Backward masking in mice requires visual cortex

Samuel D Gale, Chelsea Strawder, Corbett Bennett, Stefan Mihalas, Christof Koch*, Shawn R Olsen*

MindScope Program, Allen Institute, Seattle, WA, 98109

*Corresponding authors: christofk@alleninstitute.org, shawno@alleninstitute.org

Abstract

Visual masking is used extensively to infer the timescale of conscious perception in humans; yet the underlying circuit mechanisms are not understood. We describe a robust backward masking paradigm in mice, in which the location of a briefly flashed grating is effectively masked within a 50 ms window after stimulus onset. Optogenetic silencing of visual cortex likewise reduces performance in this window, but response rates and accuracy do not match masking, demonstrating cortical silencing and masking are distinct phenomena. Spiking responses recorded in primary visual cortex (V1) are consistent with masked behavior when quantified over long, but not short, time windows, indicating masking involves further downstream processing. Accuracy and performance can be quantitatively recapitulated by a dual accumulator model constrained by V1 activity. The model and the animal's performance for the earliest decisions imply that the initial spike or two arriving from the periphery trigger a correct response, but subsequent V1 spikes, evoked by the mask, degrade performance for later decisions. To test the necessity of visual cortex for backward masking, we optogenetically silenced mask-evoked cortical activity which fully restored discrimination of target location. Together, these results demonstrate that mice, like humans, are susceptible to backward visual masking and that visual cortex causally contributes to this process.

Introduction

Visual masking has been used to probe the temporal dynamics of conscious perception and unconscious processing for well over a century (Bachmann and Francis, 2013; Breitmeyer and Ogmen, 2006; Dehaene, 2014). In backward masking, the visibility of a target stimulus is suppressed by a second stimulus—the mask—that follows the target in time. This retroactive influence of the mask can be accounted for by multiple mechanisms, including lateral inhibitory interactions between feedforward circuits in the visual pathway (Alpern, 1953; Battersby and Sturr, 1970; Breitmeyer and Ogmen, 2006; Macknik and Martinez-Conde, 2004a; Öğmen, 1993), disruption of recurrent cortical feedback (Lamme et al., 2002; Ro et al., 2003), or long temporal integration of feedforward sensory signals causing the perception of target and mask to merge (Eriksen and Hoffman, 1963; Schultz and Eriksen, 1977; Thompson, 1966). The inter-ocular transfer of backward masking under some conditions (Kinsbourne and Warrington, 1962a; Schiller, 1965) suggests the involvement of V1, where signals from both eyes are first processed. Depending on the exact elements constituting the stimulus and the mask, the use of bulk-tissue methods (such as fMRI and EEG) have implicated a variety of brain regions in masking, including lower and higher visual cortical areas (Green et al., 2005; Haynes et al., 2005; Tse et al., 2005) parietal cortex, anterior cingulate, and thalamus (Green et al., 2005). Single unit recordings in macaque monkeys and cats show correlates of masking in retinal ganglion cells (Coenen and Eijkman, 1972), lateral geniculate nucleus (LGN) (Coenen and Eijkman, 1972; Macknik and Martinez-Conde, 2004b; Schiller, 1968), V1 (Bridgeman, 1980; Macknik and Livingstone, 1998), V2/V3 (Maeda et al., 2010), and inferotemporal areas (Kovács et al., 1995; Rolls et al., 1999; Rolls and Tovee, 1994). However, none of these studies use perturbation methods to infer circuit causality. Thus, the mechanistic determinants of masking remain largely unknown.

Mice are an essential species for dissecting brain circuits. However, whereas mice have been successfully trained in visual contrast detection (Burgess et al., 2017; Busse et al., 2011), change detection (Garrett et al., 2020; Glickfeld et al., 2013), and orientation discrimination tasks (Andermann et al., 2010; Resulaj et al., 2018), no study has yet demonstrated backward visual masking. A rat study was inconclusive regarding the feasibility of a masking paradigm (Dell et al., 2018; but see Watanabe et al., 2014), particularly because the stimulus regime used in masking (very brief, low contrast stimuli) promoted guessing behavior. We here demonstrate a robust mouse-based masking paradigm, enabling use of genetic tools to quantitatively dissect the circuits necessary for masking.

Main text

We modified a 2-alternative choice task (Burgess et al., 2017) in which mice turn a wheel left or right to indicate whether a vertical grating, the target, was presented on the right or left side, respectively, of a visual display in front of the mouse (Fig. 1A). Performance was quantified by the probability of a response in either direction (*response rate*) and the fraction of responses in the correct direction (*accuracy*). We tested animals' sensitivity to target duration and contrast (Fig. S1). For all subsequent masking and optogenetic experiments, we chose a combination of parameters for the stimulus (17 ms duration, 40% contrast) at which performance was near maximum but approached the steep portion of the psychophysical curve where response rate and accuracy decreased, and reaction times increased.

On masking trials, the mask, a vertical-horizontal plaid of the same size, location, and contrast as the target, was presented for 200 ms on both sides of the screen at variable intervals following target onset. Mice had high response rates on masked trials (Fig. 1B). We also evaluated the response rate on trials in which only the mask was presented and on “catch” trials with no visual stimulus (see Methods); mice never received rewards on these two trial types.

We found a monotonic relationship between accuracy and the time of mask onset relative to target onset (Fig. 1C). For the earliest mask onset (17 ms), the mask appeared immediately after the target and accuracy was strongly, and highly significantly (Fig. S2A,B), reduced. Accuracy recovered to near non-masked levels when mask onset was delayed by 50 ms following target onset. Across mice, there was a significant difference in accuracy across all mask onsets (Fig. S2B). 15/16 mice showed a significant Spearman correlation ($r>0.9$, $p<0.05$) between mask onset and accuracy. The response rate on target-only trials was reduced relative to trials with a mask (Fig. 1B, S2A) and to sessions that did not include masks (Fig. S1A,D); yet, high accuracy on target-only trials persisted (Fig. 1C). Results were similar in wild-type and VGAT-ChR2 mice (Fig. S2C,D).

These results indicate that the mask disrupts target localization when presented within ~50 ms after target onset. This time course is remarkably like backward visual masking in humans and monkeys (see Discussion).

Previous experiments using a similar task, but without masking, showed that target detection depended on visual cortex (Burgess et al., 2017). To test whether disruption of ongoing target-driven activity in visual cortex is sufficient to explain the effects of masking on behavior, we replaced the mask with bilateral optogenetic inhibition of visual cortex at various times relative to target onset. We photo-suppressed excitatory neurons using VGAT-ChR2 mice in which GABAergic interneurons in cortex express channelrhodopsin (Li et al., 2019; Resulaj et al., 2018). Recording of spiking activity across all V1 layers with Neuropixels probes verified that inhibition of putative excitatory neurons was rapid (~12 ms after light onset) and sustained (Fig. S4A), as previously reported (Bennett et al., 2019; Li et al., 2019; Resulaj et al., 2018). The response rate to the target was reduced to chance levels (Fig. 2A) when visual cortex was inhibited within the time it takes V1 to respond to the target (~45 ms; Fig. S3C,D). The response rate increased when cortical inhibition began as early as 17 ms after the visual response latency (*i.e.,* 50 ms after target onset) and rose to levels observed in the absence of cortical inhibition within ~50 ms of the visual response latency (Fig. 2A, S5A,B). Accuracy on responsive trials was greater than 75% even when the response rate was only slightly above chance (Fig. 2B,C). Unilateral inhibition revealed that the effect of cortical inhibition was specific to the hemisphere contralateral to the target (Fig. S5C). No effect on response rate or accuracy was observed when blue light was delivered above visual cortex in wild-type mice (Fig. S5D,E).

Taken together, our cortical inhibition and masking experiments suggest that a window of target-evoked activity in contralateral visual cortex, lasting approximately 50 ms, is critical for detecting the presence and location of the target. However, masking and optogenetic suppression have distinct effects on behavior. Optogenetically inhibiting bilateral visual cortical activity reduces the probability of responding to the target with only a modest effect on response accuracy (Fig. 2B,C). In contrast, stimulating activity with the bilateral mask (non-informative with respect to target location) decreases the accuracy, but not the rate, of responses. Thus, the

effects of visual masking are not recapitulated by direct suppression of target-driven spiking in visual cortex.

These observations suggested a simple model in which the total bilateral activity driven by the target and mask determines the response rate of mice, while the difference in activity between the contralateral and ipsilateral visual cortex determines the accuracy of their responses. To explore this idea, we recorded spiking responses of V1 neurons in wild-type mice to the target and mask for the mask onset times shown in Fig. 1B,C. As expected, we observed responses to a contralateral but not ipsilateral target. The response to a contralateral target in the presence of a mask was not a simple sum of the responses to each of these stimuli presented alone (Fig. 3A). Instead, the earliest mask onsets appeared to accelerate the spiking response to the mask, while later mask onsets led to a reduced response to the mask. In correspondence with the response rate of mice (Fig. 1B), the cumulative number of spikes evoked by the target plus mask was larger than the response to the target alone and reached a similar level for all mask onsets of 1.3 spikes per neuron on average (range 0.1-4.1; Fig. 3B), a remarkably small number.

To better understand the extent to which V1 population activity is constraining accuracy, we considered an ideal observer with access to V1 activity. We trained linear decoders to classify target location using single trial spike counts of V1 neurons recorded contralateral or ipsilateral to the target. We tested how decoding accuracy depended on the spike integration window by gradually lengthening the time from target onset over which spikes were counted. For short integration windows, the decoder discriminated target location with high accuracy, regardless of whether and when a mask appeared (Fig. 3C). As the integration window extended over longer durations, decoder accuracy declined to levels consistent with the relative accuracy of mice for different mask onsets (Fig. 1C, 3C). Thus, while the initial wave of stimulus-evoked V1 activity is highly informative of target location, spikes evoked by the mask obscure target location over longer integration windows.

Why do mice fail to discriminate the location of the target when followed by a mask if, in principle, they could perform equally well on target-only and masking trials using the earliest wave of V1 activity? Our analysis of V1 population activity suggests that masking occurs when activity is integrated over longer time scales associated with a degraded representation of target location. This predicts that rapid decisions (and corresponding behavioral responses) will tend to be accurate, whereas delayed decisions will more likely be erroneous. Consistent with this prediction, median reaction times were much longer on incorrect trials compared to correct trials for all mask onset times and target-only trials (Fig. 3D, S6,B); the speed of wheel movements, once initiated, did not differ (Fig. S6C,D). Moreover, pooling data across mice revealed an inverse relationship between reaction time and accuracy (Fig. 3E). Across conditions, accuracy was highest for the earliest responses and declined to chance with time. While the distribution of reaction times was similar for all mask onset times (Fig. S6B), the decline in accuracy with respect to reaction time began sooner with earlier mask onsets (Fig. 3E); hence, a greater proportion of choices for early mask onsets were mistakes (Fig. 1C).

Drift diffusion and accumulator models of perceptual decision-making account for performance and reaction times in a variety of tasks (Gold and Shadlen, 2007; Smith and Ratcliff, 2004). We studied whether such models could explain the response rate, accuracy, and reaction times in our visual masking paradigm. To capture the left/right decision of the mouse, we used a dual accumulator model in which evidence supporting each choice accumulates

separately amidst noise towards a threshold; choice is determined by the first accumulator to reach threshold (Fig. 3F). V1 population activity during trials with a contralateral or ipsilateral target (Fig. 3A) served as the input signal to the two accumulators after divisive normalization (see Methods). We fit the model using a grid search to minimize the sum-of-square error between the response rate and accuracy of the model and mice. The model, with six degrees of freedom, closely recapitulates the response rate and accuracy of mice (Fig. 3G,H). The relationship between decision times, which were not explicitly fit, and model accuracy showed the same inverse relationship as observed for reaction times in mice (Fig. 3I,J, S6E).

Our data and model suggest that high accuracy on target-only trials can be explained by the relative difference in cortical spiking across the two hemispheres, and that the bilateral mask reduces accuracy by narrowing this difference. If the critical site where mask-evoked activity disrupts the target signal is in visual cortex, as opposed to upstream circuitry (retina, thalamus) or alternative visual pathways (e.g., superior colliculus), our results predict that selectively suppressing mask-evoked activity in visual cortex should rescue response accuracy (*i.e.,* prevent backward masking). To test this prediction, we inhibited cortical activity bilaterally at various times relative to the mask (using the earliest and most effective mask onset time, 17 ms; Fig. S4B). For the earliest onset of visual cortex inhibition that allowed a response rate greater than chance, response accuracy on masking trials was increased to a level not significantly different from accuracy on target-only trials (Fig. 4A,B, S8A,B). This improvement in accuracy decreased monotonically as cortical inhibition was delayed relative to target onset, allowing more mask-evoked spikes. The optogenetic light stimulus had no effect in wild-type mice (Fig. S8C,D). Simulated inhibition of the cortical input signals to our model (Fig. 3F) likewise resulted in a recovery of accuracy during masking trials with a similar time dependence as observed in mice (Fig. 4C,D). Thus, appropriately timed inhibition of visual cortex that preserves the earliest target-evoked spikes but suppresses mask-evoked activity prevents backward visual masking.

Discussion

We successfully developed a backward visual masking paradigm to explore the precise timing and circuits underlying visual perception in a model organism amenable to mechanistic circuit dissection. Our masking paradigm uses a variant of a 2-alternative choice task developed by the International Brain Laboratory to study spatial decision-making (Burgess et al., 2017; IBL et al., 2021). We modified the task to use targets of very short duration presented at one of two possible locations, combined with a bilateral mask. Variations of this task could study other forms of masking, such as metacontrast masking (Alpern, 1953) in which the spatial location of the mask does not overlap the target.

The target was susceptible to masking for ~50 ms after stimulus onset, which is remarkably similar to the timescale of masking in humans (Alpern, 1953; Bachmann and Francis, 2013; Bacon-Macé et al., 2005; Breitmeyer and Ogmen, 2006; Dehaene, 2014; Eriksen and Lappin, 1964; Kinsbourne and Warrington, 1962b; Kolers, 1962; Schiller, 1965; Sperling, 1965) and monkeys (Kovács et al., 1995; Macknik and Livingstone, 1998; Rolls et al., 1999). The effect of the mask decreased monotonically with duration from target onset, typical for our parameters in which the mask spatially overlaps the target and is presented for a relatively longer duration (Bachmann and Francis, 2013; Breitmeyer and Ogmen, 2006). This shared

phenomenology across mammalian species suggests our masking paradigm can be a powerful tool for dissecting circuit mechanisms.

By silencing visual cortical activity at precise time points following the mask, we could ameliorate the behavioral effects of masking. This demonstrates a causal role of visual cortex and indicates that target-mask interactions in the retina, thalamus, and/or superior colliculus are insufficient on their own to generate masking in our task. Studies in humans have reported release from masking effects. The visibility of the mask stimulus itself can be suppressed with a second mask, leading to recovery of target perception (Dember and Purcell, 1967; Robinson, 1966). In addition, transcranial magnetic stimulation of visual cortex following the mask can increase target detection (Amassian et al., 1993; Ro et al., 2003). However, the exact cortical perturbation caused by TMS in masking experiments is not well understood and could involve inhibition of neurons, added activity that produces noise, or phosphene generation that serves as a mask (Bachmann and Francis, 2013). In contrast, we directly measured and suppressed mask-evoked spiking in visual cortex in a well-controlled manner to show it is necessary for backward masking.

Our data is compatible with the hypothesis that the initial one or two spikes in the wave of target-evoked activity arriving in visual cortex are sufficient for an ideal observer to decode the target with high accuracy, consistent with results in humans (Thorpe et al., 1996), monkeys (Kovács et al., 1995; Rolls and Tovee, 1994) and mice (Resulaj et al., 2018). Indeed, our decoding analysis indicates that, on average, neurons respond to the target stimulus with less than one action potential in the relevant time window (Fig. 3B,C). In the presence of a mask, the target location becomes ambiguous as downstream decision stages integrate over longer durations. This likely reflects sensory integration mechanisms and decision thresholds tuned for noise reduction rather than resistance to masking. The brain regions performing this temporal integration of visual cortical signals remain unknown.

Interestingly, even in the face of a mask, accuracy on the fastest reaction time trials can be very high (~90%). Studies in humans have also shown that the fastest behavioral responses can evade masking in tasks using either saccadic (Crouzet et al., 2014) or manual responses (Bacon-Macé et al., 2005; Van Rullen and Koch, 2003); however, absolute accuracy in these studies is much lower on the fastest reaction time trials (e.g. ~60% in Crouzet et al. 2014), which could reflect shorter neural processing times for the simpler discrimination in our task. Escape from masking has been interpreted as evidence that the mask disrupts reentrant feedback processing while sparing the initial feedforward wave (Crouzet et al., 2017, 2014; Van Rullen and Koch, 2003). Our dual accumulator model suggests that the lack of integration of feedforward target and mask signals before reaching a decision threshold can also explain the similar accuracy on fast reaction time trials for target-only and masking trials. Similar models have been successful in a variety of contexts (Gold and Shadlen, 2007; Smith and Ratcliff, 2004), but rarely have been applied to visual masking (Ratcliff and Rouder, 2000; Vorberg et al., 2003), and never with a neuronally constrained model. Our work suggests that such models can help bridge perceptual decision-making and backward visual masking.

In humans, visual masking goes hand in hand with a loss of perceptual visibility of the target. Whether the same phenomenology exists in mice, in particular at this fast time scale during which processing is usually unconscious in humans (Van Rullen and Koch, 2003), remains unanswered for now. It may be possible to adapt our task to include measures of subjective confidence (Kepecs et al., 2008; Kepecs and Mainen, 2012).

Acknowledgements We thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support. We thank the Tiny Blue Dot Foundation for providing funding for this study. We thank the Neurosurgery and Behavior team for help with mouse surgeries. We thank Sarah Naylor for support with program management. We thank Masataka Watanabe and Doug Ollerenshaw for discussion of masking paradigms in rodents.

Author contributions SDG, CB, CK, and SRO conceived the project. SDG, CS, and CB developed the experimental methodology. SDG and CS performed the experiments. SDG and CB analyzed the data. SM provided advice on modeling and data analysis. CK and SRO supervised all aspects of the project. All authors contributed to writing the manuscript.

Competing interests The authors declare no competing interests.

Methods

Mice

All experiments were performed at the Allen Institute using protocols approved by the Allen Institute's Institutional Animal Care and Use Committee. We used 8 wild-type mice (C57BL6J; 3 female, 5 male) and 10 VGAT-ChR2 mice (6 female, 4 male) (Zhao et al., 2011) for experiments. Two of the VGAT-ChR2 mice were used for the cortical inhibition experiment shown in Fig. 2 but not for masking experiments. An additional 5 wildtype mice and 3 VGAT-ChR2 mice were not used for experiments due to poor performance during training.

Mice were single-housed and maintained on a reverse 12-hour light cycle. All experiments were performed during the dark phase. Prior to the start of training on the behavioral task, mice were anesthetized with isoflurane and a headpost was attached to the skull with clear dental acrylic. After one week of recovery, mice were provided daily with an amount of water required to maintain 85% of their initial body weight, with continued unrestricted access to food. Mice were 67-145 days old (median 89) at the start of training.

Behavior Task

The behavior task was adapted from Burgess et al. 2017. Mice were head-fixed on a platform with their front paws resting on a 6 cm diameter wheel. Visual stimuli were controlled using Psychopy version 3.2.4 (Peirce, 2009) and presented at 120 frames/s on a monitor (ASUS VG248QZ; 1920 x 1080 pixels, 53.3 cm wide) centered 21.6 cm in front of the mouse. The bottom right corner of the screen (50 x 50 pixels) flipped between black and white each frame and was covered with a photodiode. The signal from the photodiode was used to determine frame times. The target stimulus was a circular vertical grating (25° diameter, 0.08 cycles/°) positioned at the center of either the left or right half of the screen (31.7° horizontal and 0° vertical from the mouse). The mask was a pair of vertical-horizontal plaids of the same size, spatial frequency, and position (both left and right sides) of the target. The outer 8% of the target and mask were cosine blurred. The luminance of the blank/gray screen was 60 cd/m^2. The target was 100% contrast during training and for testing sensitivity to target duration (Fig. S1A-C), variable contrast for testing contrast sensitivity (Fig. S1D-F), and 40% contrast for masking and cortical inhibition experiments. The mask was 40% contrast.

Each trial began with a blank, gray screen for a fixed 3 s followed by a variable duration drawn from an exponential distribution (mean 1 s, max 5 s) before the target appeared. Mice were rewarded with a drop of water (~2.5 µL) delivered from a lick spout below the mouth for rotating the wheel left when the target was on the right side of the screen, and vice versa. During initial training, the target persisted on the screen and movement of the wheel was coupled to movement of the target in the same direction. Wheel movement during the first 133 ms after target onset was ignored. The movement threshold for reward was initially 3-10 mm along the circumference of the wheel (adjusted for each mouse at the trainer's discretion). The gain between wheel and target movement was set such that the target was at the center of the screen when the reward threshold was reached. Mice were initially given a response window of up to 30 s after target onset to move the wheel past threshold. Movement past threshold in either direction (correct or incorrect) during the response window ended the trial. Lack of

movement past threshold within the response window was deemed “no response.” Trials with either an incorrect or no response were repeated up to 5 times during training.

As mice learned to turn the wheel in response to the target, the response window was gradually decreased to a final value of 650 ms from target onset. Three additional task features were introduced: (1) “catch” trials (15% of trials during training) with no visual stimulus or reward were used to monitor the rate of non-visually driven movements past threshold during the response window; (2) a quiescent period in which movement of the wheel >1mm in either direction during the last 500 ms before target onset restarted the variable gray screen period with a new random duration; and (3) and incorrect response was followed by a 1 s white noise burst and an additional 3-6 s timeout period (during training only). For some training sessions, the probability that the target appeared on the right or left was biased from the normally equal probability to correct movement biases for individual mice.

After mice were consistently performing the task at >90% accuracy with wheel and target movement coupled (16-51 training sessions; median 32), mice were trained to respond to flashed target stimuli with fixed position and duration of 33-200 ms for 1-5 sessions. The reward threshold was reduced to 2 mm. Following these sessions, sensitivity to target duration and contrast were tested (Fig. S1). During these experiments, the set of possible target durations (8, 17,33 and 100 ms) or contrasts (20, 40, 60, and 100%), and target sides (left, right), were randomly sampled before repetition. The catch trial probability was set to 11% such that there were approximately equal numbers of target-left and target-right trials (for each duration or contrast) and catch trials (Table S1).

For masking experiments (Fig. 1), on each trial there was a 60% probability that either the target (17 ms, 40% contrast) was followed by the mask (40% contrast, 200 ms) after a variable interval relative to target onset, or the mask was presented alone (mask-only trials). The set of possible mask onset times (17, 25, 33, and 50 ms relative to target onset) and target sides were sampled in random order before repetition. The other ~40% of trials were target-only trials or catch trials with no visual stimulus. Rewards were never delivered on mask-only or catch trials. The probability that a trial included no target presentation (mask-only and catch trails) was set to 11% such that there were approximately equal numbers of target-left and target-right trials (for each mask onset) and no-target trials (Table S1).

Optogenetic inhibition of visual cortex

For optogenetic inhibition of visual cortex in VGAT-ChR2 mice, or control experiments in wild-type mice, optical fibers (200 μm diameter tip) were placed bilaterally over the skull 2.7 mm lateral and 0.5 mm anterior of lambda (the skull was covered by a thin layer of clear dental acrylic during surgery). A black plastic cone was positioned above the headpost to block stray light. For all trials, there was 60% probability that blue light (470 nm, 2.5 mW) was delivered from one or both of the optical fibers. The set of possible optogenetic light onset times (17, 33, 50, 67, and 83 ms relative to target onset; Figs. 2 and 4) or brain hemispheres (Fig. S4C) and target sides were sampled in random order before repetition. The blue light persisted until the end of the behavioral response window (650 ms after target onset) and then linearly ramped off over 100 ms. The probability that a trial included no target presentation (mask-only or catch trails) was set to 33% such that there were approximately equal numbers of target-left, target-right, and no-target trials for each optogenetic light onset (Table S1).

Electrophysiology

Putative V1 recordings were made during separate experiments/days after the behavioral sessions analyzed for Figures 1, 2, and 4. The day before recording, or the same day at least 4 hours before recording, mice were anesthetized with isoflurane and a 1 mm diameter craniotomy was made 2.7 mm lateral and 0.5 mm anterior of lambda above the left hemisphere. During recordings in VGAT-ChR2 mice, the left optical fiber was placed above the skull at the medial edge of the craniotomy. Neurons were recorded with Phase 2 Neuropixels probes (128 channels arranged in two columns, with 20 μm between each recording site) (Jun et al., 2017) throughout all cortical layers. Data were acquired at 30 kHz using the Open Ephys acquisition board and GUI (Siegle et al., 2017).

At the beginning of each recording, receptive fields were coarsely mapped by presenting a 25° diameter vertical-horizontal plaid (100% contrast, 0.08 cycles/°) centered at one of 9 equally spaced positions that formed a grid across the right half of the display. The stimulus was presented for 50 ms at all 9 positions in random order (with 800 ms between presentations) before repetition. After 25-30 repeats, spikes were detected across all channels using an amplitude threshold, and post-stimulus time histograms were calculated for each stimulus position. If the peak response was greatest at the center position (where the target/mask would be presented), we continued to record during the masking task. Otherwise, we moved the probe and repeated this process, or ended the experiment.

Analysis of behavior data

Data for each experiment (Fig. 1,2,4) are based on a single session from each mouse. Trials were excluded if an unexpectedly long frame interval disrupted the timing of stimulus onset or duration (0.4% of trials across all sessions), the mouse moved the wheel >1 mm during the first 125 ms after target onset (3.4% of trials), or the mouse did not respond during the previous 10 trials (not including catch trials or trials with optogenetic inhibition; 1.2% of trials). Data from target-left and target-right trials were pooled, except for the comparison of performance on these trial types shown in Fig. S2E,F. Reaction times were calculated as the last time point after target onset that wheel displacement was greater than a movement initiation threshold of 0.2 mm before crossing the reward threshold (2 mm) in the same direction. Movement speed was calculated as the difference between the reward and movement initiation thresholds (1.8 mm) divided by the difference between the times these thresholds were crossed.

Analysis of electrophysiology data

Spikes were sorted automatically by Kilosort2 (Stringer et al., 2019). Putative single units and multi-units were accepted for further analysis if the average firing rate over the entire session was greater than 0.1 spikes/s and the majority of spikes were from a single neuron (false-positive rate less than 0.5) as estimated by comparing the rate of refractory period (1.5 ms) violations to the overall spike rate (Fig. S3A) (Hill et al., 2011; Siegle et al., 2021). Units were defined as fast-spiking (FS; putative inhibitory interneurons) if the peak-to-trough duration of the average spike waveform was less than 0.4, or otherwise as regular spiking (RS; putative

pyramidal neurons; Fig. S3B). Time to first spike (Fig. S3D) (Siegle et al., 2021) for each unit was calculated as the median across trials of the time of the first spike 30 ms or later after visual stimulus onset; trials with no spike within 150 ms of stimulus onset were excluded.

Post-stimulus time histograms (PSTHs) of neural activity aligned to visual or optogenetic stimulus onset were calculated using 8.33 ms bins (the monitor frame duration). Units were considered visually responsive if they had a significant response to either the target or the mask. A response was considered significant if (1) the peak of the PSTH after visual stimulus onset was greater than 5 times the standard deviation of the PSTH during the last 500 ms before stimulus onset, and (2) a Wilcoxon signed-rank test comparing the spike rate 40-100 ms after stimulus onset and the spike rate before stimulus onset resulted in a p-value less than 0.05.

Based on responses to the optogenetic light stimulus (Fig. S3A), units recorded in VGAT-ChR2 mice (n=267 units from 4 mice) were classified as (1) excited if the mean of the PSTH 100-500 ms after light onset was greater than the standard deviation of the PSTH during the last 500 ms before light onset; (2) transiently excited if the peak of the PSTH during the first 100 ms after light onset was greater than 5 times the standard deviation of the PSTH before light onset; or (3) inhibited if there was no transient excitation and the mean of the PSTH 100-500 ms after light onset was less than the standard deviation of the PSTH before light onset. 74 visually-responsive units that were not excited or transiently excited by the optogenetic light stimulus were analyzed for Fig. S4B.

58/276 units from 5 wild-type mice were visually responsive RS units and contribute to the data shown in Fig. 3A-C. For the decoding analysis shown in Fig. 3C, neurons were pooled across mice. For each mouse/session, trials for each target side (contralateral or ipsilateral to the recording) and mask onset condition were evenly spilt at random into training and testing groups. We then created 100 training and testing pseudo-trials for each condition by randomly selecting a trial from the training or testing group for each neuron. A linear support vector machine (SVM) was trained to classify the target side for each pseudo-trial from the training group using the spike count from each neuron as input. Decoder accuracy was then determined from the decoder's performance on the pseudo-trials from the testing group. This process was repeated for 100 separate splits of training and testing data to determine the average decoder accuracy for each duration over which spikes counts were quantified (x-axis of Fig. 3C).

Dual accumulator model

The model of mouse behavior shown in Fig. 3F is a race-to-threshold between two accumulators integrating activity from the population of V1 neurons on trials with a contralateral or ipsilateral target, respectively. The output of the accumulators, L and R, is determined by the following equations,

$$\tau_A \frac{dL}{dt} = \frac{S_{contra}}{\alpha + I_L} - L - \omega R + noise$$

$$\tau_I \frac{dI_L}{dt} = S_{contra} - I_L$$

$$\tau_A \frac{dR}{dt} = \frac{S_{ipsi}}{\alpha + I_R} - R - \omega L + noise$$

$$\tau_I \frac{dI_R}{dt} = S_{ipsi} - I_R$$

S_{contra} and S_{ipsi} are trial-averaged V1 population activity on trials with a contralateral or ipsilateral target (Fig. 3A) and provide input to the accumulators L and R, respectively. I_L and I_R integrate the same V1 activity and divisively normalize the input to L and R, respectively. The 6 free parameters are the time constants of L and R (τ_A) and I_L and I_R (τ_I), the half-saturation constant for divisive normalization (α), the noise standard deviation (σ), the strength of mutual inhibition (ω), and the decision threshold. These parameters were the same for both accumulators but the input and noise on each time step was independent. The time step, dt, was 8.33 ms. On each trial, the first accumulator to reach threshold determined the decision of the model (contralateral or ipsilateral target). If neither accumulator reached threshold within 200 ms the outcome was "no response."

The model parameters were fit via a brute force grid search to minimize the sum-of-square error between the response rate and accuracy of the model and mice (Fig. 3G,H). The parameters obtained from this fit were used to test model behavior when V1 input to the model was "inhibited" (by setting it to zero) at the time points shown in Fig. 4C,D. Model decision time (Fig. 3I,J, S6E), which was not explicitly fit, was the time on each trial at which one of the accumulators reached the decision threshold.

Similar results were observed when we used an out-of-sample/leave-one-out fitting procedure in which the response rate and fraction correct values for each of the 7 conditions on the x-axis of Fig. 3G,H were obtained by fitting the model to mouse data for the other 6 conditions. Compared to the fit to data for all 7 conditions simultaneously (Fig. 3G,H), the model response rate with leave-one-out fits was lower for the target-only condition and higher for the no-stimulus condition (Fig. S7).

450 **References**

451 Alpern M. 1953. Metacontrast. *J Opt Soc Am* **43**:648–657. doi:10.1364/JOSA.43.000648

452 Amassian VE, Cracco RQ, Maccabee PJ, Cracco JB, Rudell AP, Eberle L. 1993. Unmasking
453 human visual perception with the magnetic coil and its relationship to hemispheric
454 asymmetry. *Brain Res* **605**:312–316. doi:10.1016/0006-8993(93)91757-J

455 Andermann ML, Kerlin AM, Reid RC. 2010. Chronic Cellular Imaging of Mouse Visual Cortex
456 During Operant Behavior and Passive Viewing. *Front Cell Neurosci* **4**:3.
457 doi:10.3389/fncel.2010.00003

458 Bachmann T, Francis G. 2013. Visual masking: Studying perception, attention, and
459 consciousness. Academic Press.

460 Bacon-Macé N, Macé MJM, Fabre-Thorpe M, Thorpe SJ. 2005. The time course of visual
461 processing: Backward masking and natural scene categorisation. *Vision Res* **45**:1459–
462 1469. doi:10.1016/j.visres.2005.01.004

463 Battersby W, Sturr J. 1970. Neural Limitations of Visual Excitability. *J Opt Soc Amer* **60**:121–
464 126.

465 Bennett C, Gale SD, Garrett ME, Newton ML, Callaway EM, Murphy GJ, Olsen SR. 2019.
466 Higher-Order Thalamic Circuits Channel Parallel Streams of Visual Information in Mice.
467 *Neuron* **102**:477-492.e5. doi:10.1016/j.neuron.2019.02.010

468 Breitmeyer B, Ogmen H. 2006. Visual masking: Time slices through conscious and unconscious
469 vision. Oxford University Press.

470 Bridgeman B. 1980. Temporal response characteristics of cells in monkey striate cortex
471 measured with metacontrast masking and brightness discrimination. *Brain Res* **196**:347–
472 364. doi:10.1016/0006-8993(80)90400-X

473 Burgess CP, Lak A, Steinmetz NA, Zatka-Haas P, Bai Reddy C, Jacobs EAK, Linden JF, Paton
474 JJ, Ranson A, Schröder S, Soares S, Wells MJ, Wool LE, Harris KD, Carandini M. 2017.
475 High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed
476 Mice. *Cell Rep* **20**:2513–2524. doi:10.1016/j.celrep.2017.08.047

477 Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Schölvinck ML, Zaharia AD, Carandini M.
478 2011. The detection of visual contrast in the behaving mouse. *J Neurosci* **31**:11351–61.
479 doi:10.1523/JNEUROSCI.6689-10.2011

480 Coenen AML, Eijkman EGJ. 1972. Cat optic tract and geniculate unit responses corresponding
481 to human visual masking effects. *Exp brain Res* **15**:441.

482 Crouzet SM, Kovalenko LY, del Pin SH, Overgaard M, Busch NA. 2017. Early visual processing
483 allows for selective behavior, shifts of attention, and conscious visual experience in spite of
484 masking. *Conscious Cogn* **54**:89–100. doi:10.1016/j.concog.2017.01.021

485 Crouzet SM, Overgaard M, Busch NA. 2014. The fastest saccadic responses escape visual
486 masking. *PLoS One* **9**. doi:10.1371/journal.pone.0087418

487 Dehaene S. 2014. Consciousness and the brain: Deciphering how the brain codes our thoughts.
488 Penguin.

489 Dell KL, Arabzadeh E, Chiang NS, Id P. 2018. Human-like perceptual masking is difficult to
490 observe in rats performing an orientation discrimination task 1–23.

doi:10.1371/journal.pone.0207179

Dember WN, Purcell DG. 1967. Recovery of masked visual targets by inhibition of the masking stimulus. *Science (80-)* **157**:1335–1336.

Eriksen CW, Hoffman M. 1963. Form recognition at brief durations as a function of adapting field and interval between stimulations. *J Exp Psychol* **66**:485–499. doi:10.1037/h0042228

Eriksen CW, Lappin JS. 1964. Luminance summation-contrast reduction as a basis for certain forward and backward masking effects. *Psychon Sci* **1**:313–314. doi:10.3758/bf03342929

Garrett ME, Manavi S, Roll K, Ollerenshaw DR, Groblewski PA, Kiggins J, Jia X, Casal L, Mace K, Williford A, Leon A, Mihalas S, Olsen SR. 2020. Experience shapes activity dynamics and stimulus coding of VIP inhibitory and excitatory cells in visual cortex. *Elife*.

Glickfeld LL, Histed MH, Maunsell JHR. 2013. Mouse primary visual cortex is used to detect both orientation and contrast changes. *J Neurosci* **33**:19416–22. doi:10.1523/JNEUROSCI.3560-13.2013

Gold JI, Shadlen MN. 2007. The neural basis of decision making. *Annu Rev Neurosci* **30**:535–574. doi:10.1146/annurev.neuro.29.051605.113038

Green MF, Glahn D, Engel SA, Nuechterlein KH, Sabb F, Strojwas M, Cohen MS. 2005. Regional brain activity associated with visual backward masking. *J Cogn Neurosci* **17**:13–23. doi:10.1162/0898929052880011

Haynes JD, Driver J, Rees G. 2005. Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron* **46**:811–821. doi:10.1016/j.neuron.2005.05.012

Hill DN, Mehta SB, Kleinfeld D. 2011. Quality metrics to accompany spike sorting of extracellular signals. *J Neurosci* **31**:8699–8705. doi:10.1523/JNEUROSCI.0971-11.2011

IBL, Aguillon-Rodriguez V, Angelaki D, Bayer H, Bonacchi N, Carandini M, Cazettes F, Chapuis G, Churchland AK, Dan Y, Dewitt E, Faulkner M, Forrest H, Haetzel L, Häusser M, Hofer SB, Hu F, Khanal A, Krasniak C, Laranjeira I, Mainen ZF, Meijer G, Miska NJ, Mrsic-Flogel TD, Murakami M, Noel JP, Pan-Vazquez A, Rossant C, Sanders J, Socha K, Terry R, Urai AE, Vergara H, Wells M, Wilson CJ, Witten IB, Wool LE, Zador AM. 2021. Standardized and reproducible measurement of decision-making in mice. *Elife* **10**:1–28. doi:10.7554/eLife.63711

Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, Lee AK, Anastassiou CA, Andrei A, Aydin Ç, Barbic M, Blanche TJ, Bonin V, Couto J, Dutta B, Gratiy SL, Gutnisky DA, Häusser M, Karsh B, Ledochowitsch P, Lopez CM, Mitelut C, Musa S, Okun M, Pachitariu M, Putzeys J, Rich PD, Rossant C, Sun WL, Svoboda K, Carandini M, Harris KD, Koch C, O'Keefe J, Harris TD. 2017. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**:232–236. doi:10.1038/nature24636

Kepecs A, Mainen ZF. 2012. A computational framework for the study of confidence in humans and animals. *Philos Trans R Soc Lond B Biol Sci* **367**:1322–37. doi:10.1098/rstb.2012.0037

Kepecs A, Uchida N, Zariwala HA, Mainen ZF. 2008. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**:227–231.

Kinsbourne M, Warrington EK. 1962a. Further Studies on the Masking of brief Visual Stimuli by a Random Pattern. *Q J Exp Psychol* **14**:235–245. doi:10.1080/17470216208416541

Kinsbourne M, Warrington EK. 1962b. The Effect of an After-coming Random Pattern on the Perception of Brief Visual Stimuli. *Q J Exp Psychol* **14**:223–234. doi:10.1080/17470216208416540

Kolers PA. 1962. Intensity and contour effects in visual masking. *Vision Res* **2**:277–294. doi:10.1016/0042-6989(62)90037-8

Kovács G, Vogels R, Orban G a. 1995. Cortical correlate of pattern backward masking. *Proc Natl Acad Sci U S A* **92**:5587–5591. doi:10.1073/pnas.92.12.5587

Lamme V a. F, Zipser K, Spekreijse H. 2002. Masking interrupts figure-ground signals in V1. *J Cogn Neurosci* **14**:1044–1053. doi:10.1167/1.3.32

Li N, Chen S, Guo Z V., Chen H, Huo Y, Inagaki HK, Chen G, Davis C, Hansel D, Guo C, Svoboda K. 2019. Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *Elife* **8**. doi:10.7554/eLife.48622

Macknik SL, Livingstone MS. 1998. Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci* **1**:144–149. doi:10.1038/393

Macknik SL, Martinez-Conde S. 2004a. The spatial and temporal effects of lateral inhibitory networks and their relevance to the visibility of spatiotemporal edges. *Neurocomputing* **58–60**:775–782. doi:10.1016/j.neucom.2004.01.126

Macknik SL, Martinez-Conde S. 2004b. Dichoptic visual masking reveals that early binocular neurons exhibit weak interocular suppression: Implications for binocular vision and visual awareness. *J Cogn Neurosci* **16**:1049–1059. doi:10.1162/0898929041502788

Maeda K, Yamamoto H, Fukunaga M, Umeda M, Tanaka C, Ejima Y. 2010. Neural correlates of color-selective metacontrast in human early retinotopic areas. *J Neurophysiol* **104**:2291–2301. doi:10.1152/jn.00923.2009

Öğmen H. 1993. A neural theory of retino-cortical dynamics. *Neural Networks* **6**:245–273. doi:10.1016/0893-6080(93)90020-W

Peirce JW. 2009. Generating stimuli for neuroscience using PsychoPy. *Front Neuroinform*. doi:10.3389/neuro.11.010.2008

Ratcliff R, Rouder JN. 2000. A diffusion model account of masking in two-choice letter identification. *J Exp Psychol Hum Percept Perform* **26**:127–140. doi:10.1037/0096-1523.26.1.127

Resulaj A, Ruediger S, Olsen SR, Scanziani M. 2018. First spikes in visual cortex enable perceptual discrimination. *Elife* **7**:1–22. doi:10.7554/eLife.34044

Ro T, Breitmeyer B, Burton P, Singhal NS, Lane D. 2003. Feedback contributions to visual awareness in human occipital cortex. *Curr Biol* **13**:1038–1041.

Robinson DN. 1966. Disinhibition of Visually Masked Stimuli. *Science (80-)* **154**:157–158. doi:10.1126/science.154.3745.157

Rolls ET, Tovee MJ. 1994. Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proc R Soc London Ser B Biol Sci* **257**:9–15.

Rolls ET, Tovée MJ, Panzeri S. 1999. The neurophysiology of backward visual masking: Information analysis. *J Cogn Neurosci* **11**:300–311. doi:10.1162/089892999563409

Schiller PH. 1968. Single unit analysis of backward visual masking and metacontrast in the cat lateral geniculate nucleus. *Vision Res* **8**:855–866. doi:10.1016/0042-6989(68)90135-1

Schiller PH. 1965. Monoptic and dichoptic visual masking by patterns and flashes. *J Exp Psychol* **69**:193–199. doi:10.1037/h0021574

Schultz DW, Eriksen CW. 1977. Do noise masks terminate target processing? *Mem Cognit* **5**:90–96. doi:10.3758/BF03209198

Siegle JH, Jia X, Durand S, Gale S, Bennett C, Graddis N, Heller G, Ramirez TK, Choi H, Luviano JA, Groblewski PA, Ahmed R, Arkhipov A, Bernard A, Billeh YN, Brown D, Buice MA, Cain N, Caldejon S, Casal L, Cho A, Chvilicek M, Cox TC, Dai K, Denman DJ, de Vries SEJ, Dietzman R, Esposito L, Farrell C, Feng D, Galbraith J, Garrett M, Gelfand EC, Hancock N, Harris JA, Howard R, Hu B, Hytnen R, Iyer R, Jessett E, Johnson K, Kato I, Kiggins J, Lambert S, Lecoq J, Ledochowitsch P, Lee JH, Leon A, Li Y, Liang E, Long F, Mace K, Melchior J, Millman D, Mollenkopf T, Nayan C, Ng L, Ngo K, Nguyen T, Nicovich PR, North K, Ocker GK, Ollerenshaw D, Oliver M, Pachitariu M, Perkins J, Reding M, Reid D, Robertson M, Ronellenfitch K, Seid S, Slaughterbeck C, Stoecklin M, Sullivan D, Sutton B, Swapp J, Thompson C, Turner K, Wakeman W, Whitesell JD, Williams D, Williford A, Young R, Zeng H, Naylor S, Phillips JW, Reid RC, Mihalas S, Olsen SR, Koch C. 2021. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*. doi:10.1038/s41586-020-03171-x

Siegle JH, López AC, Patel YA, Abramov K, Ohayon S, Voigts J. 2017. Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *J Neural Eng* **14**:045003. doi:10.1088/1741-2552/aa5eea

Smith PL, Ratcliff R. 2004. Psychology and neurobiology of simple decisions. *Trends Neurosci* **27**:161–168. doi:10.1016/j.tins.2004.01.006

Sperling G. 1965. Temporal and Spatial Visual Masking I Masking by Impulse Flashes. *J Opt Soc Am* **55**:541. doi:10.1364/josa.55.000541

Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. 2019. Spontaneous behaviors drive multidimensional, brainwide activity. *Science (80-)* **364**. doi:10.1126/science.aav7893

Thompson JH. 1966. What happens to the stimulus in backward masking? *J Exp Psychol* **71**:580–586. doi:10.1037/h0022957

Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Am J Ophthalmol*. doi:10.1016/s0002-9394(14)72148-8

Tse PU, Martinez-Conde S, Schlegel AA, Macknik SL. 2005. Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proc Natl Acad Sci U S A* **102**:17178–17183. doi:10.1073/pnas.0508010102

Van Rullen R, Koch C. 2003. Visual selective behavior can be triggered by a feed-forward process. *J Cogn Neurosci* **15**:209–217. doi:10.1162/089892903321208141

Vorberg D, Mattler U, Heinecke A, Schmidt T, Schwarzbach J. 2003. Different time courses for visual perception and action priming. *Proc Natl Acad Sci U S A* **100**:6275–6280. doi:10.1073/pnas.0931489100\n0931489100 [pii]

Watanabe M, Totah N, Löwe S, Kaiser K, Logothetis N. 2014. Visual backward masking in rats:

616 a behavioral task for studying the neural mechanisms of visual awareness.Society for
617 Neuroscience Meeting. doi:10.13140/2.1.2519.2648

618 Zhao S, Ting JT, Atallah HE, Qiu L, Tan J, Gloss B, Augustine GJ, Deisseroth K, Luo M,
619 Graybiel AM, Feng G. 2011. Cell type–specific channelrhodopsin-2 transgenic mice for
620 optogenetic dissection of neural circuitry function. *Nat Methods* **8**:745–52.

621

622

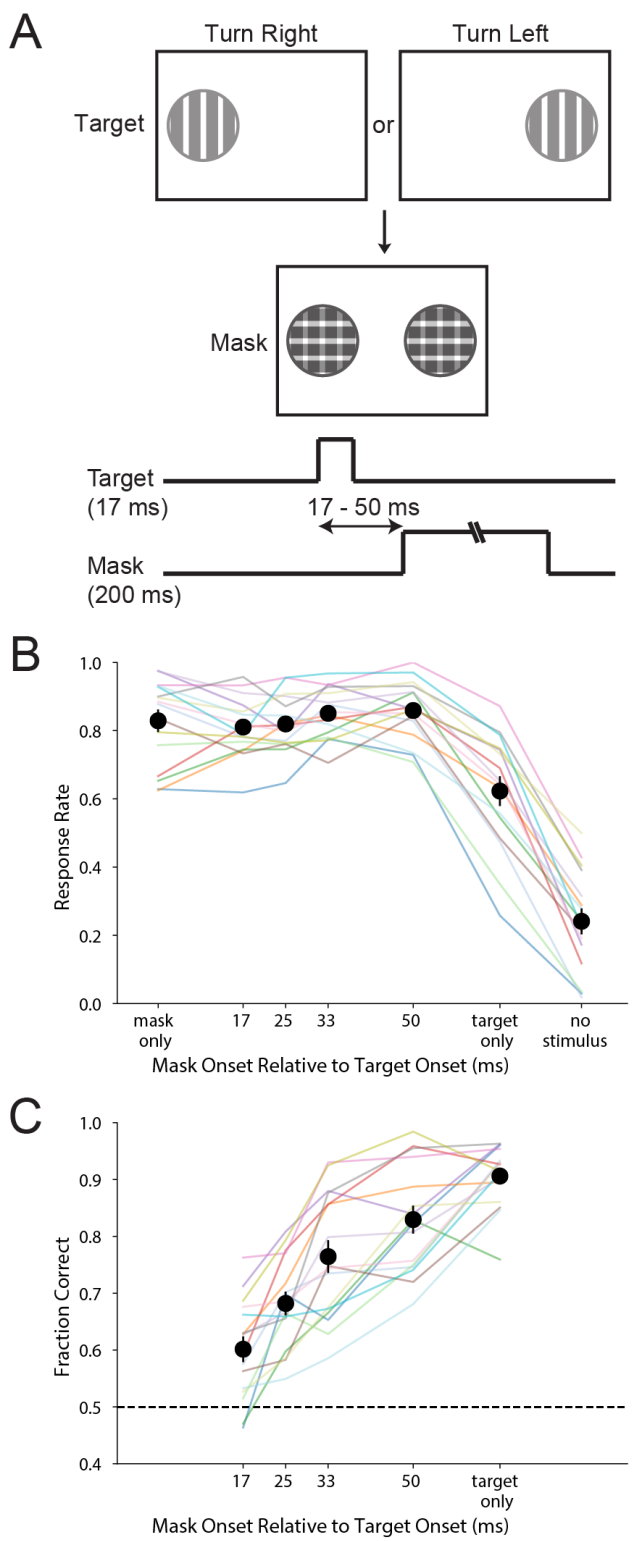

Figure 1. Backward visual masking impairs discrimination of target location. **(A)** A target appeared on the left or right side of a screen in front of the mouse for 17 ms, and water rewards were earned by rotating a wheel right or left, respectively. On some trials, the target was followed by a mask presented on both sides of the screen for 200 ms with variable onset

relative to target onset. Both target and mask were 40% contrast. **(B)** Fraction of trials on which mice responded in either direction (response rate). Colored lines are data from single sessions from individual mice (n=8 wild-type and 8 VGAT-ChR2 mice). Black symbols are means across mice; error bars represent standard error of the mean. **(C)** Fraction of responses in the correct direction. Chance accuracy is 0.5 (dashed line). For the earliest mask onset (17 ms), the accuracy of 9/16 mice was significantly greater than chance (binomial test $p<0.05$). Statistical comparisons between conditions in B and C are shown in Fig. S2A,B.

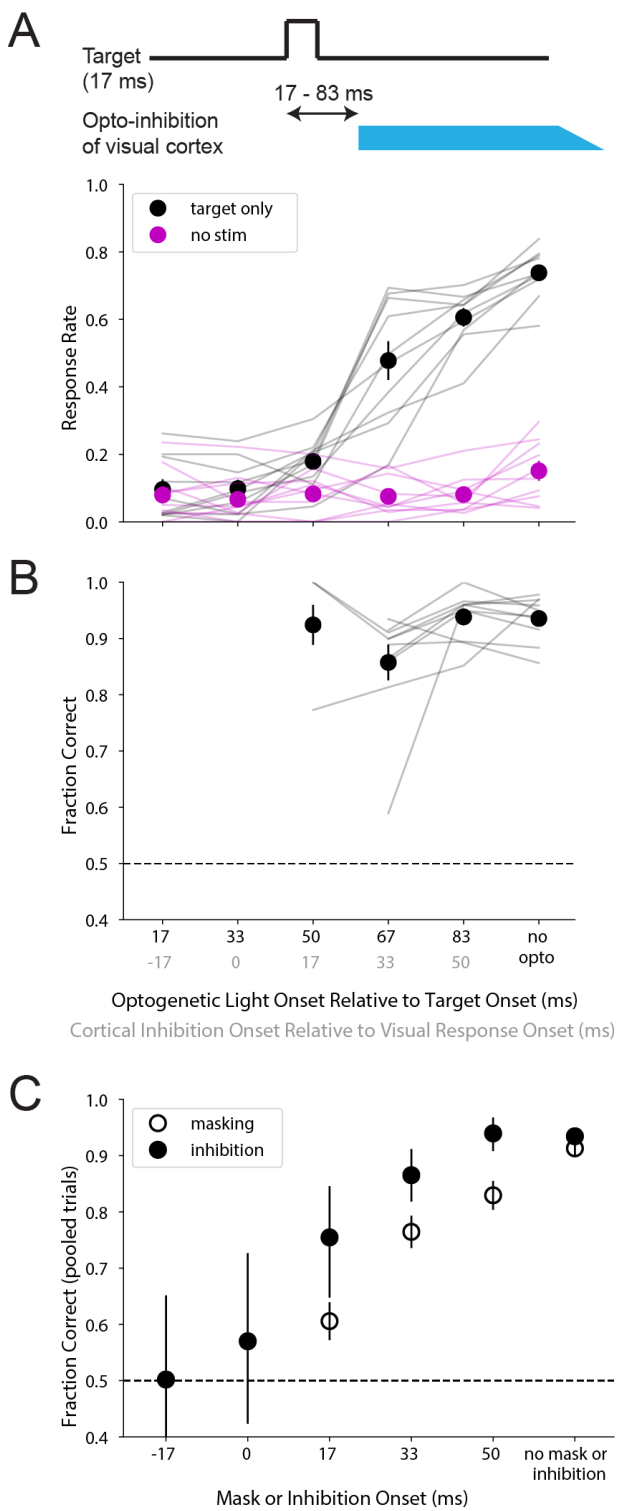

Figure 2. Visual cortical activity during a brief window is required for target detection. **(A)** Response rate and **(B)** accuracy for sessions during which the onset of bilateral optogenetic inhibition of visual cortex was varied in time relative to the onset of the target. The blue light persisted for 650 ms after target onset and then linearly ramped off over 100 ms. The target was presented for 17 ms at 40% contrast. Onset of cortical inhibition relative to the start of the visual

response (gray values on x-axis) was estimated by subtracting the visual response latency of ~45 ms (Fig. S3C,D) from the optogenetic light onset and adding the cortical inhibition latency of ~12 ms (Fig. S4A), resulting in an offset of 33 ms (4 visual stimulus frames) between the two sets of x-axis values. Thin lines represent data from single sessions from 10 VGAT-ChR2 mice. Black symbols are means across mice; error bars represent standard error of the mean. Statistical comparisons between conditions are shown in Fig. S5A,B. Magenta lines denote "catch" trials with no visual stimulus, used to estimate the chance response rate. Fraction correct values for each mouse in (B) are only shown for conditions on the x-axis with response rates above chance (>95th percentile of the binomial distribution given the response probability on catch trials and the number of trials; n = 3, 10, 10, and 10 mice for the conditions from left to right). **(C)** Trials were pooled across mice to estimate accuracy when response rates were low (filled symbols). Error bars are the 95% confidence interval from the binomial distribution given the fraction correct values and the total number of responsive trials (n = 40, 40, 77, 211, 255, and 1247 trials for the conditions from left to right). For comparison, accuracy on masking trials (Fig. 1) was similarly computed (pooling trials across mice) and plotted (open symbols) for corresponding times of mask onset (relative to target onset) or cortical inhibition (relative to cortical visual response latency).

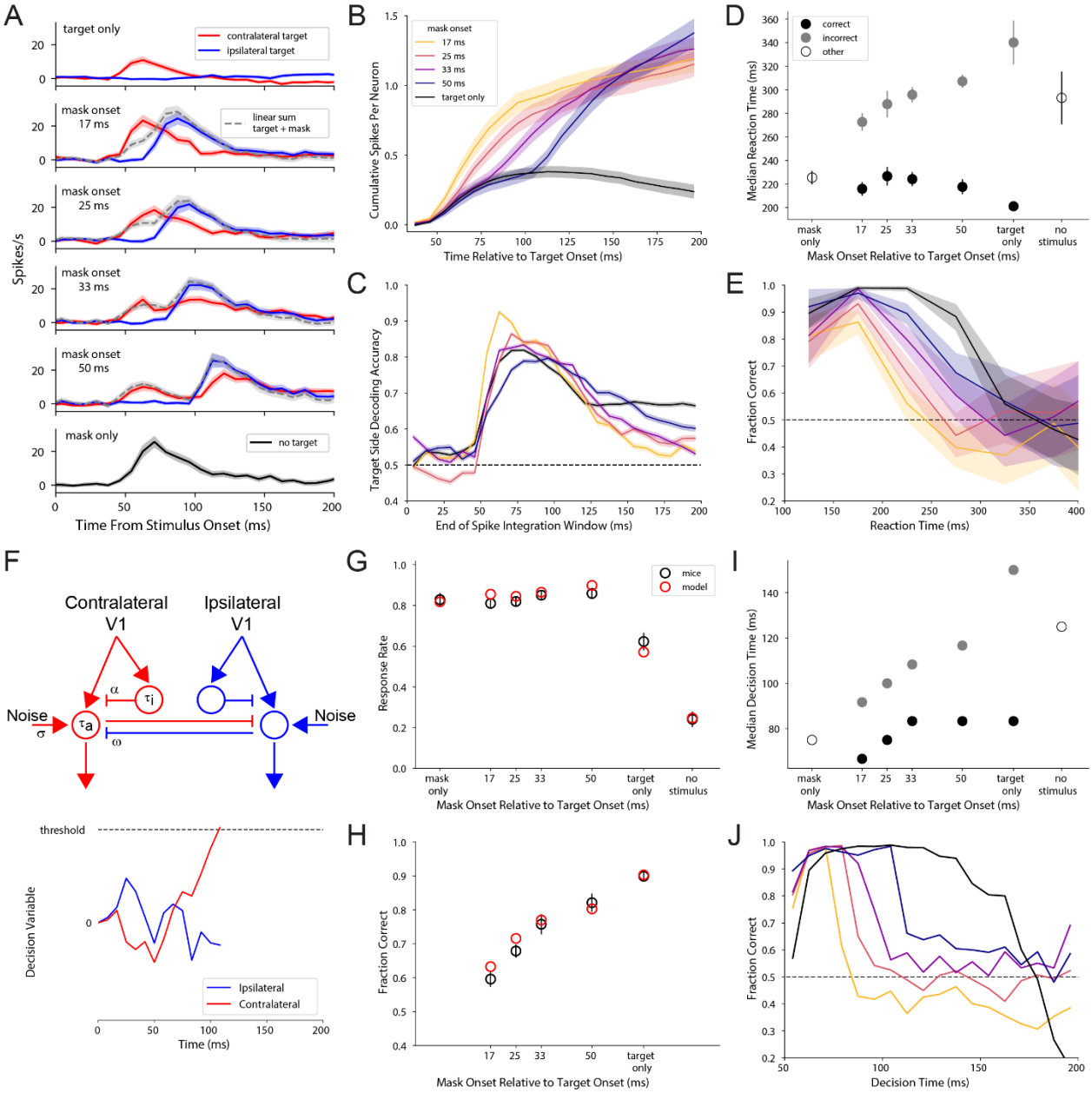

Figure 3. A dual accumulator model replicates the response rate and accuracy of mice as a function of mask timing as well as the inverse relationship between reaction times and accuracy. **(A)** Mean response of V1 neurons (n=58 visually responsive putative pyramidal neurons from 5 wild-type mice; see Methods) to contralateral (red) and ipsilateral (blue) targets and the mask for the mask onset times in Fig. 1B,C. Mask onset was 0 ms on the x-axis on mask-only trials. Shaded region surrounding each line represents the standard error of the mean. Gray, dashed lines are the sum of responses on contralateral target-only trials and masking trials with an ipsilateral target. An ipsilateral target did not evoke a spiking response on its own (top row). **(B)** Cumulative stimulus-evoked spikes per neuron following onset of a contralateral target. Spike number was calculated by integrating the trial-averaged post-stimulus time histogram after baseline subtraction for each neuron, and then averaging across neurons. Shaded region surrounding each line represents the standard error of the mean. **(C)** Decoding of target location

using V1 activity. Linear support vector machines were trained to classify target side using single trial spike counts from the population of V1 neurons. The duration of the integration window, starting at target onset, over which spike counts were calculated varied as indicated by the x-axis. Each line is the mean and standard error across decoders trained and tested using different random subsets of trials (see Methods). **(D)** Median reaction times (separately for correct and incorrect trials when applicable) from the 16 mice used for Fig. 1B,C. Symbols are means across mice and error bars represent standard error of the mean. **(E)** Fraction correct versus reaction time after pooling reaction times across all mice in 50 ms bins. Shaded region surrounding each line is the 95% confidence interval given the fraction correct values and the number of pooled trials for each bin (the median number of trials across mask onset conditions was 66, 262, 259, 150, 75, and 39 for the time bins from left to right). **(F)** Schematic representation of a dual accumulator model fit to mouse behavior data. Two leaky, mutually-inhibiting accumulators integrate divisively-normalized V1 population activity from contralateral and ipsilateral hemispheres (together with independent noise) in a race to threshold to determine the response direction. The 6 free parameters (same values for both "hemispheres") were the time constants of the accumulator (τ_a) and the unit mediating divisive normalization (τ_i), the half-saturation constant for divisive normalization (α), the noise standard deviation (σ), the strength of mutual inhibition (ω), and the decision threshold (see Methods). **(G,H)** Response rate and accuracy of the model (red) fit to mouse behavior data from Fig. 1B,C (black). **(I,J)** Median decision times and accuracy versus decision time for the model, corresponding to (D,E) for mice. Panels (C, E, J) follow the same color as (B).

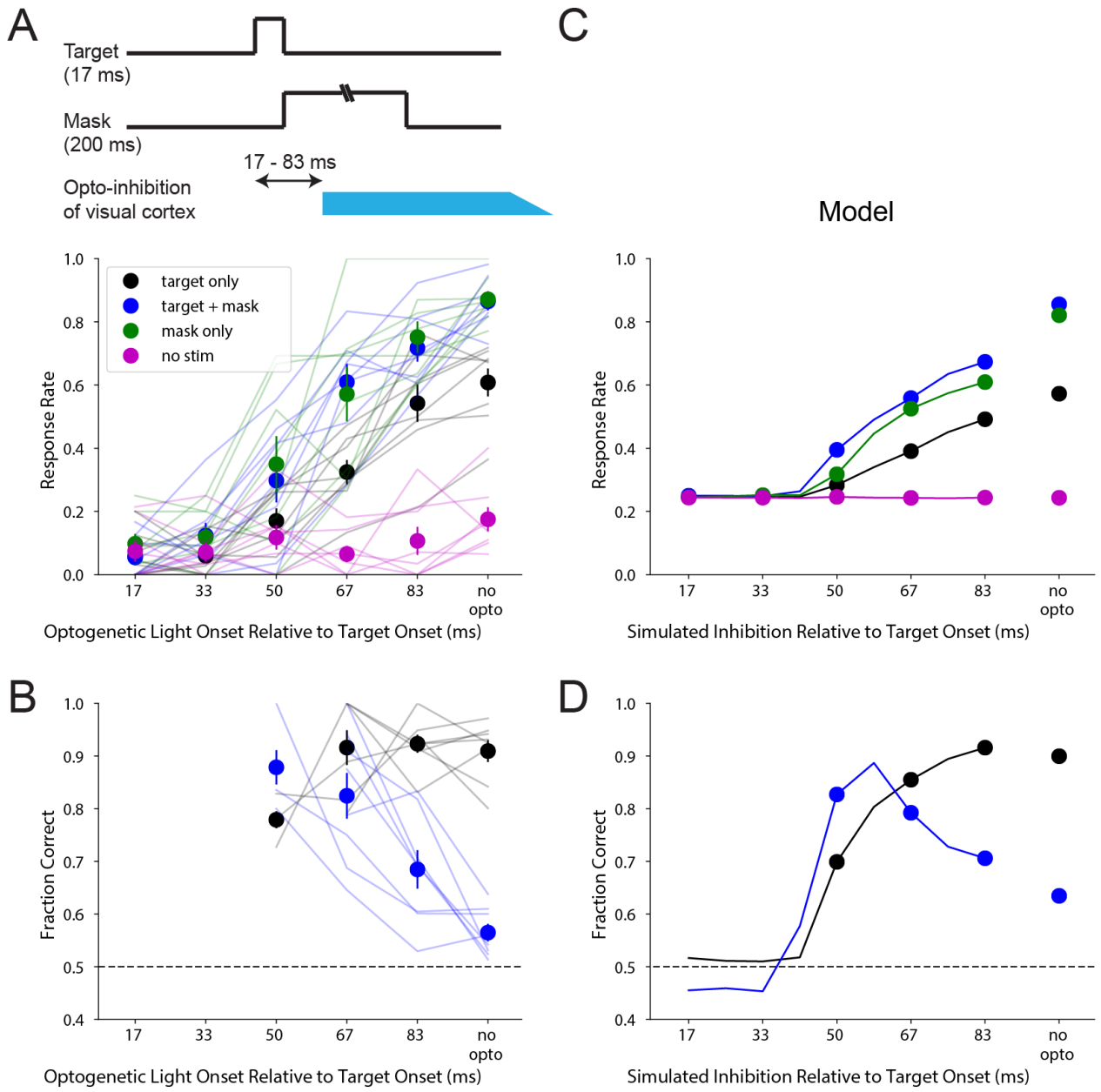

Figure 4. Optogenetic inhibition of mask-evoked activity in visual cortex rescues accuracy of behavioral responses. **(A)** Response rate and **(B)** accuracy for sessions during which the onset of bilateral optogenetic silencing of visual cortex was varied in time relative to the onset of the target and mask. The target stimulus was presented for 17 ms and, on masking trials, was immediately followed by the mask. Thin lines represent data from single sessions from 8 VGAT-ChR2 mice. Fraction correct values for each mouse in (B) are only shown for conditions on the x-axis with response rates above chance, as described for Fig. 2B (n = 3, 6, 8, and 8 mice for target-only trials and 3, 8, 8, and 8 mice for mask trials for the light onsets from left to right). Black symbols are means across mice; error bars represent standard error of the mean. Statistical comparisons between conditions are shown in Fig. S8A,B. **(C,D)** Performance of the model in Fig. 3F, after fitting to mouse behavior data in Fig. 3G,H, in simulated V1 inactivation

experiments similar to (A,B). Symbols show model output for time points tested in mice (A,B); lines show model output for additional time points.

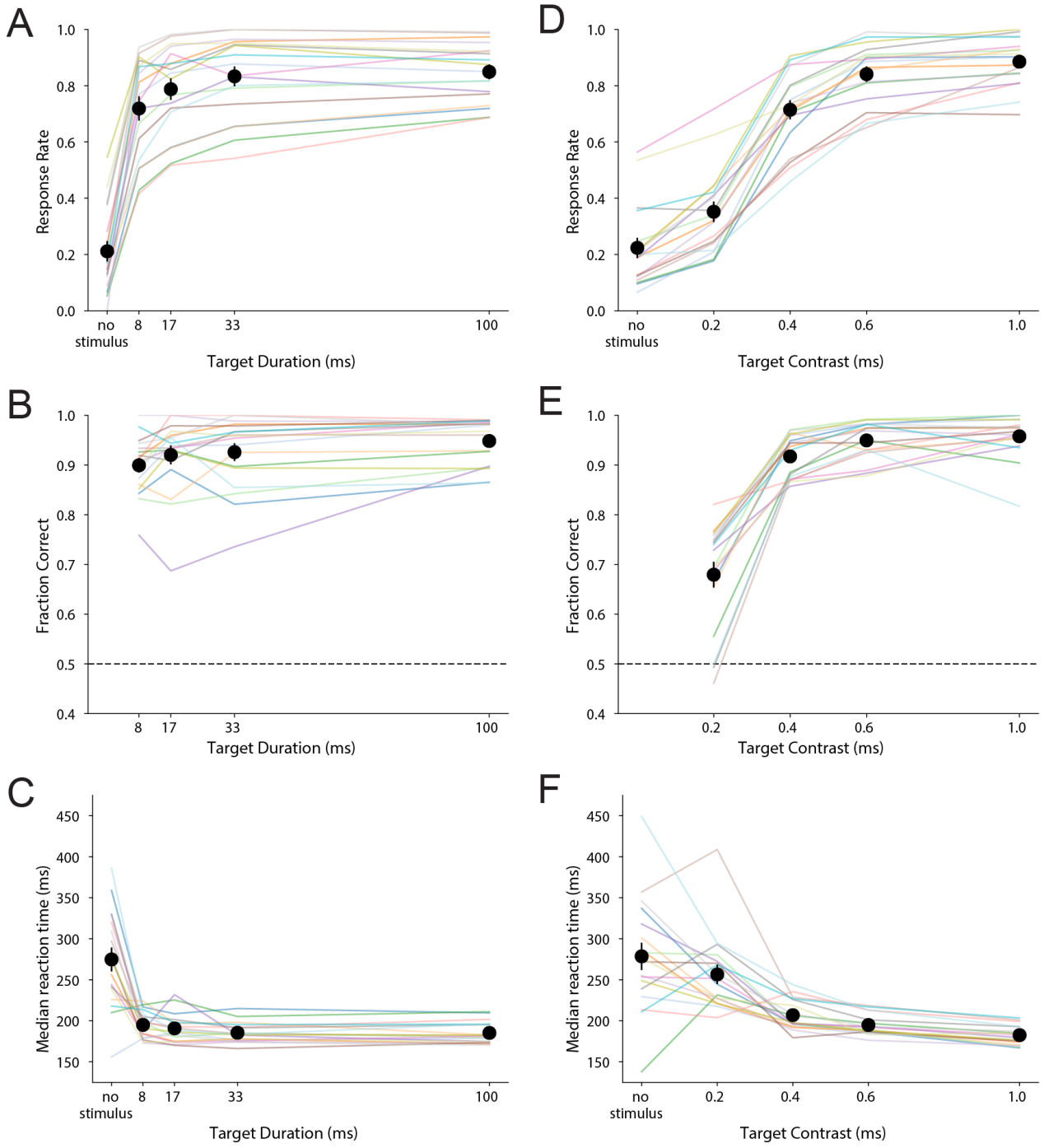

Figure S1. Sensitivity of task performance to target duration **(A-C)** and contrast **(D-F)** in the absence of a mask. Target contrast was 100% for experiments varying target duration (A-C). Target duration was 17 ms for experiments varying target contrast (D-F). Colored lines are data from single sessions from individual mice (n=8 wild-type mice, 10 VGAT-ChR2 mice). Black symbols are means across mice; error bars represent standard error of the mean.

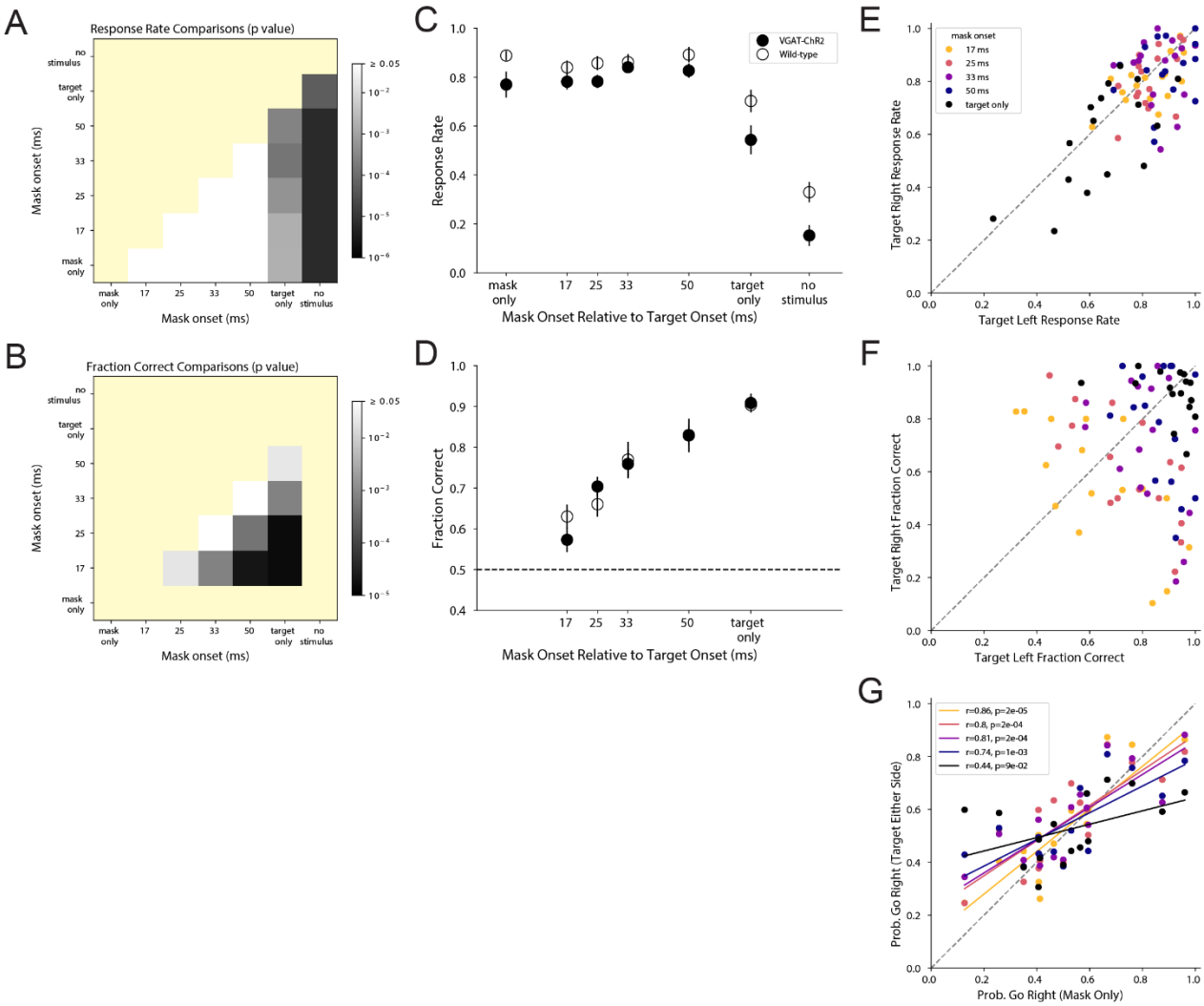

Figure S2. **(A,B)** p-values from comparison of response rates (A) and accuracy (B) for the conditions shown in Fig. 1B,C. Distributions were compared with a Kruskal-Wallis test followed by pairwise Wilcoxon rank sum tests. p-values were corrected for multiple comparisons using the Benjamini-Hochberg method. Color scale is $\log_{10}$ based. Non-significant p-values (≥0.05) are white. Redundant or non-tested comparisons are yellow. **(C,D)** Comparison of response rate and accuracy for the 8 wild-type and 8 VGAT-ChR2 mice pooled in Fig. 1B,C. **(E,F)** Comparison of response rates (E) and accuracy (F) for target-left and target-right trials. Each symbol corresponds to one animal. **(G)** Probability of rightward movement on trials with a target (either side, with or without a mask) on the y-axis versus mask-only trials on the x-axis. Legend indicates the Pearson correlation of these values for each condition. In some mice, response direction on trials with earlier mask onsets was biased in a similar manner to mask-only trials.

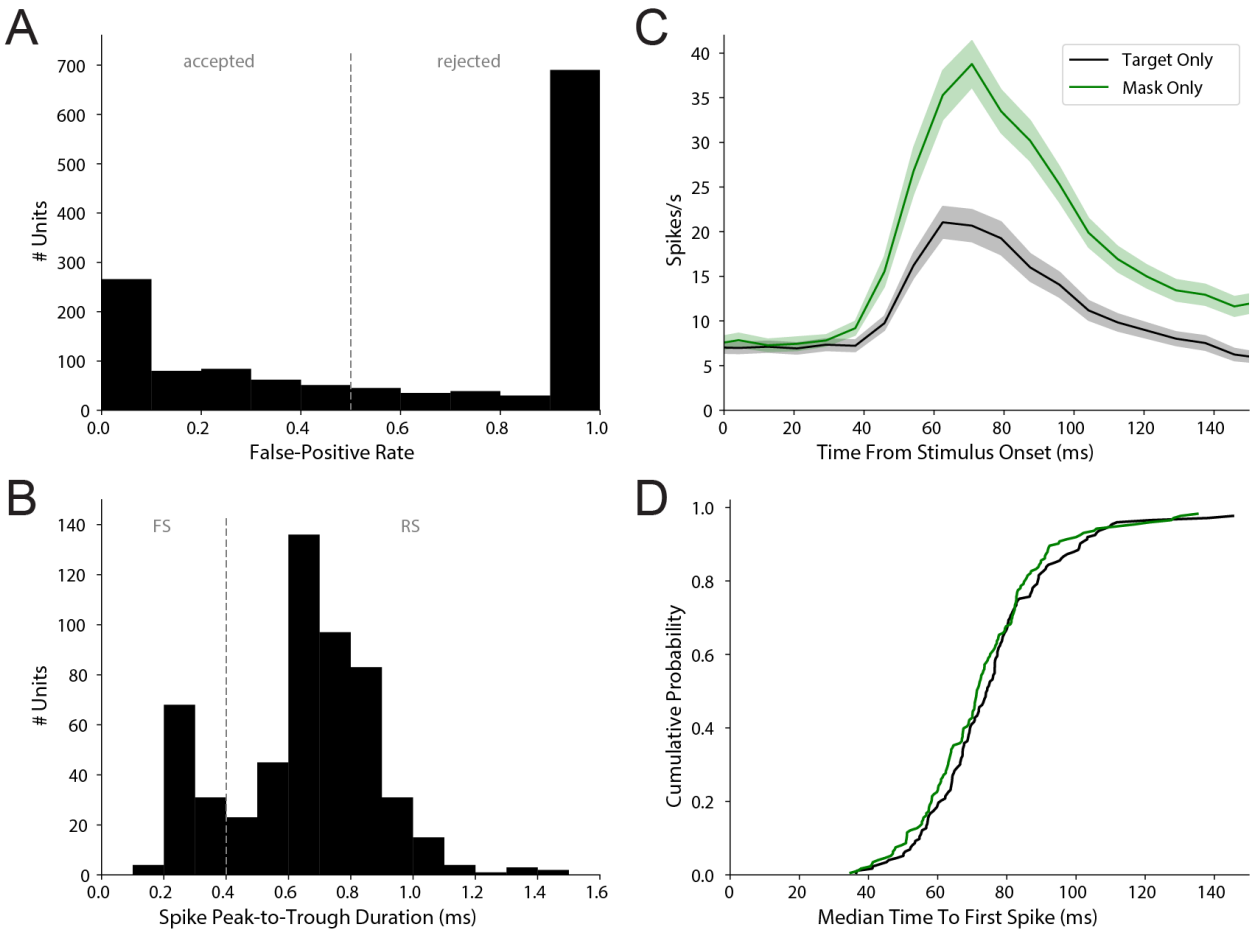

Figure S3. **(A)** False-positive rate for spikes recorded from 1382 units from 5 wild-type and 4 VGAT-ChR2 mice. The false-positive rate was defined as the refractory-period violation rate (number of spikes within 1.5 ms refractory periods divided by the total duration of refractory periods) divided by the total spike rate (see Methods). Units with a false-positive rate greater than one by this definition are displayed at one in the histogram. 543 units with false-positive rates less than 0.5 were considered for further analysis in this study. **(B)** Peak-to-trough duration of the average spike waveforms (n=543). Units with peak-to-trough duration less than 0.4 were classified as fast spiking (FS; putative inhibitory interneurons); all other units were classified as regular spiking (RS; putative pyramidal neurons). **(C)** Average firing rate of 173 visually responsive units (see Methods) following presentation of the target (17 ms duration) or mask (200 ms) starting at 0 ms on the x-axis. Shaded area is the standard error of the mean. **(D)** Cumulative distribution of the median time to first spike (see Methods) of the 173 visually responsive units. There was not a significant difference in the time to first spike between target-only (median across units 73.9 ms) and mask-only (71.3 ms) trials (p=0.22, Wilcoxon rank sum test). There was also not a significant difference between the time to first spike of RS (n=128) and FS (n=45) units (target only: 75.1 vs. 68.1 ms, p=0.08; mask only: 71.4 vs. 68.9 ms, p=0.39), which were combined for the data plotted.

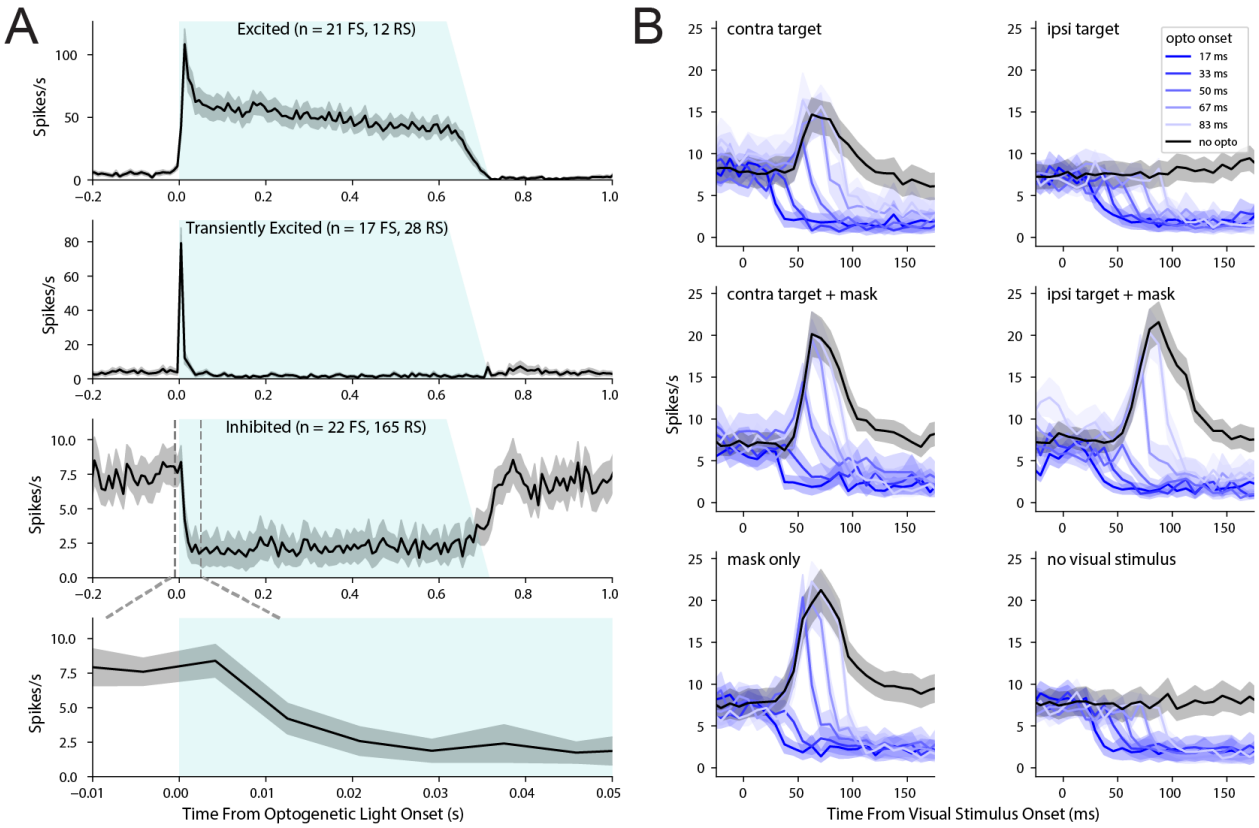

Figure S4. **(A)** Mean response of V1 neurons from 4 VGAT-ChR2 mice to stimulation with blue light. Gray, shaded region indicates standard error of the mean. Fast-spiking (FS, putative interneurons) and regular spiking (RS, putative pyramidal neurons) were defined by the peak-to-trough duration of their spike waveforms. Neurons were classified as excited, transiently excited, or inhibited (see Methods). Two neurons were non-responsive (not shown). The bottom panel expands the time scale (indicated by dashed lines) of the inhibited cells. **(B)** Mean response of 74 visually-responsive V1 neurons (see Methods) from the same mice in (A) to the visual and optogenetic stimuli used for the behavior data in Fig. 4A,B. Neurons that were excited or transiently excited by the optogenetic light were excluded. Optogenetic light onset was varied relative to the onset of the target presented contralateral or ipsilateral to the recorded neurons. On a subset of trials, the target (17 ms duration) was immediately followed by a bilateral mask for 200 ms.

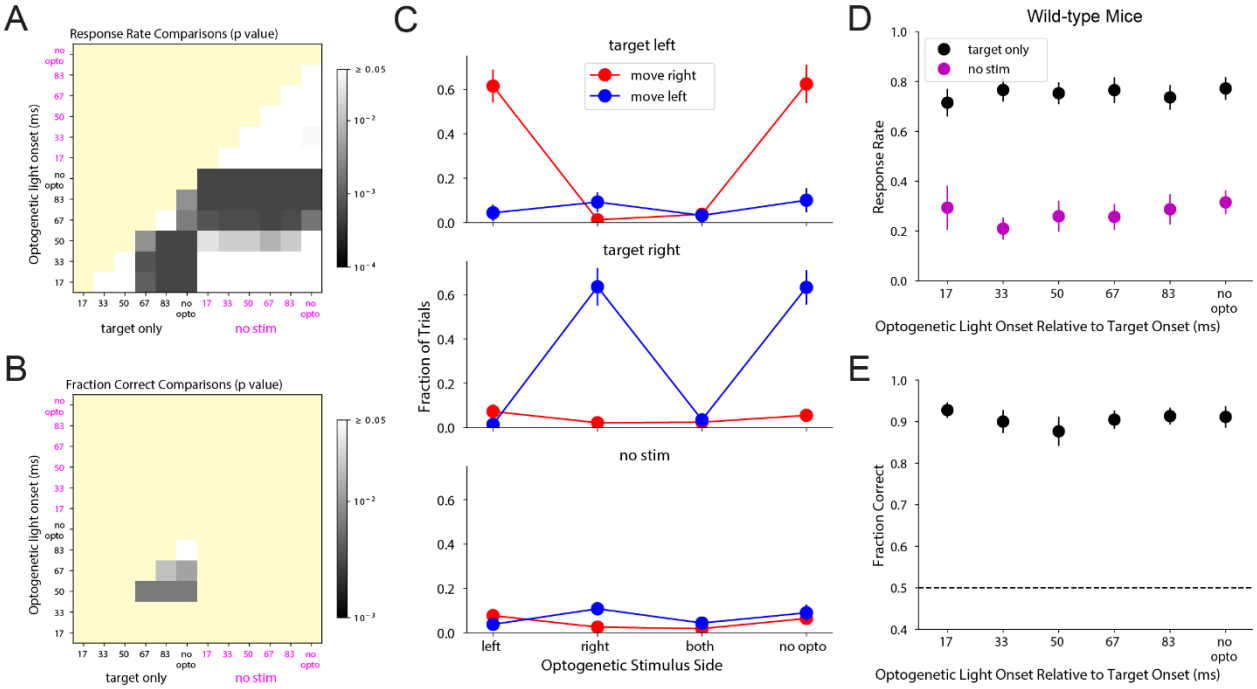

Figure S5. **(A,B)** p-values from comparison of response rates (A) and accuracy (B) for the conditions shown in Fig. 2A,B, using statistical procedures described in the legend for Fig. S2 A,B. **(C)** Effect of unilateral cortical inhibition on target detection (n=8 VGAT-ChR2 mice). The inhibited hemisphere(s) are indicated on the x-axis. Optogenetic light onset was 17 ms before target onset. The top plot shows, for trials on which the target was left, the fraction of trials that mice moved right (red, correct) or left (blue, incorrect). The sum of the red and blue data points for each condition is the response rate. The middle and bottom plots show similar data for target-right and no visual stimulus trials, respectively. **(D,E)** Effect of optogenetic light stimulus on response rate and accuracy in wild-type mice (n=8) for the same conditions shown in Fig. 2A,B.

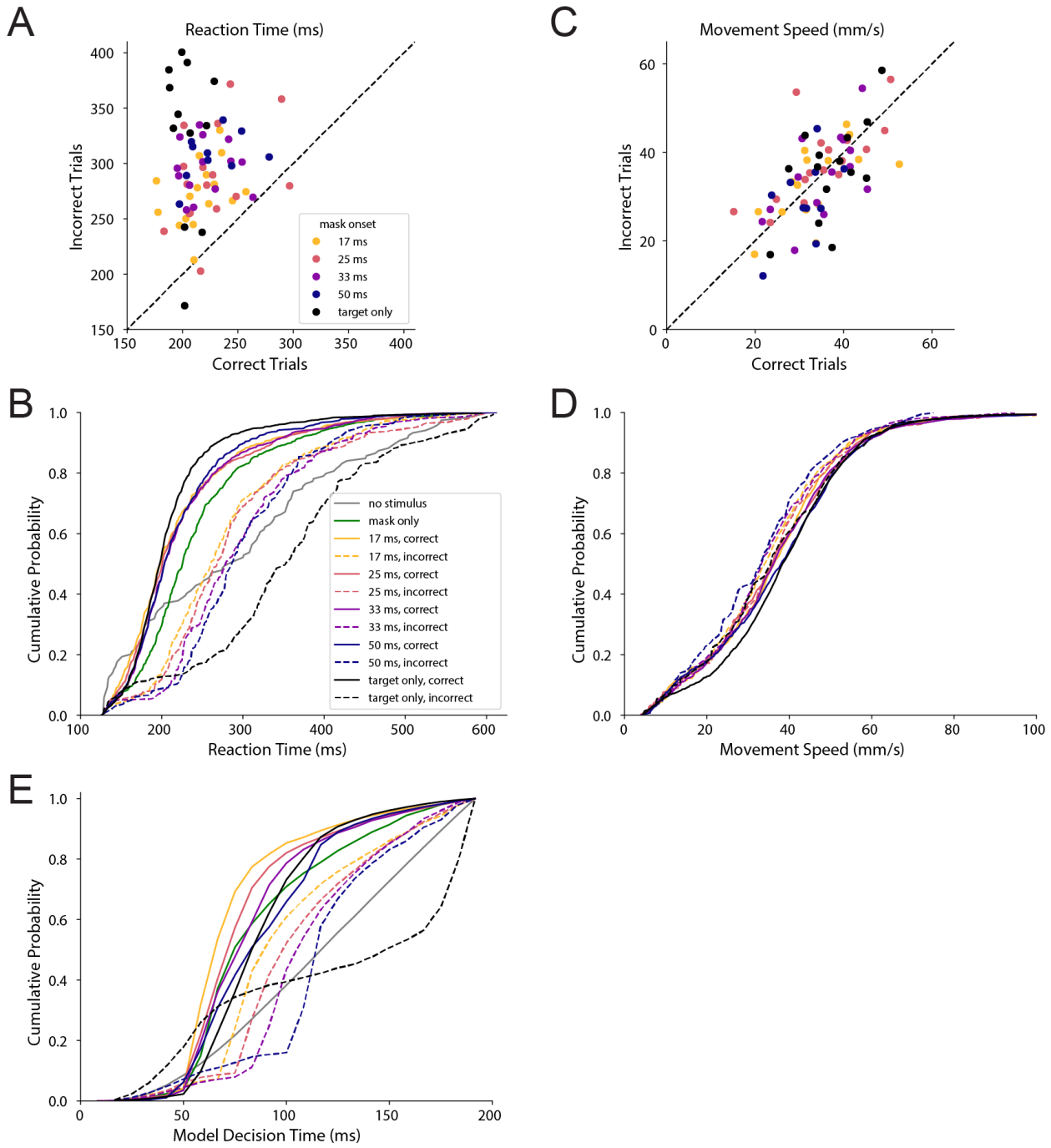

Figure S6. **(A)** Comparison of median reaction times for correct and incorrect responses for each mouse and mask onset condition for the 16 mice from Fig. 1B,C. **(B)** Cumulative probability distributions of reaction times pooled across the 16 mice for correct (solid lines) and incorrect (dashed lines) responses. Reaction times on mask-only and catch trials are also shown. **(C,D)** Same as (A,B) for movement speed. **(E)** Same as (B) for the decision times of the model shown in Fig. 3F-J.

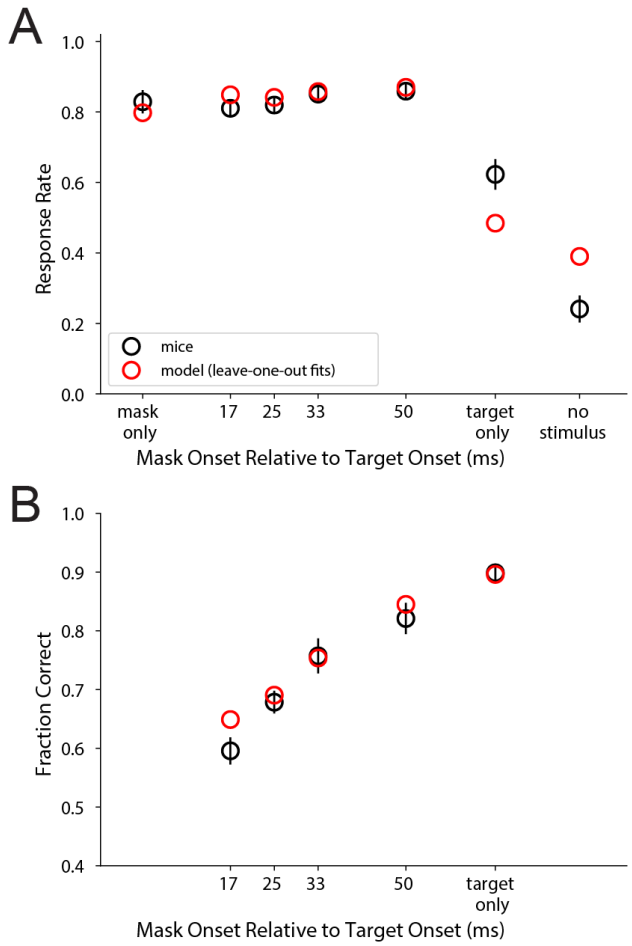

Figure S7. **(A,B)** Out-of-sample fits of the model in Fig. 3F to mouse behavior data from Fig. 1B,C (black). Instead of fitting the response rate and fraction correct (where applicable) values to mouse data for all 7 conditions on the x-axis simultaneously (as in Fig. 3G,H), the values for each condition were obtained by fitting the model to mouse data for the other 6 conditions ("leave one out").

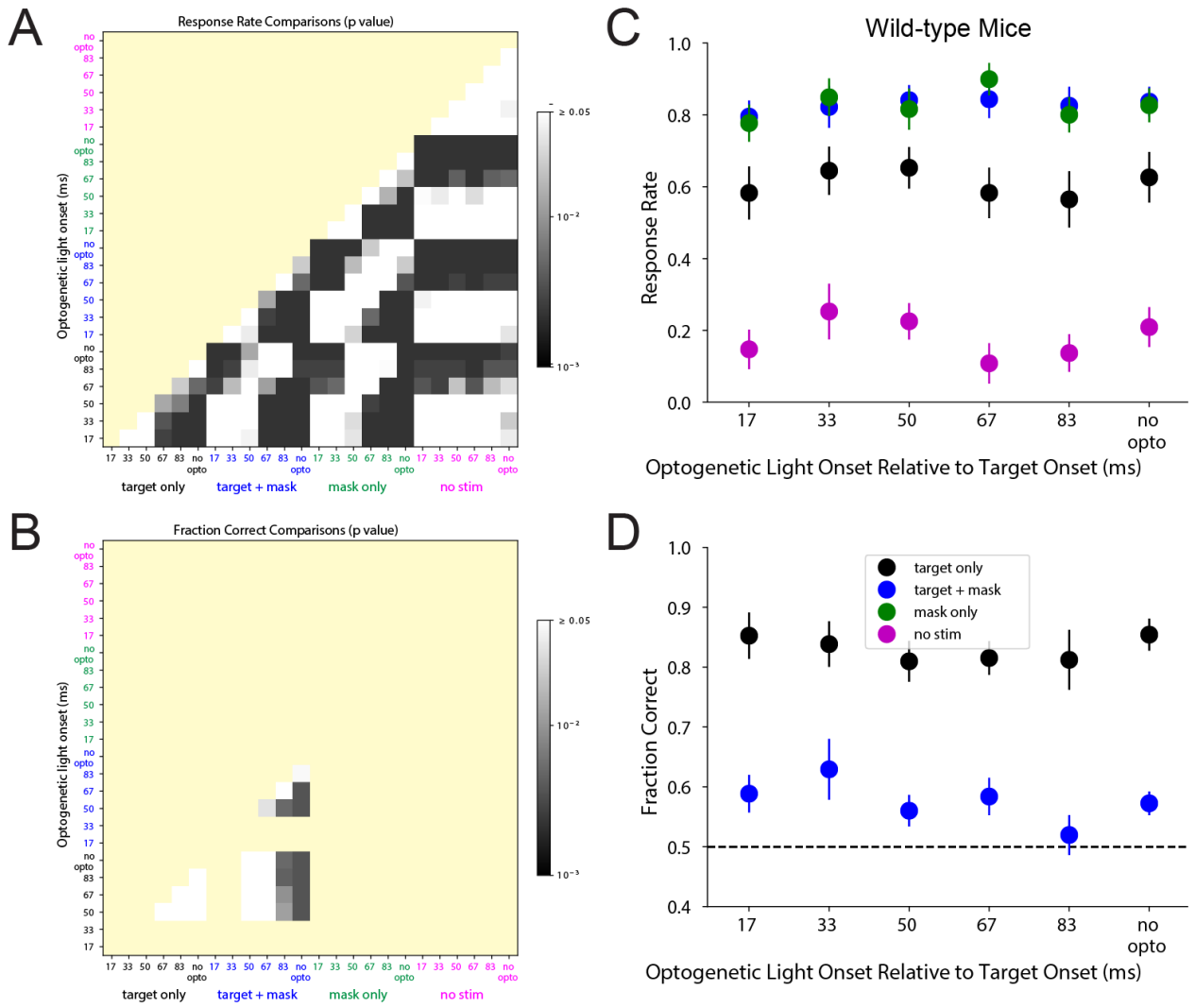

Figure S8. **(A,B)** p-values from comparison of response rates (A) and accuracy (B) for the conditions shown in Fig. 4A,B, using statistical procedures described in the legend for Fig. S2 A,B. **(C,D)** Effect of optogenetic light stimulus on response rate and accuracy in wild-type mice (n=8) for the same conditions shown in Fig. 4A,B.

Table S1

Median number of trials (and range) per mouse/session

								# wild-type mice	# VGAT-ChR2 mice	total trials per session	session duration (minutes)
Target duration (Fig. S1A-C)	**no stim**	**8 ms**	**17 ms**	**33 ms**	**100 ms**			8	10	608 (332-803)	60 (31-69)
target left	64 (33-149)	56 (35-81)	58 (28-81)	58 (35-80)	56 (37-81)						
target right	64 (33-149)	57 (30-79)	56 (34-82)	56 (35-80)	56 (32-80)						
Target contrast (Fig. S1D-F)	**no stim**	**0.2**	**0.4**	**0.6**	**1**			8	10	618 (398-775)	61 (49-69)
target left	58 (39-136)	61 (40-74)	60 (40-74)	60 (41-73)	60 (39-74)						
target right	58 (39-136)	60 (41-74)	59 (40-75)	59 (36-74)	59 (40-74)						
Mask onset (Fig. 1)	**mask only**	**17 ms**	**25 ms**	**33 ms**	**50 ms**	**target only**	**no stim**	8	8	659 (395-977)	65 (53-72)
target left	43 (29-105)	35 (23-48)	36 (22-46)	36 (23-48)	36 (22-48)	110 (61-149)	31 (14-68)				
target right	43 (29-105)	36 (21-47)	36 (23-48)	34 (22-48)	36 (23-46)	110 (63-152)	31 (14-68)				
Optogenetic light onset (Fig. 2; S5D,E)	**17 ms**	**33 ms**	**50 ms**	**67 ms**	**83 ms**	**no opto**		8	10	780 (595-1183)	64 (51-81)
target left	21 (16-32)	20 (15-30)	20 (16-32)	21 (16-31)	21 (16-32)	81 (55-111)					
target right	20 (17-32)	20 (15-31)	20 (17-30)	20 (16-32)	21 (16-30)	78 (55-111)					
no visual stimulus	22 (16-44)	22 (17-45)	23 (14-45)	22 (14-44)	23 (16-44)	88 (69-175)					
Optogenetic light onset (Fig. 4; S8C,D)	**17 ms**	**33 ms**	**50 ms**	**67 ms**	**83 ms**	**no opto**		8	8	776 (615-1523)	67 (50-100)
target only (left)	11 (8-17)	11 (8-18)	11 (8-17)	12 (8-17)	11 (9-17)	44 (35-70)					
target only (right)	11 (8-17)	11 (8-17)	12 (9-18)	11 (8-17)	10 (8-17)	44 (36-71)					
target left + mask	10 (8-16)	10 (8-16)	10 (7-16)	12 (8-16)	10 (8-16)	42 (24-68)					
target right + mask	10 (8-16)	10 (8-16)	11 (7-15)	10 (9-16)	11 (6-16)	42 (23-66)					
mask only	11 (7-30)	12 (7-27)	12 (7-26)	10 (7-28)	11 (8-29)	44 (25-110)					
no visual stimulus	11 (7-30)	12 (7-27)	12 (7-26)	10 (7-28)	11 (8-29)	44 (25-110)					
Optogenetic stimulus side (Fig. S5C)	**left**	**right**	**both**	**no opto**				0	8	733 (470-954)	70 (60-85)
target left	42 (28-56)	42 (31-58)	44 (30-57)	94 (75-140)							
target right	43 (28-58)	43 (30-58)	44 (28-57)	94 (74-138)							
no visual stimulus	52 (28-72)	51 (26-70)	51 (27-73)	102 (65-171)							