

Neuroscience Area - PhD course in
Cognitive Neuroscience

**Individual strategies for ignoring
irrelevant information are reflected in distinct
behavioral and neural signatures
during temporal attention**

Candidate:

Marlen Gironimi

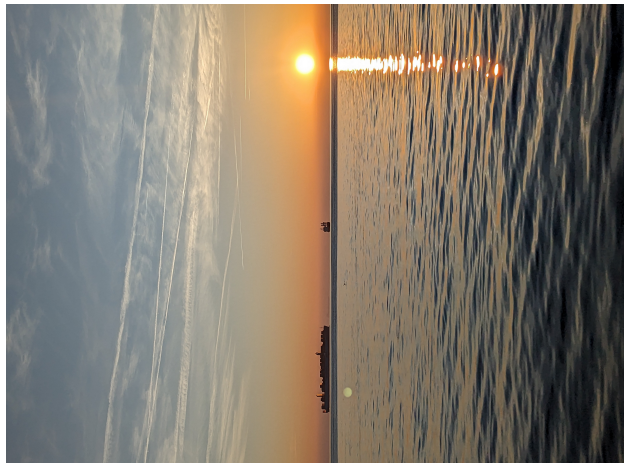
Advisor:

Mathew E. Diamond

Academic Year 2025-26



*To all the outliers and outcasts,
for you will realize that a different wiring is still a beautiful wiring.*



"Because what shapes you is your own mind and the world around you."

Gendo Ikari, *Neon Genesis Evangelion*,

ep. 26, "Take Care of Yourself"

Contents

Abstract	iv
Attentional mechanisms through space, time, and sensory modalities	1
Filtering relevant information in a sensory-rich world	1
What is known about selective attention and what remains unexplored . . .	1
Temporal attention in the tactile system	2
A new behavioral paradigm to study temporal selection	2
Attention as selective weighting of sensory information	4
Neural representations of relevant and irrelevant information	5
A cross-species approach	7
Significance and broader implications	7
Individual differences in rats reflect multiple mechanisms of temporal attention	8
Abstract	8
Introduction	10
Results	13
Irrelevant stimuli systematically bias perceptual judgments	13
Rats rely on temporal structure to infer stimulus relevance	19
Limited flexibility in allocating attention across time	21
Learning reduces the influence of irrelevant stimuli when task structure is stable	25
Neural representations of relevant and irrelevant stimuli	31
Inter-areal interactions during task performance	40
Discussion	46
Methods	53

Human behavioral patterns reflect diversity in behavioral strategies during	
temporal attention	76
Abstract	76
Introduction	78
Results	81
Temporal attention in the tactile modality of humans	81
Effect of irrelevant information on humans' perceptual judgments	84
Behavioral clustering reveals diversity in attentional strategies	89
Relationship between behavioral performance and introspection	98
Correlation between neuro-divergent traits and attentional control	102
Discussion	104
Methods	107
General conclusions	115
From learning what is irrelevant to flexible selective weighting	115
Neural representation does not necessarily determine behavioral influence	115
Rats and human participants share individual variability	116
Limitations and future directions	117
References	119
Supplementary Material	128

Abstract

Executive functions enable the nervous system to regulate the continuous stream of incoming information by selecting behaviorally relevant signals and filtering out distracting inputs. Through studies in rats and humans, this thesis seeks to advance our understanding of the neural mechanisms that support the selective processing of relevant information.

To investigate attentional control in rats, we developed a behavioral paradigm in which animals learned to judge a relevant tactile stimulus while ignoring an irrelevant one. The perceptual judgment required animals to categorize the relevant vibrissal vibration as weak or strong according to its intensity. Rats achieved proficient performance, defined as assigning greater decisional weight to the relevant stimulus than to the irrelevant stimulus. Nevertheless, the irrelevant stimulus exerted an attractive bias on judgments of the relevant stimulus. A subset of rats subsequently mastered the task and progressed from proficient to expert performance, defined as the ability to ignore the irrelevant stimulus completely.

In this initial task, the relevant stimulus was always presented in the second temporal position. To test attentional flexibility, we developed a second, dynamic version of the task in which the relevant stimulus could appear in either the first or the second position. In this variant, rats had difficulty flexibly allocating attention across time.

Electrophysiological recordings from two expert rats implanted in primary and secondary motor cortex (M1/M2) while performing the initial task revealed distinct neural representations of relevant and irrelevant stimuli. In addition, recordings from a proficient rat implanted in both primary vibrissal somatosensory cortex (vS1) and M1/M2 suggested that outcome-dependent attentional modulation was localized to vS1, whereas M1/M2 transformed weak sensory signals into robust categorical representations while partially suppressing irrelevant information.

Analyses of local field potentials further indicated that accurate stimulus representation was supported by high-gamma synchronization between vS1 and M1/M2. Moreover, M1/M2 appeared to influence vS1 through top-down modulation in the low-beta frequency range. Together, these behavioral and neural findings demonstrate that learning to ignore irrelevant information reshapes how neuronal populations in vS1 and M1/M2 interact and encode task-relevant variables.

To assess the cross-species generalizability of these behavioral dynamics, we tested human participants using a task analogous to the dynamic paradigm used in rats. As expected, humans achieved higher overall performance. A cluster-based analysis revealed a striking divergence between species: whereas rats exhibited an attractive bias induced by the irrelevant stimulus, a subgroup of human participants displayed a repulsive bias. This finding suggests that the two species may rely on different cognitive strategies to solve the task.

Overall, these findings provide novel, multiscale evidence of temporal attentional abilities in the rat tactile system and broaden our understanding of temporal tactile attention in humans.

Attentional mechanisms through space, time, and sensory modalities

Filtering relevant information in a sensory-rich world

Organisms are continuously exposed to immeasurably large amounts of sensory information [1]; therefore, efficient behavioral strategies are essential for survival and optimal interaction with the environment. Executive functions support this flexible adaptation by guiding goal-directed actions through key processes such as the selection of relevant information over irrelevant signals. This adaptive behavior is expressed through attentional and inhibitory control, by prioritizing behaviorally relevant stimuli and suppressing distracting sensory signals. These high-level functions are core components of this filtering process, allowing the nervous system to select and process only what is behaviorally relevant to the current goals.

What is known about selective attention and what remains unexplored

Selective attention - the ability to prioritize behaviorally relevant information for sensory processing - has been studied extensively in the visuospatial domain, in humans and non-human primates, contributing to the understanding of how goal-oriented actions and sensory-perceptual processing interact. For example, past studies have highlighted the role of neural oscillations during selective attention [2, 3], whereas other studies focused on describing the interaction between cortical networks involved in attentional control [4]. While previous research has shown that mice express visual selective attention [5] using a Posner-style cueing task, and that rats are able to select behaviorally relevant olfactory stim-

uli [6], less is known about how rodents assign selective attention across time in the tactile modality. Studies of temporal attention in humans have predominantly focused on visual processing, describing how selecting information across time affects perception [7, 8], and other cognitive domains [9]. On the other hand, research on tactile attention has largely examined its spatial dimensions, identifying mechanisms common also to the auditory and visual modalities [10, 11].

Temporal attention in the tactile system

Therefore, a key question remains: how do animals, and humans, allocate attention across time, in the tactile domain? To provide an initial answer to this, the present thesis examines whether rats can selectively attend to tactile information presented at different moments in time. The main focus lies on the behavioral and neural mechanisms that support this ability.

An important question that the thesis addresses is whether selective attention can be directed voluntarily across time, in a goal-directed manner, rather than fluctuating spontaneously. Previous studies have found that attention follows this latter pattern, in mice, yielding a high detection rate of a vibrissal deflection in alert states and a high miss-rate in unalert states [12]. Here, we attempt to bring the modulation of attention under temporal control, rather than leaving it spontaneous. The thesis expands the findings identified in rats, by examining whether principles of selective attention also generalize to humans.

A new behavioral paradigm to study temporal selection

To study temporal attention in the tactile domain, we developed a new behavioral paradigm to study temporal attention: the Irrelevant Stimulus Paradigm. This novel task addresses these questions by presenting two vibrotactile stimuli sequentially, separated by a variable

inter-stimulus interval. In the standard version of the task, an acoustic cue indicates the stimulus that is behaviorally relevant, which is always presented in the second temporal position. Rats must ignore the first, non-cued stimulus (i.e., the irrelevant stimulus), and categorize the second, cued stimulus, as weak or strong. Therefore, the perceptual judgment required animals to categorize the relevant vibrissal vibration as *weak* or *strong* according to its intensity.

Behavioral consequences of irrelevant sensory information

We first characterized the behavioral consequences of irrelevant sensory information by using psychometric measures established in previous studies [13–15]. Specifically, we investigated whether an irrelevant tactile stimulus influences judgments about a relevant stimulus, and whether rats can successfully ignore distracting information presented only moments earlier. Together, these questions address the limits of selective attention across time, in a sensory modality that has been less studied than the visual and auditory domains.

Learning to suppress distraction

We first investigated how rats learned to suppress distraction, by quantifying the initial effects of irrelevant stimuli on perceptual judgments, and whether this interference decreased with training. An important goal was to determine whether expert animals could fully suppress irrelevant information, by determining its influence at the behavioral level.

Temporal predictability versus cue-guided attention

In the standard version of the task, the relevant stimulus always occurred second, raising the possibility that animals relied only on temporal order, rather than the auditory cue. We

sought to understand whether the rats exploited the temporal predictability of the relevant stimulus, or if attention was implemented in a cue-guided manner, by further developing a control task, in which the acoustic cue was removed. This allowed us to distinguish between strategies (i.e., temporal predictability or cue-guided attention). The results indicated that the acoustic cue facilitated selective attention but was not its predominant determinant, suggesting that the animals relied primarily on temporal order.

The flexibility of temporal attention

To determine whether rats could flexibly allocate attention across time, we developed a more demanding version of the task, that allowed the relevant stimulus to occur in either temporal position (first or second). This dynamic task allowed us to directly test the ability of rats to rely only on the acoustic cue to assess relevance, rather than relying on a fixed temporal rule. The results, reporting reduced performance, revealed limitations in flexible attentional allocation across time.

Attention as selective weighting of sensory information

Intuitively, attention implies the ignoring, or elimination, of irrelevant information. Past studies have argued about the locus of this mechanism, either before sensory encoding [16], or before downstream decisional processes [17]. However, irrelevant information may still be encoded and represented by sensory and frontal cortices, despite minimal influence on behavioral choices.

Selective weighting can provide an alternative explanation to traditional accounts of attentional control. Rather than suppressing or erasing irrelevant information, selective weighting argues that stimuli receive different weights based on their relevance.

The current thesis presents evidence favoring this view. Expert behavioral performance was accompanied by a reduced influence of irrelevant information, and when the task required additional flexibility in allocating attentional resources, relevant and irrelevant stimuli were also weighted differently.

This framework is related to classical attenuation accounts of selective attention, in which irrelevant signals are not completely eliminated from processing, but instead exert a reduced influence relative to relevant information [18].

Neural representations of relevant and irrelevant information

After characterizing behavior, the thesis investigates how the brain represents relevant and irrelevant tactile stimuli. Particular emphasis is placed on the interactions between sensory and frontal cortical networks.

Top-down modulation of sensory processing

A substantial body of research has shown that frontal cortical regions exert top-down influences on early sensory areas, shaping sensory processing in favor of behaviorally relevant information.

In [19], the authors proposed that multiple representations compete for processing, however, the relevant representation increasingly dominates processing by means of a change in its competitive weight, determined by behavioral relevance. Top-down signals can influence the processing and suppression of irrelevant information by selectively amplifying the relevant representation, modulating the relative weight of competing sensory signals [20]. Previous studies found that the modulation exerted by frontal cortices suppress sensory transmission from early sensory areas to high-level regions through the thalamus, describing

an anatomical pathway extending beyond cortical areas [21]. These findings support the framework of selective weighting present in the current thesis, arguing that attentional control can be achieved through assigning differential weight to sensory information.

Neural recordings were obtained while animals performed the task, spanning different stages of learning and expertise. More specifically, we collected neuronal data from two rats performing the standard task, that showed suppression of irrelevant information at the behavioral level, and from an animal that retained residual bias. In addition, we recorded from an animal performing the dynamic task. Population decoding and local field potential analyses were used to investigate neural coding and inter-areal communication.

Key neural questions

The key questions addressed by the thesis concern how, at the neural level, relevant and irrelevant stimuli are represented in frontal motor areas (M1/M2) and in the primary somatosensory cortex (vS1), and whether these representations changed as animals became experts. Furthermore, the thesis characterizes the interaction between sensory and frontal areas and the emergence of stimulus representation, and outlines the neural mechanisms that support communication between these regions.

Predictions about attentional modulation

Based on previous findings from studies of attention and cortical control [22], we predicted that frontal cortical regions would distinguish between relevant and irrelevant information. Similar attentional effects may also be evident in barrel cortex (i.e., vS1),

A cross-species approach

The behavioral paradigm developed for rodents was adapted for human participants. This allowed a direct comparison between rodents and humans, allowing us to evaluate the generalizability of findings from one species to the other. Specifically, the study asks whether similar strategies and attentional mechanisms operate across the two species, when rats and human participants performed the dynamic task.

Significance and broader implications

The present thesis examines how sensory and frontal cortical circuits interact during selective attention, expressed through the temporal domain. The rodent and human studies extend the existing attention research beyond traditional visuospatial paradigms, by connecting the established literature with emerging research on other forms of attention - including temporal attention - across a different sensory modality and in both humans and rodents. More broadly, it contributes to our understanding of temporal attention, tactile cognition, and shared mechanisms of cognitive control across species.

Individual differences in rats reflect multiple mechanisms of temporal attention

Abstract

Attentional functions enable the nervous system to regulate incoming information, selecting what is behaviorally relevant while filtering out distractions. To investigate attentional control in rats, we developed a paradigm in which animals were required to ignore an irrelevant tactile stimulus (a vibration) and categorize the relevant one as weak or strong. Stimuli were separated in time but delivered to the same set of vibrissae, making this a temporal rather than spatial attention task. In the first task version, the irrelevant stimulus was presented first and the relevant second. Across animals, we observed substantial variability in how this task was learned and performed, consistent with emerging work showing that learning unfolds along idiosyncratic trajectories rather than converging on a single behavioral solution. While the irrelevant stimulus typically exerted an attractive bias on judgments, some rats progressively reduced this influence and achieved near-complete suppression, transitioning from a proficient to an expert stage of performance. Other animals, however, adopted alternative stable strategies and did not exhibit comparable levels of attentional filtering.

To probe behavioral flexibility, we designed a version of the task in which the relevant stimulus could appear in one or the other of two sequential temporal positions. Under these conditions, rats generally struggled to flexibly allocate attention across time, again with substantial variability across individuals. Importantly, behavioral analyses indicate that this variability was not random but reflected a small number of reproducible strategy classes,

characterized by distinct patterns of sensitivity to stimulus order. These findings suggest that attentional control in this task did not rely on a single canonical algorithm, but instead emerged from a constrained set of alternative solutions.

Electrophysiological recordings from a limited number of animals revealed neural correlates consistent with these behavioral differences. In expert animals performing the original task, neuronal populations in motor cortex showed differential encoding of the irrelevant and relevant stimuli. In a proficient animal implanted in both vS1 and M1/M2, outcome-dependent modulation was prominent in vS1, whereas M1/M2 transformed sensory inputs into categorical representations with partial suppression of the irrelevant stimulus. Local field potential analyses further indicated that effective task performance was associated with increased high-gamma synchronization between vS1 and M1/M2, alongside low-beta modulation consistent with top-down interactions.

Taken together, these results suggest that learning to ignore irrelevant information reshaped interactions between sensory and motor cortices, but that this process unfolded heterogeneously across individuals. Rather than reflecting a single mechanism of attentional control, the data support a framework in which multiple, identifiable strategies coexist within a shared task structure, underscoring the importance of individual differences in understanding the neural basis of attention.

Introduction

We are constantly flooded with streams of sensory information. Executive functions enable goal-directed behavior, including the ability to ignore stimuli that are irrelevant to current objectives and to selectively attend to those that are informative. While visuospatial attention has been extensively studied in humans and non-human primates [2–4, 23], less is known about how rodents allocate attention within a single sensory modality.

In the current study, we investigated how rats learn to ignore irrelevant vibrotactile information, focusing on both the behavioral and neural dynamics of temporal attention. In our paradigm, the relevant and irrelevant stimuli were physically indistinguishable; as a result, distinguishing relevant from irrelevant inputs was therefore not an innate capacity but an acquired one, requiring prolonged training.

This intensive, longitudinal interaction allowed us to ask whether animals converged on a single solution to the task or instead adopted individual-specific strategies. More generally, recent work has emphasized that learning and behavior do not necessarily converge onto a single canonical solution, but may instead unfold along distinct trajectories, with different animals adopting internally consistent strategies [24, 25]. We therefore asked whether attentional control in this paradigm would be homogeneous or heterogeneous across animals. Pooling subjects may obscure such heterogeneity when behavioral outcomes are similar. For this reason, we adopted an individual-subject approach. Similar behavioral outcomes can arise from different underlying strategies, which are not necessarily accompanied by changes in overall performance: an animal may maintain a similar level of accuracy while weighting task information differently. A stable strategy can therefore be defined by a consistent pattern of information use across time, rather than by stability of performance itself.

To address these questions, we developed the Irrelevant Stimulus Paradigm, a task in which two vibrotactile stimuli were delivered sequentially, separated by a variable inter-stimulus interval. In the first version of the task, the relevant stimulus was tagged by an acoustic cue and always appeared in the second temporal position. Rats were required to ignore the first, non-tagged stimulus and report the intensity of the relevant one. By delivering both stimuli to the same set of vibrissae, the task isolated temporal attention rather than spatial selection.

This design allowed us to examine how the irrelevant stimulus biased judgments about the relevant one, as well as how such biases were modulated through training. Because the relevant stimulus always appeared in the second temporal position in the standard task, animals could, in principle, rely on stimulus order rather than attentional selection. To control for this possibility, we included tests in which the acoustic cue was removed, allowing us to dissociate the contribution of temporal position and cue-based selection.

To further assess behavioral flexibility, we introduced a second version of the task in which the relevant stimulus could occur in either temporal position, forcing animals to rely exclusively on the acoustic cue. This manipulation enabled us to test whether rats can flexibly allocate attention across time, rather than adopting fixed strategies tied to stimulus order.

In parallel, we investigated how sensory and motor cortical areas represented relevant and irrelevant information and how their interaction evolved with learning and expertise. Electrophysiological recordings were obtained in rats performing both the standard and dynamic tasks at different stages of training.

Frontal areas are known to exert top-down influences on early sensory cortex, shaping activity to enhance the processing of behaviorally relevant stimuli [19–22]. Here, we asked whether similar mechanisms operated in the rodent somatosensory system during temporal

attention. Using population decoding and local field potential analyses, we characterized (i) how relevant and irrelevant stimuli were represented in vS1 and M1/M2, (ii) how these representations depended on behavioral performance and training stage, and (iii) how interactions between these areas supported stimulus selection.

Results

Irrelevant stimuli systematically bias perceptual judgments

Effects of irrelevant stimuli on perceptual judgments

We investigated whether rats could ignore task-irrelevant sensory inputs when forming perceptual judgments. To this end, we developed the Irrelevant Stimulus Paradigm (ISP), in which animals judged the intensity of a vibrotactile target stimulus (REL) while ignoring a temporally separated, non-cued distractor stimulus (IRR). In the standard version of the task, the relevant stimulus always occurred in the second temporal position (Figure 1).

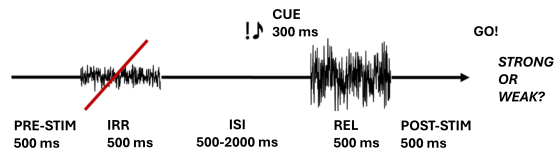


Figure 1: Irrelevant Stimulus Paradigm. The first variant of the *Irrelevant Stimulus Paradigm*, called standard task.

Each trial began with nose-poke fixation, followed by presentation of IRR, a variable inter-stimulus interval (ISI; 500–2000 ms), and then REL, whose onset was signaled by an acoustic cue. Rats reported whether REL was “strong” or “weak” via their choice to lick at the L or R spout (Figure 2). Stimulus intensity was defined as the nominal mean velocity of a sequence of vibration samples drawn from Gaussian distributions, yielding nine discrete intensity levels for both REL and IRR.

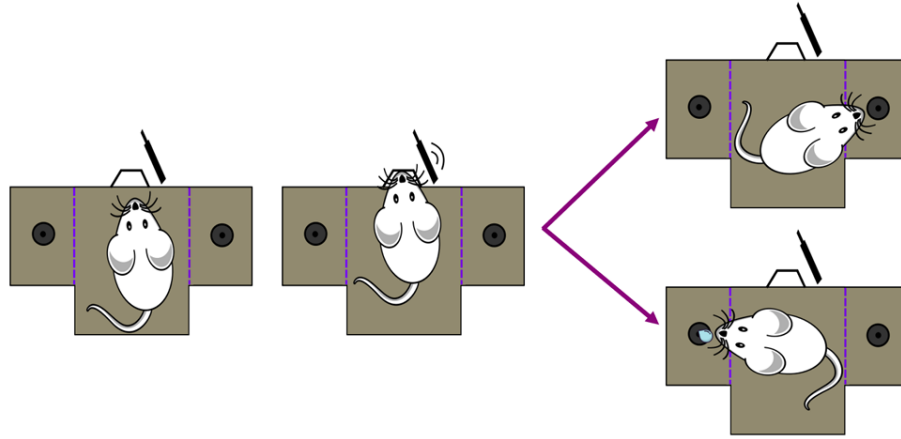


Figure 2: *Experimental setup* Schematization of the experimental setup used in the rat study. The animal is self-paced and freely moving, and can initiate the trial at its will by entering the nosepoke with the snout. The vibration is delivered via the plate onto the whiskers. To make a perceptual judgment, the animal moves to the right or left reward spout. In this example, going to the right side corresponds to a wrong response, whereas turning left represents the correct choice. Correct choices are rewarded with water.

Approximately 20% of the trials within each training session consisted of *single-stimulus* trials, in which only the relevant stimulus was delivered. These trials were introduced to ensure that rats remained in the nosepoke during presentation of the first stimulus, and to exclude the possibility that they remained in the nose poke and ignored events for a fixed period of time (see Supplementary Figures S1, S2, S3).

All rats acquired stable performance, as indicated by steep psychometric functions (Figure 3, top), consistent with the use of reliable stimulus–response strategies. However, performance was systematically biased by the irrelevant stimulus. When conditioned on IRR intensities, psychometric functions shifted toward the category of the distractor, indicating an attractive bias: REL was more likely to be judged “strong” when preceded by a strong IRR, and “weak” when preceded by a weak IRR (Figure 3A). To quantify this bias, we used the

distance measure D_{IRR} (see Methods). Positive and negative values indicate biases in opposite choice directions, whereas values close to zero indicate little or no IRR-related bias (Figure 3B).

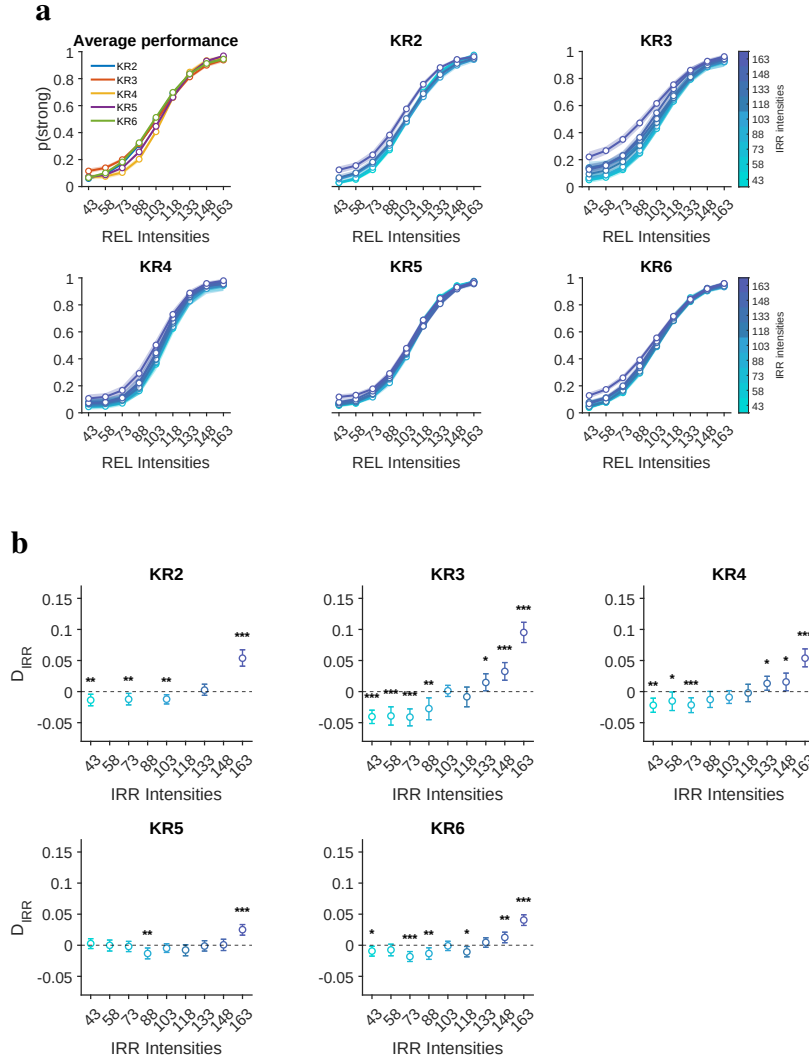


Figure 3: Overall performance and IRR bias, for individual rats. a. Overall behavioral performance for individual rats (top left), and IRR-induced bias. Individual rats exhibited varying degrees of the magnitude of IRR-induced bias. Dots represent empirical probabilities at each stimulus intensity, and lines connect consecutive fitted probabilities. Shaded areas represent 95% bootstrapped confidence intervals. **b.** IRR-induced bias as a function of irrelevant-stimulus intensity for each animal. D_{IRR} indicates the signed IRR-related bias, with $D_{IRR} = 0$ corresponding to no bias. Circles show the observed values and error bars the 95% bootstrapped confidence intervals. Significance was assessed independently at each intensity using two-sided permutation tests. Dashed lines indicate zero bias. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This effect showed two notable features. First, the bias was asymmetric, primarily affecting weak REL stimuli, whereas strong REL stimuli were largely resistant to interference. Second, its magnitude varied substantially across individuals, ranging from negligible (e.g. KR5) to pronounced (e.g. KR3), indicating marked inter-individual variability.

The temporal separation between IRR and REL modulated the strength of this bias. Short ISIs enhanced the effect of strong IRR stimuli, whereas longer intervals reduced it (Figure 4). In contrast, the influence of weak IRR stimuli was more heterogeneous across rats, with some individuals showing little dependence on ISI duration. The effect of ISI was quantified as the difference in IRR magnitude between short- and long-ISI trials, $\Delta M_{\text{IRR}} = M_{\text{Short}} - M_{\text{Long}}$, such that positive values indicate a stronger influence of the IRR stimulus at short ISIs (Figure 4).

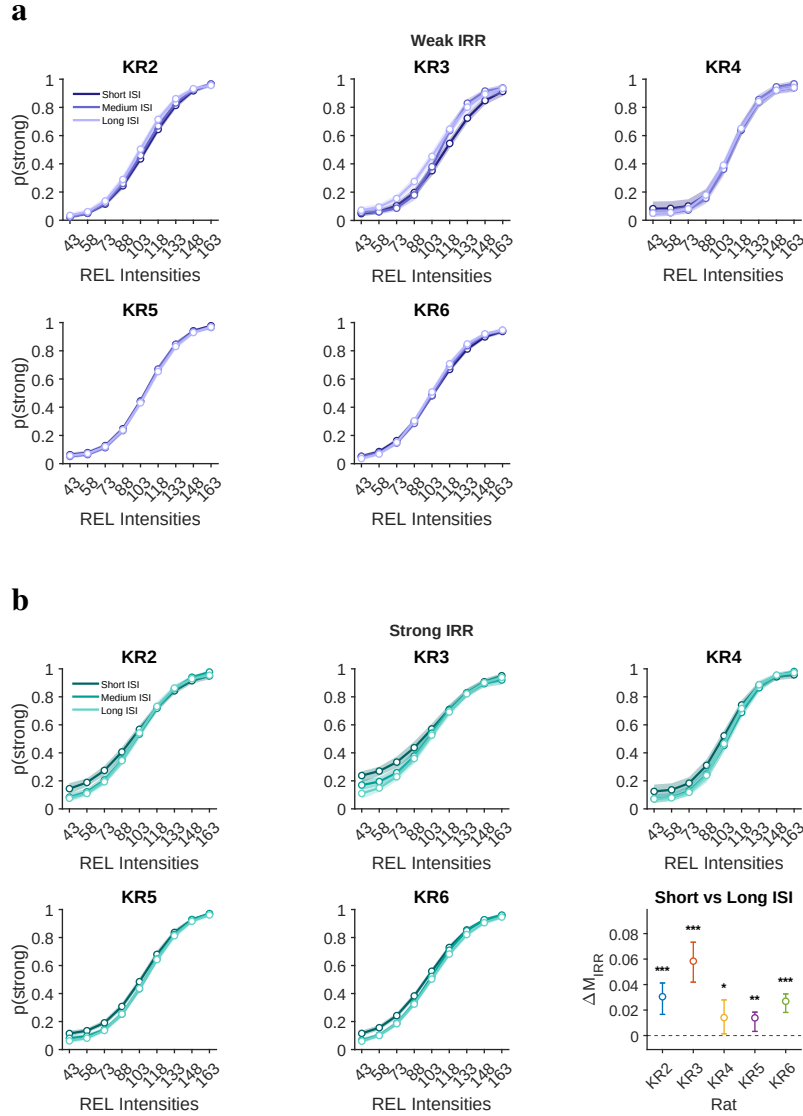


Figure 4: Effect of the ISI duration on the IRR bias, for individual rats. a. Effect of ISI duration when IRR was weak. **b.** Effect of ISI duration when IRR was strong. Bottom right: Effect of inter-stimulus interval (ISI) on the influence of the irrelevant stimulus. Shaded areas represent the 95% bootstrapped confidence intervals. For each animal, the ISI effect was quantified as $\Delta M_{IRR} = M_{Short} - M_{Long}$, with positive values indicating a stronger IRR effect at short ISIs. Circles show the observed differences and error bars the 95% bootstrapped confidence intervals. Significance was assessed using two-sided permutation tests. Dashed line indicates no ISI effect. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

The number of sessions and trials for each animal performing the standard task is reported in Supplementary Table S1.

Rats rely on temporal structure to infer stimulus relevance

Effect of cue removal on performance

In the standard task, the relevant stimulus was always presented second. This raises the possibility that rats used temporal position, rather than the acoustic cue, to determine stimulus relevance. To test this, we removed the acoustic cue in a modified version of the task (Figure 5) and compared performance before and after cue removal. In this version, approximately 40% of trials were constituted by single-stimulus trials (see Supplementary Figures S1, S2, S3).

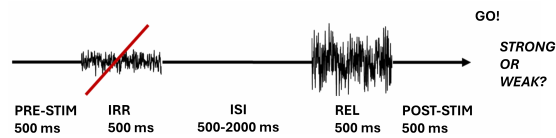


Figure 5: ISP modification: control task. Schematization of the control task, used to test whether rats also used the acoustic cue to assess relevance.

Psychometric functions changed following removal of the cue (Figure 6). With the exception of KR4, animals showed moderate changes in psychometric width, while the PSE decreased in all animals except KR3. Despite these changes, overall psychometric performance remained largely preserved in the absence of the acoustic cue. These results indicate that rats might have primarily relied on the temporal structure of the stimulus sequence to infer

stimulus relevance, while the acoustic cue provided an additional, but limited, contribution for assessing relevance.

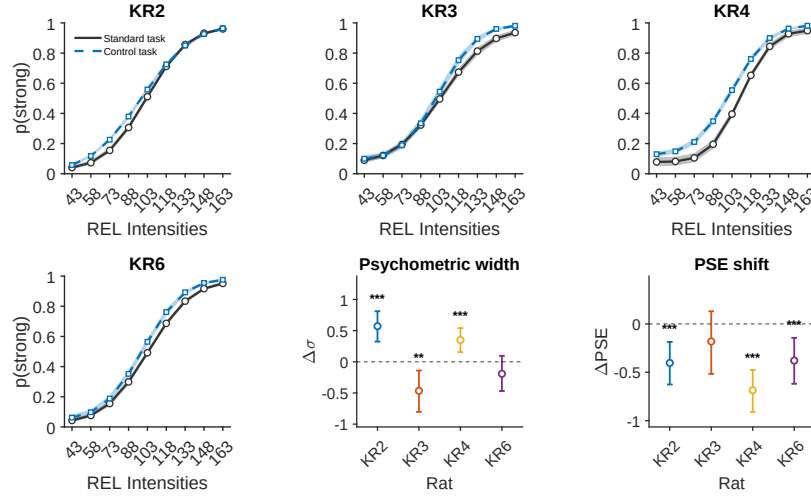


Figure 6: Average performance during the standard task and control task. Psychometric functions for the standard task (gray) and cue-removal control task (blue) are shown for each animal using matched numbers of sessions. Shaded areas indicate 95% bootstrapped confidence intervals. Bottom panels show the change in psychometric width ($\Delta\sigma = \sigma_{\text{Control}} - \sigma_{\text{Standard}}$) and PSE ($\Delta\text{PSE} = \mu_{\text{Control}} - \mu_{\text{Standard}}$). Error bars indicate 95% bootstrapped confidence intervals. Statistical significance was assessed using two-sided permutation tests in which trials were randomly reassigned to pseudo-standard and pseudo-control conditions while preserving the original number of trials in each condition. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

The number of sessions and trials for each animal performing the standard task is reported in Supplementary Table S2.

Limited flexibility in allocating attention across time

Performance in the dynamic task

We next asked whether rats could flexibly allocate attention when stimulus relevance varied across trials. In a further variant of the ISP, the position of the relevant stimulus was randomized within each session: half of the trials contained REL in the second position (IRR–REL), and half in the first position (REL–IRR) (Figure 7). Animals trained in this task variant were exposed to REL-IRR trials and IRR-REL trials since the early training phases.

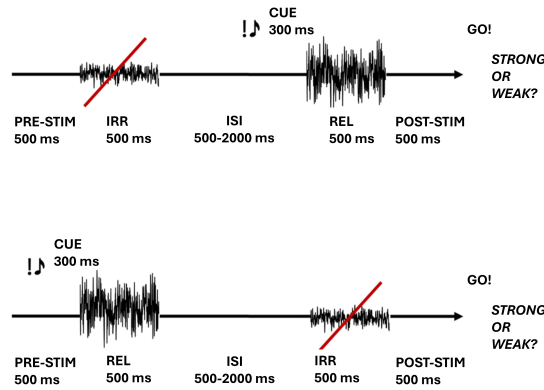


Figure 7: ISP modification: dynamic task. Schematization of the dynamic task, used to test the attentional flexibility of rats across time.

Across all animals, performance was higher when the relevant stimulus appeared in the second position (IRR–REL trials) than when it appeared first (REL–IRR trials) (Figure 8). The attractive bias induced by IRR was present in both trial types, but appeared stronger when REL occurred first (Figure 9 and Figure 10), highlighting a possible tendency to overweight later stimuli in the sequence.

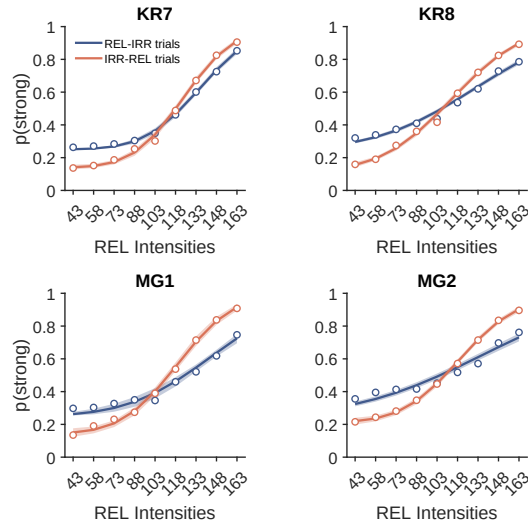


Figure 8: Overall performance during the dynamic task. Psychometric functions showing the overall performance for each trial type, for single subjects. Shaded areas represent the 95% bootstrapped confidence intervals.

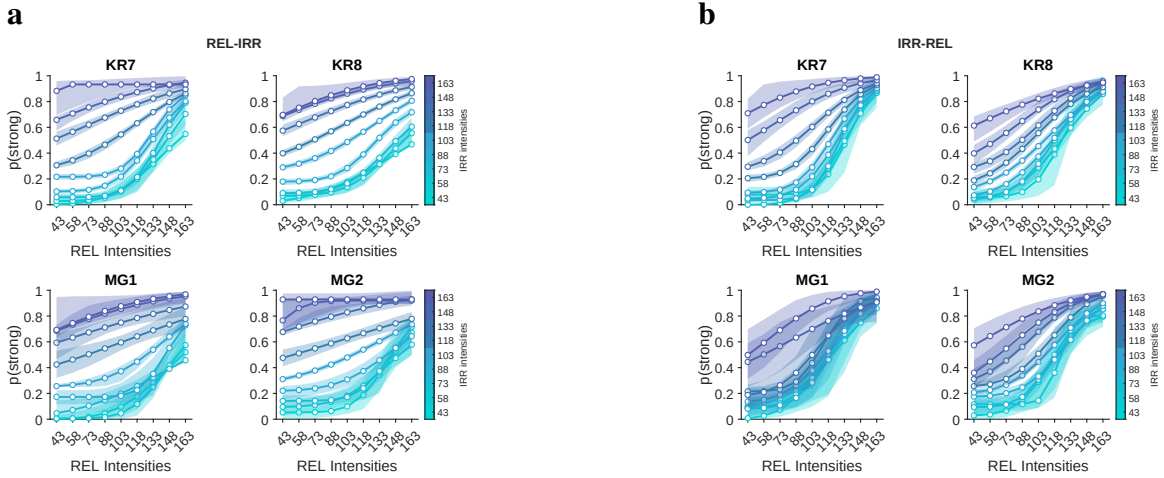


Figure 9: Psychometric functions describing the IRR effect. **a.** Psychometric functions showing the IRR-induced bias for REL-IRR trials, for single subjects. **b.** Same, for IRR-REL trials. Shaded areas represent the 95% bootstrapped confidence intervals.

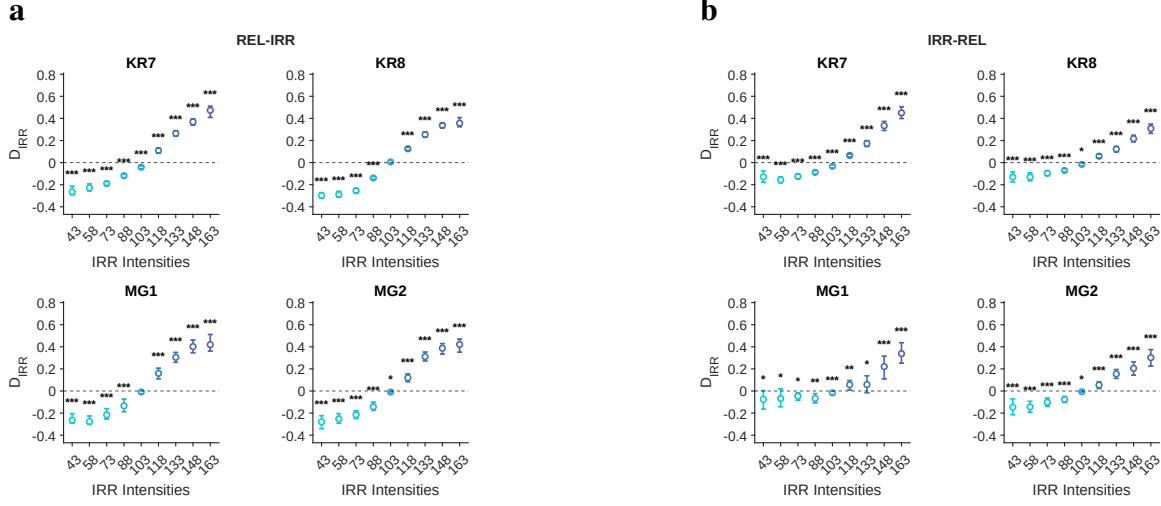


Figure 10: IRR-induced bias for REL-IRR trials and IRR-REL trials.

a. IRR distance for REL-IRR trials. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. **b.** Same, for IRR-REL trials.

Behavioral patterns compatible with a second-stimulus heuristic

To further describe this, we recomputed psychometric functions by plotting choice as a function of IRR rather than REL. Now, performance improved selectively in conditions in which IRR was in the second position as compared to IRR in the first position (Figure 11 and Figure 12). This is compatible with the possible use of a simple heuristic favoring the use of the second stimulus, irrespective of its actual relevance according to the task rule.

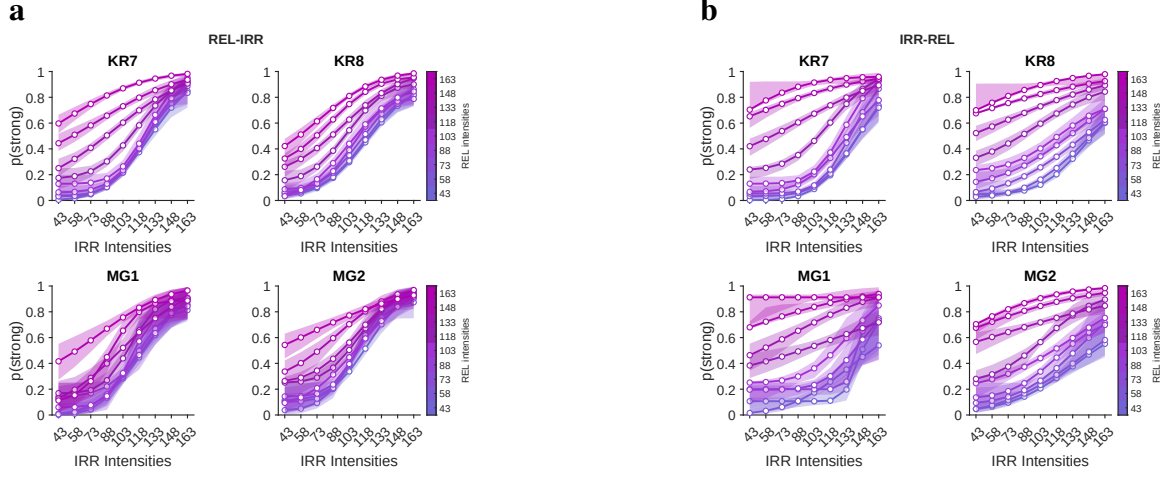


Figure 11: Psychometric functions assuming rats judge IRR. a. Psychometric functions showing the REL-induced bias for REL-IRR trials, for single subjects. **b.** Same, for IRR-REL trials. Shaded areas represent the 95% bootstrapped confidence intervals.

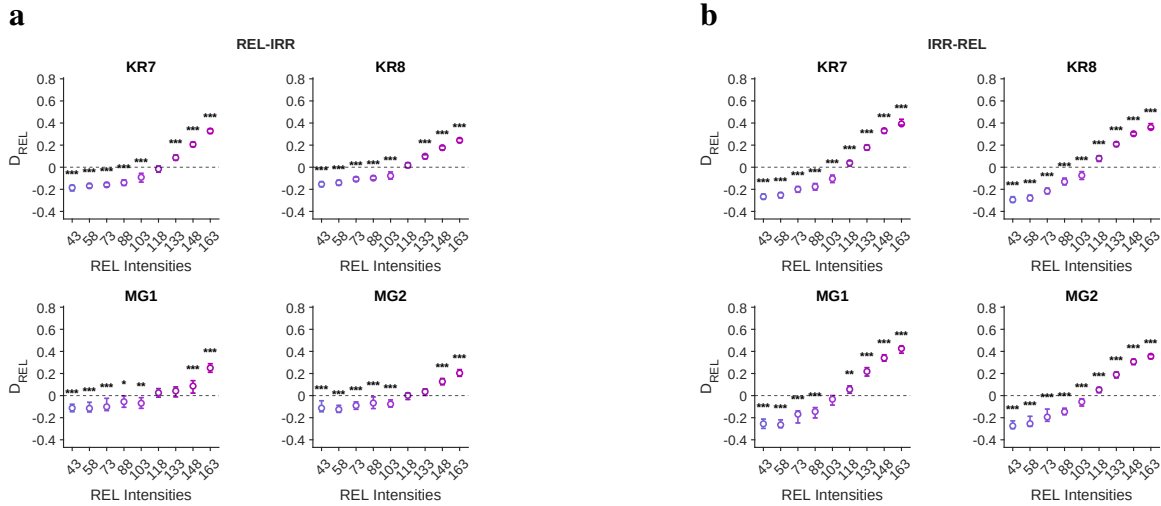


Figure 12: REL-induced bias for REL-IRR trials and IRR-REL trials. a. REL distance for REL-IRR trials. For each animal, D_{REL} quantifies the deviation of the REL-conditioned IRR psychometric function from the overall IRR psychometric function at each REL intensity. Circles show the observed distances and error bars the 95% bootstrapped confidence intervals. Dashed lines indicate $D_{REL} = 0$. Significance was assessed using two-sided permutation tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **b.** Same, for IRR-REL trials.

Learning reduces the influence of irrelevant stimuli when task structure is stable

We next examined how sensitivity to irrelevant information evolved with training, relative to the standard task. For each subject, we estimated the contribution of REL and IRR to choice using a generalized linear model (GLM) applied to successive bins of training sessions, with the goal of describing how IRR influence changed over time.

Rats exhibited distinct learning trajectories. In some animals (KR2, KR3, KR4), the weight assigned to IRR progressively decreased over training, approaching zero, consistent with a gradual suppression of irrelevant information (Figure 13). In others (KR5, KR6), IRR weights were near zero from early in training, suggesting a rapid acquisition of an effective strategy.

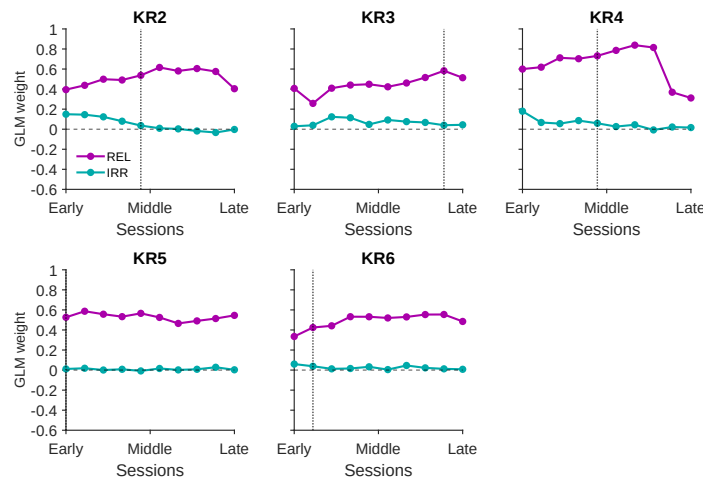


Figure 13: Progression of learning (standard task). Contribution of REL and IRR stimuli to the choice, as described by the GLM weights. All training sessions were binned into 10 groups, and GLM coefficients were computed for each group. The vertical line indicates the transition between proficient and expert stages.

Based on these dynamics, we defined two operational regimes to better describe these changes: a proficient stage, in which rats performed the task successfully but remained susceptible to IRR-induced bias, and an expert stage, in which IRR exerted little to no influence on behavior. When the relative ratio between REL and IRR coefficients was below 0.1 in at least two consecutive session bins, we considered the animal to have transitioned into the expert stage (Figure 14)

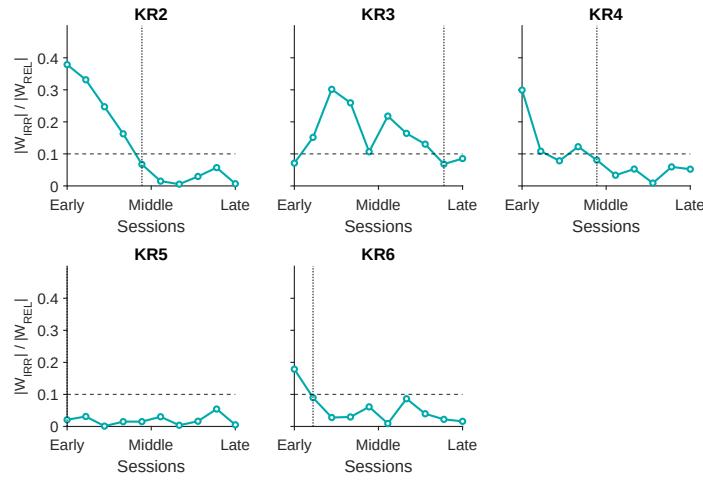


Figure 14: IRR/REL ratio dynamics across time. The relative contribution of the irrelevant and relevant stimulus to the choice was quantified as $R_{IRR} = |W_{IRR}|/|W_{REL}|$, using the GLM weights previously computed for learning trajectories. Lower values indicate a weaker influence of the IRR stimulus relative to the REL stimulus. Vertical lines indicate the transition between proficient and expert stages.

In rats showing gradual learning, the transition between these stages was accompanied by a clear reduction in bias magnitude (Figure 15 and Figure 17), whereas in rapidly learning rats no such transition was apparent (Figure 16 and Figure 17).

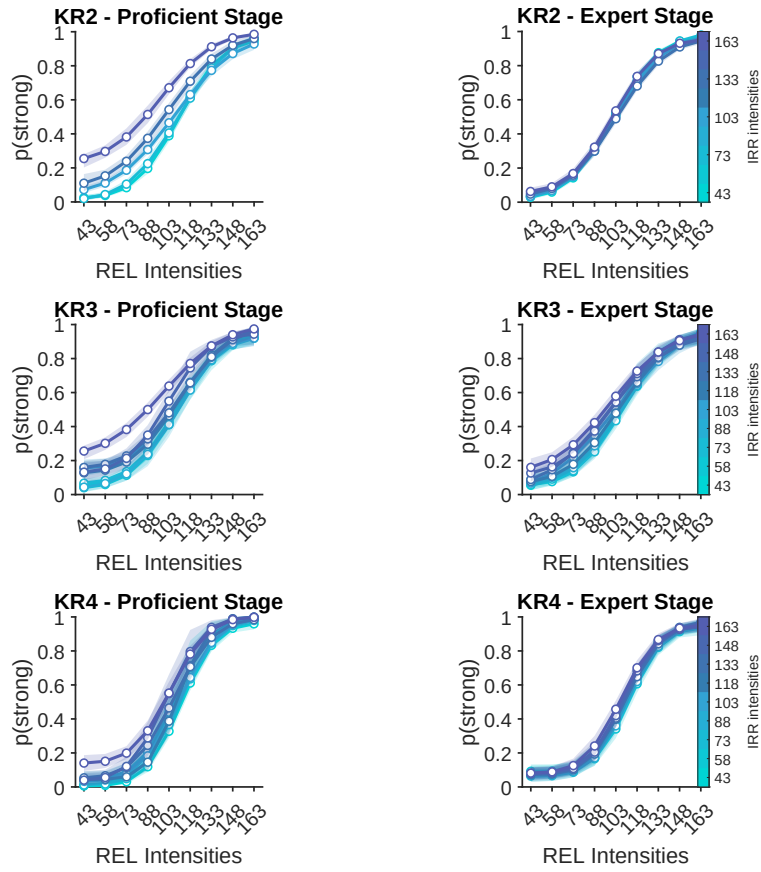


Figure 15: Psychometric functions identifying the proficient stage and expert stage. Difference in IRR bias during the two training stages, for rats KR2, KR3, KR4 performing the standard task. Shaded areas represent the 95% bootstrapped confidence intervals.

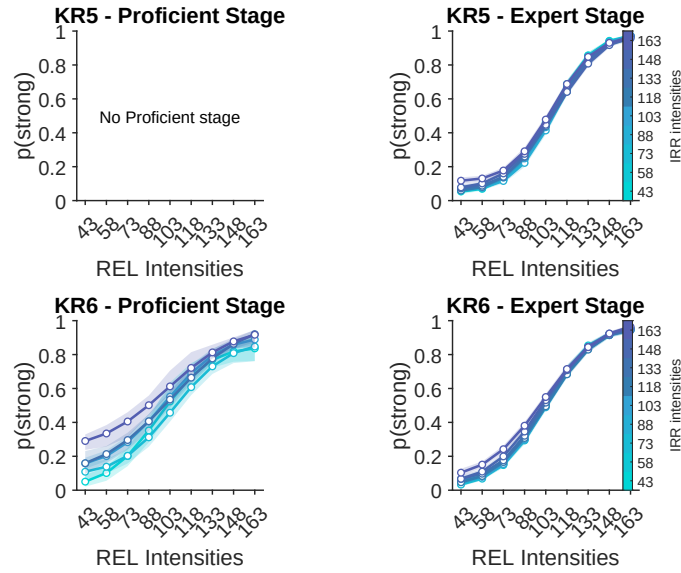


Figure 16: Psychometric functions identifying the proficient stage and expert stage. Difference in IRR bias during the two training stages, for rats KR5 and KR6 performing the standard task. Shaded areas represent the 95% bootstrapped confidence intervals.

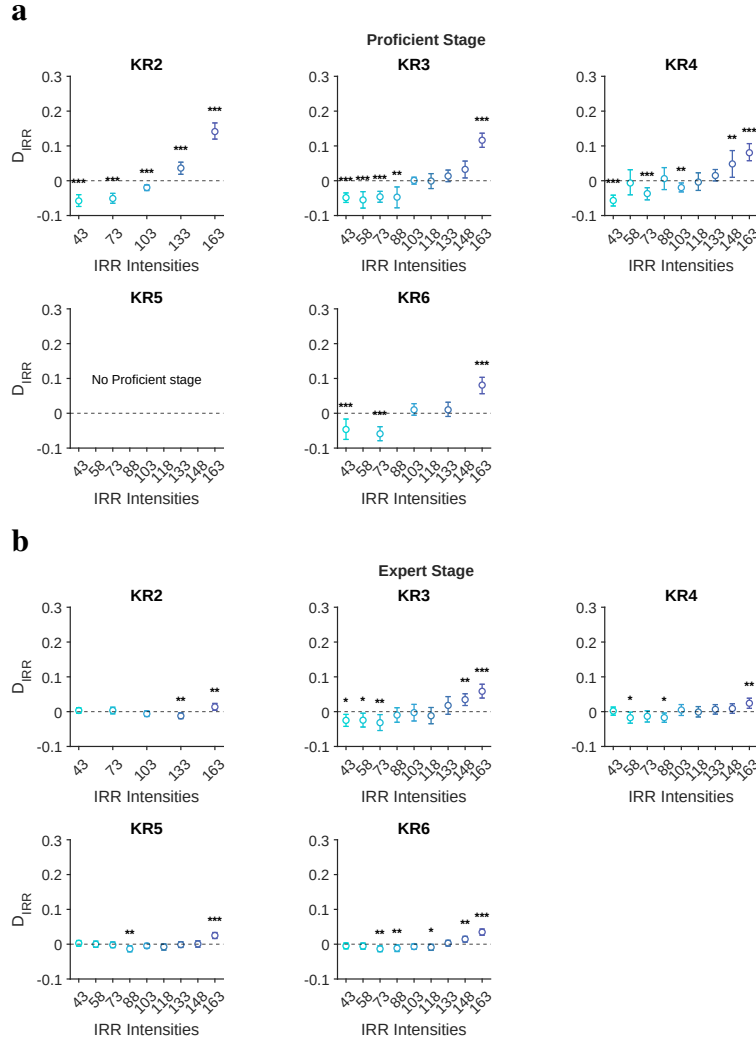


Figure 17: IRR-induced bias during the proficient and expert stages.
a. IRR distance for rats KR2, KR3, KR4, and KR6 during the proficient stage. Rat KR5 had no proficient stage. **b.** IRR distance for rats KR2, KR3, KR4, KR5, and KR6 during the expert stage. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Engagement and disengagement states can be an alternative explanation to the reduction of the IRR bias across time. Increased engagement towards task demands might have improved performance and the ability to suppress IRR. To address this possibility, we considered how performance (i.e., proportion of correct responses) changed between the first and second

halves of each experimental session (Figure 18). In all rats, performance remained relatively stable within and across experimental sessions, suggesting that the reduction of the IRR bias was more likely due to an effective attentional strategy, rather than variability in engagement states.

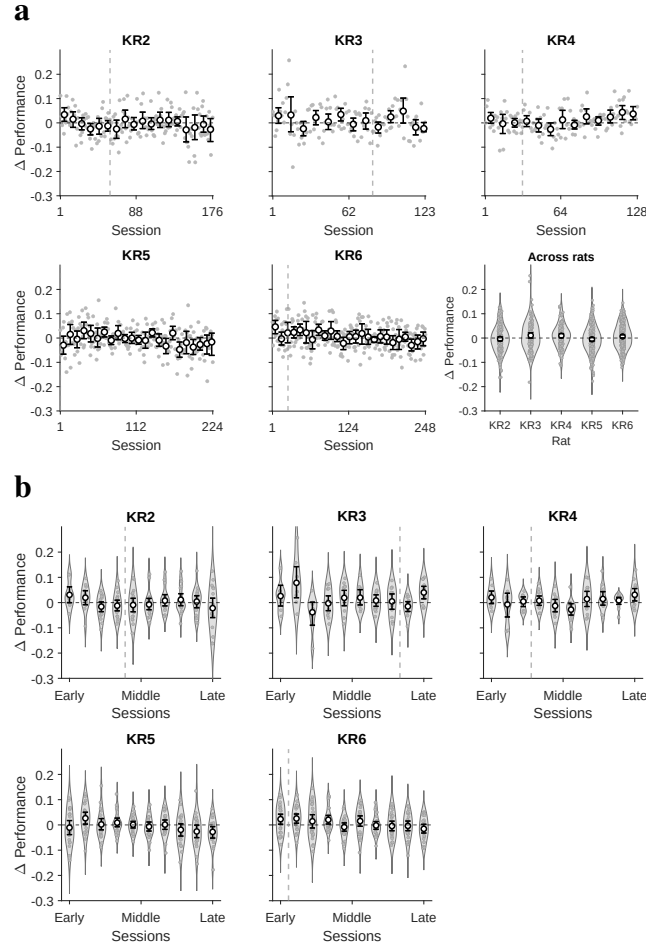


Figure 18: Stability of within-session performance. **a.** Difference between second- and first-half performance for each session (Δ performance). Gray dots indicate individual sessions, circles and error bars show the mean and bootstrapped 95% confidence interval across successive bins. Vertical dashed lines indicate expert stage onset. Bottom right: overall distribution of session wise Δ performance across animals. **b.** Distribution of Δ performance across the 10 bins used for the GLM analysis. Violins show the session-wise distributions, circles and error bars indicate the mean and bootstrapped 95% confidence interval.

Video inspection during the experiment, and recorded video clips, showed that rats' whiskers were in contact with the plate during both IRR and REL, meaning that the reduced weight of the IRR did not arise from preventing the vibration from being transduced by the sensory system.

Together, these results indicate that rats could learn to ignore irrelevant information when task structure is stable, but exhibited limited flexibility and stable learning trajectories when relevance must be rapidly reassigned across time (see Supplementary Figure S4).

Neural representations of relevant and irrelevant stimuli

Encoding of relevant and irrelevant stimuli in expert animals

To relate neural activity to behavioral performance, we recorded extracellular activity from motor (M1/M2) and somatosensory (vS1) cortices in four rats performing the task. These animals spanned different behavioral regimes: two expert rats (KR2, KR3), one proficient rat (MG1), and one rat trained in the dynamic task (KR8). This allowed us to relate neural dynamics to behavioral strategy and performance.

We first examined neural representations in expert rats performing the standard task. Behavioral analyses during recording sessions confirmed that these animals had effectively suppressed the influence of the irrelevant stimulus (Figure 19).

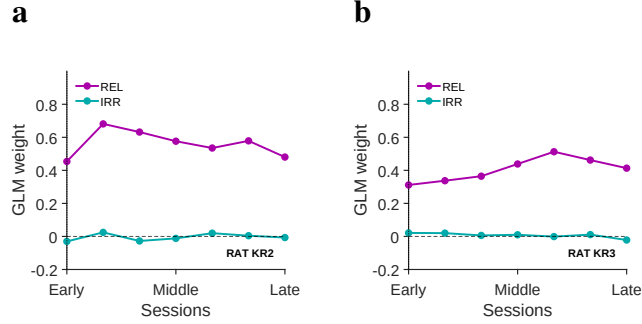


Figure 19: GLM weights during the recording sessions (standard task), for the two M1/M2-implanted expert rats. a. Weights of REL and IRR during the recording sessions, for rat KR2. **b.** Weights of REL and IRR during the recording sessions, for rat KR3.

Single-unit activity in M1/M2 showed modulation by both stimulus relevance (REL vs IRR) and intensity category (Figure 20), indicating sensitivity to task variables. At the population level, decoding analyses revealed that the intensity category of REL and the animal's choice could be robustly predicted from M1/M2 activity in both rats (Figure 21). In contrast, the irrelevant stimulus (IRR) was either not decodable (KR2) or only weakly represented (KR3), particularly during the inter-stimulus interval. However, both rats exhibited similar behavioral performance despite encoding IRR differently (see Supplementary Figure S5).

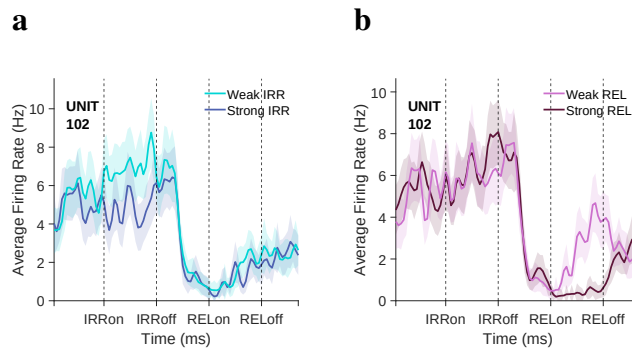


Figure 20: Example single-unit firing rate a. Average firing rate of example unit 102 during the entire trial, conditioned on IRR intensity category. **b.** Average firing rate conditioned on REL intensity category.

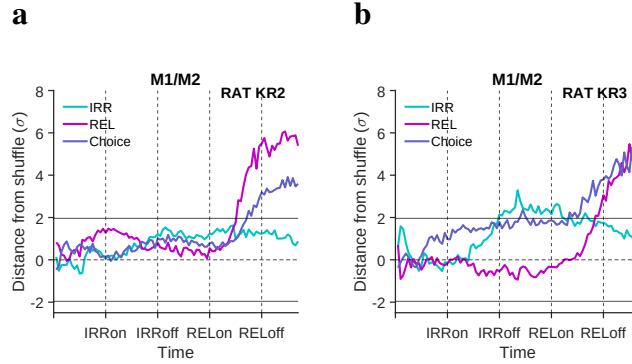


Figure 21: IRR, REL, and choice decoding in M1/M2, for rats KR2 and KR3. a. Decoding results for rat KR2. **b.** Decoding results for rat KR3.

Supplementary Figure S8 and Supplementary Figure S9 show the histological verification of electrode positioning for rats KR2 and KR3.

Neural correlates of residual bias in a proficient animal

We next analyzed a proficient animal (MG1), characterized behaviorally by accurate judgment of REL in addition to IRR-induced bias (Figure 22). The GLM analysis across the recording period indicated a transition toward the expert stage, with the influence of IRR becoming minimal during later sessions. However, because the animal did not exhibit an expert behavioral profile throughout the entire recording period, we described it as proficient. Recordings in both vS1 and M1/M2 nevertheless revealed substantial differences from those observed in expert animals.

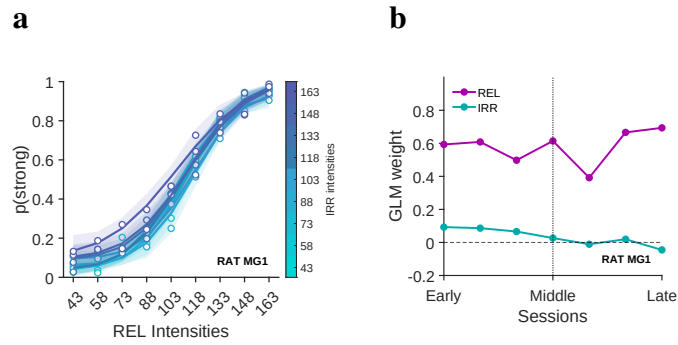


Figure 22: Behavioral performance of proficient rat MG1 during the recording period. a. Psychometric functions describing the IRR attractive bias. **b.** GLM weights of REL and IRR.

In vS1, both REL and IRR were significantly decodable (Figure 23a), indicating that irrelevant information was still encoded at the sensory level. In M1/M2, IRR remained decodable but showed partial suppression following REL onset, while REL was robustly represented (Figure 23b). Choice signals, however, were weak or absent compared to expert animals, suggesting incomplete transformation of sensory inputs into categorical decisions. Comparing decoding magnitudes across areas revealed stronger signals in M1/M2 than in vS1, consistent with the emergence of more abstract representations in frontal cortex [26,27]. Temporal analyses indicated that both areas encoded stimuli on similar time scales, with no consistent lag between them (Figure 23c-d), suggesting coordinated rather than strictly sequential processing.

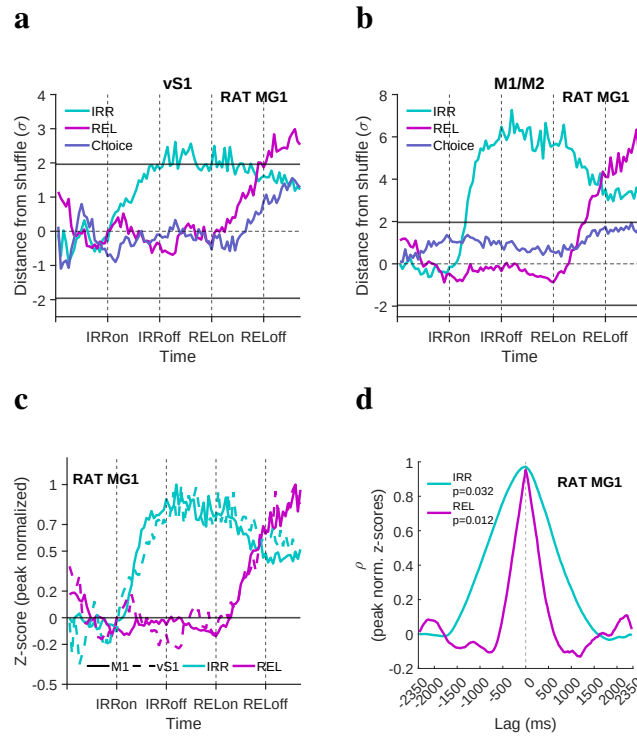


Figure 23: IRR, REL, and choice decoding in vS1 and M1/M2, for proficient rat MG1. **a.** Decoding in vS1. **b.** Decoding in M1/M2. **c.** Peak-normalized z-scores of IRR and REL decoding in both vS1 and M1/M2. **d.** Lag analysis (cross-correlation of peak-normalized z-scores of vS1 and M1/M2).

Supplementary Figure S10 shows the histological verification of electrode positioning for rat MG1.

Outcome-dependent modulation of stimulus representations

To probe how neural representations relate to behavioral success, we compared decoding during correct and error trials.

In expert animals, IRR decoding was minimal or transient during correct trials but could occasionally emerge during error trials (Figure 24). In contrast, REL remained decodable

in both outcomes, but its representation differed qualitatively, suggesting that errors may reflect misrepresentation rather than absence of encoding.

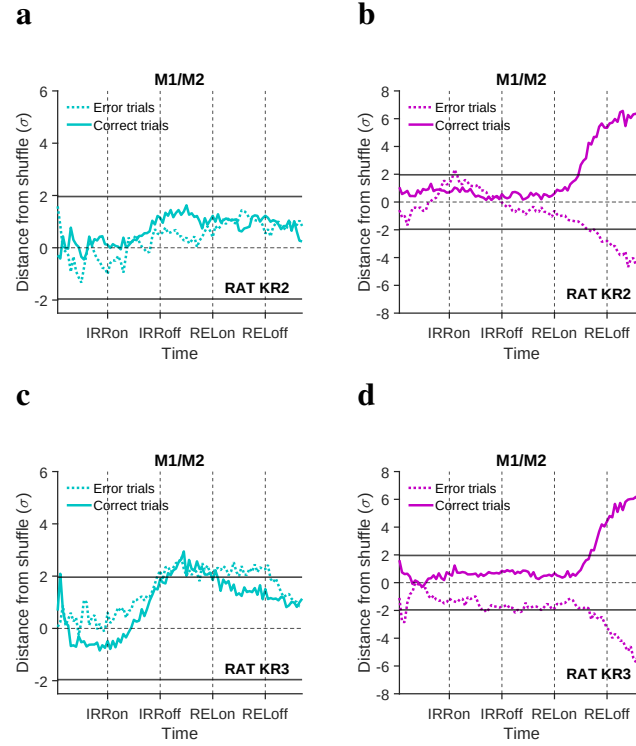


Figure 24: Decoding REL and IRR during correct and error trials, for rats KR2 and KR3. a. IRR decoding for rat KR2. **b.** REL decoding for rat KR2. **c.** IRR decoding for rat KR3. **d.** REL decoding for rat KR3.

In the proficient animal, outcome-dependent effects were more pronounced (Figure 25). In vS1, IRR was significantly decoded during error trials but suppressed during correct trials, consistent with attentional modulation at the sensory level. In M1/M2, both REL and IRR were generally represented, with stronger engagement during correct trials. However, during errors, both areas exhibited weaker and less consistent encoding of REL, indicating a broader failure of stimulus processing.

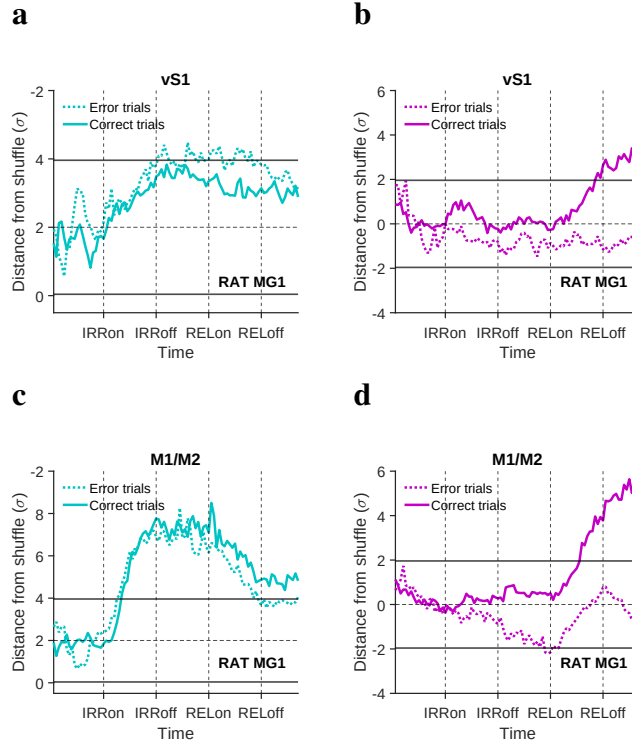


Figure 25: IRR and REL decoding in vS1 and M1/M2, during error and correct trials, for rat MG1. a. IRR decoding in vS1. b. REL decoding in vS1. c. IRR decoding in M1/M2. d. REL decoding in M1/M2.

Together, these results suggest that proficient performance was associated with partial suppression of irrelevant information and that errors reflected a broad disruption in stimulus representation and attentional modulation.

Misalignment between neural encoding and behavior in the dynamic task

Finally, we examined neural representations in a rat performing the dynamic task (KR8), which behaviorally relied on a suboptimal heuristic favoring the second stimulus (Figure 26).

Decoding analyses revealed a striking dissociation between neural representations and behavior (Figure 27). In both vS1 and M1/M2, the first stimulus of the sequence —

regardless of its relevance — was robustly encoded, whereas the second stimulus was only weakly or inconsistently represented. Choice signals were also weak and did not reach significance.

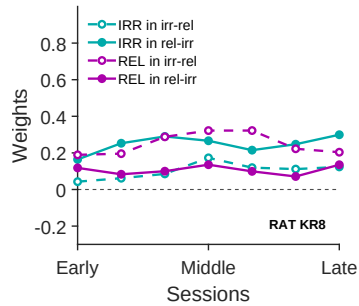


Figure 26: GLM weights during the recording sessions of rat KR8. GLM analysis of IRR and REL weights for REL-IRR trials and IRR-REL trials.

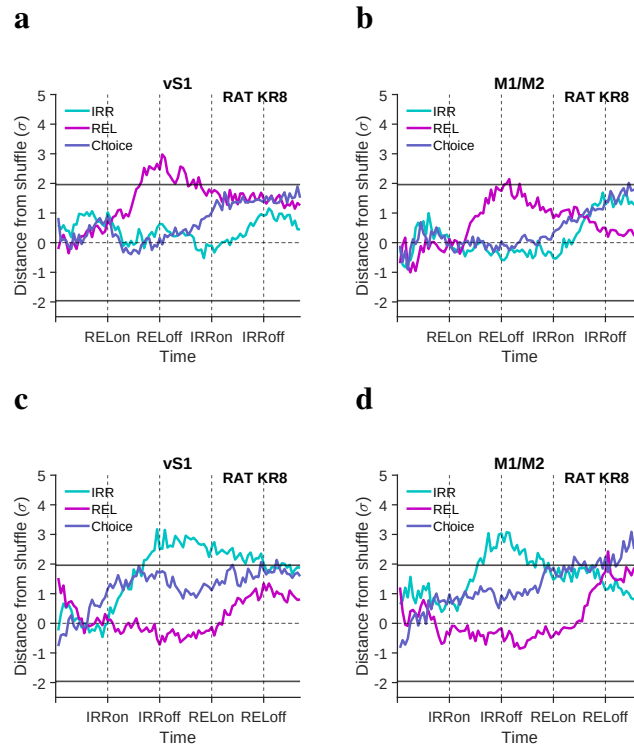


Figure 27: Decoding results in vS1 and M1/M2, for REL-IRR and IRR-REL trials, for rat KR8. a. Decoding in vS1, for REL-IRR trials. **b.** Decoding in M1/M2, for REL-IRR trials. **c.** Decoding results in vS1, for IRR-REL trials. **d.** Decoding results in M1/M2, for IRR-REL trials.

This pattern indicates a misalignment between neural encoding and behavioral strategy: while neuronal populations preferentially represented the first stimulus, behavior was biased toward the second. These findings highlight that suboptimal behavioral strategies were associated with misaligned or weaker neural representations.

Supplementary Figure S11 shows the histological verification of electrode positioning for rat KR8.

Inter-areal interactions during task performance

So far, we have shown that neural representations of relevant and irrelevant stimuli differed across behavioral regimes. We next asked how interactions between sensory and frontal cortices contributed to this process, focusing on inter-areal synchronization and the direction of information flow.

To investigate how interactions between sensory and frontal cortices supported task performance, we analyzed local field potentials (LFPs) in vS1 and M1/M2, focusing on synchronization and directional coupling.

Coherence reflects effective stimulus processing

In the proficient animal performing the standard task, vS1 and M1/M2 showed strong coherence across multiple frequency bands during stimulus processing (Figure 28b,d). In particular, gamma-band synchronization increased during stimulus epochs and was stronger for the relevant stimulus, consistent with a role in information transfer.

Interestingly, coherence increased further during error trials (Figure 28a,c), particularly in the gamma range. This hyper-synchronization may have reflected inefficient or uncontrolled network interactions, indicating that stronger coupling was not necessarily beneficial for performance.

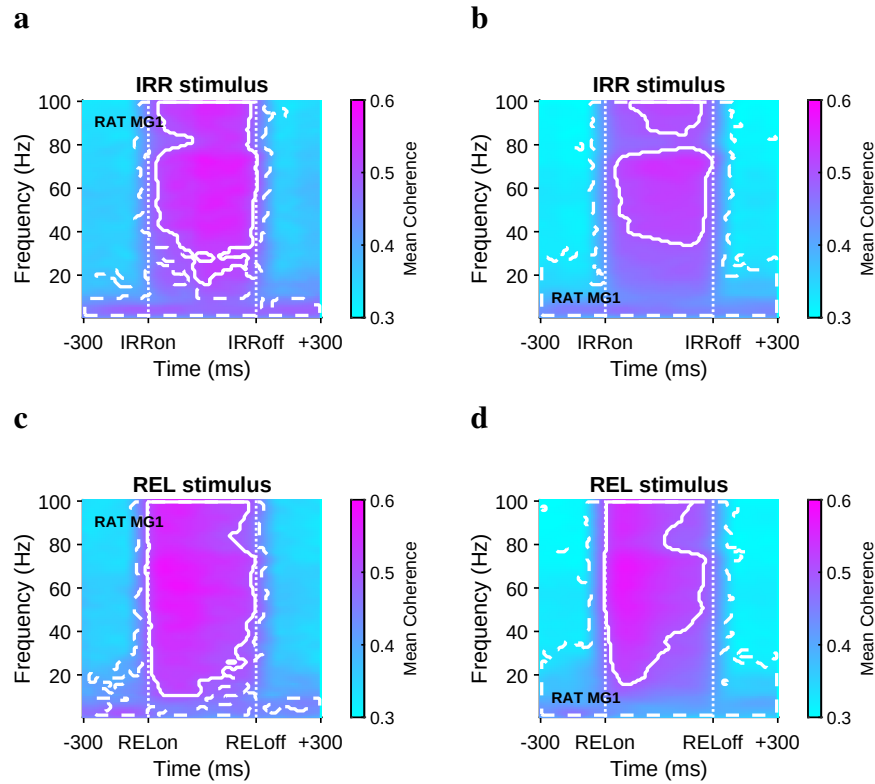


Figure 28: Coherograms during IRR and REL stimulus epochs, for incorrect and correct trials, for rat MG1. **a.** Coherogram for IRR stimulus epoch during incorrect trials (including 300 ms before stimulus onset, and 300 ms after stimulus offset). Dashed lines represent frequency ranges and time bins during which coherence was statistically significant. Solid lines represent frequency ranges and time bins during which coherence was higher than 0.5. **b.** Coherogram for IRR stimulus, during correct trials. **c.** Coherogram for REL stimulus, during incorrect trials. **d.** Coherogram for REL stimulus, during correct trials.

In contrast, in the dynamic task animal, coherence between areas was markedly reduced (Figure 29). This reduction paralleled weaker decoding of both stimuli, supporting the idea that inter-areal synchronization contributed to effective stimulus representation.

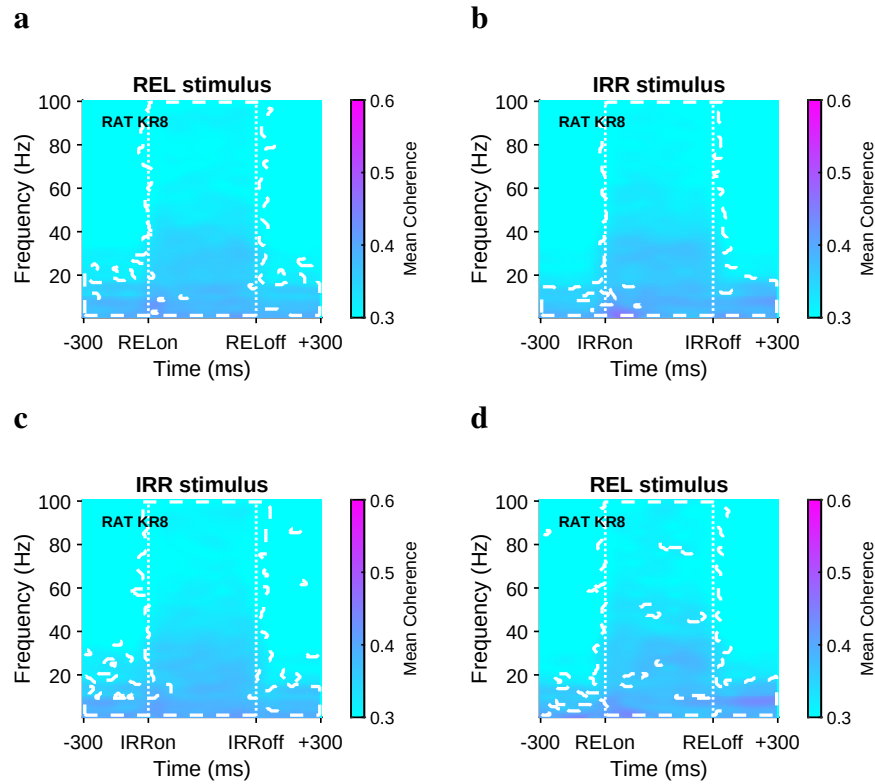


Figure 29: Coherograms during IRR and REL stimulus epochs, for REL-IRR and IRR-REL trials, for rat KR8. **a.** Coherogram for REL stimulus epoch during REL-IRR trials (including 300 ms before stimulus onset, and 300 ms after stimulus offset). Dashed lines represent frequency ranges and time bins during which coherence was statistically significant. **b.** Same, for IRR stimulus. **c.** Coherogram for IRR stimulus, during IRR-REL trials. **d.** Same, for REL stimulus.

Directionality of information flow: top-down modulation

We next examined the direction of information flow using Granger causality analysis. In the proficient animal, outcome-dependent analyses revealed that, during correct trials, both stimulus epochs were characterized by significant top-down modulation from M1/M2 to vS1 across a broad frequency range, particularly in beta and gamma bands (Figure 30b,d). This modulation during the stimulus presentation suggests that frontal cortex exerted control over sensory processing. By contrast, during error trials, top-down modulation was reduced

(Figure 30a,c), especially in the beta band, indicating that effective top-down influence was associated with successful performance. Supplementary Figure Figure S6 shows the time-frequency Granger causality results, pooling correct and error trials.

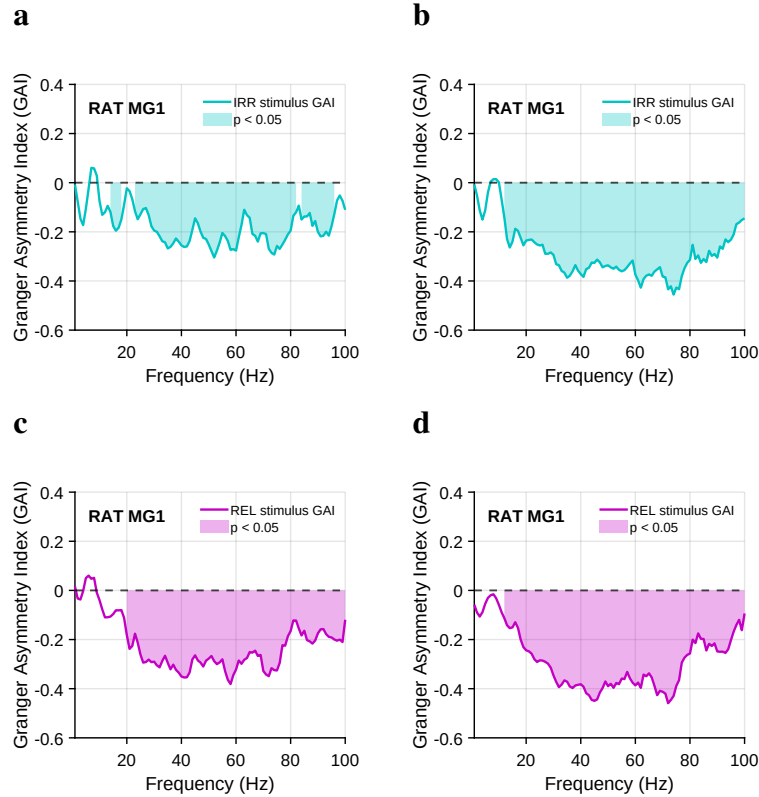


Figure 30: Granger Causality between vS1 and M1/M2, for both REL and IRR epochs, during error and correct trials, for rat MG1.
a. Frequency-domain Granger Causality during IRR epoch, for error trials. Shaded areas represent frequencies in which GAI values were statistically significant ($p < 0.05$, cluster-based permutation test). **b.** Same, for correct trials. **c.** Frequency-domain Granger Causality during REL stimulus epoch, for incorrect trials. **d.** Same, for correct trials.

In the dynamic task animal, top-down modulation was present but altered (Figure 31). Beta-band influence emerged primarily during the second stimulus epoch, aligning with the animal's behavioral bias toward the second stimulus. However, this modulation failed to support effective stimulus representation, as gamma synchronization remained weak.

Supplementary Figure S7 shows the time-frequency domain Granger causality for both trial types.

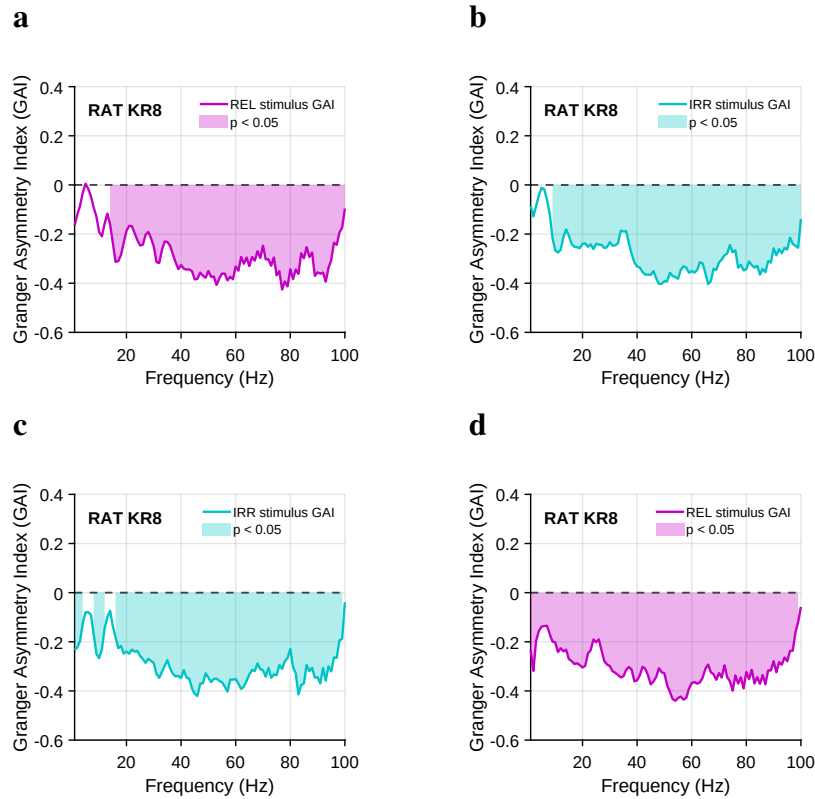


Figure 31: Granger Causality between vS1 and M1/M2, for both REL and IRR epochs, during REL-IRR and IRR-REL trials, for rat KR8. **a.** Frequency-domain Granger Causality during REL epoch, for REL-IRR trials. Shaded areas represent frequencies in which GAI values were statistically significant ($p < 0.05$, cluster-based permutation test). **b.** Same, for IRR stimulus. **c.** Frequency-domain Granger Causality during IRR epoch, for IRR-REL trials. **d.** Same, for REL stimulus.

Summary

Across animals, neural representations of relevant and irrelevant stimuli closely tracked behavioral strategies and levels of expertise. Expert performance was associated with strong encoding of relevant information and effective suppression of irrelevant inputs, whereas pro-

ficient performance reflected partial suppression and increased variability. In the dynamic task, suboptimal strategies were accompanied by a marked dissociation between neural encoding and behavior. At the circuit level, these differences were paralleled by changes in inter-areal synchronization and top-down modulation between frontal and sensory cortices. Given the small sample size, differences observed across brain areas and learning stages should be considered preliminary. Expertise, recording area, configuration, and task were not fully balanced, preventing these factors from being completely dissociated. The observed differences identify candidate relationships between learning and neural population dynamics, rather than definitive regional or learning effects.

Discussion

Irrelevant information biases perceptual judgments

In this study, we examined how rats allocate attention across time when required to ignore an irrelevant stimulus. Although animals successfully learned the task, their judgments were systematically biased by the irrelevant stimulus, which exerted an attractive effect on perception. This bias was strongest for weak relevant stimuli and decreased as a function of inter-stimulus interval (ISI), indicating that the interaction between stimuli is temporally constrained.

The asymmetry of the effect, with weaker relevant stimuli being more susceptible to interference, is consistent with differences in signal reliability. Weak stimuli generate less robust representations and are therefore more vulnerable to contamination by competing inputs [28, 29].

Learning reduces – but does not abolish – irrelevant information

Rats trained on the standard task progressively reduced the influence of the irrelevant stimulus. This is evident from the decrease in its contribution to behavioral choice over training. This reduction does not imply that irrelevant information disappeared at the sensory level. Rather, the data suggest that its impact on behavior was attenuated.

Across animals, learning followed different trajectories. Some rats gradually reduced the influence of the irrelevant stimulus, while others showed little susceptibility from early in training. We therefore define two operational regimes: a proficient regime, in which behavior was accurate but still biased, and an expert regime, in which the bias was minimal.

In contrast, rats trained on the dynamic task did not achieve comparable suppression. When relevance varied across trials, animals did not rely on the acoustic cue but instead adopted a simpler rule, favoring the second stimulus. This suggests that while rats could learn to allocate attention across time under stable conditions, they exhibited limited flexibility when relevance had to be reassigned dynamically.

Behavioral variability reflects constrained solutions

Despite variability across individuals, behavior was not random [30,31]. Instead, it clustered into a small number of reproducible patterns. In the standard task, animals differed in the degree to which irrelevant information influenced their choices. In the dynamic task, all animals converged on a similar heuristic based on stimulus order.

These observations suggest that attentional control in this paradigm did not rely on a single solution, but instead emerged from a constrained set of strategies shaped by task structure [32–34]. Temporal order appeared to act as a strong organizing principle, even when it was not explicitly informative.

The distinction between proficient and expert performance should be understood as an operational one. The data did not indicate discrete categories, but rather regimes within a continuum of behavior. In addition, the psychometric functions describing performance during the proficient and expert stages should be considered as an internal consistency check, and not independent validation of the learning trajectories.

Performance stability and alternative explanations

An alternative interpretation of the behavioral variability observed in the current study is that individual differences reflected different occupancy of shared latent behavioral states,

rather than qualitatively distinct strategies across animals [35]. Pooling sessions together could make an alternation between engaged and disengaged states appear as a stable trait. We cannot exclude this possibility without an explicit latent-state model.

Nevertheless, animals showed a relatively stable performance both across experimental sessions and within individual sessions. Rats that exhibited a decreasing influence of IRR, while retaining a stable performance across time, suggest a pattern more consistent with a gradual learning-related change in how information is weighted than with simple alternation between periods of engagement and disengagement.

Application of approaches such as GLM-HMMs could determine whether these trajectories reflected changes in the occupancy of shared latent states or more persistent subject-specific strategies.

Neural representations dissociate encoding from readout

To relate behavioral variability to neural dynamics, we recorded activity in M1/M2 and vS1 across animals expressing different behavioral regimes. Across these recordings, a consistent dissociation emerged between sensory encoding and behavioral readout.

In expert animals, M1/M2 robustly encoded the relevant stimulus and predicted choice, while the irrelevant stimulus was weakly represented or not decodable. However, the two expert rats differed in the extent to which irrelevant information remained detectable. One animal showed no clear representation of the irrelevant stimulus, whereas the other retained a residual encoding despite the absence of behavioral influence.

This dissociation suggests that suppression of irrelevant information operated primarily at the level of readout rather than through complete elimination of sensory encoding. In other words, irrelevant stimuli may still have been represented neurally, but their influence on

choice was reduced. Accordingly, similar behavioral outcomes could arise from different underlying neural states. Linear decodability indicates, as in our case, that information was present and encoded in the population, but it does not demonstrate that this information was used to guide behavior, as in the case of the expert rat retaining residual encoding of the irrelevant stimulus. Information may remain represented, but have little influence on downstream decision processes. Context-dependent population dynamics may determine which components of the representation are read out for guiding behavior [36].

The distinction between encoding and readout is also relevant to the early rise of choice-related signals observed in rats KR3 and KR8, where choice could be decoded before the REL representation emerged. Such choice activity does not necessarily imply that the decision had been formed, but may have reflected biases, history effects, anticipatory activity, or other internal states correlating with the choice itself.

While these patterns are consistent with different computational solutions, such as selection-like versus integration-like processing, they do not uniquely specify the underlying mechanisms. Given the limited number of subjects and the indirect nature of decoding analyses, alternative explanations remain possible.

In the proficient animal, both relevant and irrelevant stimuli were robustly encoded in vS1 and M1/M2. Partial suppression of the irrelevant stimulus emerged only after the onset of the relevant one, and choice-related signals were weaker and less stable than in expert animals. This suggests that proficient performance reflected a regime in which irrelevant information was still processed and only incompletely filtered.

Error analyses support this view. Failures were associated with both degraded representation of relevant information and inconsistent attenuation of irrelevant input, indicating that errors reflected a distributed breakdown of processing rather than a single deficit.

Suboptimal strategies reveal dissociations between encoding and decision

In the dynamic-task animal, neural and behavioral patterns diverged in a striking way. Neural activity in both vS1 and M1/M2 preferentially encoded the first stimulus, whereas behavior was biased toward the second. This revealed a misalignment between early sensory processing and downstream decision mechanisms.

One possible interpretation is that sensory encoding was largely driven by stimulus order or salience, whereas decision processes applied an additional temporal weighting favoring more recent inputs. Under this account, the observed heuristic arose at the level of readout rather than encoding.

More generally, this finding suggests that suboptimal strategies were not simply weaker versions of optimal ones. Instead, they may have reflected qualitatively different mappings between sensory representations and behavior.

Nevertheless, these observations should be considered preliminary. The limited number of animals, together with differences in expertise, task configuration, and recorded areas across animals, prevent these factors from being fully dissociated. These findings therefore represent candidate patterns that will require corroboration in a larger, more balanced dataset.

Inter-areal interactions support – but do not determine – performance

Efficient task performance was associated with increased coordination between vS1 and M1/M2. In the proficient animal, coherence between these regions increased during stimulus processing, particularly in the gamma band. However, synchronization was even stronger

during error trials, indicating that more coupling is not necessarily beneficial [37, 38].

These results suggest a non-monotonic relationship between inter-areal coupling and performance, in which both insufficient and excessive synchronization could have been suboptimal. In the dynamic-task animal, reduced coherence was accompanied by weaker stimulus encoding, consistent with the idea that coordinated activity supports effective processing. However, these observations remain correlational.

Analyses of directionality revealed top-down influences from M1/M2 to vS1, particularly during stimulus epochs. This influence was reduced during error trials, especially in the beta band, and may have reflected engagement of task-related control processes [39, 40]. However, given the limited dataset and the nature of the analysis, this interpretation should be treated with caution.

Interestingly, in the dynamic-task animal, beta modulation was aligned with the second stimulus, mirroring the behavioral bias, yet did not improve representation. This suggests that patterns of inter-areal communication can reflect the adopted strategy, even when that strategy is suboptimal.

Adaptation, attention, or memory interference?

The behavioral and neural results are consistent with an attentional interpretation, where the influence of the first stimulus may reflect incomplete filtering of irrelevant information. However, alternative mechanisms could also explain this effect. One possibility is sensory adaptation, whereby the processing of the first stimulus can alter the neural response to the subsequent relevant stimulus. Therefore IRR could modulate activity in sensory and frontal areas, influencing the encoding of REL, without necessarily reflecting a failure of attentional mechanisms. Nevertheless, a short-lived adaptation phenomenon appears less

likely in our case, given that the two stimuli were separated by an interval of at least 500 ms. Moreover, the persistence of IRR-related neural encoding beyond stimulus presentation suggests that the observed effect may have involved mechanisms beyond immediate sensory adaptation. Although adaptation cannot be fully excluded, the results are more consistent with the influence of irrelevant information on subsequent processing.

Memory interference can also explain the results. The representation of IRR may have persisted across the ISI and interfered with the subsequent encoding or readout of REL. This interpretation is compatible with the behavioral influence of IRR during the proficient stage. Nevertheless, this influence became minimal in expert rats, despite the presence of IRR encoding in the frontal cortex. Attentional control may therefore act by reducing this interference on the processing or downstream readout of REL. In this sense, memory interference and attentional control are not mutually exclusive, but rather two complementary processes acting simultaneously.

Methods

All protocols conformed to international norms and were approved by the Ethics Committee of SISSA and by the Italian Health Ministry.

Rat subjects

9 Wistar male rats (Envigo) were housed in pairs and examined periodically by a veterinarian. They were regularly checked for their health and welfare conditions, and provided daily with environmental and social enrichment. They were maintained on a 12/12-hour light/dark cycle. The rats began the training protocols at the age of 8 weeks. To maintain motivation during the behavioral training, animals were water-deprived. The behavioral training protocol required approximately 1 hour of training on each week-day. During weekends, the rats had access to water for around 30 hours.

Experimental apparatus

The experimental setup consisted of a T-shaped Plexiglas box, measuring 25 x 25 x 38 cm. In the front wall, a hole allowed the rat to extend its head into the stimulus delivery port. Here, a nosepoke detected the snout by an infrared detector, while a blue LED light signaled to the rat that a new trial could begin. The setup was positioned into a Faraday cage, and the experiments were run in the dark. The only light provided was a red LED, and was used by the experimenter to monitor the rats during the task, via a webcam (Logitech HD Webcam C310).

Two reward spouts, on 8 cm high pedestals, were placed to the right and left compared to the nosepoke. Two custom made infrared sensors, one for each spout, detected the rat's position

and activated reward delivery accordingly. Water was delivered by two syringe-pumps, one per side, controlled by an AVR32 board (National Instruments, Austin, TX).

Three audio speakers were placed on the three walls of the setup, one above each spout and one at the back. The one at the back delivered the go cue, while the lateral ones gave a reward delivery cue, upon correct response, to act as reinforcement. Two additional audio speakers were placed on the front external sides of the setup: these delivered the acoustic cue used to tag the relevant vibrotactile stimulus. Experiments were controlled using in-house LabVIEW software (National Instruments).

Vibrotactile stimulus generation

The stimulation medium consisted of a rectangular plate measuring 20 x 30 mm connected to a motorized shaker (Bruel and Kjar, type 4808), to which velocity values were sent as analog signals, moving the plate along the rostro-caudal axis. Stimuli were vibrations of the plate made of low-pass filtered Gaussian white noise, obtained by stringing together randomly sampled velocities. Velocities were obtained by sampling a normal distribution with 0 mean and defined by the standard deviation σ , ranging from 2.84 to 14.85. There were 50 seeds available for each velocity standard deviation. The noise was low-passed through a Butterworth filter with 150 Hz cutoff, amplified, and sent as voltage input to the shaker motor. It is possible to convert the intensities values defined by σ in mean speed values (sp, measured in mm/s) by using the following equation:

$$sp_{avg} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} dv |v| \exp\left(-\frac{v^2}{2\sigma^2}\right) = \sqrt{\frac{2}{\pi}}\sigma \quad (1)$$

In every session, there were 9 or 5 linearly-spaced values of sp.

Behavioral training procedure

Once rats arrived in the laboratory, they were placed in their new cages with ad libitum water and food, to get habituated to the new environment. After a few days of familiarization, the training procedure started.

Handling

This stage is aimed at getting the rats familiarized with human manipulation. They were handled daily for one week, until they felt comfortable being manipulated and in contact with the experimenter. Handling was accomplished by letting the rats smell and touch the experimenter's hand, and by progressively increasing contact. To reinforce them, food was given. After this first step, water was removed and administered only in the presence of the experimenter, so that the rats could form a positive association between water availability and human handling. Once this stage was completed, training could start. Handling procedures were maintained with a daily frequency, both before and after the training sessions.

Familiarization and nosepoking

Rats were placed in the setup's chamber to familiarize themselves with it, by exploring the space. Once they were habituated to this new environment, the nosepoking phase began. Rats were encouraged to approach the head window and the nosepoke using a cotton swab soaked with water. If they poked for at least 10 ms, water was given randomly at the left or right lick ports. Once they learned the association between nosepoking and water reward, the nosepoking time was progressively increased until 2000 ms.

Passive stage

When rats successfully learned the nosepoking and reached a waiting time of 2000 ms, they entered the passive stage. In this phase, vibrotactile stimuli were introduced by progressively increasing the motor shaker's gain and the duration. The ISI and stimulus duration were initially kept short, and progressively increased as the rats waited the correct amount of time in the nosepoke. The REL acoustic cue and go cue were subsequently introduced. During the passive stage, IRR stimuli were kept at the neutral, boundary intensity while REL stimuli were gradually introduced starting from the most extreme values. Rats received the reward for both correct and incorrect responses, from the licking port corresponding to the correct reward side. The rats trained in the dynamic task (KR7, KR8, MG1, and MG2) were introduced to both REL-IRR and IRR-REL trials during this stage. We gradually increased the frequency of trial types, following a one-week, three-days, one-day schedule, until both trial types were delivered within the same training session.

Active stage

When the animals reached a stable performance during the passive stage, approximately 65-70% of correct responses, they entered the active stage. During this stage, all the intensities of the IRR stimulus set were introduced, and only correct responses were rewarded.

Inclusion criteria

For the behavioral analyses, only data collected during the active stage was retained, and sessions with an average performance below 50% of correct responses were discarded.

Experimental design: Irrelevant Stimulus Paradigm

Before every trial, the blue light turned on to signal to the rat that a new trial could begin. The rat had to insert its head in the hole and keep its nose in the nosepoke, triggering the delivery of stimuli. The rat had to remain in the nosepoke for the entire trial, before a go cue signaled the rat to make a choice, by exiting the nosepoke and turning to the left or right sides of the setup, based on the rule that was reinforced during learning. Half of the rats, randomly chosen, were trained to go to the left spout if the relevant stimulus was stronger than the boundary value, the other half were trained following the reversed rule.

The task was designed in the following way: pre-stimulus delay (500 ms), irrelevant stimulus IRR (500 ms), inter-stimulus interval (ranging from 500 ms to 2000 ms), relevant stimulus REL (500 ms), post-stimulus delay (500 ms). The acoustic cue tagging the relevant stimulus lasted 300 ms and overlapped the last 300 ms of the ISI. In some instances (e.g., early withdrawals from the nosepoke in highly trained rats), the post-stimulus delay was reduced to 350 ms, to reduce the number of early withdrawals and the number of trials excluded from the analysis. This also helped the animals to be less nervous and calmer during the behavioral training.

9 stimulus intensities were employed for both the relevant and irrelevant stimuli, and could be combined in 81 possible combinations. Only one rat (KR2) had a different stimulus set (9 intensities for REL, 5 intensities for IRR), as it was initially tested as a pilot subject. Behavioral and neuronal results of this rat were consistent with the other subjects, excluding the possibility that a different stimulus set led to different results.

Standard task

The standard task represented the first implementation of the Irrelevant Stimulus Paradigm described above. In the standard task, the first stimulus of the sequence is IRR, and the second is REL.

Rats KR2, KR3, KR4, KR5, and KR6 were trained performing the standard task.

Control task

In this task variant the acoustic cue tagging the relevant stimulus was removed.

Rats KR2, KR3, KR4, and KR6 performed the control task.

Dynamic task

The dynamic task was analogous to the standard task, however, approximately half of the trials within a single training session were characterized by REL in the second temporal position, and the remaining approximately half of trials had REL in the first temporal position. The order of presentation of such trial types (IRR-REL and REL-IRR) was random.

Rats KR7, KR8, and MG2 were exclusively trained in the dynamic task. Rat MG1 was initially trained in the dynamic task, and then trained exclusively with the standard task, approximately 1.5 months before performing neuronal recordings. This ensured that the rat achieved a proficient level of performance in the standard task.

Behavioral analysis

Psychometric curve estimation

All analyses were performed in MATLAB (MathWorks, Natick MA) using custom code. To estimate psychometric functions, we computed for each stimulus vibration sp , the proportion of trials in which the rat responded by going to the side corresponding to the "strong" category. We fitted the data with a cumulative Gaussian function, including asymmetric lapse parameters:

$$p(sp \in strong) = \gamma + (1 - \gamma - \lambda) \frac{1 + \operatorname{erf}\left(\frac{sp - \mu}{\sqrt{2}\sigma}\right)}{2}. \quad (2)$$

where sp is the stimulus intensity, μ is the midpoint parameter, σ is the width parameter, γ and λ the lower and higher asymptotes of the function, respectively, corresponding to lapse rates. Parameter values were estimated by MLE using the MATLAB functions `fmincon` and `fminsearch`.

Bootstrap procedure and confidence intervals estimation

Confidence intervals were used as measure of uncertainty. We estimated confidence intervals using a non-parametric session bootstrap procedure. For each rat, training sessions were sampled with replacement, all trials from the selected sessions were retained, and the full psychometric function was recomputed for each resample. This procedure was iterated 2000 times. At each stimulus intensity, the 2.5th and 97.5th percentiles of the fitted probabilities were used as the pointwise 95% confidence interval.

Permutation test procedure for behavioral analyses

We used non-parametric permutation tests to assess the statistical significance of the behavioral distance measures. In all cases, we compared the observed statistic against a null distribution obtained from 5000 random permutations, performed independently from the whole-session bootstrap procedure. We computed two-sided p-values as the proportion of permutations in which the absolute value of the null distribution was equal to or greater than the absolute observed value, with the standard +1 correction. Statistical significance was set at $p < 0.05$.

Effect of inter-stimulus interval

To assess whether the behavioral influence of the irrelevant stimulus depended on the inter-stimulus interval (ISI), trials were divided into short (500-1000 ms), medium (1000-1500 ms), and long (1500-2000 ms) ISI categories. Psychometric functions were computed separately for each ISI category. Within each category, additional psychometric functions were obtained by conditioning trials on weak and strong IRR intensity category.

To quantify the ISI-dependent effect of irrelevant information, we focused on the short and long ISI conditions. For each of these conditions, the displacement of the psychometric function conditioned on weak IRR stimuli from the corresponding overall psychometric function was defined as

$$D_{\text{Weak,ISI}} = \langle P_{\text{Weak,ISI}}(x) - P_{\text{ISI}}(x) \rangle_x, \quad (3)$$

whereas the displacement produced by strong IRR stimuli was defined as

$$D_{\text{Strong,ISI}} = \langle P_{\text{Strong,ISI}}(x) - P_{\text{ISI}}(x) \rangle_x. \quad (4)$$

Here, $P_{\text{ISI}}(x)$ denotes the overall psychometric function for the corresponding ISI condition, and $\langle \cdot \rangle_x$ denotes the average displacement across relevant-stimulus intensities.

Because weak and strong IRR stimuli were expected to shift the psychometric function in opposite directions, their absolute displacements were combined to obtain an overall measure of IRR influence:

$$M_{\text{IRR,ISI}} = \frac{|D_{\text{Weak,ISI}}| + |D_{\text{Strong,ISI}}|}{2}, \quad (5)$$

with $\text{ISI} \in \{\text{Short}, \text{Long}\}$.

The dependence of the IRR effect on ISI was then quantified by comparing its magnitude between short and long ISIs:

$$\Delta M_{\text{IRR}} = M_{\text{IRR,Short}} - M_{\text{IRR,Long}}. \quad (6)$$

Positive values of ΔM_{IRR} indicate a stronger behavioral influence of the irrelevant stimulus at short ISIs, whereas negative values indicate a stronger influence at long ISIs.

The 95% confidence interval of ΔM_{IRR} was defined by the 2.5th and 97.5th percentiles of its bootstrap distribution.

We assessed statistical significance of the short-versus-long ISI contrast using a permutation test. The null distribution was generated by randomly shuffling ISI durations across trials within each session. For each permutation, trials were reassigned to the Short, Medium, and Long ISI categories according to their shuffled ISI values, and the corresponding psycho-

metric functions, IRR-effect magnitudes, and Short-minus-Long contrast were recomputed.

Effect induced by the irrelevant stimulus

To quantify the bias induced by the irrelevant stimulus, we computed psychometric functions separately for each IRR intensity and we compared them against the overall psychometric curve. The displacement induced by each IRR condition was defined as the mean point-wise difference between the corresponding IRR-conditioned psychometric function and the overall psychometric function across REL stimulus intensities:

$$D_{\text{IRR}_j} = \langle P_{\text{IRR}_j}(x) - P_{\text{Overall}}(x) \rangle_x, \quad (7)$$

where IRR_j denotes the considered IRR intensity and $\langle \cdot \rangle_x$ denotes the average across relevant stimulus intensities. Positive values indicate an upward displacement of the IRR-conditioned psychometric function relative to the overall curve, whereas negative values indicate a downward displacement.

The 95% confidence intervals were defined by the 2.5th and 97.5th percentiles of the corresponding bootstrap distributions.

We assessed statistical significance using a permutation test. To generate the null distribution, we randomly shuffled IRR stimulus labels across trials within each session, thereby disrupting the association between IRR intensity and behavioral choice while preserving the session structure. The psychometric functions and corresponding distance measures were recomputed after each permutation.

Effect of cue removal

To assess whether removal of the REL acoustic cue affected behavioral performance, we compared the parameters of the overall psychometric function between the standard task and the control task. The experimental sessions used for this analysis included the last N sessions corresponding to the expert stage, and the first N experimental sessions performed during the control task. More specifically, we considered the parameter σ , describing the spread of the psychometric function, and the point of subjective equality (PSE), corresponding to the parameter μ .

For each animal, the overall psychometric function was computed from the total N sessions included in the operational expert stage (see Generalized Linear Model estimation) and from the corresponding N sessions acquired during training in the control task. The effect of cue removal on the two psychometric parameters was quantified as

$$\Delta\sigma = \sigma_{\text{Control}} - \sigma_{\text{Standard}}, \quad (8)$$

and

$$\Delta\mu = \mu_{\text{Control}} - \mu_{\text{Standard}}. \quad (9)$$

$\Delta\sigma$ quantified the change in the spread of the psychometric function between standard and control conditions, whereas $\Delta\mu$ quantified the corresponding shift in the PSE.

The 95% confidence intervals were defined by the 2.5th and 97.5th percentiles of the corresponding bootstrap distributions.

We assessed statistical significance using a permutation test. To generate the null dis-

tribution, we pooled all trials from the matched standard and control datasets. On every permutation, we randomly reassigned trials to standard or control labels. Psychometric functions were refitted for the resulting pseudo-standard and pseudo-control datasets and the corresponding $\Delta\sigma$ and $\Delta\mu$ values were recomputed after each permutation.

% For each rat, all trials from the matched Standard and Control% datasets are pooled. On every permutation, trials are randomly% reassigned to:% pseudo-Standard% pseudo-Control

Generalized Linear Model estimation

GLM analyses were implemented using custom-made MATLAB code, fitting free parameters with the following equation:

$$Y = \gamma + (1 - \gamma - \lambda) \frac{1}{1 + \exp[-(W_1x + W_2z + b)]}. \quad (10)$$

Where Y is the predicted rat's choice, γ and λ are, respectively, the lower and upper lapse rates, W_1 is the weight for the REL stimulus, W_2 is the weight for the IRR stimulus, x is the REL stimulus intensity, and z is the IRR stimulus intensity. Parameters values were estimated by MLE using the MATLAB functions `fmincon` and `fminsearch`.

The purpose of the GLM analysis was to quantify the relative influence of the REL and IRR stimuli on perceptual judgments and to characterize how this influence changed over time. Because the analysis was descriptive rather than aimed at predicting choices in unseen data, out-of-sample validation was not performed.

Proficient and expert stages To describe the evolution of the IRR-induced bias across training, we defined two operational stages: a proficient stage and an expert stage. For each

animal, we divided the complete behavioral dataset of N sessions into 10 successive bins, and we fitted a GLM independently to each bin.

The relative contribution of the irrelevant stimulus was quantified as

$$R_{\text{IRR}} = \frac{|W_{\text{IRR}}|}{|W_{\text{REL}}|}. \quad (11)$$

The expert stage was operationally defined as beginning when $R_{\text{IRR}} < 0.1$ for at least two consecutive temporal bins. The criterion identified periods in which the magnitude of the IRR weight was less than 10% of the magnitude of the REL weight. Bins preceding this criterion were considered part of the proficient stage.

Within-session performance variability across time

To assess whether behavioral performance changed over the course of individual sessions, we divided each session into two halves according to trial order. For sessions containing an odd number of trials, we assigned the additional trial to the second half. We calculated performance separately for the first and second halves as the proportion of correct choices.

For each session, the within-session change in performance was quantified as

$$\Delta P_s = P_{\text{Second half},s} - P_{\text{First half},s}. \quad (12)$$

Positive values of ΔP_s indicated higher performance during the second half of the session, whereas negative values indicated a decline in performance.

For each animal, the observed effect was defined as the mean session-level difference

$$\Delta P_{\text{obs}} = \frac{1}{N} \sum_{s=1}^N \Delta P_s, \quad (13)$$

where N denotes the total number of sessions included for that animal.

We estimated uncertainty in ΔP_{obs} using 2000 session-level bootstrap resamples. Sessions were sampled with replacement, and the mean ΔP was recomputed for each bootstrap sample. The 95% confidence intervals were defined by the 2.5th and 97.5th percentiles of the resulting bootstrap distribution.

Surgery protocol

For implanting electrodes into left M1/M2 and left vS1, both contralateral to the right side vibrissal stimulation (M1/M2 only rats $n=2$, M1/M2 and vS1 rats $n=2$), animals were anesthetized with Isoflurane through a plastic snout mask, administered via a veterinary apparatus (V-1 Tabletop, VetEquip Inc.). The concentration of the agent was variable, based on the depth of required anesthesia and subject differences. To induce anesthesia, the rat was placed in a box and gradually anesthetized with Isoflurane until reaching an agent level of 2%-2.5% mixed with oxygen (O₂ 98%), as calculated according to the MAC value (Minimum Alveolar Concentration: 1.35% $\pm$ 0.10%). Once immobile, a mask was positioned on the rat's snout to provide a constant administration of Isoflurane with concentration of 2.5%, and a flow rate of 1 L/min. This established a state characterized by analgesia, muscle relaxation, and loss of reflexes (absence of a reaction to tail, hind paw pinches, and corneal air puff). This state was maintained across the entire surgical procedure (opening of skin and bone). Blood oxygen saturation was at least 98% and cardiac frequency was 350 $\pm$ 25 BPM.

Once the craniotomy/craniotomies was/were completed, provided that the rat continued

to show no reflexes or signs of discomfort, Isoflurane concentration was decreased to 2% in 1 L/min of oxygen to prevent gas accumulation and allow a more rapid postoperative recovery. This same flow rate was maintained during the electrophysiological recordings used to verify electrode positions. At the end of the surgical procedure the Isoflurane vaporizer was turned off and the rat was allowed to recover while receiving Isoflurane 1 L/min of oxygen. During the entire procedure and until the rat completely recovered from the anesthesia, body temperature, blood oxygenation, breathing frequency, and heart rate were continuously monitored. The absence of reflexes was regularly tested. In case of bleeding, fluids were reintroduced by subcutaneous injection to ensure proper hydration. Target cortical regions were accessed by craniotomies, using the Paxinos Rat Brain Atlas' stereotaxic coordinates [41], and standard stereotaxic techniques. The dura mater was removed over the entire craniotomy with a small needle.

The electrode arrays were inserted by slowly advancing a Narishige automatic micro-manipulator. After inserting the array, the remaining exposed cortex was covered with biocompatible silicone (KwikSil; World Precision Instruments). In all rats, at least 4 small screws were fixed in the skull as support for dental cement. All screws served as ground and reference electrodes.

Before the conclusion of the surgical procedure, animals were given antibiotic (Baytril; 5 mg/kg; i.p) and analgesic (Rimadyl; 2.5 mg/kg; i.m.). In addition, both the antibiotic and the analgesic were delivered through a water bottle for 24 hours after the completion of the surgery. During this recovery time, rats had ad libitum access to water and food.

For the M1/M2-vS1 implanted rats, we decided not to proceed with the mapping of vS1 using a single probe, for the following reasons: i) extended surgical times pose a health risk to the animal, and should be kept as short as possible, ii) the microdrive array was designed in such a way to cover the greatest cortical area possible, to maximize the probability of

placing individual electrodes in individual barrels, iii) the task requires the rat to be freely moving and not head fixed, making it not possible to determine with certainty whether the stimulated whiskers were those always in contact with the stimulation plate.

Microdrives configuration

Rat KR2 was implanted with a microdrive array constituted by 16 electrodes, arranged in 2 rows of 8 electrodes, with the long side placed along the medio-lateral axis, targeting M1/M2. The stereotaxic coordinates of the craniotomy were 1 mm to 5.5 mm AP from bregma, and 0.5 mm to 4.25 mm ML from bregma.

Rat KR3 was implanted with a microdrive array constituted by 32 electrodes, arranged in 4 rows of 8 electrodes each, with the long side placed along the antero-posterior axis, targeting M1/M2. The stereotaxic coordinates of the craniotomy were 0.5 mm to 3.5 mm AP from bregma, and 0.5 mm to 3.5 mm ML from bregma.

Rats KR8 and MG1 were both implanted with two custom-made microdrive arrays constituted by 12 electrodes each, arranged in 3 rows of 4 electrodes. Both arrays had movable electrodes, meaning it was possible to move them deeper or shallower during the electrophysiological recordings period, to improve recording quality. The array targeting M1/M2 was placed with the long side aligned to the antero-posterior axis, whereas the vS1 array with the long side aligned to the medio-lateral axis. The stereotaxic coordinates for the craniotomy in M1/M2 were 0.5 mm to 2.5 mm AP from bregma, and 0.5 mm to 2 mm ML from bregma. For the vS1 craniotomy, the stereotaxic coordinates we used were -1 mm to -3.5 mm AP from bregma, and 4 mm to 7 mm ML from bregma.

Neuronal analysis

Neuronal recordings preprocessing

Neuronal recordings were spike-sorted manually using Wave Clus [42]. Single units were selected according to spike shape consistency and inter-spike interval. To be included in the analyses, units had to meet two criteria: i) at least 2000 total spikes, and ii) firing in at least two trials per REL-IRR intensities combinations.

Firing rate

Firing rates were computed by binning spikes in non-overlapping 25 ms time-bins. A Gaussian smoothing over a 500 ms window with a standard deviation of 100 ms was then applied. Since trials had variable ISI duration, we only included the first 200 ms and last 300 ms of the ISI, overlapping with the delivery of the acoustic cue. Therefore, trials included pre-stimulus delay, IRR/REL, first 200 ms of ISI, last 300 ms of ISI, REL/IRR, post-stimulus delay.

Decoding

Pseudopopulation construction The decoding analysis was performed using a pseudopopulation approach, since units were recorded across multiple sessions. We ran separate decoders for stimuli and choice. When decoding REL and IRR, for each recording session, trials were categorized into four conditions based on the factorial combinations of REL and IRR intensity categories. When decoding choice, trials were instead categorized according to the factorial combinations of REL intensity category and choice category. After labeling trials, we randomly split them into a training set and a test set (80% and 20%, for each

condition). To account for unequal trial counts, and to diminish potential noise correlations, we used a bootstrapping approach, resampling with replacement 800 trials from the training set, and 200 trials from the test set. This procedure was performed independently for each unit and each session. Each resulting pseudosession of N units was concatenated to build a pseudopopulation where each condition was perfectly balanced. We started by including all trials (correct and incorrect trials), before running separate decoders for miss and hit trials.

Classifier Before training the Linear Discriminant Analysis classifier (LDA), we projected the neuronal activity of the training set and test set onto their stimulus (or choice) first principal component, separately computed using dPCA [43]. LDA identifies a weight vector that maximizes the separation between condition means, while minimizing within-condition variance: we trained the classifier on the projected neural activity belonging to the training set, and then tested the model on the unseen test activity. Next, we obtained raw accuracy scores, and we assessed the classifier’s performance by computing the decoder’s score

$$s(x_i) = w \cdot x_i + b \quad (14)$$

where w is the weight vector, x_i is the vector of neural activity for trial i , and b is the bias term. The final classification for each test trial was determined by the sign of the score $s(x_i)$.

A separate LDA was trained and tested on projected neuronal activity with shuffled trial labels, to obtain a control distribution and shuffled accuracy scores.

This whole procedure was iterated 100 times, each time ensuring a different train/test split.

Statistical significance of decoding results To test whether accuracy scores were statistically significant (i.e., statistically different from shuffled accuracy), we z-scored the real accuracies against the shuffled accuracies, similarly to [44]:

$$\text{Distance}(t) = \frac{\text{Accuracy}_{\text{Real}}(t) - \text{Accuracy}_{\text{Shuffle}}(t)}{\sigma_{\text{Shuffle}}(t)} \quad (15)$$

where $\text{Distance}(t)$ is the z-score accuracy score for time t , $\text{Accuracy}_{\text{Real}}$ is the average accuracy over all iterations of cross-validation for correctly ordered labels, $\text{Accuracy}_{\text{Shuffle}}$ is the average accuracy over iterations for shuffled labels, and σ_{Shuffle} is the standard deviation over all iterations for shuffled labels.

Relating classifier’s score and rat choices For each iteration of the decoder, we stored the rat’s choice (i.e., answering "strong" or "weak") to REL and IRR for each held-out trial, along with the classifier’s score, computed as described above. Next, we computed $p(\text{strong})$ for REL and IRR as proportion of "strong" choices, obtaining one probability value for each of the 100 iterations of the decoder. We averaged the single-trial classifier’s scores, for each iteration. Next, we chose the time-bin with the highest accuracy score for REL and IRR intensity category prediction, and plotted the averaged classifier’s score for strong/weak REL and IRR, against rat’s choices ($p(\text{strong})$ to REL and IRR).

Classifier for correct and incorrect trials To visualize differences in stimulus processing related to correct or incorrect trials (i.e., trials in which the rat’s choice was rewarded or not), we performed a separate decoding analysis. The analysis differed only in the trial pool used to train and test the classifier. When decoding hit trials, only hit trials were pooled together, and then split into training/test. This ensured that the resulting accuracies reflected exclusively neuronal activity underlying optimal behavioral performance. When

decoding miss trials, the training set was built by randomly sampling 80% of the hit trials. The test set was built using the entire pool of miss trials. Given that error trials constitute approximately 25% of the whole trial set, we decided to use all error trials for the test set, in order to avoid underfitting.

Cross-correlation We cross-correlated the vS1 and M1/M2 peak-normalized decoding z-scores using the MATLAB `xcorr` function to identify the optimal lag and the maximum correlation coefficient (R_{obs}). Statistical significance was assessed using a permutation test ($n = 1000$ iterations). The null distribution was constructed by circularly shifting the vS1 z-scores while keeping the M1/M2 sequence fixed. This preserved the signal's intrinsic autocorrelation structure while destroying the temporal alignment between areas. The p-value was defined as the proportion of permutations where $R_{null} \geq R_{obs}$.

Local Field Potential acquisition and analysis

We performed inter-areal synchronization analysis and Granger causality analysis using the Fieldtrip toolbox [45] and custom-built MATLAB codes.

LFP acquisition and preprocessing We obtained LFPs offline using neuronal data yielded by the same electrodes used for the decoding analysis, to ensure consistency and sufficient quality of recordings. Raw signals were downsampled to 1000 Hz, and bandpass filtered between 1 and 300 Hz. Since recordings were carried out within a Faraday cage (see Experimental apparatus), we abolished significant power-line interference, and therefore, chose not to apply a notch filter, preserving the integrity of the signal within the 50-60 Hz frequency range. We divided continuously recorded data in stimulus epochs corresponding to REL and IRR delivery, storing the timestamps of nosepoke entry and exit. Since the

ISI was variable, we could not segment the data in trials of identical duration, hence the stimulus epoch segmentation strategy. By storing the timestamps of nosepoke entry/exit, we were able to include or exclude for further analyses activity occurring before or after the stimulus delivery.

Coherograms To visualize the synchronized activity of vS1 and M1/M2, we computed time-frequency resolved coherograms for REL and IRR stimulus epochs, including 300 ms before stimulus onset, and 300 ms after stimulus offset. We used a Hanning taper to estimate power and cross-spectral density for frequencies ranging from 1 to 100 Hz, in 1 Hz steps. For all frequencies, we used a fixed 300 ms sliding window, which shifted in 10 ms increments across the considered epoch. We employed the 300 ms sliding window as it provided an optimal compromise between temporal and frequency resolution across the entire frequency range. It allowed us to capture multiple cycles of low-frequency oscillations (θ and β), while maintaining sufficient temporal precision to track the dynamics of high-frequency ranges (γ), minimizing spectral leakage. We first computed the coherence for every possible pair-wise combination of electrodes between the two areas, and then averaged these individual channel-pair coherograms to produce a grand-averaged coherogram, providing a robust measure of the global synchronization between vS1 and M1/M2.

Frequency-resolved Granger causality To identify the direction of information flow between vS1 and M1/M2, we ran frequency-resolved Granger causality. We included 500 ms from stimulus onset until stimulus offset (to include only stimulus epoch), for both REL and IRR. We performed spectral decomposition using the Multitaper Fast Fourier Transform (mtmfft) with a Hanning taper, including frequencies ranging from 1 Hz to 100 Hz, in 1 Hz steps. To achieve a consistent frequency resolution of 0.25 Hz, we zero-padded all data segments to a 4 s duration prior to the FFT. We then estimated Granger causality for both

directions ($vS1 \rightarrow M1/M2$ and $M1/M2 \rightarrow vS1$). To quantify the direction of information flow we used the Granger Asymmetry Index (GAI), which is the normalized difference between Granger Causality scores (GC) for each direction:

$$GAI(f) = \frac{GC_{vS1 \rightarrow M1/M2}(f) - GC_{M1/M2 \rightarrow vS1}(f)}{GC_{vS1 \rightarrow M1/M2}(f) + GC_{M1/M2 \rightarrow vS1}(f)} \quad (16)$$

Time-frequency resolved Granger causality We computed time-frequency resolved Granger causality using a sliding window with a duration of 300 ms, shifted in 10 ms increments across the epoch (from -300 ms to +800 ms relative to stimulus onset). Data segments were extracted with an additional 150 ms of padding on either side of each window, to ensure that all time-frequency estimates within the analyzed period were based on complete and non-truncated activity windows. We performed the spectral decomposition for each temporal window, by estimating Fourier coefficients using the Multitaper Fast Fourier Transform (mtmfft) with a Hanning taper, including frequencies between 1 Hz and 100 Hz, in 1 Hz steps. Each 300 ms segment was zero padded to 4 seconds prior to the FFT, to ensure a consistent frequency resolution of 0.25 Hz. We then computed the GAI:

$$GAI(t, f) = \frac{GC_{vS1 \rightarrow M1/M2}(t, f) - GC_{M1/M2 \rightarrow vS1}(t, f)}{GC_{vS1 \rightarrow M1/M2}(t, f) + GC_{M1/M2 \rightarrow vS1}(t, f)} \quad (17)$$

Cluster-based permutation tests We used a non-parametric cluster-based permutation test to assess statistical significance when computing coherograms. The observed coherence values were compared against a null distribution generated by shuffling trial identities. Initial clusters were formed by grouping adjacent time-frequency bins exceeding a t-threshold corresponding to a $p < 0.001$ significance level. The significance of these clusters was then evaluated by comparing their summated t-values against a distribution created from 5000 random permutations. The final significance for a cluster was set at $p < 0.05$. This method

controls for multiple comparisons.

For the frequency-resolved Granger causality, we used a similar approach. To assess whether the directionality was statistically significant against a null hypothesis of no asymmetry, we used a non-parametric cluster-based permutation test across sessions, using a dependent-samples t-statistic. For each session, the observed GAI was compared to a null distribution of zeros. Initial clusters were formed by grouping adjacent frequency bins with a threshold of $p < 0.05$. We then determined the significance of these clusters by comparing their summated t-statistic against a null distribution generated from 5000 random permutations, to identify frequency bands in which information flow was significantly directional.

Human behavioral patterns reflect diversity in behavioral strategies during temporal attention

Abstract

To make adaptive perceptual judgments, the nervous system must selectively process behaviorally relevant sensory information while filtering out competing distractions. Although attentional control has been extensively studied in the visual domain, particularly in the context of spatial selection, considerably less is known about how attention operates in the tactile modality and across time rather than space. Here, we adapted for human participants a paradigm originally developed to investigate temporal tactile attention in rodents, enabling a direct cross-species comparison of attentional behavior and underlying behavioral strategies.

Participants were instructed to categorize the intensity of a task-relevant tactile stimulus delivered to the fingertip while ignoring an irrelevant tactile stimulus. A visual cue indicated which of two sequentially presented stimuli was relevant on each trial. In addition, participants completed self-report questionnaires assessing autistic traits and aberrant salience (the tendency to assign significance to otherwise neutral stimuli or events).

Across subjects, participants performed the task with high accuracy. However, clustering analyses based on behavioral features revealed diverse response profiles. One cluster did not exhibit biases induced by the irrelevant stimulus, consistent with previous findings in expert rats. In contrast, a third cluster displayed a repulsive effect of the irrelevant stimulus and showed lower overall performance. These behavioral phenotypes were also reflected, to some extent, in differences in learning trajectories across training sessions.

We further explored participants' metacognitive awareness through a post-experiment questionnaire assessing subjective evaluations of task difficulty and performance. Although exploratory, the results suggest a relationship between metacognitive reports, behavioral strategies, and objective task performance. In contrast, neither autistic traits nor aberrant salience scores were associated with performance measures or behavioral phenotypes.

Together, these findings extend our previous rodent results to humans and support the view that attentional control is implemented through multiple, individualistic behavioral strategies rather than a single mechanism pertaining to all. The convergence of results across species suggests that task structure interacts with individual predispositions to shape various modes of attentional control.

Introduction

Goal-directed behavior depends on a set of higher-order cognitive command processes collectively referred to as executive functions. Among these, attentional control enables organisms to prioritize behaviorally relevant information while suppressing competing or distracting inputs. By regulating access to limited processing resources, attention plays a fundamental role in perception, decision-making, and adaptive behavior.

Over the past decades, attentional control has been extensively investigated from both cognitive and neuroscientific perspectives. Research on attention has been largely dominated by the visual modality and by paradigms examining the allocation of attention across space [46–48]. Classical theories focused on the locus and mechanisms of attentional selection, proposing models of early selection, late selection, and attentional attenuation [16, 17, 49]. Subsequent neurophysiological and neuroimaging studies identified the role of top-down modulation from the frontoparietal attentional networks to early and extrastriate visual cortices [20, 50, 51].

Considering attention only in the context of vision and spatial selection is a narrow perspective. Organisms must often prioritize information that unfolds over time and across multiple sensory modalities. Touch represents a particularly relevant domain. Beyond its role in object exploration and manipulation [52], tactile processing is essential for social interactions [53], and becomes especially important under conditions of sensory deprivation, such as blindness, where it can support complex perceptual functions including Braille reading [54, 55]. Nevertheless, the mechanisms governing attentional control in the tactile modality remain considerably less explored than those described for vision.

In the preceding study, we investigated temporal attentional control in rats using an experimental paradigm in which animals were required to selectively report a relevant tactile

stimulus while ignoring an irrelevant one. Rather than uncovering a single attentional mechanism, we found evidence for multiple behavioral and neuronal strategies that emerged under the same task demands. Different animals exhibited distinct patterns of susceptibility to distracting information, suggesting that attentional control may have been implemented through alternative solutions shaped by both task structure and individual predispositions.

The present study extends this work to human participants. By adapting the rodent paradigm to humans, we sought to establish whether similar behavioral strategies would emerge across species and whether individual differences would contribute substantially to performance. More specifically, we asked whether variability among participants would be primarily quantitative, reflecting differences in overall performance, or whether qualitatively distinct modes of interaction with irrelevant information would also be observed.

Participants performed a task in which two sequential tactile stimuli were presented and a visual cue indicated which stimulus was relevant for the perceptual judgment. Participants were asked to categorize the intensity (weak or strong) of the relevant stimulus. Four blocks of trials were presented, distributed over two experimental sessions. This design allowed us to characterize not only steady-state performance but also learning trajectories over time, facilitating direct comparisons with the behavioral patterns previously observed in rats.

In addition to the behavioral task, participants completed self-report measures assessing autistic traits and aberrant salience. Atypical attentional and executive processes have frequently been reported in autism spectrum disorder, while aberrant salience, defined as the inappropriate assignment of significance to otherwise neutral stimuli, has been linked to altered perceptual and cognitive experiences [56, 57]. However, whether variation in these traits contributes to attentional performance in non-clinical populations remains uncertain [58]. We therefore explored whether inter-individual differences in these dimensions were associated with task performance or with specific behavioral strategies.

Finally, we examined participants' metacognitive awareness through a post-experimental questionnaire assessing subjective evaluations of task difficulty and performance. Although exploratory, this analysis allowed us to investigate whether differences in objective behavior were accompanied by corresponding differences in self-evaluation and introspective accuracy.

Overall, our aim was not only to characterize human performance in a temporal tactile attention task, but also to determine whether distinct attentional strategies emerge across individuals, as previously observed in rats. Demonstrating such convergence would provide evidence that variability in attentional control might reflect meaningful alternative behavioral solutions rather than mere noise around a common mechanism.

Results

Temporal attention in the tactile modality of humans

To assess whether the behavioral patterns observed in rats generalize across species, we tested 30 human participants in an adapted version of the dynamic task. The task required participants to selectively attend to a relevant vibrotactile stimulus (REL) while ignoring an irrelevant stimulus (IRR), allowing us to investigate how attention was flexibly allocated across time in the human tactile modality.

The human task preserved the core structure of the rodent paradigm while incorporating several modifications to accommodate human testing (Figure 32). The auditory cue signaling the relevant stimulus in the rodent task was replaced by a visual cue of identical duration, consisting of a green light displayed on a computer screen. In addition, the inter-stimulus interval was fixed at 1000 ms. Participants received vibrotactile stimuli on the tip of the right index finger and reported their perceptual judgment by pressing one of two response buttons. Visual feedback was provided after every trial.

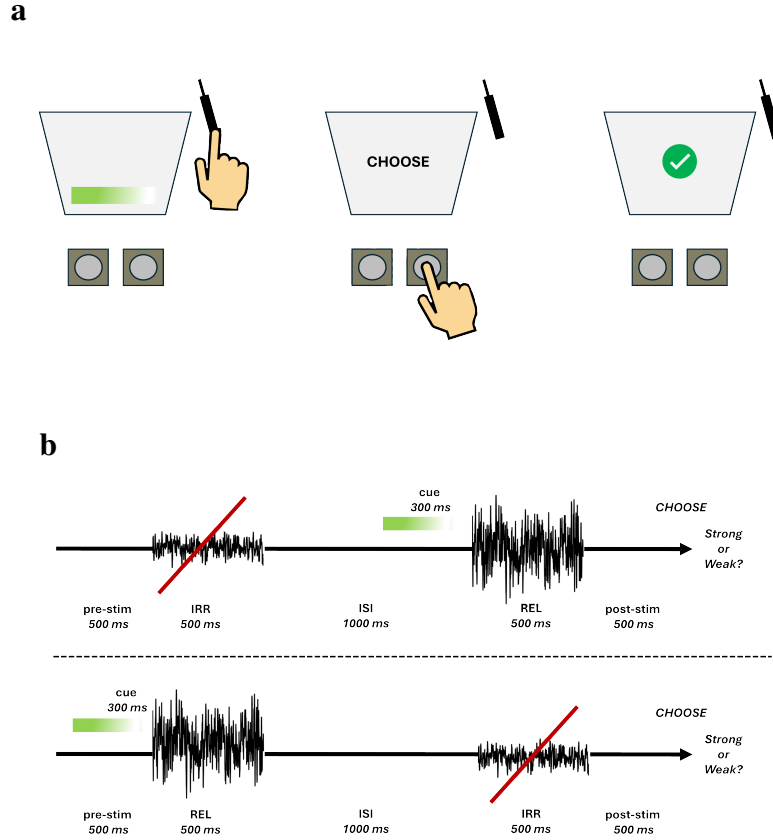


Figure 32: Setup configuration and human dynamic task. **a.** Experimental setup used for human testing. Vibrotactile stimuli were delivered to the participants' right index fingertip while a visual cue identified the relevant stimulus. Participants categorized stimulus intensity by pressing one of two response buttons and received visual feedback after each trial. **b.** Structure of the task: the auditory relevance cue used in rats was replaced by a visual cue, and the inter-stimulus interval was fixed at 1000 ms.

To reduce the duration of testing while maintaining adequate stimulus sampling, we used a reduced stimulus set consisting of seven relevant stimulus intensities and five irrelevant stimulus intensities. As in the rat experiments, each vibration (see Methods) consisted of a sequence of velocity values sampled from a Gaussian distribution and was defined by its

nominal speed (mm/s).

Before testing, participants completed a demographic questionnaire reporting age, gender, nationality, neuropsychiatric diagnoses, and current pharmacological treatment. Following data collection, participants reporting neuropsychiatric diagnoses or psychotropic medication use would have been excluded from the study; however, no exclusions were necessary. Participants' confidentiality was maintained through pseudonymization with the use of alphanumeric codes.

Participants also completed three self-report questionnaires: the AQ-10, the SPQ-10, and the Aberrant Salience Inventory (ASI). These measures were collected to evaluate whether variability in autistic traits and aberrant salience was associated with performance in a task requiring flexible allocation of attention across time. Previous findings linked ASD (i.e., autism spectrum disorder) traits and aberrant salience to attentional deficits [56], [59].

Testing was conducted across two experimental sessions separated by no more than one week. Each session included two testing blocks, for a total of four blocks per participant. Before the first session, participants completed a brief warm-up phase designed to familiarize them with the task and experimental setup. During this phase, they also performed 20 reference memory trials, as in [60], to ensure they understood that stimulus categorization required comparison with a reference intensity. Participants who did not demonstrate adequate understanding of the task were allowed to repeat the warm-up phase once before continuing.

The overall structure of the study is illustrated in Figure 33.

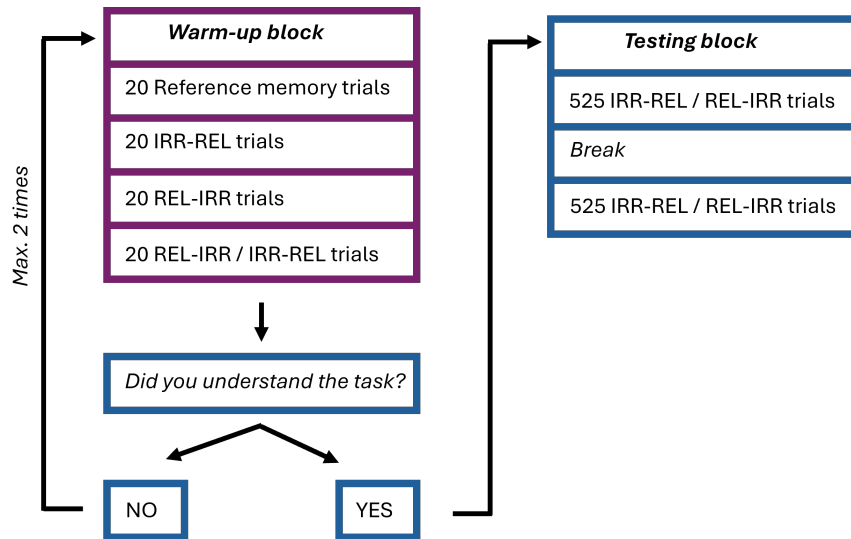


Figure 33: Structure of the human study. Participants first completed a warm-up phase that could be repeated once if necessary. After confirming task comprehension, they completed the testing blocks across two experimental sessions held on separate days. At the end of the second session, participants completed a brief post-task questionnaire designed to assess subjective perceptions of task difficulty and confidence. The questionnaire contained nine items with three response options each and evaluated perceived task difficulty, confidence in performance, the influence of stimulus position, and participants' memory of the reference stimulus.

Effect of irrelevant information on humans' perceptual judgments

Participants showed high levels of performance in both trial types and were substantially less influenced by irrelevant information than were rats. Psychometric functions revealed accurate perceptual judgments during both REL-IRR and IRR-REL trials, indicating that participants successfully used the relevant stimulus to guide their decisions independently of order (Figure 34).

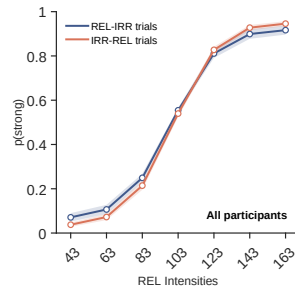


Figure 34: Across-subject overall performance during the dynamic task. Psychometric functions showing average performance across participants for REL-IRR and IRR-REL trials. Dots represent the across-subject mean probabilities obtained from individually fitted psychometric functions at each tested REL intensity. Lines connect consecutive fitted probabilities for visualization, and shaded areas represent the SEM across individually fitted probabilities.

Despite this overall high performance, conditioning behavioral responses on irrelevant-stimulus intensity revealed subtle effects of irrelevant information on perceptual judgments (Figure 35a). During REL-IRR trials, psychometric functions showed only minor differences across irrelevant-stimulus intensities, suggesting that participants largely ignored the irrelevant stimulus when the relevant stimulus appeared first. In contrast, during IRR-REL trials, the highest irrelevant-stimulus intensities produced a slight separation of the psychometric functions, whereby stronger relevant stimuli were more likely to be judged as weaker than their physical intensity would predict. This effect was largely absent during REL-IRR trials, consistent with a repulsive influence of the irrelevant stimulus.

To quantify these effects, we examined changes in the Point of Subjective Equality (PSE) as a function of irrelevant-stimulus intensity. PSE analysis revealed a positive relationship between irrelevant-stimulus intensity and PSE during IRR-REL trials (Figure 35b). As irrelevant stimuli became stronger, participants required increasingly stronger relevant stimuli to perceive them as equivalent to the reference stimulus.

This shift in PSE was consistent with the repulsive bias observed in the psychometric functions. Strong irrelevant stimuli altered perceptual judgments even when participants were instructed to ignore them, producing systematic changes in the perceived relationship between the relevant stimulus and the memorized reference value.

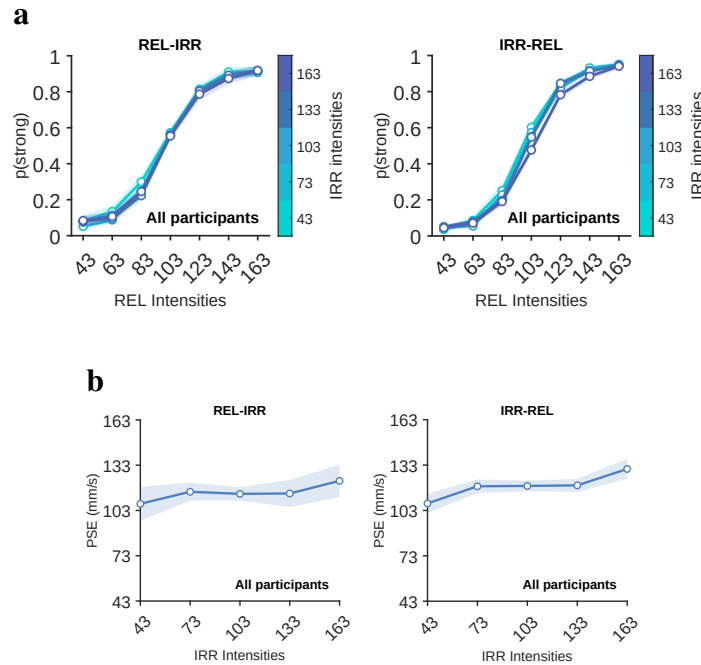


Figure 35: Across-subject irrelevant-stimulus bias during the dynamic task. **a.** Psychometric functions conditioned on irrelevant-stimulus intensity during REL-IRR trials (left) and IRR-REL trials (right). **b.** Point of Subjective Equality as a function of irrelevant-stimulus intensity during REL-IRR trials (left) and IRR-REL trials (right). Shaded regions represent SEM of fitted PSE estimates.

Overall, the effect of irrelevant information was modest across subjects. Participants generally performed the task accurately and successfully discounted distracting sensory input. Nevertheless, the presence of systematic PSE shifts and small but detectable psychometric biases indicates that irrelevant information continued to influence perceptual decisions under specific conditions, particularly when the irrelevant stimulus preceded the relevant

one.

Changes in performance across the sessions

To assess whether participants progressively increased their reliance on the relevant stimulus (REL) and reduced the influence of the irrelevant stimulus (IRR), we examined the evolution of stimulus weights across the four testing blocks using generalized linear models. Across both trial sequences, the weight assigned to the relevant stimulus increased modestly over the course of training (Figure 36). This was observed in both REL-IRR and IRR-REL trials, indicating progressively greater reliance on task-relevant sensory information when making perceptual judgments.

In contrast, the weight assigned to the irrelevant stimulus remained close to zero across all four blocks (Figure 36). This pattern was consistent across trial types and indicates that participants successfully suppressed the influence of irrelevant sensory information throughout the experiment, in contrast to the rats.

Taken together, these findings suggest that participants mastered the task rapidly and relied primarily on the REL stimulus from the earliest stages of testing. The increased weight given to the REL stimulus may correspond to a very small increase in acuity – that is, perceptual plasticity.



Figure 36: Across-subject learning progression during the dynamic task. Generalized linear model weights for relevant (REL) and irrelevant (IRR) stimuli across the four testing blocks, shown separately for REL-IRR trials (left) and IRR-REL trials (right). Shaded regions represent SEM of fitted GLM weights.

Effect of previous-trial choices on perceptual judgments

Under some conditions, past choices exert an attractive influence on current decisions [61]. If such a bias occurred in our experiments, participants would have been more likely to select “strong” on the current trial if they selected “strong” on the preceding trial, and vice versa. The data indicate that previous-trial choices exerted only a weak influence on current perceptual judgments when averaged across subjects.

Psychometric functions conditioned on the choice made in trial N-1 showed only a negligible attractive bias in both REL-IRR and IRR-REL trials (Figure 37). These results indicate that, at the population level, current perceptual judgments were driven primarily by information presented within the ongoing trial rather than by the choice made on the preceding trial.

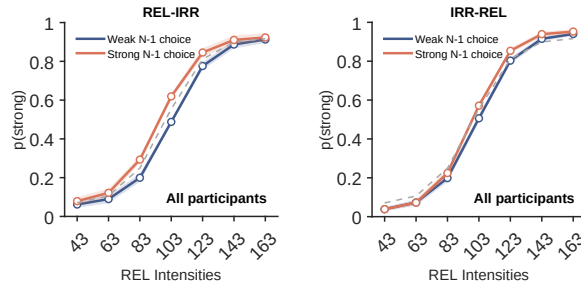


Figure 37: Effect of previous-trial choice on current perceptual judgments. Psychometric functions conditioned on the choice made in trial N-1, shown separately for REL-IRR trials (left) and IRR-REL trials (right).

Behavioral clustering reveals diversity in attentional strategies

Behavioral performance differs across clusters

When averaged across subjects, we identified effects consistent with a slight repulsive bias of IRR and a slight N-1 choice attractive bias. Because these tendencies both reflected the incorporation of irrelevant information, we posited that they may have been somehow connected in individuals. We used these two behavioral metrics as clustering features, in order to divide the participants into three subgroups. We performed a *k-medoids* clustering analysis, using $k=3$, to identify three clusters based on the GLM weights of trial N-1 choice and trial N IRR stimulus. The choice of $k = 3$ was supported by silhouette analysis, and stability was assessed using ARI (see Supplementary Figure S12)

Specifically, for each subject, we computed two separate GLMs, calculating the weight of trial N-1 choice on trial N choice, and the weight of trial N IRR stimulus on trial N choice. These weights were then used to cluster participants in three groups. In Figure 38, Cluster 3 exhibited a strong IRR repulsive bias, with a repulsive trial N-1 choice effect. Cluster 2

showed a greater trial N-1 choice attractive bias, with a slight, overall repulsive effect of IRR. Cluster 1 exhibited less trial N-1 choice bias, and mixed attractive/repulsive IRR bias. Clustering revealed substantial differences in task performance and in the influence of irrelevant information on perceptual judgments. Although all groups successfully performed the task, the three clusters exhibited diverse behavioral patterns that were not apparent at the population level.

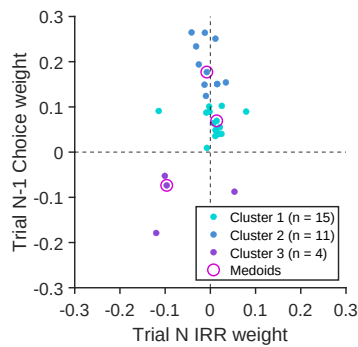


Figure 38: *k-medoids* clustering based on trial N-1 choice weight and trial N IRR weight. *k-medoids* clustering identified three distinct clusters of participants, with specific directions of trial N-1 choice and IRR bias.

Clusters differ in performance and susceptibility to irrelevant information

Overall task performance differed across clusters (Figure 39). Cluster 1 showed comparable performance in REL-IRR and IRR-REL trials, whereas Cluster 2 performed slightly better when the relevant stimulus appeared first (REL-IRR trials). Cluster 3 exhibited the lowest overall performance and, similar to the pattern observed in rats, showed reduced accuracy during REL-IRR trials relative to IRR-REL trials. These findings indicate that participants did not differ solely in the magnitude of their performance, but also in the trial types that posed the greatest difficulty.

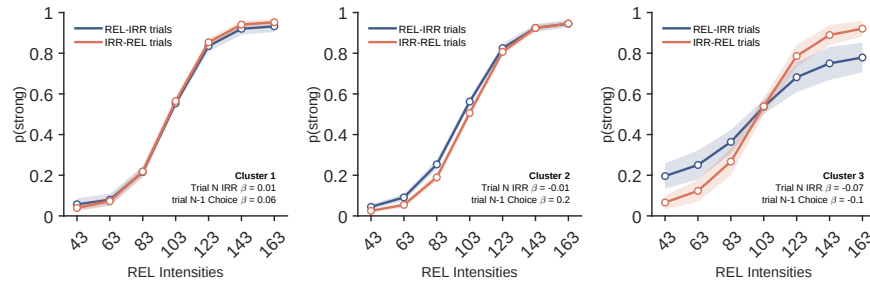


Figure 39: Overall performance at the cluster level. Psychometric functions showing average task performance for Clusters 1, 2, and 3.

The clusters also differed in the effect exerted by the IRR stimulus (Figure 40). Cluster 1 showed psychometric functions that largely overlapped during REL-IRR trials and IRR-REL trials, suggesting minimal influence of the irrelevant stimulus. This profile closely resembled the pattern observed in expert rats.

For Cluster 2, performance remained high across both trial types, but psychometric functions revealed a slight repulsive bias during IRR-REL trials. Irrelevant information tended to shift judgments away from, rather than toward, the intensity of the irrelevant stimulus.

Cluster 3 displayed a clear divergence from Clusters 1-2. Although participants in this cluster performed better during IRR-REL than REL-IRR trials, they exhibited a pronounced repulsive bias in both trial types, particularly when the relevant stimulus appeared first. This cluster therefore combined lower performance with a stronger impact of irrelevant information.

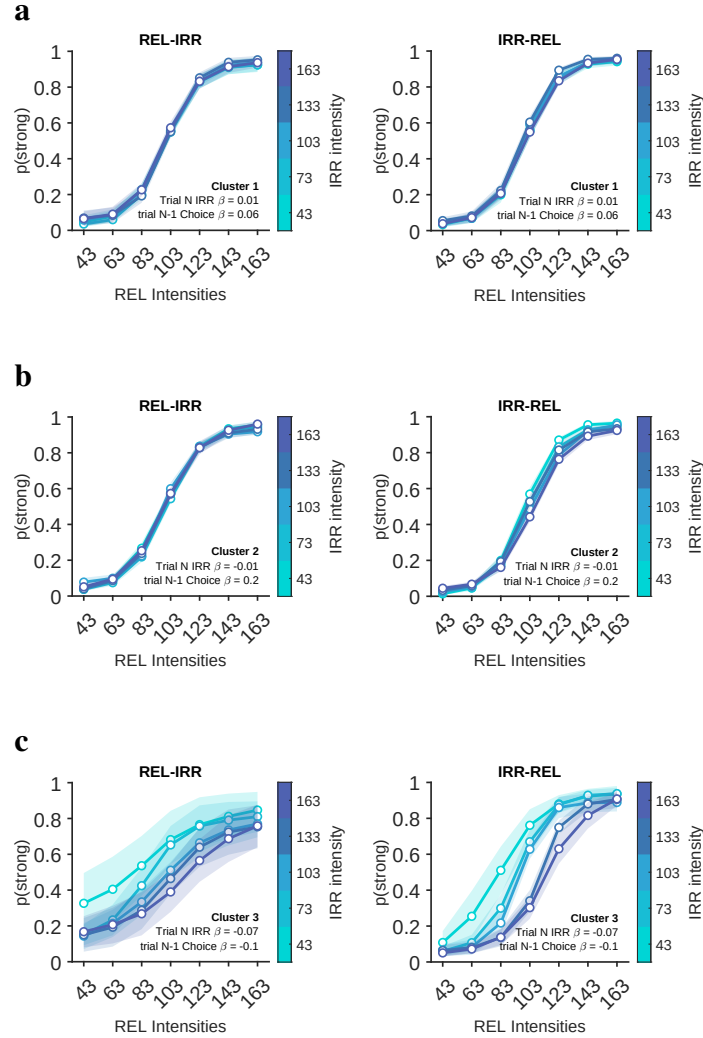


Figure 40: Irrelevant-stimulus bias at the cluster level. a. Psychometric functions conditioned on irrelevant-stimulus intensity for Cluster 1. **b.** Same analysis for Cluster 2. **c.** Same analysis for Cluster 3.

As a consequence of the psychometric curve shifts (Figure 40), the Point of Subjective Equality (PSE) also varied as a function of irrelevant-stimulus intensity (Figure 41).

Cluster 1 showed a slight decrease in PSE during REL-IRR trials and relatively stable PSE values during IRR-REL trials, consistent with the modest attractive bias observed in the psychometric functions. Cluster 2 exhibited a positive relationship between irrelevant-stimulus

intensity and PSE during IRR-REL trials, in agreement with the repulsive behavioral bias identified in this condition. Cluster 3 showed the largest PSE shifts, particularly during REL-IRR trials, mirroring the strong repulsive bias observed in the psychometric data.

The PSE analysis therefore provides a convergent summary of the cluster-specific psychometric shifts.

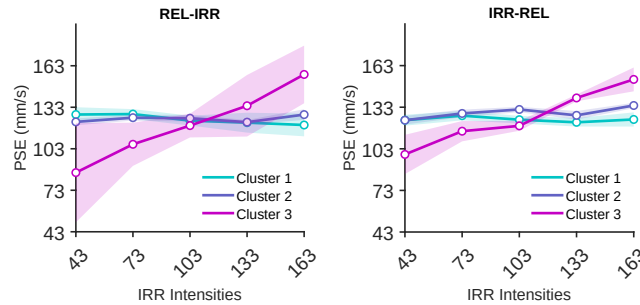


Figure 41: PSE variation as a function of irrelevant-stimulus intensity. PSE values are shown separately for REL-IRR (left) and IRR-REL (right) trials.

Clusters appear to differ in history dependence and learning dynamics

The cluster-specific psychometric functions also reflected the differences in previous-trial choice dependence captured by the GLM-based clustering (Figure 42).

Cluster 2 showed a modest attractive previous-trial choice effect, particularly during REL-IRR trials, consistent with the feature used for clustering. In contrast, Cluster 1 and Cluster 3 showed little evidence of systematic dependence on the immediately preceding choice. Thus, previous-trial history contributed differentially across participant groups and was not uniformly expressed throughout the sample.

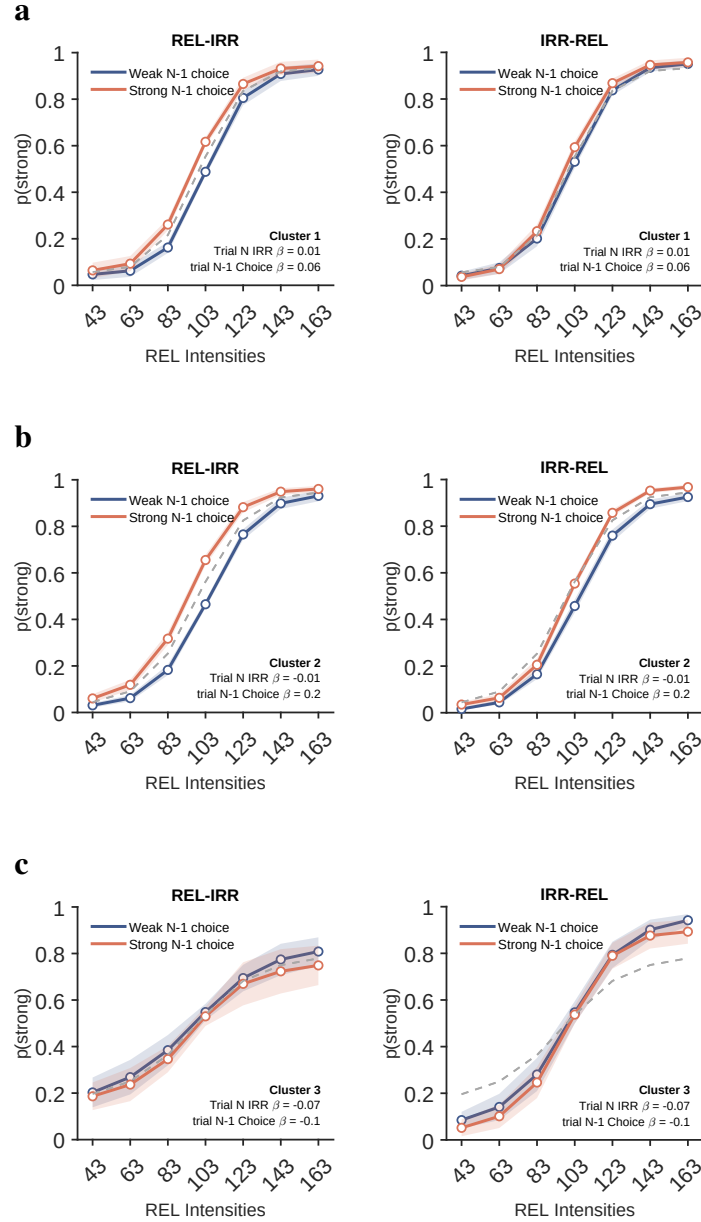


Figure 42: Effect of previous-trial choice at the cluster level. a. Psychometric functions conditioned on previous-trial choice for Cluster 1. b. Same analysis for Cluster 2. c. Same analysis for Cluster 3.

Overall, these results show that the apparent homogeneity observed at the population level concealed diverse behavioral profiles. The three clusters exhibited various patterns of

susceptibility to irrelevant information, ranging from minimal IRR influence in Cluster 1 to increasingly pronounced repulsive effects in Clusters 2 and 3. The third cluster also was characterized by lower task performance. These findings reveal substantial individual differences in how participants filtered irrelevant information and performed perceptual judgments during the dynamic task.

The influence of previous-trial choices differed across the three behavioral clusters. While group-level analyses suggested only a weak effect of trial history on current decisions, cluster-level analyses revealed that this apparent homogeneity concealed marked individual differences.

Cluster 2 exhibited the strongest previous-trial choice effect (Figure 42b). In this group, psychometric functions conditioned on the choice made in trial N-1 showed a modest attractive bias, indicating that participants were more likely to repeat the category selected on the preceding trial. This effect seemed present in both trial types and was more pronounced during REL-IRR trials.

In contrast, Cluster 1 showed little evidence of a systematic influence of previous-trial choice. Psychometric functions corresponding to the two previous-choice conditions largely overlapped, suggesting that current perceptual judgments were determined primarily by information available within the current trial rather than by recent decision history.

Cluster 3 similarly exhibited only weak history dependence. Despite displaying the strongest effects of irrelevant sensory information in other analyses, participants belonging to this cluster showed minimal shifts in psychometric functions as a function of previous-trial choice, indicating that trial history exerted only a limited influence on their decisions.

These results are consistent with the clustering analysis itself, which identified previous-trial choice weight as one of the principal dimensions separating participants. Importantly, the history effects were not distributed uniformly across the population. Instead, they were

driven largely by Cluster 2, whereas Clusters 1 and 3 showed little evidence of systematic attraction toward previous choices.

Taken together, these findings demonstrate that history-dependent effects were not a universal feature of task performance. Rather, sensitivity to previous choices characterized a specific subset of participants, suggesting that the contribution of trial history to perceptual decision-making varied substantially across individuals.

Learning dynamics differed across the three behavioral clusters, indicating that participants not only reached different behavioral states, but also followed different paths to achieve them (Figure 43).

At the population level, participants showed little evidence of learning because performance was already high at the beginning of training. However, separating participants according to behavioral cluster revealed distinct trajectories in the weighting of relevant and irrelevant information across the four testing blocks.

Cluster 1 showed a gradual increase in the weight assigned to the relevant stimulus in both trial types. In contrast, the weight assigned to the irrelevant stimulus remained close to zero throughout training, particularly during IRR-REL trials. These results indicate a progressive improvement in the use of task-relevant information while maintaining effective suppression of irrelevant inputs.

Cluster 2 exhibited the most stable behavioral profile. Across the four blocks, both relevant- and irrelevant-stimulus weights remained largely unchanged, with irrelevant-stimulus weights remaining close to zero in both trial types. This stability suggests that participants in this cluster adopted an effective strategy early in training and showed little subsequent adjustment in their use of task variables.

Cluster 3 displayed the most distinctive learning trajectory. In IRR-REL trials, the irrelevant-

stimulus weights were initially negative, consistent with the repulsive biases observed in the psychometric analyses. Across training, these weights remained relatively stable. In the case of REL-IRR trials, the irrelevant stimulus weights were initially close to zero, and became increasingly negative over time, indicating a gradual rise in the influence of irrelevant information.

At the same time, the weight assigned to the relevant stimulus increased, particularly during IRR-REL trials. Thus, participants in this cluster showed evidence of increased reliance on task-relevant information, but inconsistent reduction of the influence of irrelevant stimuli.

Despite these differences, all clusters converged toward greater weighting of the relevant stimulus over the course of training. This convergence suggests that participants progressively refined their use of task-relevant information, although the magnitude and temporal profile of learning varied substantially across clusters.

A modest reduction in relevant-stimulus weights was observed in the final training block across all clusters. Because this effect appeared consistently across groups, it may have reflected non-specific factors such as fatigue or reduced engagement during the final phase of the experiment.

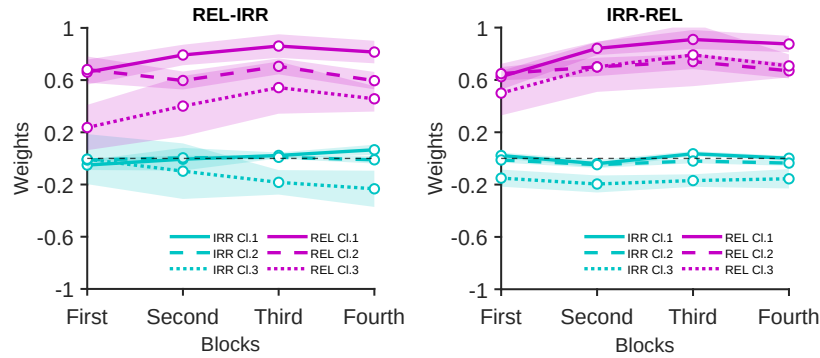


Figure 43: Learning trajectories at the cluster level. GLM weights for relevant (REL) and irrelevant (IRR) stimuli across the four testing blocks, shown separately for REL-IRR and IRR-REL trials.

Nevertheless, the differences among participants emerging from the clustering should be considered as exploratory profiles, rather than distinct, reproducible strategies. The individual variability may also reflect a continuum across participants, rather than discrete categories. Because the IRR and N-1 GLM coefficients were used to define the clusters, the resulting psychometric functions should be regarded as an internal consistency check, rather than independent validation of the clustering solution. See Supplementary Table S6 for participants' GLM coefficients.

Relationship between behavioral performance and introspection

To assess participants' subjective experience of the task, we administered a brief post-task questionnaire at the end of the second experimental session. The questionnaire evaluated perceived task difficulty, confidence in performance, the influence of stimulus position, and memory of the reference stimulus (see Supplementary Figure S13). We then compared questionnaire responses with objective behavioral measures, including hit rate, previous-trial choice weight, and irrelevant-stimulus weight.

Behavioral performance showed systematic relationships with subjective evaluations of the task (Figure 44). Participants who reported that the task was difficult tended to achieve higher hit rates than participants who described the task as easy. Likewise, participants who reported that trials were more difficult when the relevant stimulus appeared first generally exhibited higher behavioral performance. In contrast, lower performance was associated with uncertainty regarding the intensity of the reference stimulus.

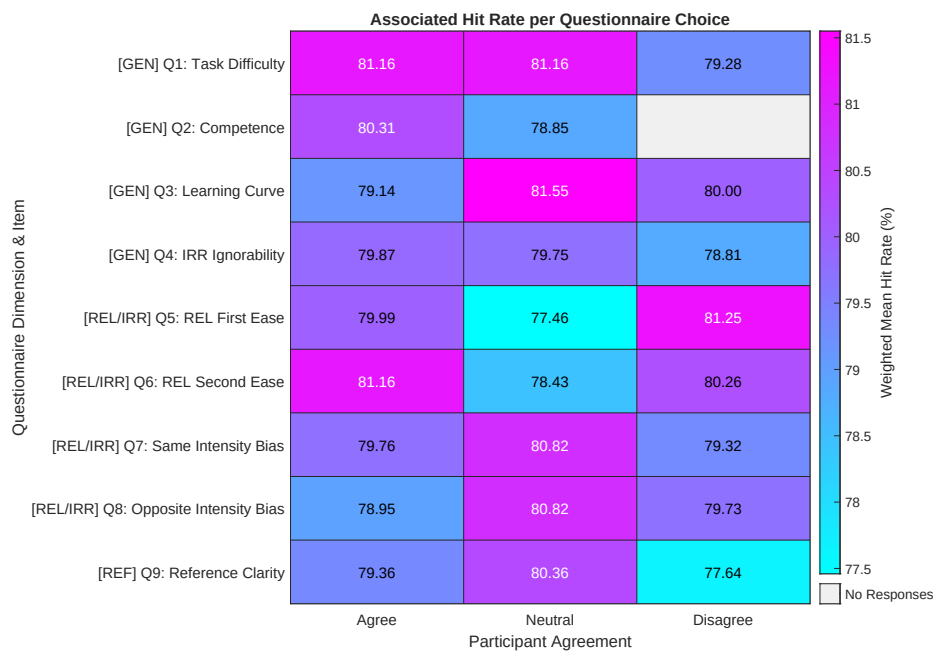


Figure 44: Relationship between task performance and questionnaire responses. Heatmap showing the association between questionnaire responses and average hit rate. Each cell represents a normalized cluster-weighted behavioral measure associated with a specific question-and-answer combination.

A similar pattern emerged when questionnaire responses were examined as a function of previous-trial choice effects (Figure 45). Participants showing stronger previous-trial choice weights were more likely to report that the task was difficult and to identify differences

in difficulty between trial types. Conversely, weaker previous-trial choice effects were associated with uncertainty regarding the reference stimulus.

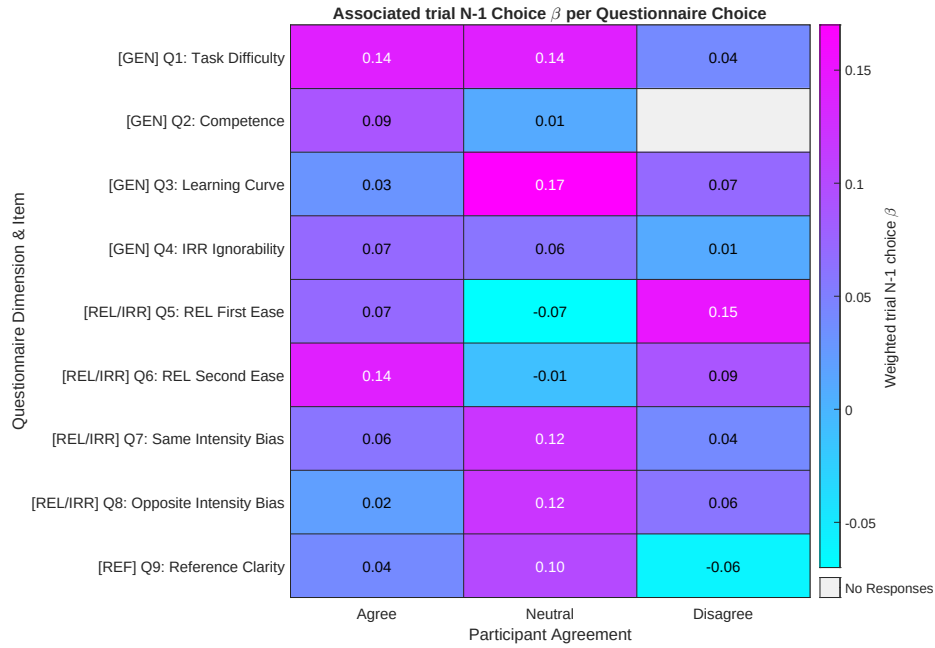


Figure 45: Relationship between previous-trial choice effects and questionnaire responses. Heatmap showing the association between questionnaire responses and average previous-trial choice GLM weight.

Questionnaire responses were also related to the influence of irrelevant information on perceptual judgments (Figure 46). Participants exhibiting smaller irrelevant stimulus weights generally reported greater task difficulty and were more likely to distinguish between the two trial types. Conversely, participants who reported uncertainty regarding the reference stimulus tended to exhibit larger irrelevant-stimulus effects.

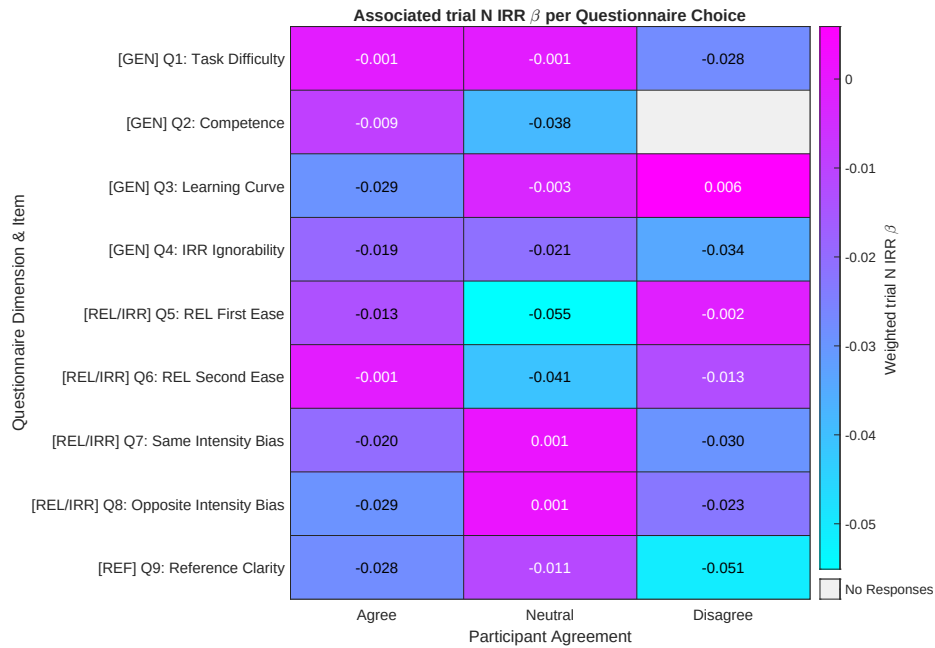


Figure 46: Relationship between irrelevant stimulus effects and questionnaire responses. Heatmap showing the association between questionnaire responses and average irrelevant stimulus GLM weight.

Although these analyses are descriptive and exploratory, they reveal a consistent relationship between subjective task awareness and objective behavioral performance. Participants who demonstrated stronger performance, lower susceptibility to irrelevant information, and clearer task representations were also more likely to report greater awareness of task demands and stimulus structure.

Taken together, these findings suggest that individual differences in task performance are reflected not only in behavioral measures but also in participants' subjective evaluations of the task.

Correlation between neurodivergent traits and attentional control

To investigate whether individual differences in attentional performance were associated with neurodivergent traits, participants completed the AQ-10, SPQ-10, and Aberrant Salience Inventory (ASI) before behavioral testing. We then examined the relationship between questionnaire scores and overall task performance.

No significant correlation was observed between task performance and scores on any of the three questionnaires (Figure 47). This absence of association was evident for measures of autistic traits, sensory-perceptual autistic characteristics, and aberrant salience tendencies.

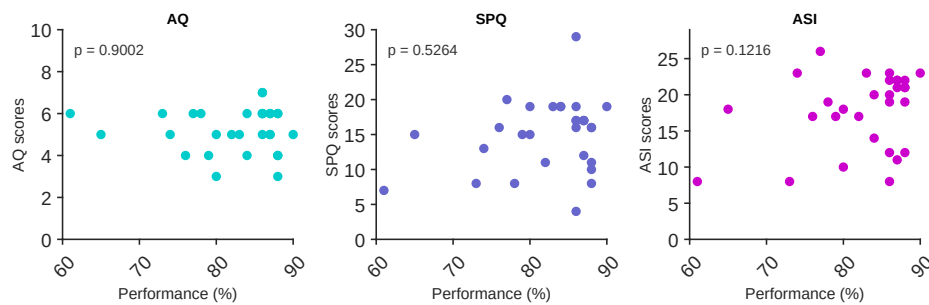


Figure 47: Relationship between questionnaire scores and behavioral performance. Scatterplots showing the relationship between AQ-10 scores (left), SPQ-10 scores (center), ASI scores (right), and average task performance. Each point represents an individual participant. No significant correlations were observed.

These results indicate that variation in questionnaire scores within this sample did not predict success in the dynamic task. Because the present sample consisted exclusively of individuals without reported neuropsychiatric diagnoses, these results should be interpreted as evidence that normal variation in these traits was not associated with meaningful differences in task performance within a non-clinical population.

Overall, the behavioral variability observed throughout the chapter was more strongly

explained by differences in task strategy than by variation in questionnaire-derived measures of autistic traits or aberrant salience.

Discussion

The present study investigated temporal attentional control in the tactile modality using a behavioral paradigm adapted for humans from our previous work in rats. Participants were required to selectively report a relevant tactile stimulus while ignoring an irrelevant tactile distractor presented within the same trial. On average, performance was high. However, analyses conducted at the individual level revealed substantial heterogeneity in behavioral strategies, suggesting that attentional control is not implemented through one single uniform mechanism.

Individual differences and cross-species similarities in attentional control

A central finding is that participants could be separated into three clusters based on their interaction with irrelevant sensory information. One cluster exhibited slight repulsive biases, where judgments of the relevant stimulus were systematically pulled away from the intensity of the distractor. Another cluster instead showed near-zero influence of irrelevant information. In contrast, a third cluster displayed a strong repulsive influence of the distractor alongside lower overall performance. These results indicate that the same task demands can be solved through seemingly different behavioral strategies.

These observations parallel the findings reported in rats. In the rodent study, animals exposed to the same task structure exhibited different forms of attentional control, characterized by different behavioral and neuronal signatures. The present findings suggest that this variability was not restricted to rodents but may have reflected a more general principle governing attentional behavior across species. Rather than representing noise around a common optimal solution, individual differences may correspond to stable alternative strategies for managing competing streams of sensory information.

Alternative strategies for processing irrelevant information

The residual influence of irrelevant stimuli observed across participants is consistent with an incomplete filtering account of attentional control [62,63]. Even when explicitly instructed to ignore a stimulus, participants appeared unable to eliminate its influence entirely. Instead, information from both tactile events contributed to the final perceptual judgment. Such a mechanism may be advantageous in natural environments, where signals initially considered irrelevant can suddenly become behaviorally meaningful. This scenario suggests that irrelevant stimuli should not be completely suppressed, but should be represented with some salience just below the threshold for action. From this perspective, attentional selection may operate more as a weighting process than as an all-or-none filtering mechanism [64].

The repulsive pattern observed in the third cluster is more difficult to interpret. One possibility is that these participants adopted an active suppression strategy, exaggerating the distinction between relevant and irrelevant information. Such an approach could have reduced susceptibility to distraction but may also have introduced systematic distortions into the perceptual decision process. Alternatively, the repulsive effect may have reflected a less stable learning strategy or a misunderstanding of task contingencies. The fact that this cluster also showed lower average performance favors the possibility that repulsion emerged from a suboptimal solution to the task rather than a superior distractor-suppression mechanism. Nevertheless, because the present study was designed primarily to identify behavioral phenotypes, the cognitive and neural processes underlying these differences remain to be clarified.

Individual traits and metacognitive awareness

One of the motivations of this study was to investigate whether variability in attentional performance could be related to individual differences in autistic traits or aberrant salience [57,65–67]. Contrary to our expectations, neither measure was significantly associated with overall performance nor with the behavioral clusters identified in the task. This absence of effect should be interpreted cautiously. First, the present sample consisted of neurotypical individuals, limiting the range of scores available for detecting meaningful relationships. Second, attentional control is influenced by multiple cognitive factors, and the contribution of broad self-reported traits may have been relatively small compared to task-specific strategies. Although null results cannot exclude a relationship between these dimensions and attentional behavior, they suggest that the behavioral heterogeneity observed here was not straightforwardly explained by variation in autistic traits or aberrant salience within a non-clinical population.

We also explored participants' metacognitive awareness by administering a post-experimental questionnaire assessing subjective evaluations of task difficulty and performance. While these analyses were exploratory, the results suggest that participants possessed varying levels of insight into their own behavior. Individuals reporting greater task difficulty tended to exhibit behavioral patterns consistent with higher objective performance, whereas other participants appeared to overestimate their success. These observations raise the possibility that metacognitive monitoring captured meaningful aspects of attentional strategy selection [68]. Future studies employing established confidence-rating procedures could investigate this relationship more systematically.

Methods

Human subjects

30 healthy adult volunteers were recruited to participate in the study. Participants reported normal or corrected-to-normal vision, and no known neurological disorders, as assessed by the demographic questionnaire reporting age, gender, nationality, neurological diagnoses, and current pharmacological treatment. All participants provided informed consent prior to participation. All participants performed two experimental sessions, each lasting approximately one and a half hours, and were paid for their participation, after study completion, on a performance basis. The experimental procedures were conducted in accordance with institutional ethical guidelines and the principles outlined in the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of the International School for Advanced Studies (SISSA).

In addition to the behavioral experiment, participants were invited to complete a series of self-report questionnaires assessing autistic traits, aberrant salience, and post-task metacognitive evaluations.

Experimental apparatus

All experiments took place in the SENSEs Lab, at SISSA. Participants were seated comfortably in front of a computer monitor in a quiet testing environment. Their right arm rested on a pillow placed on a desk. Tactile stimuli were delivered to the fingertip through a vibrotactile stimulator consisting of a plastic probe, attached to a shaker motor (type 4808; Bruel and Kjaer) producing vibrations in the horizontal direction, perpendicular to the fingertip. Visual cues were presented on a computer display positioned directly in front of the

participant. Stimulus presentation, response collection, and data storage were controlled using custom LabVIEW software (National Instruments).

Vibrotactile stimulus generation

Vibrotactile stimuli consisted of vibrations, produced in the same manner as those described in the rat study (see Methods).

Experimental design

The experiment was based on an adaptation of the Irrelevant Stimulus Paradigm previously developed for rodents. The task performed by human subjects was analogous to the rat dynamic task, with the following modifications. The inter-stimulus interval was kept fixed at 1000 ms. The stimulus set was reduced to 7 intensities for the relevant stimulus, and 5 intensities for the irrelevant stimulus. These modifications provided a compromise between sufficient trial count and total study duration.

Each trial consisted of two sequential tactile stimuli presented to the same fingertip. One stimulus was designated as task-relevant, whereas the other served as an irrelevant distractor. The order of relevant and irrelevant stimuli varied across trials. The relevant stimulus was tagged with a visual cue, delivered immediately before stimulus presentation. Participants were instructed to ignore the uncued stimulus and categorize the intensity of the cued stimulus as either weak or strong. The visual cue consisted of an on-screen green light, turning on for the same duration as the acoustic cue used in the rat study (300 ms).

The intensity of both relevant and irrelevant stimuli varied across trials according to the predefined stimulus sets. Participants responded after the presentation of both tactile stimuli by selecting the category corresponding to the perceived intensity of the relevant stimulus.

Experimental procedure

The experiment consisted of four test blocks distributed across two experimental sessions. Sessions were separated in time to allow the assessment of learning trajectories and performance changes across training.

Prior to the start of the experiment, participants received verbal and written instructions explaining the task structure. Practice trials were administered to ensure that participants understood the response requirements and the meaning of the visual cues.

During the test blocks, participants completed multiple trials in which relevant stimulus position, stimulus intensity, and distractor intensity were systematically varied.

Behavioral analysis

Psychometric curve estimation

Psychometric curves were estimated using the same approach described in the rat study. Specifically, psychometric functions were fitted separately for each participant. For across-subject plots, fitted probabilities at each tested stimulus intensity were first computed for individual subjects, and then averaged across participants. SEM was computed across the individually fitted probabilities. Behavioral performance was quantified as the proportion of correct responses across trials. Psychometric analyses were performed to evaluate the influence of stimulus intensity and distractor magnitude on perceptual judgments.

Generalized Linear Model estimation

GLM weights were estimated following the same approach described in the rat study. For the learning trajectory analyses, GLM weights were estimated separately for each participant and

experimental block, and subsequently averaged across subjects. The purpose of the GLM analysis was to quantify the relative influence of the REL and IRR stimuli on perceptual judgments and to characterize how this influence changed over time. Participant-level estimates were therefore available and are reported in Supplementary Table S6. Out-of-sample validation was not performed, as the GLM analysis was used to describe the influence of stimuli on choice, rather than to evaluate predictive performance on unseen trials.

Point of Subjective Equality estimation

We estimated the Point of Subjective Equality (PSE) by computing the psychometric curves and storing the parameter μ . PSE was defined as the parameter μ of the fitted cumulative Gaussian psychometric function, and expressed in the corresponding stimulus velocity.

Clustering analysis

To investigate individual differences in task performance, participants were clustered according to two behavioral features quantifying the influence on current perceptual judgments: the effect of the irrelevant stimulus presented in trial N and the effect of the choice made in trial $N - 1$.

For each participant, the two clustering features were estimated using separate binomial Generalized Linear Models with a probit link, implemented with the MATLAB `fitglm` function. The trial N irrelevant-stimulus weight was obtained by modeling current-trial choice as a function of relevant- and irrelevant-stimulus intensities, whereas previous-choice dependence was estimated by modeling current-trial choice as a function of the choice made on trial $N - 1$. The corresponding GLM coefficients were retained as the trial N IRR and trial $N - 1$ choice features, respectively.

Each participant was therefore represented by a two-dimensional feature vector containing the trial N IRR weight and the trial $N - 1$ choice weight. The resulting 30×2 feature matrix was clustered using the *k-medoids* algorithm with Manhattan distance (also known as city-block distance).

The *k-medoids* clustering minimizes the within-cluster distance between each participant and the medoid of the cluster to which that participant is assigned:

$$J = \sum_{k=1}^K \sum_{i \in C_k} d(x_i, m_k) \quad (18)$$

where K is the total number of clusters ($K = 3$), C_k is the set of participants assigned to cluster k , x_i is the two-dimensional feature vector describing participant i , and m_k is the medoid of cluster k , corresponding to an actual participant in the dataset.

Manhattan distance was used as the distance metric:

$$d(x_i, m_k) = |x_{i,1} - m_{k,1}| + |x_{i,2} - m_{k,2}| \quad (19)$$

where $x_{i,1}$ and $m_{k,1}$ correspond to the trial N IRR GLM weights of participant i and the cluster medoid, respectively, whereas $x_{i,2}$ and $m_{k,2}$ correspond to their trial $N - 1$ choice GLM weights.

To evaluate the number of clusters, a silhouette analysis was performed for *k-medoids* solutions with $k = 2$ to $k = 5$. For each participant, the silhouette coefficient quantified the relative similarity to observations within the assigned cluster compared with the nearest alternative cluster. The coefficient was computed as

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (20)$$

where $a(i)$ is the mean city-block distance between participant i and the other participants in the same cluster, and $b(i)$ is the minimum mean distance between participant i and participants in any other cluster. Silhouette coefficients range from -1 to 1 , with larger values indicating greater within-cluster cohesion and between-cluster separation. The mean silhouette coefficient across participants was used to compare candidate values of k . We chose $k = 3$ as it corresponded to the highest mean silhouette coefficient.

Cluster stability was assessed using repeated participant subsampling. Across 1000 iterations, 80% of participants were randomly selected and reclustered using $k = 3$ and city-block distance. The resulting partition was compared with the corresponding original participant assignments from the full-sample clustering using the Adjusted Rand Index (ARI). The ARI measures similarity between two cluster partitions based on pairwise agreement, while correcting for agreement expected by chance. Values approaching 1 indicate highly similar partitions, whereas values around 0 indicate agreement close to chance. Negative values indicate agreement below chance expectation.

Cluster assignments were subsequently used to characterize performance, learning trajectories, and cluster-specific psychometric patterns.

Questionnaire Measures

Autism Spectrum Quotient and Sensory-Perception Quotient

Participants completed the AQ-10 to assess autistic traits and the SPQ-10 to assess sensory-perceptual characteristics associated with autism ([69]). Total scores were calculated according to the corresponding scoring procedures and used as continuous measures of trait variation within the sample. Total questionnaire scores were correlated with the behavioral performance (i.e., proportion of hit trials) using Spearman's correlation, with $\alpha = 0.05$.

Aberrant Salience

Participants completed a questionnaire assessing aberrant salience tendencies, defined as the propensity to attribute significance to otherwise neutral events or stimuli (ASI questionnaire, [70]). Total questionnaire scores were correlated with the behavioral performance (i.e., proportion of hit trials) using Spearman's correlation, with $\alpha = 0.05$.

Metacognitive Assessment

Following completion of the experimental sessions, participants completed a brief questionnaire evaluating their subjective experience of the task. Items assessed perceived task difficulty, confidence in performance, and general impressions regarding attentional demands.

These measures were analyzed in an exploratory manner to examine possible relationships between self-evaluation and objective task performance. The procedure involved computing a weighted average of the behavioral metrics across a cluster-based feature space.

First, the distribution of question-and-answer responses across the three clusters was tabulated: to prevent cluster size imbalances from biasing the results, these raw counts were normalized by total cluster size to yield a within-cluster response proportion. Next, these normalized proportions were used as weights to calculate a weighted average of the three behavioral metrics (hit rate, trial N IRR β , trial N-1 choice β) for each individual cell.

Each cell in the resulting heatmaps reflects the normalized, cluster-weighted behavioral metric associated with a specific item and answer of the questionnaire. The equation below describes the measure contained in each cell:

$$\text{Cell Value} = \frac{\sum_{c=1}^3 \left(\text{Cluster Mean}_{m,c} \times \text{Normalized Weight}_{c,q,a} \right)}{\sum_{c=1}^3 \text{Normalized Weight}_{c,q,a}} \quad (21)$$

Where c represents the cluster index ($c \in \{1, 2, 3\}$). m represents the specific behavioral or computational metric evaluated (Hit Rate, trial N IRR β , or trial $N - 1$ choice β). q represents the specific item (question) from the questionnaire ($q \in \{1, \dots, 9\}$). a represents the participant's level of agreement (answer) for that item ($a \in \{\text{Agree, Neutral, Disagree}\}$). $\text{Cluster Mean}_{m,c}$ is the mean value of metric m within cluster c . $\text{Normalized Weight}_{c,q,a}$ is the response proportion of cluster c associated with item q and answer a , computed as the raw count of participants in cluster c who chose answer a divided by the total number of individuals assigned to that cluster.

General Conclusions

From learning what is irrelevant to flexible selective weighting

When the task structure was stable, rats were able to learn to ignore irrelevant information, progressively decreasing its influence by assigning less weight. However, when task structure became unpredictable and relevance had to be assigned flexibly across time, animals showed difficulty in filtering out irrelevant stimuli.

This difference can be understood at the level of task demands. The standard task provided a stable mapping that supported learning of a consistent policy. The dynamic task introduced a conflict between temporal structure and task rule, increasing computational demands [71]. Under these conditions, animals appeared to adopt a simpler temporal strategy that was only partially effective, favoring the last stimulus of the sequence.

Overall, the current thesis characterized both the development and the limitations of attentional control in rats. A key finding is that the ability to ignore irrelevant information is not all-or-none, but changes with learning and depends strongly on task demands. These results also support a view of attention as the differential weighting of sensory information, rather than the complete elimination of irrelevant signals.

Neural representation does not necessarily determine behavioral influence

Neural recordings revealed how relevant and irrelevant information were represented across different behavioral stages. In an expert animal, irrelevant signals remained represented within the neural population despite exerting no influence on behavioral choices. Similarly,

in some cases, choice-related signals could be detected before the overt decision, further indicating that the presence of a neural representation does not necessarily imply a direct influence on behavior.

The clearest example of this dissociation between sensory encoding and downstream read-out was observed in the animal performing the dynamic task. Both sensory and frontal cortices represented the first stimulus, while showing no clear representation of the second stimulus. Nevertheless, behavioral choices were biased toward the second stimulus. Thus, the information most strongly represented at the neural level was not necessarily the information exerting the strongest influence on choice.

The proficient animal provided an intermediate case between sensory encoding and decision-related processing. Irrelevant information remained represented but was partially suppressed as relevant signals emerged, suggesting a possible mechanism through which sensory representations may be progressively transformed before influencing behavior.

Overall, the present thesis provides evidence that neural populations can preferentially represent one source of sensory information while behavioral processes assign greater weight to another. This dissociation between neural representation and decisional influence is consistent with the view that attentional control can be implemented through selective weighting, rather than through the complete elimination of irrelevant sensory information.

Rats and human participants share individual variability

Although rats showed individual variability, two broad behavioral patterns emerged across task structures. In the standard task, animals consistently judged the relevant stimulus but differed in the extent to which irrelevant information continued to influence their choices. In contrast, animals performing the dynamic task exhibited limited flexibility and an overall

tendency to favor the second stimulus of the trial.

Individual variability was also evident in human participants. Exploratory clustering identified behavioral profiles characterized by different contributions of irrelevant information and previous choices, while learning trajectories further showed that improvements across sessions were not uniform across clusters. Thus, despite undergoing the same training procedure, participants differed in how task-relevant and task-irrelevant information contributed to their behavior.

Overall, the present thesis suggests that attentional control does not emerge through a single, canonical process. Across both rats and human participants, behavior reflected different degrees and forms of selective weighting, shaped by task structure, learning, and individual differences.

Future work employing trial-by-trial computational modeling may help determine whether variability across rats and human participants arises from differences in sensory encoding, selective weighting of information, decision processes, or learning rates.

Limitations and future directions

Several limitations of the rat study should be acknowledged. The number of subjects, particularly for electrophysiological recordings, is limited, and individual animals should be considered case studies rather than representatives of a broader population. The absence of vS1 recordings in expert animals further restricts interpretation across processing stages.

In addition, inferences about underlying strategies are indirect. Neural decoding reveals what information is represented, but not how it is used to guide behavior. As a result, different neural patterns may reflect multiple plausible computational mechanisms.

Future work should track learning within subjects, increase sample size, and incorporate

causal manipulations to directly test the role of specific circuits and frequency bands in attentional control.

In relation to the human study, several limitations should be acknowledged as well. First, the sample size may have constrained the detection of subtle associations between questionnaire measures and behavioral performance. Second, the clustering analyses identify reproducible behavioral phenotypes but do not directly explain the underlying cognitive mechanisms. Third, unlike the rodent study, the present work relied exclusively on behavioral measurements and self-report data, preventing direct examination of the neural processes supporting the observed strategies.

References

- [1] Jieyu Zheng and Markus Meister. The unbearable slowness of being: Why do we live at 10 bits/s? *Neuron*, 113(2):192–204, 2025.
- [2] Ruobing Xia, Xiaomo Chen, Tatiana A Engel, and Tirin Moore. Common and distinct neural mechanisms of attention. *Trends in Cognitive Sciences*, 28(6):554–567, 2024.
- [3] Tirin Moore and Marc Zirnsak. Neural mechanisms of selective visual attention. *Annual review of psychology*, 68:47–72, 2017.
- [4] Annalisa Tosoni, Paolo Capotosto, Antonello Baldassarre, Sara Spadone, and Carlo Sestieri. Neuroimaging evidence supporting a dual-network architecture for the control of visuospatial attention in the human brain: a mini review. *Frontiers in Human Neuroscience*, 17:1250096, 2023.
- [5] Lupeng Wang and Richard J Krauzlis. Visual selective attention in mice. *Current Biology*, 28(5):676–685, 2018.
- [6] Kaitlin S Carlson, Marie A Gadziola, Emma S Dauster, and Daniel W Wesson. Selective attention controls olfactory decisions and the neural encoding of odors. *Current Biology*, 28(14):2195–2205, 2018.
- [7] Rachel N Denison. Visual temporal attention from perception to computation. *Nature Reviews Psychology*, 3(4):261–274, 2024.
- [8] Verena C Seibold, Janina Balke, and Bettina Rolke. Temporal attention. *Frontiers in Cognition*, 2:1168320, 2023.
- [9] Anna C Nobre and Freek Van Ede. Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience*, 19(1):34–48, 2018.

- [10] Manuel Gomez-Ramirez, Kristjana Hysaj, and Ernst Niebur. Neural mechanisms of selective attention in the somatosensory system. *Journal of neurophysiology*, 116(3):1218–1231, 2016.
- [11] Heidi Johansen-Berg, Donna M Lloyd, et al. The physiology and psychology of selective attention to touch. *Front Biosci*, 5:D894–D904, 2000.
- [12] Conrad CY Lee, Ehsan Kheradpezhoh, Mathew E Diamond, and Ehsan Arabzadeh. State-dependent changes in perception and coding in the mouse somatosensory cortex. *Cell Reports*, 32(13), 2020.
- [13] Athena Akrami, Charles D Kopec, Mathew E Diamond, and Carlos D Brody. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*, 554(7692):368–372, 2018.
- [14] Åke Hellström and Thomas H Rammsayer. Time-order errors and standard-position effects in duration discrimination: An experimental study and an analysis by the sensation-weighting model. *Attention, Perception, & Psychophysics*, 77(7):2409–2423, 2015.
- [15] Arash Fassihi, Athena Akrami, Vahid Esmaeili, and Mathew E Diamond. Tactile perception and working memory in rats and humans. *Proceedings of the National Academy of Sciences*, 111(6):2331–2336, 2014.
- [16] Donald Eric Broadbent. *Perception and communication*. Elsevier, 2013.
- [17] J Anthony Deutsch and Diana Deutsch. Attention: Some theoretical considerations. *Psychological review*, 70(1):80, 1963.
- [18] Anne M Treisman. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4):242–248, 1960.

- [19] Robert Desimone, John Duncan, et al. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.
- [20] Tirin Moore and Katherine M Armstrong. Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, 421(6921):370–373, 2003.
- [21] Basilis Zikopoulos and Helen Barbas. Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. *Journal of Neuroscience*, 26(28):7348–7361, 2006.
- [22] Joao Barbosa, Rémi Proville, Chris C Rodgers, Michael R DeWeese, Srdjan Ostojic, and Yves Boubenec. Early selection of task-relevant features through population gating. *Nature communications*, 14(1):6837, 2023.
- [23] Rui Bao, Shijie Chang, Ruixiang Liu, Yunning Wang, and Yifu Guan. Research status of visuospatial dysfunction and spatial navigation. *Frontiers in aging neuroscience*, 17:1609620, 2025.
- [24] Nhat Minh Le, Murat Yildirim, Yizhi Wang, Hiroki Sugihara, Mehrdad Jazayeri, and Mriganka Sur. Mixtures of strategies underlie rodent behavior during reversal learning. *PLOS Computational Biology*, 19(9):e1011430, 2023.
- [25] Mario Treviño. Stimulus similarity determines the prevalence of behavioral laterality in a visual discrimination task for mice. *Scientific reports*, 4(1):7569, 2014.
- [26] Román Rossi-Pool, Emilio Salinas, Antonio Zainos, Manuel Alvarez, José Vergara, Néstor Parga, and Ranulfo Romo. Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *Proceedings of the National Academy of Sciences*, 113(49):E7966–E7975, 2016.

- [27] Gabriel Diaz-deLeon, Manuel Alvarez, Lucas Bayones, Antonio Zainos, Jerónimo Zizumbo, Sergio Parra, Sebastián Pujalte, Ranulfo Romo, Román Rossi-Pool, and Victor De Lafuente. An abstract categorical decision code in dorsal premotor cortex. *Proceedings of the National Academy of Sciences*, 119(50):e2214562119, 2022.
- [28] W Bryan Wilent and Diego Contreras. Stimulus-dependent changes in spike threshold enhance feature selectivity in rat barrel cortex neurons. *Journal of Neuroscience*, 25(11):2983–2991, 2005.
- [29] James M Rowland, Thijs L van der Plas, Matthias Loidolt, Robert M Lees, Joshua Keeling, Jonas Dehning, Thomas Akam, Viola Priesemann, and Adam M Packer. Propagation of activity through the cortical hierarchy and perception are determined by neural variability. *Nature Neuroscience*, 26(9):1584–1594, 2023.
- [30] Louis D Matzel and Bruno Sauce. Individual differences: Case studies of rodent and primate intelligence. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43(4):325, 2017.
- [31] Valeria Mazza, Jana A Eccard, Marco Zaccaroni, Jens Jacob, and Melanie Dammhahn. The fast and the flexible: cognitive style drives individual variation in cognition in a small mammal. *Animal Behaviour*, 137:119–132, 2018.
- [32] Elsa Pittaras, Héloïse Hamelin, and Sylvie Granon. Inter-individual differences in cognitive tasks: focusing on the shaping of decision-making strategies. *Frontiers in Behavioral Neuroscience*, 16:818746, 2022.
- [33] Thomas S Wierda, Shirin Dora, Cyriel MA Pennartz, and Jorge F Mejias. Diverse and flexible behavioral strategies arise in recurrent neural networks trained on multisensory decision making. *PLOS Computational Biology*, 21(10):e1013559, 2025.

- [34] Marino Pagan, Vincent D Tang, Mikio C Aoi, Jonathan W Pillow, Valerio Mante, David Sussillo, and Carlos D Brody. Individual variability of neural computations underlying flexible decisions. *Nature*, 639(8054):421–429, 2025.
- [35] Zoe C. Ashwood, Nicholas A. Roy, Iris R. Stone, Anne E. Urai, Anne K. Churchland, Alexandre Pouget, and Jonathan W. Pillow. Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, 25(2):201–212, February 2022.
- [36] Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78–84, 2013.
- [37] Eric Drebitz, Lukas-Paul Rausch, and Andreas K Kreiter. Gamma-band synchronization between neurons in the visual cortex is causal for effective information processing and behavior. *Nature Communications*, 16(1):7380, 2025.
- [38] Pascal Fries. Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235, 2015.
- [39] Andreas K Engel and Pascal Fries. Beta-band oscillations—signalling the status quo? *Current opinion in neurobiology*, 20(2):156–165, 2010.
- [40] Mikael Lundqvist, Pawel Herman, Melissa R Warden, Scott L Brincat, and Earl K Miller. Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature communications*, 9(1):394, 2018.
- [41] George Paxinos and Charles Watson. *The rat brain in stereotaxic coordinates: hard cover edition*. Elsevier, 2006.

- [42] Fernando J Chaure, Hernan G Rey, and Rodrigo Quian Quiroga. A novel and fully automatic spike-sorting implementation with variable number of features. *Journal of neurophysiology*, 120(4):1859–1871, 2018.
- [43] Dmitry Kobak, Wieland Brendel, Christos Constantinidis, Claudia E Feierstein, Adam Kepecs, Zachary F Mainen, Xue-Lian Qi, Ranulfo Romo, Naoshige Uchida, and Christian K Machens. Demixed principal component analysis of neural population data. *elife*, 5:e10989, 2016.
- [44] Joao Barbosa, Heike Stein, Rebecca L Martinez, Adrià Galan-Gadea, Sihai Li, Josep Dalmau, Kirsten CS Adam, Josep Valls-Solé, Christos Constantinidis, and Albert Compte. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature neuroscience*, 23(8):1016–1024, 2020.
- [45] Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011(1):156869, 2011.
- [46] Michael I Posner. Orienting of attention. *Quarterly journal of experimental psychology*, 32(1):3–25, 1980.
- [47] Marisa Carrasco. Visual attention: The past 25 years. *Vision research*, 51(13):1484–1525, 2011.
- [48] Fabian Hutmacher. Why is there so much more research on vision than on any other sensory modality? *Frontiers in psychology*, 10:481030, 2019.
- [49] Anne M Treisman. Selective attention in man. *British medical bulletin*, 1964.

- [50] Maurizio Corbetta and Gordon L Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3):201–215, 2002.
- [51] Christian C Ruff, Felix Blankenburg, Otto Bjoertomt, Sven Bestmann, Elliot Freeman, John-Dylan Haynes, Geraint Rees, Oliver Josephs, Ralf Deichmann, and Jon Driver. Concurrent tms-fmri and psychophysics reveal frontal influences on human retinotopic visual cortex. *Current biology*, 16(15):1479–1488, 2006.
- [52] Susan J Lederman and Roberta L Klatzky. Hand movements: A window into haptic object recognition. *Cognitive psychology*, 19(3):342–368, 1987.
- [53] Matthew J Hertenstein, Dacher Keltner, Betsy App, Brittany A Bulleit, and Ariane R Jaskolka. Touch communicates distinct emotions. *Emotion*, 6(3):528, 2006.
- [54] Norihiro Sadato, Alvaro Pascual-Leone, Jordan Grafman, Vicente Ibañez, Marie-Pierre Deiber, George Dold, and Mark Hallett. Activation of the primary visual cortex by braille reading in blind subjects. *Nature*, 380(6574):526–528, 1996.
- [55] Daniel Goldreich and Ingrid M Kanics. Tactile acuity is enhanced in blindness. *Journal of Neuroscience*, 23(8):3439–3445, 2003.
- [56] Brandon Keehn, Ralph-Axel Müller, and Jeanne Townsend. Atypical attentional networks and the emergence of autism. *Neuroscience & Biobehavioral Reviews*, 37(2):164–183, 2013.
- [57] Eleni A Demetriou, Amit Lampit, Daniel S Quintana, Sharon L Naismith, Yun JC Song, Julia E Pye, Ian Hickie, and Adam J Guastella. Autism spectrum disorders: a meta-analysis of executive function. *Molecular psychiatry*, 23(5):1198–1204, 2018.

- [58] Charlotte A Chun, Peter Brugger, and Thomas R Kwapil. Aberrant salience across levels of processing in positive and negative schizotypy. *Frontiers in psychology*, 10:2073, 2019.
- [59] Federico Fiori Nastro, Martina Pelle, Alice Clemente, Fernando Corinto, Davide Prosperi Porta, Yael Sonnino, Carmine Gelormini, Giorgio Di Lorenzo, and Michele Ribolsi. Investigating aberrant salience in autism spectrum disorder and psychosis risk: A cross-group analysis. *Early intervention in psychiatry*, 19(10):e70099, 2025.
- [60] I Hachen, S Reinartz, R Brasselet, A Stroligo, and ME Diamond. Dynamics of history-dependent perceptual judgment. *Nature communications*, 12(1):6036, 2021.
- [61] Matthias Fritsche, Pim Mostert, and Floris P de Lange. Opposite effects of recent history on perception and decision. *Current Biology*, 27(4):590–595, 2017.
- [62] Chris Oriet and John Brand. Size averaging of irrelevant stimuli cannot be prevented. *Vision Research*, 79:8–16, 2013.
- [63] Philip T Quinlan, Dale J Cohen, and Keith Allen. Perceptual averaging on relevant and irrelevant featural dimensions. *Attention, Perception, & Psychophysics*, 87(2):698–711, 2025.
- [64] Nilli Lavie. Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, 21(3):451, 1995.
- [65] Kimberly E Bodner, Nelson Cowan, and Shawn E Christ. Contributions of filtering and attentional allocation to working memory performance in individuals with autism spectrum disorder. *Journal of Abnormal Psychology*, 128(8):881, 2019.

- [66] Y-L Chien, SS-F Gau, C-Y Shang, Y-N Chiu, W-C Tsai, and Y-Y Wu. Visual memory and sustained attention impairment in youths with autism spectrum disorders. *Psychological medicine*, 45(11):2263–2273, 2015.
- [67] Yohannes Gebreegziabhere, Kassahun Habatmu, Awoke Mihretu, Matteo Cella, and Atalay Alem. Cognitive impairment in people with schizophrenia: an umbrella review. *European archives of psychiatry and clinical neuroscience*, 272(7):1139–1155, 2022.
- [68] Chen Song, Ryota Kanai, Stephen M Fleming, Rimona S Weil, D Samuel Schwarzkopf, and Geraint Rees. Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and cognition*, 20(4):1787–1792, 2011.
- [69] David M Greenberg, Varun Warriar, Carrie Allison, and Simon Baron-Cohen. Testing the empathizing–systemizing theory of sex differences and the extreme male brain theory of autism in half a million people. *Proceedings of the National Academy of Sciences*, 115(48):12152–12157, 2018.
- [70] David C Cicero, John G Kerns, and Denis M McCarthy. The aberrant salience inventory: a new measure of psychosis proneness. *Psychological assessment*, 22(3):688, 2010.
- [71] Nilli Lavie. Attention, distraction, and cognitive control under load. *Current directions in psychological science*, 19(3):143–148, 2010.

Supplementary Material

Rat study

Data included in the analyses

Standard task		
Rat ID	Trials	Sessions
KR2	56683	176
KR3	41075	123
KR4	36332	128
KR5	70649	224
KR6	79898	248

Table S1: Summary of total number of trials and sessions included in the standard task behavioral analyses. The trials refer to both single-stimulus and double-stimulus trials.

Control task			
Rat ID	Trials (Control)	Trials (Standard)	Sessions
KR2	59618	33952	119
KR3	16248	14502	43
KR4	26863	19873	69
KR6	23904	20856	64

Table S2: Summary of total number of trials and sessions included in the control task behavioral analyses. The trials refer to both single-stimulus and double-stimulus trials.

Proficient and expert stages

Rat ID	Trials (Proficient)	Sessions (Proficient)	Trials (Expert)	Sessions (Expert)
KR2	22731	57	33952	119
KR3	26573	80	14502	43
KR4	8295	31	19873	66
KR5	/	/	70649	224
KR6	7380	25	20856	223

Table S3: Summary of total number of trials and sessions included in the operational proficient and expert stages. The trials refer to both single-stimulus and double-stimulus trials.

Dynamic task

Rat ID	Trials	Sessions
KR7	28731	99
KR8	34283	95
MG1	10291	40
MG2	36214	135

Table S4: Summary of total number of trials and sessions included in the dynamic task behavioral analyses. The trials refer to both REL-IRR and IRR-REL trials

Electrophysiological recordings

Rat ID	Trials	Sessions	M1/M2 units	vS1 units
KR2	7595	40	209	/
KR3	7736	36	326	/
MG1	4871	28	313	276
KR8	3870	31	197	214

Table S5: Summary of total number of trials, sessions, and units included in the neural data analyses. The trials refer only to double-stimulus trials (KR2, KR3, MG1), and REL-IRR trials and IRR-REL trials (KR8).

Single-stimulus trials

The training protocol of the standard task and control task contained *single-stimulus trials*: within a training session under the standard task regime, approximately 20% of trials consisted of single-stimulus trials, in which only the relevant stimulus was delivered. When rats were trained under the control task, the single-stimulus trials were approximately 40%, within a training session.



Figure S1: Single-stimulus trials. **a.** Single-stimulus trials in the standard task. **b.** Single-stimulus trials in the control task.

Single-stimulus trials were introduced to ensure that rats stayed in the nosepoke also during the presentation of the first stimulus, and to ensure that they perceived the first stimulus (which corresponded to the irrelevant one, in the case of double-stimulus trials in standard task and control task).

Average performance was lower during single-stimulus trials, for both standard task (Figure S2) and control task (Figure S3). This can be explained by the fact that rats assessed relevance by stimulus position: they learned that the relevant stimulus is the second stimulus of a sequence of two, and when asked to judge the only stimulus of a trial, they might have mistaken REL for IRR.

Despite lower performance, the steepness of the psychometric curve suggests that rats judged REL even when presented alone, excluding the hypothesis that they entered the nosepoke and ignored all events for a fixed period of time.

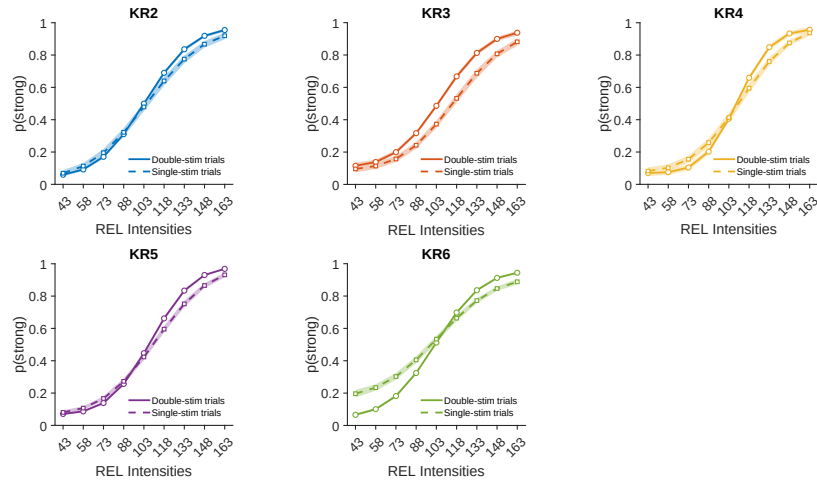


Figure S2: Single-stimulus trials. Single-stimulus trials in the standard task. Shaded areas represent the 95% bootstrapped confidence interval.

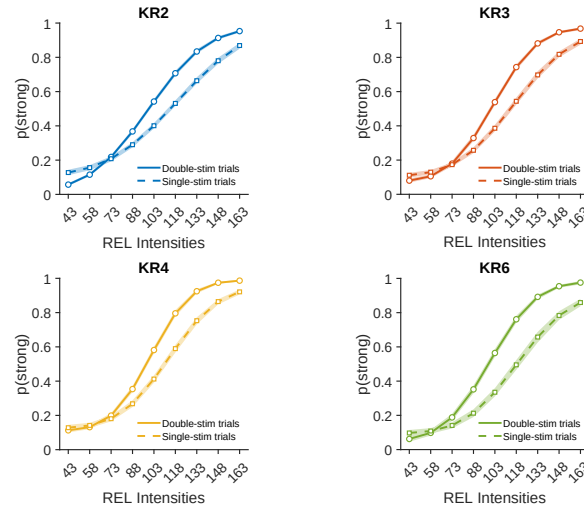


Figure S3: Single-stimulus trials. Single-stimulus trials in the control task. Shaded areas represent the 95% bootstrapped confidence interval.

Rats trained in the dynamic task do not improve over training

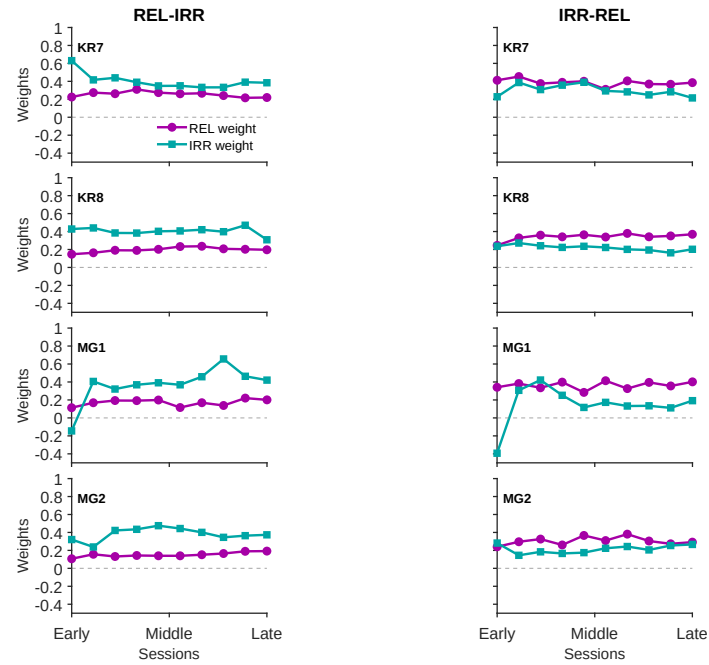


Figure S4: Learning trajectories (dynamic task). Contribution of REL and IRR stimuli to the choice, as described by the GLM weights, for REL-IRR trials (left) and IRR-REL trials (right). All training sessions were binned into successive groups of 10 sessions, and a GLM was computed for each group.

Linking neural activity to behavior in expert animals

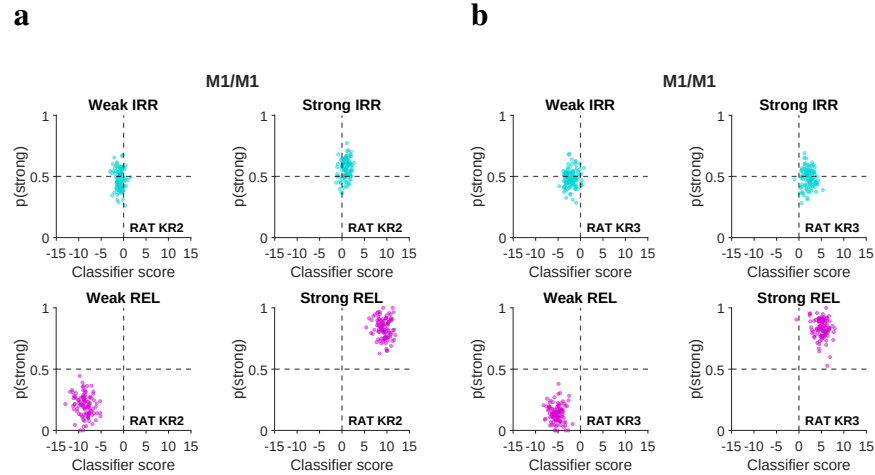


Figure S5: Relationship between decoder's performance and rat's choices, for rats KR2 and KR3. **a.** Decoder's scores plotted against behavioral choices, for each stimulus type, and for each intensity category. Each dot represents the average decoder's score and rat's choice for each iteration of the decoder (rat KR2). **b.** Same, for rat KR3.

Time-frequency domain Granger causality

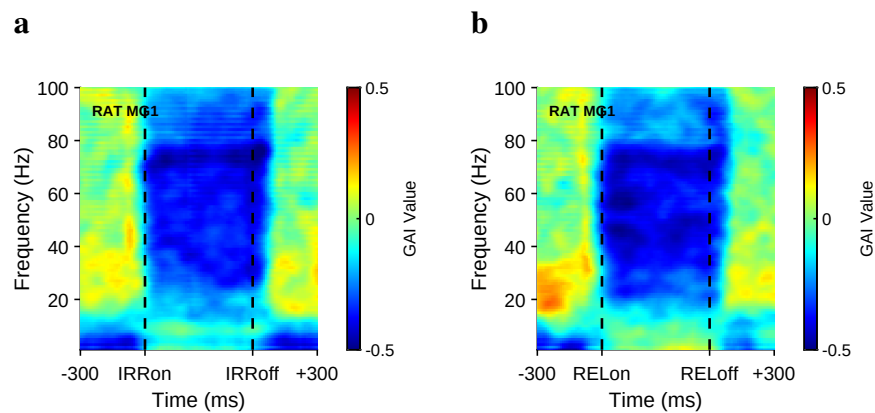


Figure S6: Granger Causality between vS1 and M1/M2, for both REL and IRR epochs, for rat MG1. **a.** Time-frequency domain Granger Causality during IRR stimulus epoch. **b.** Same, for REL stimulus epoch.

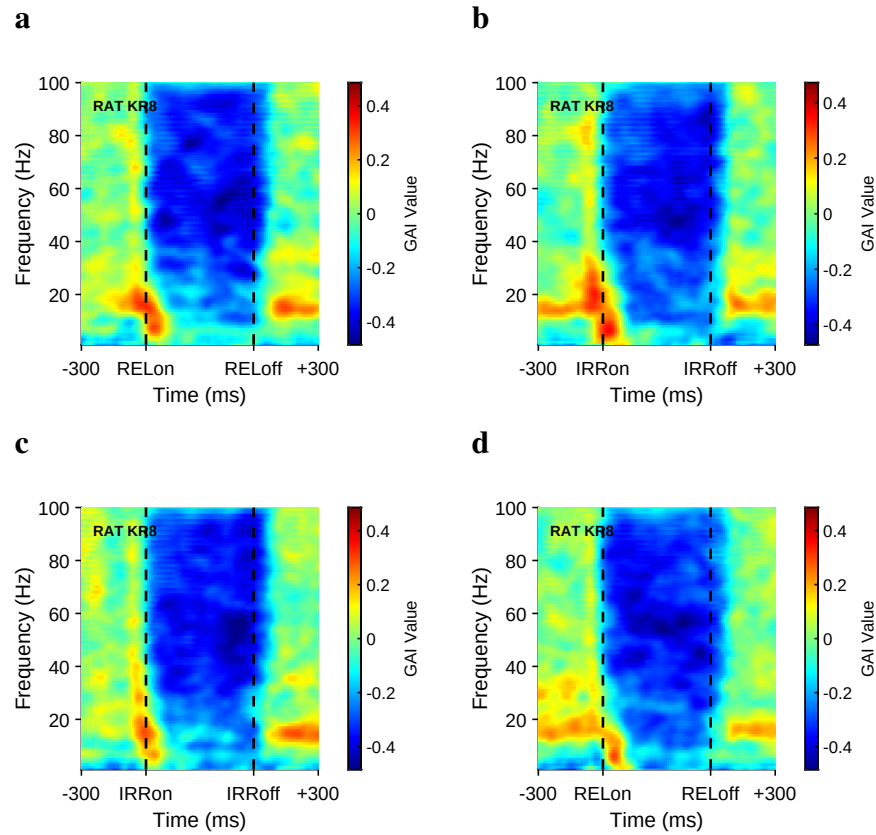


Figure S7: Granger Causality between vS1 and M1/M2, for both REL and IRR epochs, during REL-IRR and IRR-REL trials, for rat KR8. a. Time-frequency domain Granger Causality during REL stimulus epoch, for REL-IRR trials. **b.** Same, for IRR stimulus. **c.** Time-frequency domain Granger Causality during IRR stimulus epoch, for IRR-REL trials. **d.** Same, for REL stimulus.

Histology

The histological images were acquired using a Nikon Eclipse T-1 microscope mounting a Hamamatsu camera, using the software NIS-Elements.

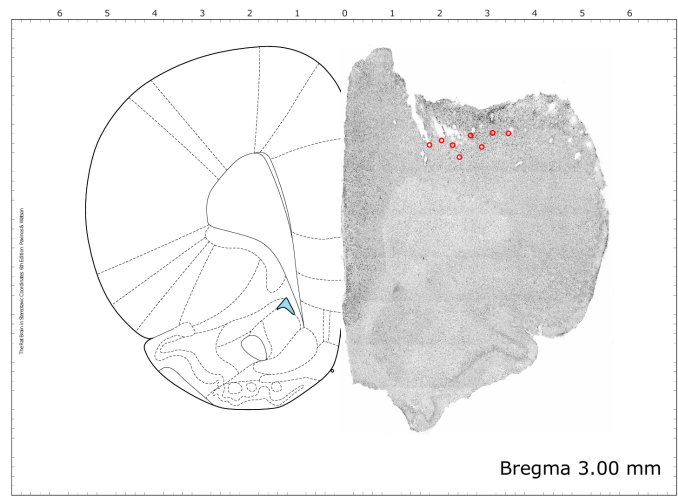


Figure S8: Reconstruction of microdrive traces on histological preparation, for rat KR2. Dots indicate microdrive traces.

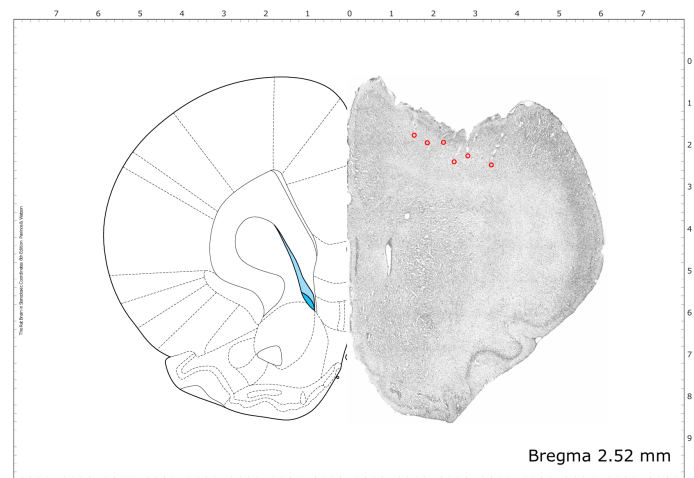


Figure S9: Reconstruction of microdrive traces on histological preparation, for rat KR3. Dots indicate microdrive traces.

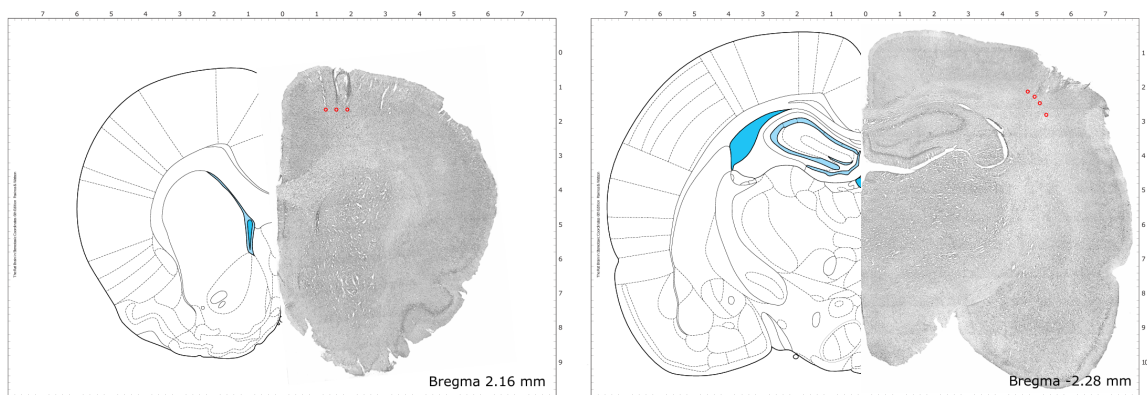


Figure S10: Reconstruction of microdrive traces on histological preparation, for rat MG1. Coronal section of the brain showing the microdrive traces in M1/M2 (left) and vS1 (right). Dots indicate microdrive traces.

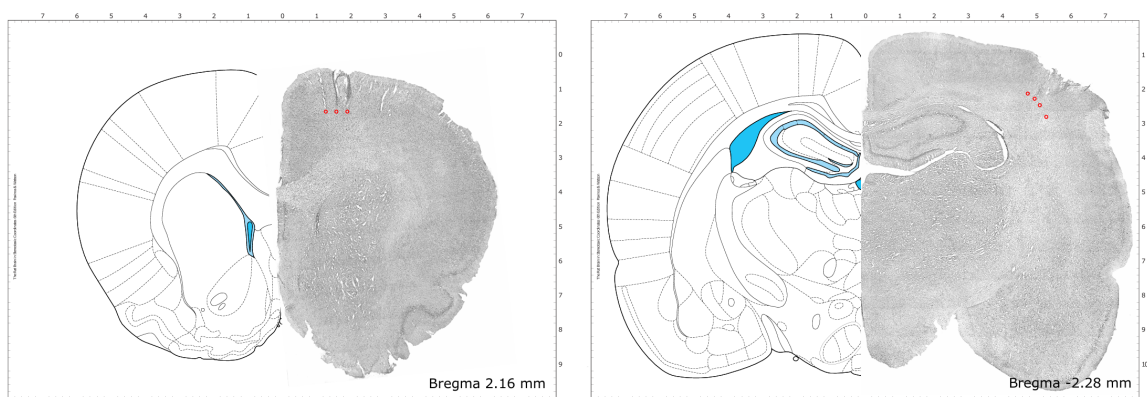


Figure S11: Reconstruction of microdrive traces on histological preparation, for rat KR8. Coronal section of the brain showing the microdrive traces in M1/M2 (left) and vS1 (right). Dots indicate microdrive traces.

Human study

Learning trajectories: GLM estimates

Human participants' GLM estimates					
Participant	Cluster	REL-IRR β_{REL}	REL-IRR β_{IRR}	IRR-REL β_{REL}	IRR-REL β_{IRR}
R1	1	0.66	0.01	0.69	-0.03
R2	3	0.24	-0.37	0.40	-0.22
R3	1	0.54	0.00	0.65	-0.06
R4	1	0.68	0.02	0.69	0.03
R5	1	1.05	0.02	0.99	0.02
R6	2	0.65	0.03	0.87	-0.08
R7	2	0.53	0.05	0.56	-0.01
R8	2	0.53	0.05	0.56	-0.01
R9	1	1.00	0.03	0.94	0.06
R10	1	0.85	0.00	0.82	0.02
R11	2	0.42	-0.04	0.58	-0.06
R12	2	0.76	0.01	0.75	-0.06
R13	3	0.20	0.26	0.60	-0.08
R14	1	0.07	0.29	0.34	0.11
R15	1	0.94	0.06	0.98	-0.01
R16	2	1.18	0.01	0.84	0.10
R17	2	0.73	-0.12	0.78	-0.04
R18	1	0.96	0.05	0.80	-0.01
R19	1	0.87	0.04	0.90	-0.04
R20	2	0.61	-0.02	0.72	-0.11
R21	1	0.64	0.01	0.72	-0.07
R22	3	0.29	-0.20	0.57	-0.14
R23	3	0.90	-0.20	1.12	-0.23
R24	2	0.74	0.00	0.74	0.04
R25	2	0.69	0.01	0.88	-0.04
R26	1	0.79	-0.02	0.80	0.11
R27	1	1.27	0.03	1.37	0.03
R28	2	0.33	0.00	0.38	-0.04
R29	1	0.61	-0.41	0.78	-0.15
R30	1	0.89	0.02	0.74	0.06

Table S6: Participant-level GLM coefficients. Values are the mean coefficients across experimental blocks for each participant.

Cluster number and stability assessment

We performed a silhouette analysis to determine the appropriate k for the clustering analysis.

Next, we assessed the stability of such clusters by computing the ARI (Adjusted Rand Index).

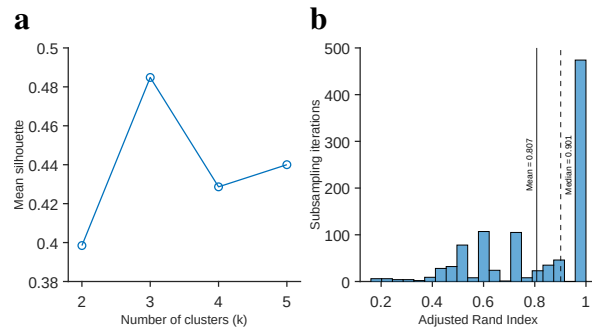


Figure S12: Supporting analyses for k-medoids clustering. a. Silhouette analysis. **b.** ARI stability assessment.

Final questionnaire

Please tick one option per question only

		<i>Agree</i>	<i>Neutral</i>	<i>Disagree</i>
1	The task was difficult.			
2	I was competent in performing the task.			
3	During the trials, I became better at evaluating the stimuli.			
4	The distractor stimulus is easy to ignore.			
5	When the stimulus to process was presented first, the task seemed easier.			
6	When the stimulus to process was presented second, the task seemed easier.			
7	I was more influenced by the distractor when it was in the same intensity direction as the stimulus to process (e.g., weak-weak / strong-strong).			
8	I was more influenced by the distractor when it was in the opposite intensity direction to the stimulus to process (e.g., weak-strong / strong-weak).			
9	In my mind, the reference stimulus intensity was clear when comparing it to the new stimuli.			

Figure S13: Final questionnaire administered to human participants.
 The participants received the Italian version of the questionnaire, since they were all native Italian speakers.