to boost perceptual and neural processing (Golisch et al., 2017; Korostenskaja et al., 2007; Light et al., 2017; Swerdlow et al., 2016, 2020). As these previous studies typically used auditory stimuli, our results with visual stimuli add to a growing body of literature indicating that memantine can enhance sensory function, instead of decreasing it, and is the first study explicitly linking memantine to recurrent processing.

At this point we can only speculate why NMDA receptor blockade via memantine increased (rather than decreased) decoding accuracy in the condition measuring illusion-specific feedback. One speculative possibility is that NMDA receptor blockade may have reduced signal variability or entropy, resulting in more stereotypical neural processing, which may have improved classifier performance. The effect of memantine on the excitation/inhibition (E/I) balance in human participants with schizophrenia, as measured noninvasively by computing the aperiodic, 1/f profile of the EEG power spectrum (Molina et al., 2020), may provide one explanation. A flatter 1/f power spectrum is thought to reflect noisier signal processing. Changes in E/I balance have been linked to variations in states of consciousness (He et al., 2010; Nir et al., 2008), and may also affect decoding performance due to changes in neural noise. As mentioned in the Introduction, this may be related to memantine modulating processing at other pre- or post-synaptic receptors present at NMDA-rich synapses, specifically affecting extrasynaptic NMDA receptors in GABAergic cells (Milnerwood et al., 2010; Yao et al., 2022). Memantine's strong effect on extrasynaptic NMDA receptors in GABAergic cells leads to increases in ACh levels, which have been shown to increase firing rates and reduce firing rate variability in macaques (Herrero et al., 2013, 2008). This may represent a mechanism through which memantine (but not ketamine or the NMDA receptor antagonists used by Self and colleagues) could boost visually evoked activity. Clearly, further research is needed to elucidate the precise mechanisms by which memantine modulates feedback processing, how EEG decoding may relate to signal variability, and to explore its implications for perception.

However, our behavioral results show that improved decoding accuracy under memantine was not associated with boosted conscious perception. There are several possible reasons for this lack of behavioral correlate. For example, EEG decoding may be a more sensitive measure of the neural effects of memantine, in particular given that perceptual sensitivity may have been at floor (masked condition, experiment 1) or ceiling (unmasked condition, experiment 1, and experiment 2). It is also

412 possible that the present decoding results are merely epiphenomenal, not mapping onto functional
413 improvements (e.g., Williams et al., 2007). However, given that we found a tight link between these
414 EEG decoding markers and behavioral performance in our previous work (Fahrenfort et al., 2017;
415 Noorman et al., 2023), it is possible that the effect of memantine was just too subtle to show up in
416 changes in overt behavior. Similarly, although we cannot fully exclude that the effect of memantine
417 reflected differential perceptual learning via memantine's impact on LTP, we did not find evidence for
418 such an effect based on perceptual learning: dividing the experiments in four time bins did not reveal
419 differences or changes over time in illusion detection between placebo and memantine (see
420 **supplementary figure S6**).

421 Interestingly, we found an effect of memantine on illusion decoding only for temporally
422 attended (experiment 1, AB task) and task-relevant Kanizsa stimuli (both experiments) that were
423 presented without a backward mask. Although these results have to be interpreted with caution
424 because the key interaction effects were not statistically significant, t his suggests that object-based
425 attention is crucial for memantine to exert its effects. Previous work has revealed that object-based
426 attention primarily enhances neuronal firing rates compared to unattended objects, especially during
427 the recurrent processing phase (Roelfsema, 2023; Thiele & Bellgrove, 2018). Attention-related neural
428 enhancement can be observed in many regions throughout cortex, e.g., in early and intermediate
429 visual regions, as well as inferotemporal, parietal, and even frontal cortices (reviewed in Roelfsema,
430 2023). It is hypothesized that when an object is attended, this facilitates the integration of its local
431 elements and that it enables figure-ground segmentation to occur, crucial for the perception of
432 coherent shapes and for the phenomenological properties of the Kanizsa illusion (Narayanan et al.,
433 2013). The observation that the effect of inhibiting NMDA receptors by memantine was only found
434 for attended objects may suggest that memantine is most effective in conditions where recurrent
435 processing is strongest. Previous work has shown that memantine also affects cholinergic receptors
436 (Glasgow et al., 2017; Johnson et al., 2015), known to be involved in attentional processes (Herrero et
437 al., 2013, 2008; Thiele & Bellgrove, 2018), which may partly explain why the effect of memantine might
438 be dependent on attentional allocation. This preliminary finding deserves future investigation to
439 further understand its mechanistic origin.

440 In conclusion, our study demonstrates that memantine exerts largely selective effects on
441 feedback processing during human perceptual inference, and these effects are contingent upon the
442 allocation of attention. These findings contribute to our understanding of the neural mechanisms
443 underpinning consciousness and highlight the importance of considering attentional factors in the
444 study of recurrent processing. Further research is needed to elucidate the precise mechanisms by
445 which memantine modulates feedback processing and to explore its implications for our
446 understanding of perceptual inference and consciousness.

447 Methods

448 Experiment 1

449 Participants

450 Thirty-three right-handed men with normal, or corrected-to-normal, vision were recruited after they
451 had passed extensive physical and mental screening (for exclusion criteria see Supplementary
452 Material). Three of them failed at the screening session's training task and did not proceed to the two
453 experimental sessions (see Procedure). One dropped out during his first experimental session because
454 of a headache, which was unrelated to memantine as he had taken the placebo. The remaining 29
455 participants (22±3 years old) were analyzed. The study was approved by the Medical Ethical
456 Committee of the Amsterdam Medical Centre (AMC) (NL64341.018.18) and the local ethics committee
457 of the University of Amsterdam (2022-BC-14355). Participants gave informed consent and received
458 monetary compensation.

Stimuli

The target stimulus set had a 2 Kanizsa illusion levels (absent/present) × 2 collinearity levels (absent/present) × 2 rotation levels (absent/present) design, resulting in eight stimuli (**Fig. 1A**). Three aligned Pac-Men induced the Kanizsa illusion. Collinearity was present when the stimuli's three other elements (the "two-legged white circles") were aligned. The controls for both the illusion and collinearity were made by rotating their elements 90 degrees. Differences in local contrast were created by rotating the entire stimulus 180 degrees. The targets spanned 7.5 degrees by 8.3 degrees of visual angle. The distance between the three Pac-Man stimuli as well as between the three aligned two-legged white circles was 2.8 degrees of visual angle. Although neuronal responses to collinearity in primary visual cortex are most robust when this distance is smaller (Kapadia et al., 1995, 2000), longer-range horizontal connections between neurons with similar orientation selectivity can span distances corresponding to visual angles considerably greater than 2.8 degrees (Bosking et al., 1997; Stettler et al., 2002).

The distractor stimulus set was the same as the target stimulus set, with two exceptions. First, the distractors were red instead of black. Second, the distractors' six elements were rotated 180 degrees relative to the targets', so neither the illusion nor collinearity were present in the distractors. Masks consisted of six differently shaped elements, all capable of covering the targets' elements. Six masks were created by rotating the original mask five times by 60 degrees. They spanned 8.5 degrees by 9.1 degrees of visual angle. Finally, the design of the fixation cross was adapted from Thaler et al. (2013).

Task designs

The training dataset for our decoding analyses was obtained from a separate "template matching" task performed in the intake session. No drugs were administered during this "localizer" task. Here, every 900-1100 ms, a target stimulus was presented for 50 ms. Participants had to indicate whether the target matched the "template", one of the eight target stimuli that was presented before the start of a block of trials. Each stimulus served as the template once. In short, the task design was 8 templates × (7 mismatches × 16 presentations + 1 match × 48 presentations), totaling 1280 trials, 3/10 of which were matches.

The main task was a rapid serial visual presentation (RSVP) task. Here, on every trial, participants were shown two targets (T1 and T2) within a RSVP of distractors (**Fig. 1B**). The targets and distractors had a stimulus onset asynchrony of 100 ms. T2 and each distractor were presented for 17 ms. With the aim of inducing an attentional blink, attending to T1 was made easier by presenting it for 67 ms. T1 was preceded by five distractors and T2 was followed by six distractors. For the first six participants, T2 and each distractor were presented for 10 ms and T1 for 40 ms (using the monitor's 100 Hz refresh rate). These durations were changed to make the task easier, so fewer volunteers would need to be excluded (see Procedure).

T2 visibility was manipulated in two ways. First, half the T2s were unmasked. The other half was immediately followed by three masks, each presented for 17 ms with an interstimulus interval of 0 ms. The three masks were selected randomly, but all differed from each other. Masking drastically reduced T2 visibility. Second, half the lags between the two targets were long: eight distractors (i.e., 900 ms). The other half was short: one or two distractors (i.e., 200 or 300 ms). The participant's short lag duration was determined beforehand, during the screening session (see Procedure). Short lags often induced an attentional blink (i.e., missing T2s). Following the RSVP sequence, participants indicated for each target whether it contained the Kanizsa illusion or not. Response screens lasted until the response. Response button mapping was counterbalanced within the task. In short, each experimental session had an 8 T1 stimuli × 8 T2 stimuli × 2 masking levels (absent/present) × 2 lags (long/short) task design, resulting in 256 conditions. Each condition was presented four times, totaling 1024 trials.

For both the RSVP and the template matching task, participants were instructed to fixate on the fixation cross.

Procedure

The experiment consisted of three separate sessions conducted on different days: a four-hour screening session and two seven-hour experimental sessions. At the beginning of the screening session, the volunteers were screened for contraindications to using memantine. If there were no contraindications, participants proceeded to performing the template matching task with EEG and training for the RSVP task. To proceed to the experimental sessions, participants had to score well enough on this training task. Three participants were excluded for failing to achieve this. The training task was also used to discover the duration of the short lag that led to the lowest score to reliably induce an attentional blink. For all participants T2 performance at short lags was more than 5% lower than at long lags.

Memantine (20 mg) and the placebo were administered in different experimental sessions, separated by at least a week. These sessions were conducted using a randomized, double-blind crossover design. The order of drug administration (placebo or memantine first) was counterbalanced across participants. Participants were required to refrain from recreational drugs four weeks before the experiment and to refrain from alcohol two days before each experimental session. Experimental sessions started at 9:00 with the pill intake and briefly practicing the RSVP task. The actual RSVP task was performed when memantine reached peak plasma level, four hours after the pharmaceutical was administered. After this task, five minutes of resting state data were collected and another task, the local-global task (Bekinschtein et al., 2009), was performed. These latter two paradigms are beyond the scope of this paper.

Before pill intake (~9:00) and before (~13:00) and after (~15:30) the tasks, participants filled in the visual analogue scale (VAS) to measure the effects of memantine on subjective state (Bond & Lader, 1974; Danion et al., 1989). Tasks were programmed in Presentation software (Neurobehavioral Systems) and displayed on a 23 inch, 60 Hz, 1920×1080 pixels monitor, viewed at an 80 cm distance with the use of a chin rest.

Behavioral analysis

To investigate behavioral performance on the RSVP task, we used perceptual sensitivity (d') and bias (criterion, see supplementary figure **S1**) derived from signal detection theory. The attentional blink requires T1 to be attended, so trials with incorrect T1s were excluded when T2 performance was investigated. To investigate the VAS measurements, the mean scores of subsets of the scales were calculated and taken as measures of alertness, contentedness, and calmness following Bond and Lader (1974) and of sedation following Danion et al. (1989). VAS scores of the two measurements after pill intake were averaged and corrected by calculating percentage change from the baseline measurement. For all behavioral analyses, repeated measures ANOVAs and double-sided (Bayesian) t-tests were used to test the differences between our conditions.

EEG recording and preprocessing

EEG was recorded at 1024 Hz using a 64 channel ActiveTwo system (BioSemi). Four electrooculographic (EOG) electrodes measured horizontal and vertical eye movements. The data were analyzed with MATLAB (MathWorks). For most of the preprocessing steps, EEGLAB was used (Delorme & Makeig, 2004). The data were re-referenced to the earlobes. Poor channels were interpolated. High-pass filtering can cause artifacts in decoding analyses, so for slow-drift removal we used trial-masked robust detrending (van Driel et al., 2021). Each trial was epoched from -250 to 1000 ms relative to target onset. To improve our independent component analysis (ICA), baseline correction was applied using the whole epoch as baseline (Groppe et al., 2009). ICA was used to remove blinks. Blink components were removed manually. Baseline correction was applied, now using a -250 to 0 ms window relative to target onset. Trials with values outside of a -300 to 300 microvolts range were

removed. We used an adapted version of FieldTrip's ft_artifact_zvalue function to detect and remove trials with muscle artifacts (Oostenveld et al., 2011). As in the behavioral analyses, T2s preceded by incorrect T1s were excluded. Finally, the data were downsampled to 128 Hz.

Multivariate pattern analysis

We used the Amsterdam Decoding and Modeling (ADAM) toolbox to perform multivariate pattern analyses (MVPAs) on the EEG data (Fahrenfort et al., 2018). Using a tenfold cross-validation scheme, we first analyzed the template matching task's data. We decoded each visual feature (e.g., the Kanizsa illusion) by training linear discriminant classifiers to discriminate between the feature's absence and presence based on the raw EEG activity [across all electrodes]. Individual participants' data were split into ten equal-sized folds after randomizing the task's trial order. A classifier was then trained on nine folds and tested on the tenth one, ensuring independence of the training and testing sets. This procedure was repeated until each fold served as the test set once. Classifier performance, here area under the receiver operating characteristic (AUC) (Hand & Till, 2001), was averaged across all ten iterations. We applied "diagonal decoding": classifier performance was calculated for every time sample and the classifiers were tested on the same time sample they had been trained on.

Each decoded visual feature showed a clear peak in classifier performance over time (**Fig. 1E**). Many of our analyses focus on the time windows encompassing these peaks: from 78 to 94 ms for local contrast, 192-292 ms for collinearity, and 223-323 ms for the illusion (following the same procedure to determine the different time windows as in our previous work, see Noorman et al., 2023). To obtain topographic maps showing the neural sources of the classifier performance, we multiplied the classifier weights with the data covariance matrix, yielding covariance/class separability maps (Haufe et al., 2014). The covariance/class separability maps of the three time windows showed that the classifiers depended on occipitoparietal electrodes during peak decoding accuracy (**Fig. 2A**). We restricted the decoding analyses to using only these electrodes to ensure that any effects we observed were not due to poor signal-to-noise ratio.

Next, classifiers were trained on the template matchings task's data and tested on each condition of the RSVP task's data. We averaged each visual feature's time window in the training set's raw EEG data to increase power. Classifiers were trained on these averages and tested on all time samples of the test set, resulting in "off-diagonal decoding" (King & Dehaene, 2014). To establish the putative markers for the different forms of neural processing, each visual feature was decoded in this manner. When decoding a visual feature, the average of the same feature's time window was used as the training set. Off-diagonal decoding allowed us to focus on the neural processes of interest. For completeness, we also performed a diagonal decoding version of this train-test scheme. Here, classifiers trained on a time sample of the template matching task's data were tested on the same time sample of the RSVP task's data. The diagonal decoding results were qualitatively similar to the off-diagonal ones (**supplementary figures S2-4**).

For the time samples from -100 to 700 ms relative to target onset, we used a double-sided *t*-test to evaluate whether classifier performance differed from chance and whether the memantine conditions' classifier performance differed from the placebo conditions'. We used cluster-based permutation testing (10,000 iterations at a threshold of 0.05) to correct for multiple comparisons (Maris & Oostenveld, 2007). As in the behavioral analyses, repeated measures ANOVAs and double-sided (Bayesian) *t*-tests—here applied to the time windows' average classifier performance—were used to test the differences between our conditions.

Experiment 2

Participants

Thirty right-handed men with normal, or corrected-to-normal, vision were recruited after they had passed extensive physical and mental screening. Two of them dropped out after the first experimental session because of personal reasons unrelated to memantine. EEG recording failed during an

experimental session of one participant. The remaining 27 participants (22±2 years old) were analyzed. The study was approved by the Medical Ethical Committee of the Amsterdam Medical Centre (AMC) and the local ethics committee of the University of Amsterdam. Participants gave informed consent and received monetary compensation.

Stimuli

Stimuli were the same as in experiment 1, except for a small change to the visual feature collinearity, for which we removed the "feet" from the two-legged white circles (**Fig. S5**).

Task designs

The training dataset for our decoding analyses was obtained from a separate "localizer" task, performed in all three sessions. In the intake, no drugs were administered. In the other sessions the localizer was performed well before memantine reached peak plasma levels. Here, in separate blocks of trials each visual feature was task-relevant. On every trial, a target stimulus was presented for 83 ms. Participants had to fixate on the fixation cross and indicate whether the current task-relevant feature (e.g., the Kanizsa illusion) was present or absent. The (first) response window started at target onset and ended after 900 ms. If no response was made, "Respond faster!" was presented until response. Trials ended with 0-200 ms jitter and at least 300 ms between response and jitter. The order of the task-relevant features was counterbalanced over participants and, because each participant performed the task multiple times, kept the same within participants. Response button mapping was counterbalanced within the task. In short, the task design was 3 task-relevant features × 8 target stimuli × 72 presentations, totaling 1728 trials. Participants completed the task three times in separate sessions, for a total of 5184 trials (1728 trials per session). All trials were used to build the classifiers.

The main task the participants performed was the "task-relevance" task (**Fig. 4A**), which was the same as the localizer task, except for the following changes. Half the trials were masked and on these trials, the target stimulus was presented for 33 ms and immediately followed by three masks, the same way as in the RSVP task. As in the localizer task each visual feature was at one point task-relevant and thus attended, while at another point task-irrelevant and thus less attended. The task also contained a base-rate manipulation, whose effects are beyond the scope of this paper. In one block of trials, the task-relevant feature was present on 75% of the trials, while on the next block, it was present on 25% of the trials, or vice versa. Overall, the task-relevant feature's was equally often present and absent, which made it possible to ignore the base-rate manipulation for the purpose of the present analyses. In short, the task design was 3 task-relevant features × 8 target stimuli × 2 masking levels (absent/present) × 80 presentations, totaling 3840 trials.

To achieve masked performance of ~75% correct (making it more comparable to a typical attentional blink performance), we staircased mask contrast using the weighted up-down method (Kaernbach, 1991). Staircasing was done when the Kanizsa illusion was task-relevant, because masking affected this feature more than the other two. Contrast levels ranged from 0 (black) to 255 (white). Mask contrast started at level 230. Each correct response made the task more difficult: masks got darker by downward step size S_{down}. Each incorrect response made the task easier: masks got lighter by upward step size S_{up}. Step sizes were determined by $S_{up} \times p = S_{down} \times (1 - p)$, where p was the desired accuracy of 0.75. The smallest step size was always 13 contrast levels. The staircase ended after 25 reversals (a mistake after a correct response, or a correct response after a mistake) . The mask contrast that was used during the experimental sessions was the average contrast level of the last 20 reversals.

Procedure

The procedure was the same as for experiment 1, except for the following changes. During the screening session, if the volunteers had no contraindications to using memantine, they now proceeded to performing the localizer task and practicing the task-relevance task. This practice included our staircase procedure. The mask strength obtained here was used during the experimental sessions. To increase the power of the decoding analyses, the pill intake at the start of the

experimental sessions was now immediately followed by another session of the localizer task. The localizer task was followed by briefly practicing the task-relevance task. The actual task-relevance task was performed four hours after the pharmaceutical was administered. No additional tasks were performed. The experiment was conducted in two different types of labs. For 18 participants, tasks were displayed on a 23 inch, 1920 × 1080 pixels monitor and for the other participants, on a 27 inch, 2560 × 1440 pixels monitor.

Behavioral analysis

Behavioral analysis was the same as for experiment 1. One of the baseline VAS forms of one participant was not filled in, so he was excluded from the VAS analysis.

EEG recording and preprocessing

EEG recording and preprocessing were the same as for experiment 1.

Multivariate pattern analysis

Multivariate pattern analysis was the same as for experiment 1, except for the three instances of the localizer task now together serving the role of the template matching task and the task-relevance task serving the role of the RSVP task. When decoding a visual feature (e.g., the Kanizsa illusion), not all of the localizer task's trials were used, only those during which the same visual feature was task-relevant. The tenfold cross-validation scheme was applied to the localizer task and each decoded visual feature again showed a clear peak in classifier performance over time (**Fig. 4D**). Here, the time windows encompassing these peaks were from 78 to 94 ms for local contrast, 247-347 ms for collinearity, and 217-317 ms for the illusion. Covariance/class separability maps again showed that the classifiers depended on occipitoparietal electrodes during peak decoding accuracy, so our decoding analyses were restricted to using only these electrodes.

References

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., Gradinaru, V., Deverman, B. E., Luo, L., & Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*, *94*(4), 891-907.e6.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677.

Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *The British Journal of Medical Psychology*, *47*(3), 211–218.

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(6), 2112–2127.

Ciceri, S., Oude Lohuis, M. N., Rottschäfer, V., Pennartz, C. M. A., Avitabile, D., van Gaal, S., & Olcese, U. (2023). The neural and computational architecture of feedback dynamics in mouse cortex during stimulus report. In *bioRxiv* (p. 2023.07.19.549692). https://doi.org/10.1101/2023.07.19.549692

Danion, J. M., Zimmermann, M. A., Willard-Schroeder, D., Grangé, D., & Singer, L. (1989). Diazepam induces a dissociation between explicit and implicit memory. *Psychopharmacology*, *99*(2), 238–243.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19*(9), 1488–1497.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2008). The spatiotemporal profile of cortical processing leading up to visual perception. *Journal of Vision*, *8*(1), 12.1-12.

Fahrenfort, J. J., Snijders, T. M., Heinen, K., van Gaal, S., Scholte, H. S., & Lamme, V. A. F. (2012). Neuronal integration in visual cortex elevates face category tuning to conscious face perception. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21504–21509.

Fahrenfort, J. J., van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA Using the Amsterdam Decoding and Modeling Toolbox (ADAM). *Frontiers in Neuroscience*, *12*, 368.

Fahrenfort, J. J., van Leeuwen, J., Olivers, C. N. L., & Hogendoorn, H. (2017). Perceptual integration without conscious access. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(14), 3744–3749.

Flohr, H., Glade, U., & Motzko, D. (1998). The role of the NMDA synapse in general anesthesia. *Toxicology Letters* , *100–101*, 23–29.

Gilbert, C. D., & Wiesel, T. N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature*, *280*(5718), 120–125.

Glasgow, N. G., Povysheva, N. V., Azofeifa, A. M., & Johnson, J. W. (2017). Memantine and Ketamine Differentially Alter NMDA Receptor Desensitization. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(40), 9686–9704.

Golisch, A., Heba, S., Glaubitz, B., Tegenthoff, M., & Lissek, S. (2017). Enhancing Effects of NMDA-Receptor Blockade on Extinction Learning and Related Brain Activation Are Modulated by BMI. *Frontiers in Behavioral Neuroscience*, *11*, 34.

Groppe, D. M., Makeig, S., & Kutas, M. (2009). Identifying reliable independent components via split-half comparisons. *NeuroImage*, *45*(4), 1199–1211.

Halgren, E., Mendola, J., Chong, C. D. R., & Dale, A. M. (2003). Cortical activation to illusory shapes as measured with magnetoencephalography. *NeuroImage*, *18*(4), 1001–1009.

Hand, D. J., & Till, R. J. (2001). A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Machine Learning*, *45*(2), 171–186.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. -D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110.

He, B. J., Zempel, J. M., Snyder, A. Z., & Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron*, *66*(3), 353–369.

Herrero, J. L., Gieselmann, M. A., Sanayei, M., & Thiele, A. (2013). Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. *Neuron*, *78*(4), 729–739.

Herrero, J. L., Roberts, M. J., Delicato, L. S., Gieselmann, M. A., Dayan, P., & Thiele, A. (2008). Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature*, *454*(7208), 1110–1114.

Johnson, J. W., Glasgow, N. G., & Povysheva, N. V. (2015). Recent insights into the mode of action of memantine and ketamine. *Current Opinion in Pharmacology*, *20*, 54–63.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, *49*(3), 227–229.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*(4), 843–856.

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, *84*(4), 2048–2062.

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210.

Kocsis, B., Brown, R. E., McCarley, R. W., & Hajos, M. (2013). Impact of ketamine on neuronal network dynamics: translational modeling of schizophrenia-relevant deficits. *CNS Neuroscience & Therapeutics*, *19*(6), 437–447.

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Current Biology: CB*, *26*(3), 371–376.

Kok, P., & de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology: CB*, *24*(13), 1531–1535.

Korostenskaja, M., Nikulin, V. V., Kicić, D., Nikulina, A. V., & Kähkönen, S. (2007). Effects of NMDA receptor antagonist memantine on mismatch negativity. *Brain Research Bulletin*, *72*(4–6), 275–283.

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*(3), 204–220.

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (1998). Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(6), 3263–3268.

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience*, *14*(7), 1044–1053.

Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 1907–1911.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, *10*(4), 903–940.

Liang, H., Gong, X., Chen, M., Yan, Y., Li, W., & Gilbert, C. D. (2017). Interactions between feedback and lateral connections in the primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(32), 8637–8642.

Light, G. A., Zhang, W., Joshi, Y. B., Bhakta, S., Talledo, J. A., & Swerdlow, N. R. (2017). Single-Dose Memantine Improves Cortical Oscillatory Response Dynamics in Patients with Schizophrenia. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *42*(13), 2633–2639.

Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., Ohkura, M., Yamanaka, A., Yanagawa, Y., Nakai, J., Hayashi, Y., Larkum, M. E., & Murayama, M. (2015). A Top-Down Cortical Circuit for Accurate Sensory Perception. *Neuron*, *86*(5), 1304–1316.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Mashour, G. A., & Hudetz, A. G. (2018). Neural Correlates of Unconsciousness in Large-Scale Brain Networks. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2018.01.003

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798.

Mathew, S. J., Shah, A., Lapidus, K., Clark, C., Jarun, N., Ostermeyer, B., & Murrough, J. W. (2012). Ketamine for treatment-resistant unipolar depression: current evidence. *CNS Drugs*, *26*(3), 189–204.

Meuwese, J. D. I., van Loon, A. M., Scholte, H. S., Lirk, P. B., Vulink, N. C. C., Hollmann, M. W., & Lamme, V. A. F. (2013). NMDA receptor antagonist ketamine impairs feature integration in visual perception. *PloS One*, *8*(11), e79326.

Milnerwood, A. J., Gladding, C. M., Pouladi, M. A., Kaufman, A. M., Hines, R. M., Boyd, J. D., Ko, R. W. Y., Vasuta, O. C., Graham, R. K., Hayden, M. R., Murphy, T. H., & Raymond, L. A. (2010). Early increase in extrasynaptic NMDA receptor signaling and expression contributes to phenotype onset in Huntington's disease mice. *Neuron*, *65*(2), 178–190.

Molina, J. L., Voytek, B., Thomas, M. L., Joshi, Y. B., Bhakta, S. G., Talledo, J. A., Swerdlow, N. R., & Light, G. A. (2020). Memantine Effects on Electroencephalographic Measures of Putative Excitatory/Inhibitory Balance in Schizophrenia. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *5*(6), 562–568.

Müller, C. P., Pum, M. E., Amato, D., Schüttler, J., Huston, J. P., & Silva, M. A. D. S. (2011). The in vivo neurochemistry of the brain during general anesthesia. *Journal of Neurochemistry*, *119*(3), 419–446.

Narayanan, B., Stevens, M. C., Jiantonio, R. E., Krystal, J. H., & Pearlson, G. D. (2013). Effects of memantine on event-related potential, oscillations, and complexity in individuals with and without family histories of alcoholism. *Journal of Studies on Alcohol and Drugs*, *74*(2), 245–257.

Nir, Y., Mukamel, R., Dinstein, I., Privman, E., Harel, M., Fisch, L., Gelbard -Sagiv, H., Kipervasser, S., Andelman, F., Neufeld, M. Y., Kramer, U., Arieli, A., Fried, I., & Malach, R. (2008). Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nature Neuroscience*, *11*(9), 1100–1108.

Noorman, S., Stein, T., Fahrenfort, J. J., & van Gaal, S. (2023). Isolating feedforward, lateral, and feedback mechanisms underlying perceptual and attentional impairments of conscious access. In *bioRxiv* (p. 2023.10.16.562523). https://doi.org/10.1101/2023.10.16.562523

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869.

Pak, A., Ryu, E., Li, C., & Chubykin, A. A. (2020). Top-Down Feedback Controls the Cortical Representation of Illusory Contours in Mouse Primary Visual Cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *40*(3), 648–660.

Rammsayer, T. H. (2001). Effects of pharmacologically induced changes in NMDA -receptor activity on long-term memory in humans. *Learning & Memory* , *8*(1), 20–25.

Rammsayer, T. H. (2006). Effects of pharmacologically induced changes in NMDA receptor activity on human timing and sensorimotor performance. *Brain Research*, *1073–1074*, 407–416.

Roelfsema, P. R. (2023). Solving the binding problem: Assemblies form when neurons enhance their firing rate-they don't need to oscillate or synchronize. *Neuron*, *111*(7), 1003–1019.

Rosburg, T., & Kreitschmann -Andermahr, I. (2016). The effects of ketamine on the mismatch negativity (MMN) in humans--a meta-analysis. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *127*(2), 1387–1394.

Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y., & Petersen, C. C. H. (2013). Membrane potential correlates of sensory perception in mouse barrel cortex. *Nature Neuroscience*, *16*(11), 1671–1677.

Schmidt, K. E., Goebel, R., Löwel, S., & Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *The European Journal of Neuroscience*, *9*(5), 1083–1089.

Scholte, H. S., Jolij, J., Fahrenfort, J. J., & Lamme, V. A. F. (2008). Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, *20*(11), 2097–2109.

Schugens, M. M., Egerter, R., Daum, I., Schepelmann, K., Klockgether, T., & Löschmann, P. A. (1997). The NMDA antagonist memantine impairs classical eyeblink conditioning in humans. *Neuroscience Letters*, *224*(1), 57–60.

Self, M. W., Kooijmans, R. N., Supèr, H., Lamme, V. A. F., & Roelfsema, P. R. (2012). Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(27), 11031–11036.

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews. Neuroscience*. https://doi.org/10.1038/s41583-022-00587-4

Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, *36*(4), 739–750.

Swerdlow, N. R., Bhakta, S., Chou, H.-H., Talledo, J. A., Balvaneda, B., & Light, G. A. (2016). Memantine Effects On Sensorimotor Gating and Mismatch Negativity in Patients with Chronic Psychosis. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *41*(2), 419–430.

Swerdlow, N. R., Bhakta, S. G., Talledo, J., Kotz, J., Roberts, B. Z., Clifford, R. E., Thomas, M. L., Joshi, Y. B., Molina, J. L., & Light, G. A. (2020). Memantine effects on auditory discrimination and training in schizophrenia patients. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *45*(13), 2180–2188.

Takahashi, N., Oertner, T. G., Hegemann, P., & Larkum, M. E. (2016). Active cortical dendrites modulate perception. *Science*, *354*(6319), 1587–1590.

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, *76*, 31–42.

Thiele, A. (2012). NMDA receptors figure it out [Review of *NMDA receptors figure it out*]. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(27), 10749–10750.

Thiele, A., & Bellgrove, M. A. (2018). Neuromodulation of Attention. *Neuron*, *97*(4), 769–785.

van Driel, J., Olivers, C. N. L., & Fahrenfort, J. J. (2021). High-pass filtering artifacts in multivariate classification of neural time series data. *Journal of Neuroscience Methods*, *352*, 109080.

Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel -Mathis, M.-A., Poort, J., Van Der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, *111*(40), 14332–14341.

van Loon, A. M., Fahrenfort, J. J., van der Velde, B., Lirk, P. B., Vulink, N. C. C., Hollmann, M. W., Scholte, H. S., & Lamme, V. A. F. (2016). NMDA Receptor Antagonist Ketamine Distorts Object Recognition by Reducing Feedback to Early Visual Cortex. *Cerebral Cortex* , *26*(5), 1986–1996.

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, *360*(6388), 537–542.

Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, *60*(2), 215–234.

Williams, M. A., Dang, S., & Kanwisher, N. G. (2007). Only some spatial patterns of fMRI response are read out in task performance. *Nature Neuroscience*, *10*(6), 685–686.

Wokke, M. E., Vandenbroucke, A. R. E., Scholte, H. S., & Lamme, V. A. F. (2013). Confuse your illusion: feedback to early visual cortex contributes to perceptual completion. *Psychological Science*, *24*(1), 63–71.

Wyatte, D., Jilk, D. J., & O'Reilly, R. C. (2014). Early recurrent feedback facilitates visual object recognition under challenging conditions. *Frontiers in Psychology*, *5*, 674.

Yao, L., Rong, Y., Ma, X., Li, H., Deng, D., Chen, Y., Yang, S., Peng, T., Ye, T., Liang, F., Xu, N., & Zhou, Q. (2022). Extrasynaptic NMDA receptors bidirectionally modulate intrinsic excitability of inhibitory neurons. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *42*(15), 3066–3079.

Zatka-Haas, P., Steinmetz, N. A., Carandini, M., & Harris, K. D. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *ELife*, *10*. https://doi.org/10.7554/eLife.63163