

Contrasting signatures of visual attention emerge from the same underlying neural dynamics

Raul Grieben

`raul.grieben@rub.de`

Ruhr University Bochum <https://orcid.org/0000-0003-1718-7679>

John Spencer

University of East Anglia <https://orcid.org/0000-0002-7320-144X>

Gregor Schöner

Ruhr University Bochum

Article

Keywords:

Posted Date: October 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11098497/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

**Contrasting signatures of visual attention emerge from the same underlying
neural dynamics**

Raul Grieben¹, John P. Spencer², and Gregor Schöner¹

¹Faculty of Computer Science, Institute for Neural Computation, Ruhr University Bochum,
Bochum, Germany

²School of Psychology, University of East Anglia, Norwich, United Kingdom

Author Note

Corresponding author: Raul Grieben, Faculty of Computer Science, Institute for
Neural Computation, Ruhr University Bochum, Bochum, Germany. Email:
`raul.grieben@rub.de`.

Abstract

Behavior is a window into cognition, yet behavioral effects rarely map one-to-one onto mechanisms: different mechanisms can produce the same effect, and the same mechanism can produce opposing effects across contexts. Visual attention provides a test case. We introduce the *Shaped Priority Account*, a neural-dynamical model in which salience and target guidance are nonlinearly transformed, unfold on different timescales, and compete within a recurrent selection field. With one fixed parameterization, the model reproduced 34 conditions from 12 experiments spanning seemingly distinct phenomena including triple-conjunction search, task-irrelevant feature effects, attentional capture, resistance to capture, and distractor-present benefits. These results show that apparently distinct behavioral signatures can emerge from the same neural dynamics rather than from distinct mechanisms. Behavioral effects are therefore not mechanistic fingerprints: before invoking distinct processes for each effect, researchers must pursue a benchmark, interrogating whether a neural process account can unify such effects under a single dynamical theory.

Contrasting signatures of visual attention emerge from the same underlying neural dynamics

A central ambition of cognitive neuroscience is to infer the brain's hidden computations from their observable consequences in behavior. Yet this enterprise often rests on a fragile assumption: that a behavioral signature uniquely identifies the mechanism invoked to explain it. This problem becomes especially acute when subtle changes in sensory input or experimental context produce seemingly contradictory effects. Such reversals are often interpreted as evidence for specialized mechanisms, latent control states or qualitative shifts in strategy. But introducing new mechanisms to accommodate each new effect risks replacing explanation with taxonomy. A more parsimonious, though more demanding, possibility is that apparently distinct behavioral signatures emerge from a common neural dynamical system.

Behavioral signatures provide essential empirical constraints, but they are not themselves mechanisms. They are compressed summaries of neural processes unfolding over time in different situations. Different neural trajectories may converge on the same behavioral mean, while the same underlying dynamics may produce different outcomes as sensory input changes. Under natural conditions, the brain is neither given the identity of an experimental condition nor instructed which control state or mechanism to deploy. Neural activity instead evolves in response to sensory input within a shared dynamical architecture, from which behavior emerges. Rather than assigning distinct explanations to each behavioral signature, the theoretical challenge is therefore to identify the minimal neural dynamics capable of generating them all, including outcomes that appear contradictory. This view offers a route toward theories that are parsimonious, mechanistically unified and able to generalize across contexts.

Visual attention provides a stringent test of this possibility. Despite decades of research under tightly controlled conditions, findings from visual search remain difficult to reconcile within a unified theoretical framework. Small changes in stimulus features or

display configuration can produce markedly different behavioral effects. A salient but task-irrelevant distractor can slow responses to a target (Theeuwes, 1992), produce little measurable cost (Bacon & Egeth, 1994), or even be associated with faster responses (Gaspelin & Luck, 2018; Gaspelin et al., 2015; Lien et al., 2022). Targets defined by multiple features can sometimes be found efficiently even when surrounding objects share many of those features, and search difficulty varies systematically with target-to-distractor similarity and distractor arrangement (Nordfang & Wolfe, 2014). Features irrelevant to the target definition can likewise alter performance by changing local contrast or display organization (Found, 1998; Proulx, 2007). These findings have motivated a wide range of proposed mechanisms, including distractor suppression (Gaspelin et al., 2015, 2025), distinct attentional-control settings such as contingent orienting, singleton-detection and feature-search modes (Bacon & Egeth, 1994; Folk et al., 1992), preattentive binding and grouping (Duncan & Humphreys, 1989; Found, 1998), clump scanning and independent parallel search (Liesefeld & Müller, 2020; Theeuwes, 2023), dimension weighting (Liesefeld & Müller, 2019), attentional-window adjustment (Theeuwes, 2023), and selection-history or implicit-learning components (Anderson et al., 2021; Gaspelin et al., 2025). The central question is whether these diverse behavioral patterns require distinct control mechanisms, or whether they can instead arise from a common set of neural dynamics.

Priority-map theories provide a natural starting point for such a unified account. Visual salience, task goals, selection history, reward, emotion, statistical learning, social relevance and related observer states can all contribute to a common spatial priority representation that guides attentional selection (Anderson et al., 2021; Awh et al., 2012; Bisley & Mirpour, 2019; Luck et al., 2021; Niu et al., 2025; Ptak, 2012; Treisman & Gelade, 1980; Wolfe, 1994, 2021; Zelinsky & Bisley, 2015). Neural dynamical implementations of attentional priority have also been developed (Deco & Rolls, 2004; Grieben et al., 2020; Hamker, 2005, 2006; Wyble et al., 2020). Yet priority-map accounts leave open a critical mechanistic question: how signals contributing to priority interact and unfold through time

to produce attentional selection and its sometimes contrasting behavioral consequences.

We addressed this question by testing whether a common neural dynamical architecture could generate seemingly disparate behavioral effects across visual-search contexts. We began with a neural priority-map model containing only two core components of visual attention: bottom-up guidance by stimulus conspicuity (Itti & Koch, 2000) and top-down guidance by target features (Wolfe, 2021; Wolfe et al., 1989). The model contained no task-specific mechanisms or parameters. We introduced only neurally motivated operations needed to account for the demanding data of Nordfang and Wolfe (2014), which systematically manipulated bottom-up and top-down influences across multiple search conditions. We then asked how far the same architecture could generalize to other challenging and apparently contradictory findings without task-specific mechanisms or retuning.

This led to the *Shaped Priority Account* which we present here, a neural dynamical model of visual search grounded in Dynamic Field Theory (Schöner et al., 2016). The model derives bottom-up salience from local feature contrast and top-down guidance from target-defining feature values. Its only inputs are the search array, represented as a static image, and the target-defining features, represented as numerical values. Crucially, the model receives neither a label identifying the experimental paradigm nor any condition-specific or benchmark-specific parameterization.

The account rests on three neural principles. First, each source of attentional priority is transformed by a source-specific nonlinearity, such that equal changes in unshaped activation need not produce equal changes in effective input. Nonlinear transformations of neural input are common in neural models; here, we formalize a general shaping operation and specify its form separately for salience and target-based guidance. Second, different priority sources have distinct temporal profiles: salience becomes effective earlier, whereas target-based guidance becomes available later, consistent with existing theoretical and empirical proposals (Donk & Van Zoest, 2008; Wolfe, 2021). Third, priority

is not represented in a separate static map. Instead, the summed input from all sources continuously drives a selective neural dynamic field defined over visual space. Activity in this field evolves through the interaction of external input, recurrent excitation and inhibition, and noise, consistent with Dynamic Field Theory (Schöner et al., 2016). Selection occurs when these dynamics stabilize a localized activity peak. Its timing is constrained by empirical estimates of covert attentional shifts, on the order of 30–50 ms (Bergen & Julesz, 1983; Buschman & Miller, 2009; Kröse & Julesz, 1989). Because attentional selection unfolds rapidly, there can be intervals in which no peak stabilizes before the next selection cycle, producing failed stabilization. Reaction time, therefore, reflects the evolving trajectory of competition until the target is found, including peak formation, failed stabilization, target verification, release, and inhibition of return.

Of these three commitments, nonlinear shaping yields the strongest qualitative predictions because it changes how canonical behavioral signatures should be interpreted. The same salient singleton that would receive uniformly high priority under a linear account can produce a response cost, little or no cost, or even facilitation in the *Shaped Priority Account*, depending on the surrounding features, despite the same nonlinear transformation being applied in every case. The same principle applies to top-down guidance: redistributing the same set of feature values across distractors can produce shallow or progressively steeper search slopes. Yet nonlinear guidance cannot eliminate all ambiguity. Items that closely resemble the target can remain viable candidates until they are selected and verified. We refer to this limitation as the *guidance diagnosticity bottleneck*: guidance can constrain the set of candidate locations, but cannot by itself identify the target before selection. Search therefore proceeds through successive verification until the target is found.

We tested this account across multiple paradigms. A single parameterization was calibrated on the triple-conjunction conditions of Nordfang and Wolfe (2014) and then held fixed across the remaining task-irrelevant feature manipulations (Found, 1998; Proulx,

2007) and capture- and suppression-related paradigms (Bacon & Egeth, 1994; Gaspelin & Luck, 2018; Gaspelin et al., 2015; Lien et al., 2022; Theeuwes, 1992). Across 12 experiments and 34 empirical conditions, the model captured the defining structure of all three benchmark families of data. Supplementary analyses used a reduced diagnostic surrogate model to isolate the contributions of nonlinear shaping, guidance timing, stochastic competition and failed stabilization.

The resulting theoretical claim is constrained but consequential: canonical signatures of capture, suppression and efficient triple-conjunction search can emerge as distinct outcomes of a single evolving neural dynamical system, without condition-specific or benchmark-specific retuning or separate mechanisms for individual effects. The account therefore replaces one-signature–one-mechanism inference with a stronger novel benchmark demand: to specify the neural dynamics that generate the behavioural signatures evident in multiple data sets before postulating additional or special-purpose mechanisms.

The Shaped Priority Account

The *Shaped Priority Account* specifies how visual input, target goals and recurrent neural interactions jointly produce attentional selection. Priority is treated as a dynamic capacity rather than a fixed item value: a location has priority to the extent that its activity can stabilize against competing activity elsewhere. Saliency and target-feature guidance bias this process, but their effects depend on nonlinear shaping, temporal availability and the current state of the selection field.

Search emerges from the continuous evolution of coupled neural fields encoding visual features, attentional selection, the active target template and the match between selected information and the target. Selection occurs when recurrent excitation and inhibition transform biased activation into a self-stabilizing localized peak. Local excitation strengthens a candidate location, broader inhibition suppresses alternatives, and noise makes stabilization probabilistic. A salient distractor may therefore perturb competition without being selected, whereas a less conspicuous but target-compatible location may

stabilize selection.

Search proceeds through successive cycles of selection and verification. Once an attentional peak stabilizes, activation from that location is gated into match fields and compared with the target template. A successful match terminates the trial; otherwise, the peak is released and an inhibitory trace reduces the likelihood of immediate reselection. Neural oscillators regulate these cycles, and some end before any usable selection emerges, producing selection failures that add processing time without selecting an item.

The term *shaped* refers to how salience and target-feature guidance become effective. Each source undergoes a source-specific nonlinear transformation before entering the attentional selection field, such that activation may have little effect, fall within a sensitive range or approach saturation. The shaped inputs also become available on different time courses, with salience becoming effective earlier than guidance.

Figure 1 summarizes the architecture. Visual input generates salience- and guidance-related signals that are shaped before entering the attentional selection field, where recurrent competition, verification and oscillatory modulation determine the trajectory toward selection, rejection or failed stabilization.

Feature Representations for Salience and Guidance

Feature maps

The visual display elicits activation profiles in neural fields, $F_f(x, y, z, t)$, defined jointly over visual space, (x, y) , and feature value, z , with field activation evolving over time, t . Separate fields encode color, orientation, size and shape, indexed by feature family f . These fields represent distributed feature information in a coarse and uncertain form, rather than already identified objects. Identification becomes possible only when an attentional peak gates one location into attended-feature and match fields, preserving the distinction between distributed feature representations and attention-dependent identification (Treisman & Gelade, 1980). Every represented feature family can contribute to salience, whereas only target-relevant feature families contribute to guidance.

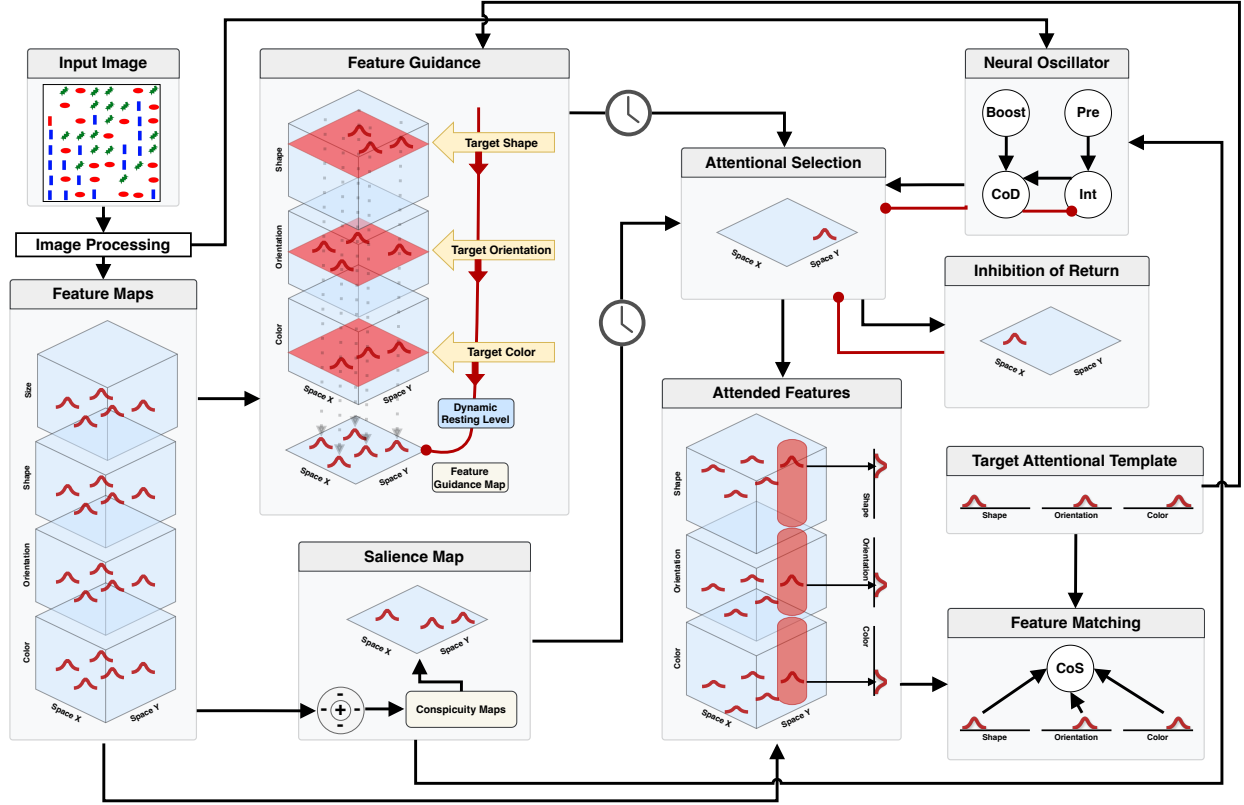


Figure 1

Architecture of the Shaped Priority Account. The visual display elicits spatially organized feature activation, which gives rise to salience- and guidance-related inputs. Nonlinear coupling shapes these inputs before they enter the attentional selection field, where recurrent interactions determine whether a localized peak stabilizes. Stabilized peaks are verified by match dynamics; failed matches release the peak and update inhibition of return, while oscillatory modulation creates successive opportunities for stabilization.

187 *Saliency map input*

188 Saliency-related activation emerges from local feature contrast within each feature
189 family. A center-surround projection produces local contrast (Itti & Koch, 2000):

$$C_f(x, y, t) = \text{ReLU} \left(\max_{z'} [F_f(x, y, z', t) * K] \right), \quad (1)$$

190 Here, $*$ denotes convolution over all non-temporal dimensions. The kernel K is a
191 Difference-of-Gaussians kernel, and $\text{ReLU}(v) = \max(v, 0)$ retains positive contrast. The
192 maximum over feature value z' then collapses the convolved feature field into the spatial

contrast map $C_f(x, y, t)$ for feature family f .

Contrast is summed across feature families, f , and scaled relative to the strongest location:

$$u_S(x, y, t) = \frac{\sum_f C_f(x, y, t)}{\max_{x', y'} \sum_f C_f(x', y', t)}. \quad (2)$$

The resulting unshaped salience, $u_S(x, y, t)$, is the normalized sum of positive local feature contrast across represented dimensions. Maximum scaling provides a compact approximation to divisive normalization (Carandini & Heeger, 2012).

Unshaped salience is a coarse retinotopic measure of conspicuity that depends on both an item’s features and the surrounding display configuration; it does not yet uniquely specify an attentional selection decision. If raw contrast magnitude were to influence selection proportionally, highly conspicuous items would dominate competition almost deterministically even when unrelated to the current search goal. Yet such locations should not be excluded from competition altogether: conspicuous events may carry behaviorally relevant information and therefore require access to attentional selection even when they are task-irrelevant in a particular search. Accordingly, $u_S(x, y, t)$ does not directly enter the attentional selection field but provides input to the salience map field, whose nonlinear sigmoid output bounds differences between strongly salient locations before entering the attentional selection field and competing with top-down feature guidance.

Feature guidance map input

Guidance-related activation, $G_f(x, y, z, t)$, emerges through coupling between the feature fields, $F_f(x, y, z, t)$, and the sustained target template, $T_f(z, t)$. Locations (x, y, z) in G_f that are consistent with the target value, t_f , in feature family f are active, whereas all other locations have zero activation (Wolfe, 1994, 2021). The unshaped feature guidance

$$u_G(x, y, t) = \sum_f \max_{z'} [G_f(x, y, z', t)]. \quad (3)$$

represents the raw summed target-feature overlap at each location. It increases linearly with the number of matching target features. This overlap does not imply that those features have already been bound into an identified object or that the location has been classified as the target. Rather, it indicates how strongly a region resembles the target description, while finer-grained object identity remains unresolved. This distinction is particularly important under natural viewing conditions in which noise, illumination changes, shadows and occlusion further reduce the diagnosticity of raw feature overlap. Accordingly, $u_G(x, y, t)$ does not directly enter the attentional selection field but provides input to the feature guidance map field, whose nonlinear sigmoid output brings the raw target-feature overlap into an operating range in which weak matches contribute little while strong and near-target matches remain competitive for selection and verification in the attentional selection field.

Priority Shaping as Nonlinear Field Output

Priority shaping determines how unshaped salience and feature guidance become effective inputs to attentional selection. Each source drives a source-specific neural map field whose sigmoid output transforms raw source activation before a delay node gates it into the attentional selection field. Thus, shaping determines which differences within each source remain consequential for recurrent competition rather than passing raw activation proportionally into selection. For compactness, this pathway can be summarized by a reduced input–output approximation that abstracts over the internal dynamics of the source-specific map fields:

$$Z_j(x, y, t) = a_j(t) w_j \frac{1}{1 + \exp[-\beta_j (\gamma_j q_j(x, y, t) - \theta_j)]}. \quad (4)$$

Here, $q_j(x, y, t)$ is unshaped source activation, $a_j(t)$ captures temporal gating, γ_j rescales source activation, θ_j locates the sensitive operating range, β_j controls its steepness, and w_j scales the resulting input to the attentional selection field.

Figure 2 illustrates how the parameters of this transformation determine which differences in raw source activation become effective for selection. The operating point determines which source values fall within the sensitive range of the sigmoid, gain rescales activation relative to that range, and slope controls how sharply differences are expressed. Activation below the sensitive range contributes little, activation within it is strongly differentiated, and sufficiently strong activation approaches saturation. Equal changes in raw source activation therefore need not produce equal changes in effective input.

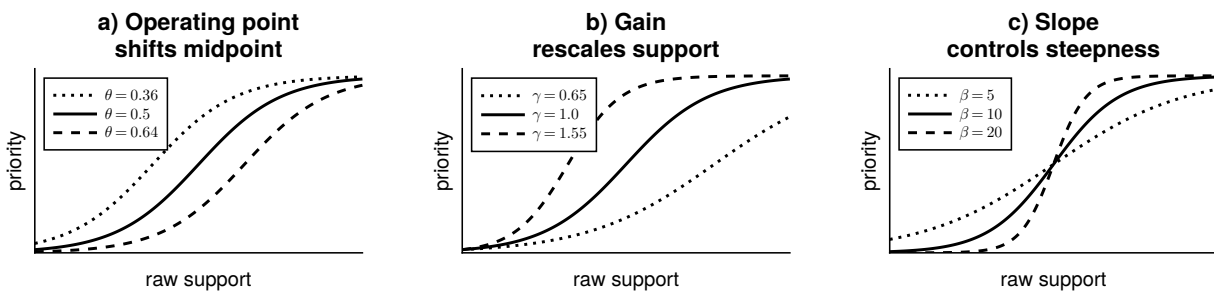


Figure 2

Generic properties of priority shaping. The operating point shifts the sensitive range of the transformation, gain rescales raw source support relative to that range, and slope determines how sharply differences within the range are expressed. Source activation may therefore have little downstream influence, fall within a sensitive regime, or approach saturation before entering attentional selection.

Shaping is mechanistically important because selection is highly sensitive to input geometry. Selection probability depends approximately linearly on the spatial support, or area, of a candidate but exponentially on its amplitude relative to competitors (see Methods and Extended Data Figure 2). Small amplitude differences can therefore make alternatives effectively unavailable. Nonlinear shaping places each source within an operating range in which it can bias recurrent competition without determining its outcome from raw source scale alone. Saliency and feature guidance instantiate this common principle differently.

254 *Shaped salience*

255 Unshaped salience is transformed as

$$Z_S(x, y, t) = a_S(t) w_S \frac{1}{1 + \exp[-\beta_S (\gamma_S u_S(x, y, t) - \theta_S)]}. \quad (5)$$

256 Salience shaping combines a fixed operating point, θ_S , with strong rescaling, γ_S , so that
 257 weak contrast can become competitive while differences among strongly salient locations
 258 are compressed by saturation. In the selection geometry described above, salience is
 259 therefore predominantly area-dominant: conspicuous items broaden the region of elevated
 260 input without acquiring a proportional height advantage. Strong local contrast can
 261 consequently influence where and when peaks stabilize without guaranteeing selection.

262 *Shaped feature guidance*

263 Unshaped feature guidance is transformed as

$$Z_G(x, y, t) = a_G(t) w_G \frac{1}{1 + \exp[-\beta_G (u_G(x, y, t) - \Theta_G)]}. \quad (6)$$

264 Unlike salience, guidance is not additionally rescaled by a separate γ_G term. Instead, its
 265 operating point depends on the number of target-defining feature families, n :

$$\Theta_G = \theta_G + n \kappa_G. \quad (7)$$

266 This template-dependent operating point maps raw target-feature overlap into the narrow
 267 amplitude range over which candidates remain competitively relevant (see Methods and
 268 Extended Data Figure 2). Weak matches receive little effective support, whereas strong
 269 and near-target matches remain competitive. Guidance is therefore predominantly
 270 amplitude-dominant: target-feature overlap produces differences in candidate heights that
 271 recurrent competition amplifies. This reduces competition from poor matches while
 272 preserving uncertainty among highly similar candidates. Under noise, occlusion or

illumination changes, such uncertainty may be beneficial because the target itself can be represented imperfectly and overly selective guidance could otherwise exclude it. We refer to this limitation as the *guidance diagnosticity bottleneck*: feature-based guidance can bias selection toward target-like locations but cannot reliably distinguish the target from sufficiently similar near-targets before attentional selection and verification.

Temporal Modulation and Attentional Peak Formation

Time-dependent coupling and oscillatory modulation

Temporal availability is implemented by delay nodes that gate the shaped outputs of the salience and guidance maps before they enter the attentional selection field. Salience-related input becomes effective before guidance-related input, moving the selection field from an early regime dominated more strongly by local feature contrast to a later regime increasingly shaped by the target template.

A neural oscillator gates selection-field excitability, alternating between windows that permit recurrent activation to form an attentional peak and windows that release it. Its timing follows empirical estimates of covert-attention shifts, approximately 30–50 ms (Bergen & Julesz, 1983; Buschman & Miller, 2009; Kröse & Julesz, 1989). The model captures this range with state-dependent rhythms: 20 Hz at lower and 33 Hz at higher feature-map activation. The oscillator does not select locations directly; its phase modulates whether recurrent excitation can overcome inhibition. Reaction time is measured from trial onset until a target-consistent attentional peak and match state stabilize.

293 *Attentional selection field*

294 The attentional selection field evolves according to

$$\begin{aligned}
 \tau \dot{u}(x, y, t) = & -u(x, y, t) + h + w_\xi \xi(x, y, t) \\
 & + Z_S(x, y, t) + Z_G(x, y, t) \\
 & - IOR(x, y, t) + Int(t) - CoD(t) \\
 & + \sigma(u(x, y, t)) * \omega.
 \end{aligned}
 \tag{8}$$

295 Here, $u(x, y, t)$ is the selection-field state, τ its time constant, h its resting level, and
 296 $w_\xi \xi(x, y, t)$ stochastic neural noise. The terms $Z_S(x, y, t)$ and $Z_G(x, y, t)$ are the shaped and
 297 temporally gated salience and guidance inputs. The remaining terms represent inhibition
 298 of return (IOR), oscillatory modulation (INT), reset-related input following an unsatisfied
 299 selection (CoD), and recurrent interactions combining local excitation with broader
 300 inhibition. Selection is not determined by the largest momentary input or by the pointwise
 301 sum of salience and guidance. A location is selected only when its activation recruits
 302 sufficient recurrent support to stabilize an above-threshold (>0) localized peak. Priority is
 303 therefore expressed as a location's capacity for stabilization within the evolving field state,
 304 rather than as a value contained in the shaped inputs themselves.

305 *Selection failures*

306 Oscillatory excitability can recede before any region acquires sufficient recurrent
 307 support to form a peak. This *selection failure* reflects unresolved field competition rather
 308 than an imposed lapse. It becomes more likely when candidates have similar activation,
 309 noise destabilizes an emerging peak, or the stabilization window is short, and it consumes
 310 time without selecting an item.

311 *Verification and inhibition of return*

312 A stable attentional peak gates activation at one location into attended-feature and
 313 feature-match fields. Target detection occurs only when the relevant match fields jointly

stabilize, generating a *Condition of Satisfaction* (CoS) that terminates the trial (Sandamirskaya & Schöner, 2010). If the selected feature pattern cannot sustain the required match, a *Condition of Dissatisfaction* (CoD) triggers reset dynamics that release the attentional peak. Verification is therefore itself a stabilization process within the coupled system rather than a symbolic comparison appended to selection. A failed verification leaves an inhibitory trace at the selected location (Klein, 1988), reducing its capacity to stabilize again. This self-sustained inhibition-of-return field provides within-trial memory without maintaining a symbolic record of previously inspected items.

Benchmark Evaluation of the Shaped Priority Account

We tested whether one fixed-parameter neural dynamic system could reproduce benchmark effects commonly attributed to distinct attentional mechanisms. Across simulations, only the reconstructed display and target template changed; equations and parameters were fixed, and the model received no information about paradigm or condition. The benchmark spanned 12 experiments across triple-conjunction search, task-irrelevant feature effects and additional-distractor paradigms. Evaluation focused on condition ordering, search slopes and distractor-present contrasts rather than absolute reaction time because modeled trajectories terminate at target-consistent match stabilization and omit detailed response, motor and participant-specific non-decision processes. Representative reconstructed displays are shown in Figure 3.

Benchmark families

Triple-conjunction search examined whether efficient search for targets defined by multiple features requires preattentive binding—the combination of features into object-level conjunctions before selective attention—or can instead arise from guidance by basic target features. The experiments of Nordfang and Wolfe (2014) varied how many of three target features each distractor shared and how distractor classes were spatially distributed. These displays require guidance to reduce support for weak matches while keeping near-target locations competitive until selection and verification.

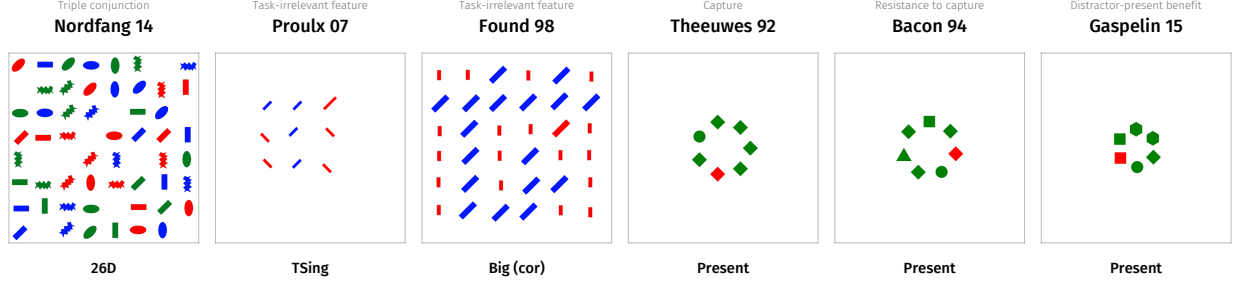


Figure 3

Representative reconstructed displays from the three benchmark families. From left to right: Nordfang and Wolfe (2014), triple conjunction (26D; red vertical target); Proulx (2007), task-irrelevant feature (TSing; red right-tilted target and size singleton); Found (1998), task-irrelevant feature (Big (cor); red right-tilted target in the size-correlated condition); Theeuwes (1992), capture (green-circle target; red-diamond color singleton); Bacon and Egeth (1994), resistance to capture (green-circle target among heterogeneous shapes; red-diamond color singleton); and Gaspelin et al. (2015), distractor-present benefit (fixed diamond- or circle-shaped target; red-square color singleton).

Task-irrelevant feature structure tested whether a feature outside the target template could nonetheless alter search. In Found (1998) and Proulx (2007), size did not define the target, but it changed local feature contrast and display organization. These studies therefore test whether salience-related input from an irrelevant dimension can reshape the selection trajectory within a concurrent guidance landscape.

Additional-distractor paradigms tested whether positive, near-zero and negative distractor effects could arise from the same dynamics. The benchmark set included classic capture costs, weak or absent costs, distractor-present advantages and nonsingleton distractor configurations (Bacon & Egeth, 1994; Gaspelin & Luck, 2018; Gaspelin et al., 2015; Lien et al., 2022; Theeuwes, 1992). For these studies, distractor effects were summarized as the difference between distractor-present and distractor-absent reaction times. Positive values indicate a distractor cost, values near zero indicate little mean difference, and negative values indicate a distractor-present advantage.

354 Fixed-parameter benchmark results

355 Figure 4 compares empirical and model-generated behavior across benchmark
 356 families. The Nordfang and Wolfe (2014) conditions were used for calibration; all

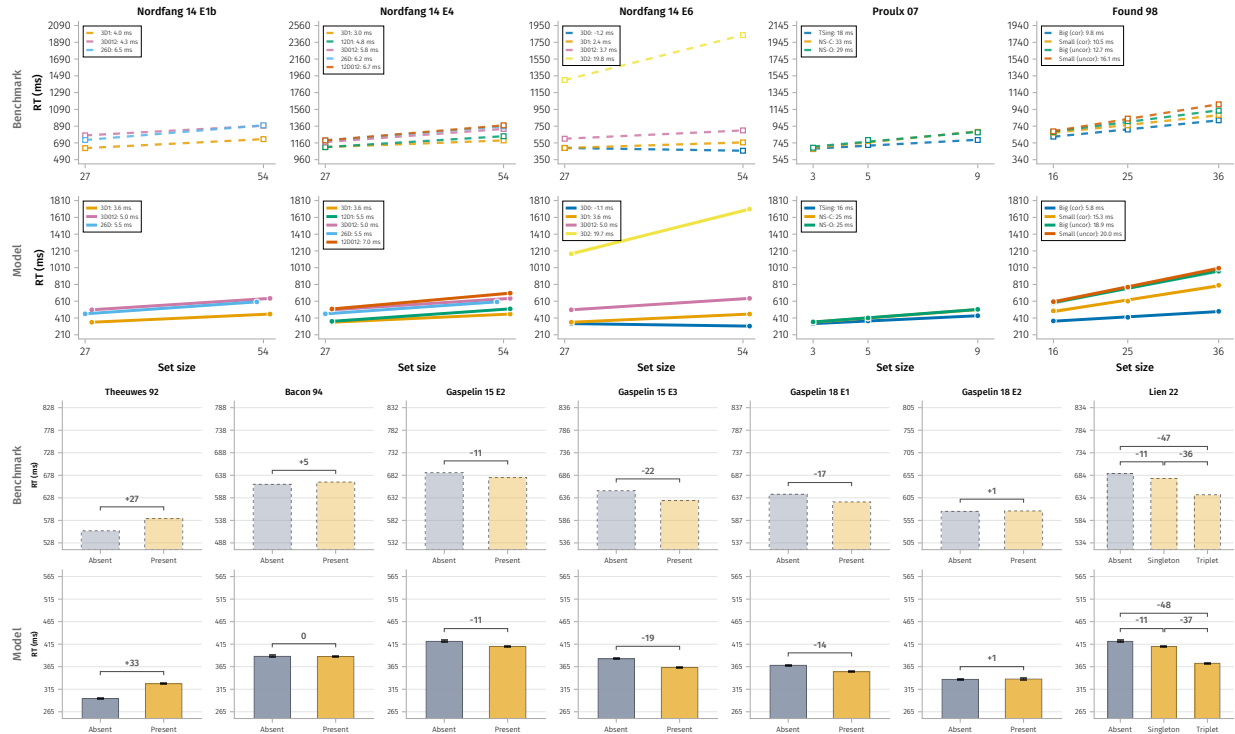


Figure 4

Fixed-parameter behavior of the Shaped Priority Account across benchmark conditions. Empirical and model-generated results are shown as set-size functions for search conditions (top) and distractor-present contrasts for additional-distractor conditions (bottom). The top panels show triple-conjunction and task-irrelevant feature conditions; the bottom panels show capture, no-capture, suppression and related distractor-interference conditions. Across 12 experiments, 34 reconstructed condition datasets comprised 60 empirical condition means and yielded 27 quantitative summaries (19 search slopes and 8 distractor-present contrasts); collapsing experimental replications left 21 unique benchmark and model conditions (seven per family). Each model cell comprised 300 reconstructed display arrays, each simulated five times with independent neural-noise seeds, yielding 1,500 stochastic trajectories per cell and 75,000 trajectories overall. 1.77% of trajectories failed to reach a target-consistent match within the simulation window and were excluded as error trials; remaining reaction times were averaged within display array before computing condition-level summaries. Shaded bands around model search regressions show 95% confidence intervals. Error bars on model reaction-time bars in the additional-distractor panels show the standard error of the mean across the 300 array-level means per cell; these quantify simulation variability across reconstructed displays, not uncertainty across human participants.

remaining conditions were generated with the same fixed equations and parameters. The model captured the principal structure of all three benchmark families: it reproduced the pattern of triple-conjunction search, including greater difficulty when all distractors shared two target features (condition 3D(2)); reproduced the direction of task-irrelevant size effects; and generated positive, near-zero and negative distractor-present contrasts across additional-distractor paradigms. Table 1 summarizes agreement between empirical and model-generated benchmark summaries, aggregated by behavioral summary type.

Table 1

Agreement between the Shaped Priority Account and empirical benchmarks. RMSE, R^2 and Pearson r are reported by benchmark family; NRMSE is computed separately for search slopes and distractor-present contrasts, with balanced NRMSE weighting these two summary classes equally.

Benchmark family	N	RMSE	R^2	Pearson r	NRMSE
Triple conjunction	12	0.8 ms/item	0.98	0.99	0.0909
Task-irrelevant feature	7	5.0 ms/item	0.65	0.81	
Additional-distractor	8	3.2 ms	0.98	0.99	0.0427
Balanced NRMSE	–	–	–	–	0.0668

Overall agreement was strong (balanced NRMSE = 0.0668), with near-perfect correspondence for triple-conjunction and additional-distractor benchmarks ($R^2 = .98$, $r = .99$ for both) and weaker but still substantial agreement for task-irrelevant feature effects ($R^2 = .65$, $r = .81$). More theoretically important, the same fixed system generated qualitatively opposing signatures without changes in equations, parameters or control states. Illustrative trials show how these contrasting outcomes emerge from different trajectories through the same underlying dynamics.

Illustrative within-trial dynamics

Figure 5 shows single stochastic trajectories for two contrasting benchmark conditions: a 3D(012) triple-conjunction display from Nordfang and Wolfe (2014) at set size 54, in which distractors share zero, one or two of the target’s three features, and a distractor-present suppression condition from Gaspelin et al. (2015) Experiment 2,

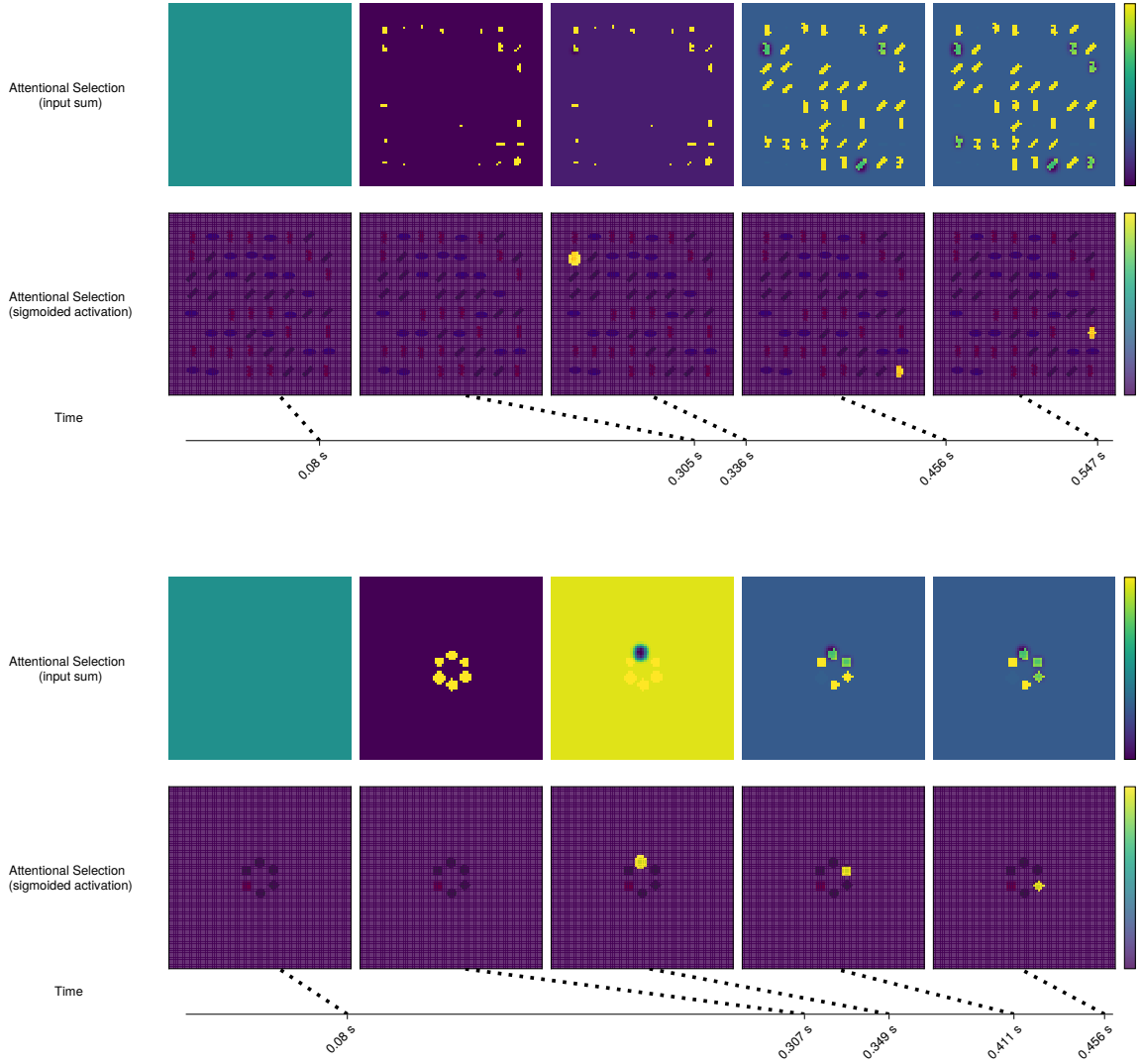


Figure 5

Illustrative single stochastic trajectories through the fixed neural dynamic system. Top: 3D(012) triple-conjunction search at set size 54. Bottom: a distractor-present trial from Gaspelin et al. (2015) Experiment 2. Five successive snapshots are shown for each condition. Upper rows show the momentary input to the attentional selection field; lower rows show the corresponding search display, with a localized activation peak marking the currently selected location. Column 1 precedes the arrival of effective salience input. In column 2, shaped bottom-up salience is the sole input to selection, producing a bottom-up selection in column 3 before target guidance is active. By column 4, delayed shaped guidance combines with salience, reducing support for locations with no target-feature overlap while retaining near-target locations as viable competitors. Subsequent stochastic selections continue until the target is found. Activation is shown using the viridis scale, from dark purple/blue (low activation), through green, to yellow (high activation). The trajectories are individual stochastic realizations rather than trial averages; repeated simulations of the same display can therefore yield different selection sequences and reaction times.

containing a salient red-square singleton that shares no target features. The trajectories are individual realizations, not trial averages. Because behavior emerges from noisy neural dynamics rather than a deterministic stimulus-to-reaction-time mapping learned through trialwise error minimization, repeated presentation of the same display can yield different selection sequences and reaction times.

For each condition, five snapshots show the evolving input to the attentional selection field (upper row) and the selected location as a localized peak on the search display (lower row). In column 1, the field receives no effective input because shaped salience is not yet available; search begins when salience arrives. In column 2, shaped bottom-up salience is the sole input, producing a purely bottom-up selection in column 3 before target guidance becomes active. By column 4, delayed shaped guidance combines with salience and reduces support for locations sharing no target features—the blue horizontal ellipses in 3D(012) and the red-square singleton in the suppression condition—while near-target distractors remain viable. Subsequent stochastic selections, biased by this evolving input landscape, continue until the target is selected and verified. Thus, the trajectories differ in their inputs and realized selection histories but follow the same dynamics: early salience-driven selection, delayed guidance and stochastic competition among remaining candidates. Nothing in the process identifies or invokes a conjunction-search- or suppression-specific mechanism; the contrasting outcomes emerge from the same fixed system operating on different inputs.

Discussion

Behavioral signatures are often used to infer the mechanisms that generated them. The present results show why this inference can be misleading. Capture-like costs, resistance to distraction, distractor-present advantages and efficient conjunction search all emerged from the same fixed neural dynamics. In the *Shaped Priority Account*, these effects arise as different trajectories through a shared recurrent selection process as visual input and target information change, rather than from switches between paradigm-specific

mechanisms.

Across 12 experiments spanning triple-conjunction search (Nordfang & Wolfe, 2014), task-irrelevant feature effects (Found, 1998; Proulx, 2007) and additional-distractor paradigms (Bacon & Egeth, 1994; Gaspelin & Luck, 2018; Gaspelin et al., 2015; Lien et al., 2022; Theeuwes, 1992), the model generated these contrasting signatures with one fixed parameterization. After calibration on the Nordfang and Wolfe (2014) conditions, only the visual display and target definition changed. Illustrative single-trial trajectories showed the common process underlying these outcomes: salience first biases competition, delayed target guidance then constrains the viable candidates, and stochastic selection continues until the target is verified. Reaction times, search slopes and distractor contrasts therefore summarize distributions of underlying neural trajectories rather than uniquely identifying the mechanisms that generated them.

The claim is one of sufficiency rather than uniqueness. Other architectures may satisfy the same constraints, and additional processes may be required in tasks designed to isolate learning, history or strategic control. The present results nevertheless shift the explanatory burden: before inferring a distinct mechanism from a behavioral signature, researchers must pursue a new benchmark — to test whether that signature can emerge from the same dynamics that explain related, and even opposing, effects. More broadly, behavioral diversity need not imply mechanistic diversity.

The additional-distractor results show how opposing behavioral effects can arise from the same process. Shaped salience biases early competition, but a salient item is selected only if activity at its location stabilizes before competing locations do. Because shaped target guidance becomes effective later, a task-irrelevant distractor can influence this early competition even when it shares no target features. Once guidance becomes available, support shifts toward target-compatible locations and the distractor becomes less competitive.

A distractor can nevertheless influence behavior without itself being selected. By

altering local contrast, it reshapes the input to competition and can change when the target is found. The same salient singleton can therefore produce a cost, little measurable effect or a benefit depending on the surrounding display. Positive, near-zero and negative distractor effects are thus compatible with different trajectories of the same dynamics and do not, by themselves, uniquely diagnose capture, absence of processing or suppression.

Triple-conjunction search exposes a complementary limitation of target guidance. Guidance reduces competition from poor matches but cannot necessarily distinguish the target from highly similar distractors before selection. These near-target candidates can therefore remain viable until they are selected and verified. This *guidance diagnosticity bottleneck* permits efficient search while preserving uncertainty among similar items, without requiring object-level identification of all items before selection. It also allows shallow, steep and even negative search slopes to emerge from the same architecture as target similarity and display structure change. Such slopes therefore need not imply distinct serial or parallel search modes.

Task-irrelevant feature effects arise through the same coupled dynamics. Features outside the target definition can alter local contrast and hence salience without contributing to target guidance. Their behavioral consequences nevertheless depend on the guidance landscape and recurrent competition into which that salience enters. Behavioral effects therefore do not isolate bottom-up and top-down influences; their consequences emerge from how those influences interact over time.

Fixed-parameter generalization is informative because the benchmark families jointly constrain the same coupled dynamics rather than independent modules. All depend on shaped salience, shaped guidance, their relative timing and recurrent competition, although different conditions expose different vulnerabilities of this shared process. Additional-distractor effects were especially sensitive to the timing of salience relative to guidance, triple-conjunction search to guidance shaping and its target-dependent operating point, and high-competition search conditions to successful stabilization among viable

457 candidates. Generalization therefore required one dynamical configuration to satisfy these
 458 interacting constraints simultaneously.

459 To identify which commitments carried this generalization, we used a reduced
 460 diagnostic surrogate of the full neural model (see Methods). The surrogate closely
 461 reproduced the benchmark-level behavior and selection geometry of the full system while
 462 replacing its continuously evolving neural fields and coupled state variables with a
 463 tractable approximation. It was not fitted as an alternative account of the behavioral data;
 464 rather, it enabled targeted ablations and systematic sensitivity analyses that would be
 465 difficult to isolate and interpret in the full neural system.

466 The ablations revealed a clear structure. Allowing salience to enter selection
 467 without nonlinear shaping produced broad failures, while removing guidance shaping
 468 collapsed most benchmark outcomes. Eliminating the delay between salience and guidance
 469 selectively disrupted the additional-distractor results, destroying the characteristic pattern
 470 of positive, near-zero and negative contrasts while leaving most conjunction and
 471 task-irrelevant feature conditions intact. Fixing the guidance operating point instead
 472 primarily disrupted triple-conjunction search, where several conditions lost their
 473 characteristic set-size functions. Removing selection failures had a narrower effect,
 474 concentrated in high-competition conditions, including the high-overlap Nordfang and
 475 Wolfe (2014) 3D(2) condition and several task-irrelevant feature conditions.

476 Sensitivity and stress tests converged on the same conclusion. Error concentrated
 477 around guidance shaping, its target-dependent operating point, guidance timing, salience
 478 shaping and competitive selection rather than being distributed diffusely across many
 479 parameters. The model therefore generalized not because different parameters could
 480 compensate freely, but because the same small set of dynamical commitments had to
 481 satisfy multiple behavioral constraints simultaneously.

482 The *Shaped Priority Account* builds on central ideas from Feature Integration
 483 Theory (Treisman & Gelade, 1980), Guided Search (Wolfe, 2021; Wolfe et al., 1989),

saliency models (Itti & Koch, 2000) and biased-competition accounts (Desimone & Duncan, 1995). Across these traditions, visual features, stimulus-driven signals and task goals have been used to explain how candidate locations become prioritized for selection. The present account specifies how these influences become effective within one neural process: salience and target guidance are nonlinearly shaped, become available on different time courses and interact through recurrent competition. Priority is therefore not a fixed score attached to an item, but a location’s capacity to stabilize selection within an evolving competitive field. The behavioral consequence of salience or target match depends not only on its raw magnitude, but on how it is shaped, when it becomes effective and the current competitive state.

Among closely related neurally inspired process models, RAGNAROC provides an informative comparison (Tam et al., 2022; Wyble et al., 2020). RAGNAROC has accounted for a subset of the phenomena included in our benchmark set, including attentional capture in Theeuwes (1992) and suppression-related findings from Gaspelin and Luck (2018) and Gaspelin et al. (2015). These simulations specify salience and task relevance at the experiment level. The *Shaped Priority Account* instead derives salience from the visual display and guidance from the active target template, allowing reaction time to emerge through successive stochastic selections. This enables a distinct generalization test: whether one fixed neural architecture can generate these effects together with qualitatively different search phenomena from changes in visual input and target definition alone.

Our conclusions are bounded by the scope of the present simulations. Displays were reconstructed rather than pixel-exact, some empirical values were digitized from published figures, and modeled trajectories terminate at target verification rather than including detailed response and motor processes. Absolute reaction time was therefore not the primary target of evaluation. The model also omits learning, reward and selection history, strategic control settings and explicit feature- or location-specific suppression. These omissions define the present sufficiency test rather than claims that such processes do not

exist. Future tests against neural time courses and manipulations of learning, history and strategic control will determine where the shared architecture is sufficient and where additional mechanisms are required.

In summary, canonical visual-search signatures are not mechanistic fingerprints. Capture-like costs, resistance to distraction, distractor-present advantages, task-irrelevant feature effects and efficient triple-conjunction search can arise from the same neural dynamics. More broadly, behavioral effects should be treated as constraints on mechanism rather than direct readouts of it. The theoretical challenge is to determine how much behavioral diversity can be generated by a common dynamical architecture before that architecture must be extended.

Methods

The Shaped Priority Account Model

The *Shaped Priority Account* was implemented as a coupled neural dynamic system in the Cedar framework (Lomp et al., 2016). The equations below specify its neural fields, nodes and couplings. Model activation and spatial quantities are expressed in dimensionless model units; time is expressed in milliseconds. Fields that differ only in feature identity are represented as indexed families. Convolution is denoted by $*$ and is evaluated over all dimensions shared by a field and its kernel; inputs without a given field dimension are broadcast uniformly along that dimension. Unless stated otherwise, maxima and integrals are taken over the full indicated domains.

Calibration and Fixed-Parameter Evaluation. Calibration used the triple-conjunction conditions from Nordfang and Wolfe (2014). Field and coupling parameters were first set manually to establish the intended dynamical regimes of the architecture: stable feature activation at stimulated locations, formation and release of localized attentional peaks, reliable stabilization of target-match states, inhibition of return following rejected selections, and repeated oscillator-regulated opportunities for attentional stabilization. This step established the qualitative operation of the coupled neural fields

rather than fitting condition-level behavioral outcomes. Six parameters were then tuned by coarse and fine grid search: two parameters of the salience Difference-of-Gaussians operation, the salience-to-selection coupling weight, the guidance-output slope, and the resting level and noise amplitude of the attentional selection field. These fitted parameters are indicated in the model equations by bold boxed values. All equations and parameters were then held fixed.

Image preprocessing. The input image is zero-padded to 800×800 pixels. Let $sat_{in}(x, y) \in [0, 255]$ and $hue(x, y) \in [0, 360]$ denote the saturation and hue of the padded image, respectively, and let $sat_{st}(s, x, y) \in [0, 1]$ denote the saturation template for stimulus s . Define

$$H_{\theta}(v) = \mathbf{1}[v \geq \theta], \quad \text{ReLU}(v) = \max(v, 0). \quad (9)$$

Let $\langle \cdot, \cdot \rangle$ and $\|\cdot\|_2$ denote the L^2 inner product and norm over (x', y') . The normalized template correlation for stimulus s is

$$\rho_s(x, y) = \frac{\langle sat_{st}(s, x', y'), sat_{in}(x + x', y + y') \rangle}{\|sat_{st}(s, x', y')\|_2 \|sat_{in}(x + x', y + y')\|_2}. \quad (10)$$

The detected spatial support of stimulus s is defined as

$$b_s(x, y) = H_{0.1} \left(\frac{[H_{0.98}(\rho_s) * sat_{st}(s)](x, y)}{\max_{x', y'} [H_{0.98}(\rho_s) * sat_{st}(s)](x', y')} \right). \quad (11)$$

For $q \in \{\text{size, shape, orientation}\}$, the corresponding input map is

$$m_q(x, y, z) = \sum_{s: q(s)=z} b_s(x, y). \quad (12)$$

For integer color bin $z \in \{0, \dots, 9\}$, define

$$p_z(x, y) = H_{100}(sat_{in}(x, y)) \mathbf{1} \left[18z \leq \frac{hue(x, y)}{2} < 18(z + 1) \right]. \quad (13)$$

553 The color input map is then

$$m_{\text{color}}(x, y, z) = H_{1.5} \left(p_z(x, y) + \max_{z'} m_{\text{size}}(x, y, z') \right). \quad (14)$$

554 The four resulting feature maps are resized to $80 \times 80 \times 10$.

555 **Compact field notation.** The activation functions used throughout the model are

$$\begin{aligned} \text{Abs}_{\beta}(v) &= \frac{1}{2} \left(1 + \frac{\beta v}{1 + \beta |v|} \right), \\ \text{Exp}_{\beta}(v) &= \frac{1}{1 + e^{-\beta v}}, \\ \text{Hvd}(v) &= \mathbf{1}[v > 0]. \end{aligned} \quad (15)$$

556 For $\mathbf{r} \in \mathbb{R}^d$, define

$$\begin{aligned} \Phi_{\sigma}(\mathbf{r}) &= \exp \left[-\frac{1}{2} \sum_{i=1}^d \frac{r_i^2}{\sigma_i^2} \right], \\ \hat{\Phi}_{\sigma}(\mathbf{r}) &= \frac{\Phi_{\sigma}(\mathbf{r})}{(2\pi)^{d/2} \prod_i \sigma_i}. \end{aligned} \quad (16)$$

557 Thus, Φ denotes an unnormalized Gaussian and $\hat{\Phi}$ denotes the corresponding unit-integral

558 Gaussian. For compactness, define the kernels

$$K_A = 2\hat{\Phi}_{(1,1,0.01)} - 9\hat{\Phi}_{(3,3,3)}.$$

559

$$K_C = \boxed{10.2} \Phi_{(1,1,0.01)} - 0.9 \Phi_{(\boxed{14.8}, \boxed{14.8}, 0.01)}.$$

560

$$K_{F2} = \Phi_{(0.01,0.01)}.$$

561

$$K_D = \hat{\Phi}_{(1,1,0.01)}.$$

562

$$K_{D4} = 0.6\Phi_{(0.01,0.01)}.$$

563

$$K_I = 4.5\Phi_{(2,2)} - 1.5.$$

564

$$K_{EI} = 1.1\hat{\Phi}_{(1,1)}.$$

$$K_E = 1.5\Phi_{(0.5,0.5)}.$$

$$K_J = 2\hat{\Phi}_{0.01}.$$

$$K_S = 1.5\hat{\Phi}_{(2,2,0.01)} - 3\hat{\Phi}_{(3,3,1.2)}.$$

$$K_K = 1.5\hat{\Phi}_{0.01} - 0.7.$$

$$K_M = \hat{\Phi}_{0.5}.$$

Feature and salience maps. Let $\mathcal{Q} = \{\text{size, shape, orientation, color}\}$ denote the

represented feature dimensions, and let $\mathcal{F} = \{\text{shape, orientation, color}\}$ denote the

dimensions that can be part of the target template. For each $q \in \mathcal{Q}$, the feature field

$u_{A_q}(x, y, z, t)$ evolves according to

$$\begin{aligned} 5\dot{u}_{A_q}(x, y, z, t) = & -u_{A_q}(x, y, z, t) - 0.9 \\ & + \text{Abs}_{1000}(u_{A_q}(x, y, z, t)) * K_A \\ & + m_q(x, y, z). \end{aligned} \tag{17}$$

The field receives the preprocessed feature input $m_q(x, y, z)$, producing a spatially and

feature-value organized representation of the current display. Local feature contrast is then

computed across the feature fields:

$$C(x, y, t) = \sum_{q \in \mathcal{Q}} \text{ReLU} \left(\max_{z'} \left[\text{Abs}_{1000}(u_{A_q}(x, y, z', t)) * K_C \right] \right). \tag{18}$$

For each feature dimension, convolution with the center-surround kernel K_C emphasizes

local contrast. The maximum over feature value collapses the feature dimension, and

positive contrast is summed across represented dimensions to form the spatial contrast

579 signal $C(x, y, t)$. The salience field $u_{F2}(x, y, t)$ evolves as

$$\begin{aligned}
 5\dot{u}_{F2}(x, y, t) = & -u_{F2}(x, y, t) - 5 \\
 & + \text{Exp}_{10}(u_{F2}(x, y, t)) * K_{F2} \\
 & + 100 \frac{C(x, y, t)}{\max_{x', y'} C(x', y', t)}.
 \end{aligned} \tag{19}$$

580 The normalized contrast signal drives the salience field, whose recurrent interaction
 581 transforms raw local contrast into the shaped salience representation that can subsequently
 582 influence attentional selection. The salience-delay node $u_{DSM}(t)$ is defined by

$$\begin{aligned}
 100\dot{u}_{DSM}(t) = & -u_{DSM}(t) - 0.9 + 0.5 \text{Hvd}(u_{DSM}(t)) \\
 & + \max_{x', y'} \text{Exp}_{10}(u_{F2}(x', y', t)).
 \end{aligned} \tag{20}$$

583 This node delays the point at which shaped salience can influence the attentional selection
 584 field. The salience peak-detection node $u_{F4}(t)$ evolves according to

$$\begin{aligned}
 \dot{u}_{F4}(t) = & -u_{F4}(t) - 0.5 + 0.5 \text{Hvd}(u_{F4}(t)) \\
 & + \max_{x', y'} \text{Exp}_{10}(u_{F2}(x', y', t)) \\
 & \cdot \text{Hvd}(u_{DSM}(t)).
 \end{aligned} \tag{21}$$

585 The node becomes active when the salience field contains sufficiently strong activation after
 586 the salience delay has elapsed. It provides a discrete control signal indicating that shaped
 587 salience is available.

588 **Feature guidance.** For each target-relevant feature dimension $f \in \mathcal{F}$, the
 589 feature-guidance cue field $u_{D_f}(x, y, z, t)$ evolves as

$$\begin{aligned}
 5\dot{u}_{D_f}(x, y, z, t) = & -u_{D_f}(x, y, z, t) - 1.7 \\
 & + \text{Abs}_{100}(u_{D_f}(x, y, z, t)) * K_D \\
 & + \text{Abs}_{1000}(u_{A_f}(x, y, z, t)) \\
 & + \text{Abs}_{100}(u_{J_f}(z, t)).
 \end{aligned} \tag{22}$$

590 The field combines the current feature representation u_{A_f} with the active target-template
 591 field u_{J_f} . Locations whose represented feature value is compatible with the target receive
 592 convergent support and therefore contribute to target-based guidance. The spatial
 593 feature-guidance field $u_{D4}(x, y, t)$ evolves according to

$$\begin{aligned}
 5\dot{u}_{D4}(x, y, t) = & -u_{D4}(x, y, t) - 5 \\
 & + \text{Exp}_{9.75}(u_{D4}(x, y, t)) * K_{D4} \\
 & + \sum_{f \in \mathcal{F}} \max_{z'} \text{Abs}_{100}(u_{D_f}(x, y, z', t)) \\
 & - 0.4 \sum_{f \in \mathcal{F}} \text{Hvd}(u_{P_f}(t)) \\
 & + 5 \text{Hvd}(u_{J7}(t)).
 \end{aligned} \tag{23}$$

594 This field collapses target-feature support across feature values and dimensions into a
 595 spatial guidance representation. The term involving u_{P_f} shifts its operating point according
 596 to the active target definition, representing the target-dependent guidance shaping
 597 described in the main text. The guidance-delay node $u_{DFG}(t)$ evolves according to

$$\begin{aligned}
 139\dot{u}_{DFG}(t) = & -u_{DFG}(t) - 0.9 + 0.5 \text{Hvd}(u_{DFG}(t)) \\
 & + \max_{x', y'} \text{Exp}_{9.75}(u_{D4}(x', y', t)).
 \end{aligned} \tag{24}$$

598 The node produces the delayed onset of target guidance relative to salience. The guidance
 599 peak-detection node $u_{D5}(t)$ is defined by

$$\begin{aligned} \dot{u}_{D5}(t) = & -u_{D5}(t) - 0.5 + 0.5 \text{Hvd}(u_{D5}(t)) \\ & + \max_{x', y'} \text{Exp}_{9.75}(u_{D4}(x', y', t)) \\ & \cdot \text{Hvd}(u_{DFG}(t)) . \end{aligned} \quad (25)$$

600 The node becomes active when the guidance field contains sufficiently strong activation
 601 after the guidance delay has elapsed. It provides a discrete control signal indicating that
 602 shaped guidance is available.

603 **Attentional selection and inhibition of return.** The attentional selection field
 604 $u_I(x, y, t)$ evolves according to

$$\begin{aligned} 2\dot{u}_I(x, y, t) = & -u_I(x, y, t) \boxed{-9.9} + \boxed{0.165} \xi_I(x, y, t) \\ & + \text{Exp}_{100}(u_I(x, y, t)) * K_I \\ & + 10 \text{Exp}_{\boxed{9.75}}(u_{D4}(x, y, t)) \text{Hvd}(u_{DFG}(t)) \\ & + \boxed{0.2} \text{Exp}_{10}(u_{F2}(x, y, t)) \text{Hvd}(u_{DSM}(t)) \\ & + 10 \text{Abs}_{500}(u_{C5}(t)) \\ & - 4 \text{Exp}_{100}(u_E(x, y, t)) * K_{EI} \\ & - 10 \text{Hvd}(u_{D5}(t)) - 200 \text{Hvd}(u_{C2}(t)) \\ & - 0.2 \text{Hvd}(u_{F4}(t)) . \end{aligned} \quad (26)$$

605 This field is the locus of attentional competition. It combines recurrent interaction and
 606 neural noise with delayed shaped guidance and salience inputs, inhibitory input from the
 607 inhibition-of-return field, and oscillator- and reset-related control signals. A location is
 608 selected when these inputs and recurrent interactions stabilize a localized activation peak.

609 The inhibition-of-return field $u_E(x, y, t)$ evolves according to

$$\begin{aligned}
 10\dot{u}_E(x, y, t) = & -u_E(x, y, t) - 0.5 \\
 & + \text{Exp}_{100}(u_E(x, y, t)) * K_E \\
 & + \text{Exp}_{100}(u_I(x, y, t)) .
 \end{aligned} \tag{27}$$

610 Activation in the selection field drives a self-sustaining inhibitory trace at previously
 611 selected locations. This trace feeds back negatively onto the attentional selection field (u_I),
 612 reducing the probability that a rejected location is immediately selected again.

613 **Target template, attended features and matching.** For each target-relevant
 614 feature dimension $f \in \mathcal{F}$, let μ_f denote the target value. The target-template field $u_{J_f}(z, t)$
 615 evolves according to

$$\begin{aligned}
 \dot{u}_{J_f}(z, t) = & -u_{J_f}(z, t) - 1 \\
 & + \text{Abs}_{100}(u_{J_f}(z, t)) * K_J \\
 & + \Phi_{0.01}(z - \mu_f).
 \end{aligned} \tag{28}$$

616 The Gaussian input centered on μ_f establishes a sustained representation of the target
 617 value in feature dimension f . The corresponding template peak-detection node $u_{P_f}(t)$
 618 evolves as

$$\begin{aligned}
 \dot{u}_{P_f}(t) = & -u_{P_f}(t) - 0.5 + 0.5 \text{Hvd}(u_{P_f}(t)) \\
 & + \max_{z'} \text{Abs}_{100}(u_{J_f}(z', t)) .
 \end{aligned} \tag{29}$$

619 This node indicates whether a stable target-template representation is active in feature
 620 dimension f . Its state also contributes to the target-dependent operating point of the
 621 guidance field. The cue-detection node $u_{J7}(t)$ evolves according to

$$\begin{aligned}
 \dot{u}_{J7}(t) = & -u_{J7}(t) - 0.5 + 0.5 \text{Hvd}(u_{J7}(t)) \\
 & + \sum_{f \in \mathcal{F}} \max_{z'} \text{Abs}_{100}(u_{J_f}(z', t)) .
 \end{aligned} \tag{30}$$

622 This node pools activation across target-template dimensions and signals whether an active
 623 target cue is present. Its input to the guidance field gates guidance formation, such that
 624 guidance can evolve only when a target cue has been established. For each feature
 625 dimension, the spatially gated attended-feature field $u_{K_f^{3D}}(x, y, z, t)$ evolves according to

$$\begin{aligned}
 \dot{u}_{K_f^{3D}}(x, y, z, t) = & -u_{K_f^{3D}}(x, y, z, t) - 1.9 \\
 & + \text{Abs}_{100}(u_{K_f^{3D}}(x, y, z, t)) * K_S \\
 & + \text{Abs}_{1000}(u_{A_f}(x, y, z, t)) \\
 & + \text{Exp}_{100}(u_I(x, y, t)).
 \end{aligned} \tag{31}$$

626 The field combines distributed feature activation with the current attentional-selection
 627 peak. Their spatial convergence gates feature information from the selected location into
 628 the attended-feature pathway. The attended-feature representation is then collapsed over
 629 visual space into $u_{K_f^{1D}}(z, t)$:

$$\begin{aligned}
 \dot{u}_{K_f^{1D}}(z, t) = & -u_{K_f^{1D}}(z, t) - 1 + 0.01\xi_{K_f^{1D}}(z, t) \\
 & + \text{Abs}_{100}(u_{K_f^{1D}}(z, t)) * K_K \\
 & + \max_{x', y'} \text{Abs}_{100}(u_{K_f^{3D}}(x', y', z, t)).
 \end{aligned} \tag{32}$$

630 This field represents the feature value extracted from the currently selected location while
 631 discarding its spatial coordinates. The feature-match field $u_{M_f}(z, t)$ evolves according to

$$\begin{aligned}
 \dot{u}_{M_f}(z, t) = & -u_{M_f}(z, t) - 1.9 \\
 & + \text{Abs}_{100}(u_{M_f}(z, t)) * K_M \\
 & + \text{Abs}_{100}(u_{K_f^{1D}}(z, t)) \\
 & + \text{Abs}_{100}(u_{J_f}(z, t)).
 \end{aligned} \tag{33}$$

632 Convergent input from the attended-feature representation and target template produces
 633 activation when the selected feature value is compatible with the target value. The
 634 corresponding match-detection node $u_{Q_f}(t)$ evolves as

$$\begin{aligned} \dot{u}_{Q_f}(t) = & -u_{Q_f}(t) - 0.5 + 0.5 \text{Hvd}(u_{Q_f}(t)) \\ & + \max_{z'} \text{Abs}_{100}(u_{M_f}(z', t)). \end{aligned} \quad (34)$$

635 This node signals that a stable match has formed in feature dimension f . The *Condition of*
 636 *Satisfaction* node $u_{C7}(t)$ integrates match signals across target-relevant dimensions:

$$\begin{aligned} \dot{u}_{C7}(t) = & -u_{C7}(t) - 0.2 + 6 \text{Abs}_{500}(u_{C7}(t)) \\ & + \sum_{f \in \mathcal{F}} \text{Hvd}(u_{Q_f}(t)) \\ & - 0.9 \sum_{f \in \mathcal{F}} \text{Hvd}(u_{P_f}(t)). \end{aligned} \quad (35)$$

637 The node compares the number of stabilized feature matches with the number of active
 638 target-template dimensions. It becomes active when the selected location satisfies the
 639 target definition and provides the target-detection signal that terminates the trial.

640 **Neural oscillator.** The oscillator is implemented by four coupled control nodes: a
 641 *Boost* node that selects the fast oscillator regime, a *Precondition* node that signals the
 642 availability of salience, an *Intention* node that transiently increases excitability of the
 643 attentional selection field, and a slower *Condition of Dissatisfaction* (CoD) node that resets
 644 the selection field and terminates the current intention state. The Boost node, $u_{C1}(t)$,
 645 evolves according to

$$\begin{aligned} \dot{u}_{C1}(t) = & -u_{C1}(t) - 120 + 0.5 \text{Hvd}(u_{C1}(t)) \\ & + \iiint m_{\text{size}}(x', y', z') dx' dy' dz'. \end{aligned} \quad (36)$$

646 The node integrates total activation in the size input map. When this activation exceeds
 647 threshold, the node becomes active and shifts the oscillator into its faster regime. The
 648 Precondition node, $u_{C4}(t)$, evolves according to

$$\begin{aligned} 5\dot{u}_{C4}(t) = & -u_{C4}(t) - 0.9 + 0.5 \text{Hvd}(u_{C4}(t)) \\ & + \text{Hvd}(u_{F4}(t)). \end{aligned} \quad (37)$$

649 The node becomes active once shaped salience is available and provides the precondition for
 650 initiating an attentional selection cycle. The Intention node, $u_{C5}(t)$, evolves according to

$$\begin{aligned} 4\dot{u}_{C5}(t) = & -u_{C5}(t) - 10.5 + 11 \text{Hvd}(u_{C4}(t)) \\ & - 10 \text{Hvd}(u_{C2}(t)) \\ & - 10 \text{Abs}_{500}(u_{C7}(t)). \end{aligned} \quad (38)$$

651 Activation of the Precondition node drives the Intention node, which transiently boosts the
 652 attentional selection field and thereby opens a window in which a localized attentional
 653 peak can stabilize. The Intention state is suppressed by the CoD node and by successful
 654 target detection. The CoD node, $u_{C2}(t)$, evolves according to

$$\begin{aligned} 20\dot{u}_{C2}(t) = & -u_{C2}(t) - 1.86 + 0.5 \text{Hvd}(u_{C2}(t)) \\ & + 0.48 \text{Hvd}(u_{C1}(t)) \\ & + 2 \text{Abs}_{500}(u_{C5}(t)). \end{aligned} \quad (39)$$

655 The CoD node evolves on a slower timescale than the Intention node. Once active, it
 656 suppresses Intention and resets the attentional selection field, terminating the current
 657 selection opportunity. Boost input shifts the oscillator into its faster regime. Together, the
 658 four nodes generate repeated windows of attentional excitability at approximately 20 Hz in
 659 the slow regime and 33 Hz when Boost is active.

Benchmark Set

The benchmark set comprised 12 experiments spanning three families of visual-search findings: triple-conjunction search, task-irrelevant feature structure and additional-distractor paradigms. The experiments were Nordfang and Wolfe (2014) Experiments 1b, 4 and 6; Found (1998); Proulx (2007); Theeuwes (1992) Experiment 1A; Bacon and Egeth (1994) Experiment 3; Gaspelin et al. (2015) Experiments 2 and 3; Gaspelin and Luck (2018) Experiments 1 and 2; and Lien et al. (2022) Experiment 1. Together, these experiments yielded 34 reconstructed condition datasets and 27 quantitative benchmark summaries: 19 search slopes and 8 distractor-present contrasts.

Triple-conjunction benchmarks tested whether the model could reproduce efficient search and increased interference from near-target distractors without preattentive target identification. Task-irrelevant feature benchmarks tested whether features outside the target template could influence search through salience-related input. Additional-distractor benchmarks tested whether positive, near-zero and negative distractor-present effects could arise from the same fixed dynamics. Probe-report and electrophysiological measures from the source studies were not fitted.

Display Reconstruction

Displays were reconstructed from published methods, figures and design constraints. Reconstructions preserved set size, reported spatial organization, target–distractor relations, grouping structure and singleton or triplet structure. Each reconstructed display was provided to the model together with the target-defining features specified to observers in the corresponding experiment. The display determined the model’s bottom-up feature input, whereas the reported target features defined the active target template. No additional information about the experimental condition or paradigm was provided. Human values were taken from published tables when available and digitized from figures otherwise. Digitized values were used to recover condition means, set-size slopes and distractor-present contrasts when numerical source data were unavailable.

Trajectory Generation

The benchmark comprised 50 model cells: 35 set-size cells and 15 additional-distractor cells. For each cell, 300 reconstructed display arrays were each simulated five times with independent neural-noise seeds, yielding 1,500 stochastic trajectories per cell and 75,000 trajectories overall. Reaction time was defined as the time from trial onset until a target-consistent match state stabilized; repeated simulations of the same array could therefore produce different selection sequences and reaction times. Trajectories were retained for analysis when reaction time was below 4,500 ms in set-size studies or 1,000 ms in additional-distractor studies. These criteria excluded 1,328 of 75,000 trajectories (1.77%). Retained reaction times were averaged within display array before condition-level summaries were computed. No display array was entirely excluded, yielding 300 usable array-level means per cell. Among the 10,500 arrays in set-size studies, 0.02%, 0.11%, 1.10%, 9.36%, and 89.41% retained one, two, three, four, or five trajectories, respectively. Among the 4,500 arrays in additional-distractor studies, 0.04%, 1.49%, and 98.47% retained three, four, or five trajectories, respectively.

Cutoff-Sensitivity Analysis. A post hoc sensitivity analysis tested whether benchmark results depended on the reaction-time exclusion thresholds. Set-size cutoffs of 3,000, 3,500, 4,000, 4,500, and 4,750 ms were crossed with additional-distractor cutoffs of 750, 875, 1,000, 1,125, 1,250, and 1,500 ms, yielding 30 combinations around the baseline thresholds of 4,500 ms and 1,000 ms. The same trajectories and aggregation procedure were used throughout. Results were stable across this range. All 15,000 display arrays remained usable, and the signs of all search slopes and distractor contrasts were unchanged. Relative to baseline, the largest absolute changes were 0.41 ms/item for search slopes and 1.34 ms for distractor contrasts. Balanced NRMSE ranged from 0.06678 to 0.07007, compared with 0.06681 at baseline. Thus, the benchmark results were insensitive to reasonable variation in the reaction-time exclusion thresholds.

713 *Behavioral Summaries and Descriptive Metrics*

714 Model-generated reaction times were summarized using the behavioral measures
 715 reported for the corresponding empirical experiments. For set-size studies, condition means
 716 were computed at each set size, and search slopes were estimated by ordinary least squares
 717 using retained array-level reaction times across set sizes. Each array contributed equally to
 718 the regression. Slopes were expressed in milliseconds per item. For additional-distractor
 719 studies, mean reaction times were summarized by the distractor-present contrast

$$\Delta_{\text{dist}} = \overline{RT}_{\text{present}} - \overline{RT}_{\text{absent}}. \quad (40)$$

720 Positive values indicate a distractor cost, values near zero indicate little mean difference,
 721 and negative values indicate a distractor-present benefit. For each summary class k ,
 722 agreement between empirical and model-generated summaries was quantified by
 723 root-mean-square error,

$$\text{RMSE}_k = \sqrt{\frac{1}{n_k} \sum_{i=1}^{n_k} (\hat{y}_{ki} - y_{ki})^2}, \quad (41)$$

724 where y_{ki} is an empirical summary, $\hat{y}_{ki}$ is the corresponding model-generated summary, and
 725 n_k is the number of comparison pairs. RMSE was normalized by the empirical range:

$$\text{NRMSE}_k = \frac{\text{RMSE}_k}{\max_i(y_{ki}) - \min_i(y_{ki})}. \quad (42)$$

726 A balanced index weighted pooled search slopes and distractor-present contrasts equally:

$$\text{NRMSE}_{\text{balanced}} = \frac{1}{2} (\text{NRMSE}_{\text{slopes}} + \text{NRMSE}_{\text{contrasts}}). \quad (43)$$

727 The slope term pooled all 19 search comparison pairs, and the contrast term included all
 728 eight capture comparison pairs. Pearson's r assessed linear association between empirical

729 and model-generated summaries:

$$r_k = \frac{\sum_{i=1}^{n_k} (y_{ki} - \bar{y}_k) (\hat{y}_{ki} - \bar{\hat{y}}_k)}{\sqrt{\sum_{i=1}^{n_k} (y_{ki} - \bar{y}_k)^2} \sqrt{\sum_{i=1}^{n_k} (\hat{y}_{ki} - \bar{\hat{y}}_k)^2}}, \quad (44)$$

730 where $\bar{y}_k$ and $\bar{\hat{y}}_k$ are the means of the empirical and model-generated summaries within
 731 class k , respectively. Agreement in their numerical values was additionally assessed using
 732 the residual-based coefficient of determination, R_k^2 :

$$R_k^2 = 1 - \frac{\sum_{i=1}^{n_k} (\hat{y}_{ki} - y_{ki})^2}{\sum_{i=1}^{n_k} (y_{ki} - \bar{y}_k)^2}, \quad \bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} y_{ki}. \quad (45)$$

733 Reduced Diagnostic Surrogate Model

734 To identify which theoretical commitments supported the fixed-parameter
 735 benchmark performance, we derived a reduced diagnostic surrogate of the full neural model.
 736 The surrogate is a process model of the algorithmic operations that approximate the neural
 737 process model: it represents each oscillator-regulated selection episode using tractable
 738 episode-level computations while preserving the full model’s core operations, including
 739 shaped salience and guidance, their temporal asymmetry, stochastic competition, inhibition
 740 of return, selection failures and target verification. It replaces the continuously evolving
 741 neural fields and coupled state variables with an explicit sequence of computational steps,
 742 making targeted ablations and systematic sensitivity analyses feasible. The surrogate was
 743 parameterized against outputs of the full neural model, not independently fitted to the
 744 human benchmark data. Extended Data Figure 1 presents the surrogate as a flow diagram
 745 showing the sequence of operations from display input through feature processing, priority
 746 computation, competitive selection, verification and trial termination or continuation.

747 **Feature maps.** The reduced model represents each display as a set of spatial
 748 feature maps. Let Ω denote visual space, and let $F_f^{\text{red}}(x)$ denote the activation of feature
 749 family f at location $x \in \Omega$. For feature families that define the target, $G_f^{\text{red}}(x)$ denotes the

750 activation at x that matches the target’s value for family f .

751 **Unshaped salience and guidance.** Let $\mathcal{F}_S = \{\text{size, shape, orientation, color}\}$

752 denote the feature families contributing to salience, and let

753 $\mathcal{F}_G = \{\text{shape, orientation, color}\}$ denote the feature families available for target guidance.

754 Let $T \subseteq \mathcal{F}_G$ be the subset defining the current target, with $n = |T|$. For $x \in \Omega$, unshaped

755 salience is computed from local feature contrast. For each $f \in \mathcal{F}_S$,

$$C_f^{\text{red}}(x) = \text{ReLU} \left(\sum_{x' \in \Omega} K(x - x') F_f^{\text{red}}(x') \right), \quad (46)$$

756 where $\text{ReLU}(z) = \max(z, 0)$. Local contrast is computed with the center-surround kernel

$$K(u) = A_+ \exp \left[-\frac{\|u\|^2}{2\sigma_+^2} \right] - A_- \exp \left[-\frac{\|u\|^2}{2\sigma_-^2} \right]. \quad (47)$$

757 Contrast is summed across feature families and normalized by the strongest location:

$$u_S^{\text{red}}(x) = \frac{\sum_{f \in \mathcal{F}_S} C_f^{\text{red}}(x)}{\max_{x' \in \Omega} \sum_{f \in \mathcal{F}_S} C_f^{\text{red}}(x')}. \quad (48)$$

758 The resulting $u_S^{\text{red}}(x)$ is the surrogate’s unshaped salience signal. Unshaped feature

759 guidance is computed by summing target-matching activation across the feature families

760 defining the target:

$$u_G^{\text{red}}(x) = \sum_{f \in T} G_f^{\text{red}}(x). \quad (49)$$

761 The resulting $u_G^{\text{red}}(x)$ increases with the amount of target-feature overlap at location x .

762 **Episode timing.** Selection unfolds in discrete episodes. A display-level control

763 signal determines whether the trial uses the fast or slow oscillator regime:

$$\Delta = \begin{cases} \Delta_{\text{fast}}, & A_{\text{ctrl}} \geq \Theta_{\text{shape}}, \\ \Delta_{\text{slow}}, & A_{\text{ctrl}} < \Theta_{\text{shape}}. \end{cases} \quad (50)$$

764 Within a trial, Δ is fixed. Episode m is evaluated at its endpoint $m\Delta$. The reduced model
 765 preserves the temporal asymmetry between salience and feature guidance. Salience is
 766 available from the first selection episode, whereas feature guidance becomes available only
 767 after a guidance delay t_G :

$$a_{S,m}^{\text{red}} = 1, \quad a_{G,m}^{\text{red}} = \mathbf{1}\{m\Delta > t_G\}. \quad (51)$$

768 The number of purely salience-driven episodes is

$$M_{\text{sal}}(\Delta) = \left\lfloor \frac{t_G}{\Delta} \right\rfloor. \quad (52)$$

769 This gives one salience-driven episode in the slow regime and three in the fast regime.

770 **Shaped salience and shaped feature guidance.** The shaped salience is

$$Z_S^{\text{red}}(x, m) = a_{S,m}^{\text{red}} w_S \frac{1}{1 + \exp[-\beta_S (\gamma_S u_S^{\text{red}}(x) - \theta_S)]}. \quad (53)$$

771 The shaped feature guidance is

$$Z_G^{\text{red}}(x, m) = a_{G,m}^{\text{red}} w_G \frac{1}{1 + \exp[-\beta_G (u_G^{\text{red}}(x) - \Theta_G)]}. \quad (54)$$

772 The guidance operating point increases with the number of target-defining feature families:

$$\Theta_G = \theta_G + n\kappa_G. \quad (55)$$

773 This preserves the *guidance diagnosticity bottleneck* in the reduced model.

774 **Effective priority map.** The effective priority map is an episode-level
 775 approximation of the momentary competitive landscape in the full attentional selection
 776 field; it is not an additional neural representation in the Shaped Priority Account. Before

inhibition of return is applied, the combined priority input is

$$P_m^{\text{red}}(x) = Z_S^{\text{red}}(x, m) + Z_G^{\text{red}}(x, m). \quad (56)$$

The trial is initialized as

$$m = 1, \quad I_1(x) = 0 \quad \text{for all } x \in \Omega. \quad (57)$$

The effective priority map is

$$E_m^{\text{red}}(x) = (1 - I_m(x)) P_m^{\text{red}}(x). \quad (58)$$

The reduced model approximates inhibition of return with a within-trial rejection mask.

Candidate regions and competitive sampling. At episode m , candidate regions are the item-centered connected support regions in the effective priority map:

$$\mathcal{C}_m = \{\mathcal{C}_{1,m}, \dots, \mathcal{C}_{K_m,m}\}. \quad (59)$$

Let

$$M_m^{\text{red}} = \max_{\substack{k \in \{1, \dots, K_m\} \\ x \in \mathcal{C}_{k,m}}} E_m^{\text{red}}(x) \quad (60)$$

be the maximum effective priority across candidate locations. The unnormalized selection weight of candidate component k is

$$\alpha_{k,m}^{\text{red}} = \sum_{x \in \mathcal{C}_{k,m}} \exp\left[\lambda \left(E_m^{\text{red}}(x) - M_m^{\text{red}}\right)\right], \quad (61)$$

and the corresponding selection probability is

$$\pi_{k,m}^{\text{red}} = \frac{\alpha_{k,m}^{\text{red}}}{\sum_{\ell=1}^{K_m} \alpha_{\ell,m}^{\text{red}}}. \quad (62)$$

The parameter λ controls how decisively differences in effective priority influence selection. This sampling rule is not proposed as an additional mechanism of attention; it provides a compact approximation to the distribution of winners produced by stochastic recurrent competition in the full attentional selection field.

Selection failures. The reduced model approximates the non-stabilization event in the attentional selection field with a compact failure score. Let

$$\mathcal{D}_m = \bigcup_{k=1}^{K_m} \mathcal{C}_{k,m} \quad (63)$$

be the union of candidate support regions, and let

$$H_m = \sum_{x \in \mathcal{D}_m} E_m^{\text{red}}(x) \quad (64)$$

be the total effective priority over the candidate domain. For any criterion τ , define

$$N_m(\tau) = \left| \left\{ k : \max_{x \in \mathcal{C}_{k,m}} E_m^{\text{red}}(x) > \tau \right\} \right| \quad (65)$$

as the number of candidate components whose peak priority exceeds τ . The failure score is

$$R_m = \eta_1 \frac{N_m(\tau_{\text{pk}})}{\max\{1, N_m(\tau_{\text{hi}})\}} + \eta_2 \frac{H_m}{\xi}. \quad (66)$$

The failure probability is

$$p_{\text{fail},m} = \begin{cases} \min(p_{\text{max}}, R_m), & \Delta = \Delta_{\text{fast}}, \\ 0, & \Delta = \Delta_{\text{slow}}. \end{cases} \quad (67)$$

This captures the fact that selection failures are confined to the fast regime, where the field has less time to stabilize before the oscillator resets the episode. Let $Y_m \in \mathcal{C}_m \cup \{\emptyset\}$ denote

799 the outcome of episode m , where $Y_m = \emptyset$ indicates a selection failure. Then

$$\Pr(Y_m = \emptyset \mid E_m^{\text{red}}) = p_{\text{fail},m}, \quad (68)$$

800 and, for candidate component k ,

$$\Pr(Y_m = \mathcal{C}_{k,m} \mid E_m^{\text{red}}) = (1 - p_{\text{fail},m}) \pi_{k,m}^{\text{red}}. \quad (69)$$

801 A selection failure consumes one episode but does not trigger verification, rejection, or
802 inhibition of return.

803 **Verification and inhibition of return.** After a component is selected, the model
804 verifies whether it contains the target-defining features. A successful verification terminates
805 the trial. Otherwise, the component is rejected, inhibition of return is updated,

$$I_{m+1}(x) = \max \{I_m(x), \mathbf{1}\{x \in \mathcal{C}_{k^*,m}\}\}, \quad (70)$$

806 and the episode counter advances:

$$m \leftarrow m + 1. \quad (71)$$

807 The next episode then begins with the updated effective priority map. If the target is
808 verified on episode M , the model-predicted reaction time is

$$RT = M\Delta. \quad (72)$$

809 *Parameterization*

810 The reduced model combined parameters inherited directly from the full neural
811 model, fixed structural quantities, and parameters optimized to reproduce the full model's
812 benchmark outputs (Table 2). Some optimized parameters have direct counterparts in the
813 full model, but required re-estimation because the surrogate replaces recurrent neural-field

Table 2
Parameters of the reduced diagnostic surrogate model.

Parameter	Symbol	Value	Source
Saliency DoG excitation	A_+	10.2	neural model
Saliency DoG excitation spread	σ_+	1.0	neural model
Saliency DoG inhibition	A_-	0.9	neural model
Saliency DoG inhibition spread	σ_-	14.8	neural model
Saliency normalization gain	γ_S	100	neural model
Saliency sigmoid operating point	θ_S	4.55	bounded multistart BOBYQA
Saliency sigmoid slope	β_S	28.2	bounded multistart BOBYQA
Guidance feature-count shift	κ_G	0.4	neural model
Guidance bias	θ_G	0.394	bounded multistart BOBYQA
Guidance sigmoid slope	β_G	20.0	bounded multistart BOBYQA
Guidance onset	t_G	95 ms	bounded multistart BOBYQA
Saliency weight in priority map	w_S	0.2	neural model
Guidance weight in priority map	w_G	10	neural model
Slow cycle duration	Δ_{slow}	50 ms	literature and neural model
Fast cycle duration	Δ_{fast}	30 ms	literature and neural model
Fast-cycle criterion	Θ_{shape}	120	neural model
Competition exponent	λ	45.7	bounded multistart BOBYQA
High-priority criterion	τ_{hi}	10.10	fixed
Peak criterion	τ_{pk}	10.198	fixed
Failure ratio term	η_1	0.23	bounded multistart BOBYQA
Failure mass term	η_2	0.15	bounded multistart BOBYQA
Failure mass scale	ξ	7000	fixed
Maximum failure probability	p_{max}	0.9	fixed

interactions with simplified episode-level computations, changing their effective operating values. The remaining optimized parameters are specific to the surrogate formulation. All optimized parameters were estimated using a random multistart strategy with the bound-constrained, derivative-free BOBYQA algorithm as the local optimizer (Powell, 2009). Parameters were linearly transformed to the unit hypercube, and independent BOBYQA runs were initialized from a projected baseline parameter vector and from uniformly sampled pseudorandom points within the parameter bounds. The parameter vector yielding the lowest objective value across all runs was retained.

Validation against the full neural model

The surrogate is informative only to the extent that it preserves the benchmark structure generated by the full neural model. After parameterization against the full-model outputs, all surrogate parameters were fixed. Agreement was then assessed for the same behavioral summaries used in the main analysis: set-size slopes for search benchmarks and distractor-present contrasts for additional-distractor benchmarks. Agreement with the full neural model was high. For search slopes, $r = .991$, concordance correlation coefficient $= .990$, RMSE = 1.2 and bias = -0.3 ; for distractor contrasts, $r = .992$, concordance correlation coefficient $= .990$, RMSE = 2.9 and bias = -0.4 . Extended Data Figure 3 shows that this summary-level agreement was accompanied by preservation of the broader benchmark functions and condition ordering. As an external reference, the fixed surrogate also retained close agreement with the human benchmark structure (balanced NRMSE $= 0.0673$). This agreement is not an independent fit to the behavioral data: it follows from approximating a neural model that was itself calibrated and evaluated against those benchmarks. The relevant validation for the diagnostic analyses is the surrogate’s close correspondence with the full neural model. This correspondence supports its use for targeted ablations and parameter-sensitivity analyses.

Selection-geometry analysis

The surrogate’s selection rule provides an analytically tractable approximation for examining how the geometry of effective input influences competition (Extended Data Figure 2). Suppose effective priority is approximately constant within candidate component k , such that $E_m^{\text{red}}(x) \approx h_{k,m}$ for $x \in \mathcal{C}_{k,m}$, and let $|\mathcal{C}_{k,m}|$ denote its spatial support. Then

$$\alpha_{k,m}^{\text{red}} \approx |\mathcal{C}_{k,m}| \exp\left[\lambda \left(h_{k,m} - M_m^{\text{red}}\right)\right]. \quad (73)$$

For two candidate components k and ℓ , the selection-odds ratio is therefore approximately

$$\frac{\pi_{k,m}^{\text{red}}}{\pi_{\ell,m}^{\text{red}}} \approx \frac{|\mathcal{C}_{k,m}|}{|\mathcal{C}_{\ell,m}|} \exp[\lambda (h_{k,m} - h_{\ell,m})]. \quad (74)$$

Spatial support therefore affects selection multiplicatively, whereas differences in effective-priority height are amplified exponentially. A candidate can gain an advantage by occupying a broader region of elevated input, by reaching a greater amplitude, or by both. In the surrogate, shaped salience acts predominantly through spatial support because saturation compresses height differences among strong contrasts. Shaped guidance acts predominantly through amplitude because target-feature overlap differentiates candidate heights, while its feature-count-sensitive operating point keeps near-target matches within the competitive range. These tendencies are relative rather than exclusive and provide a diagnostic description of the selection geometry generated by the full model.

The exponential dependence on height also produces an effective competitive cutoff. For two components with equal spatial support, if one peaks $d \geq 0$ below the other, its selection probability is

$$\pi_{\text{low}}(d; \lambda) = \frac{1}{1 + \exp(\lambda d)}. \quad (75)$$

For a criterion ε defining effective noncompetition, the corresponding priority-distance

858 threshold is

$$d_\varepsilon(\lambda) = \frac{1}{\lambda} \log\left(\frac{1-\varepsilon}{\varepsilon}\right). \quad (76)$$

859 For the reduced-model value $\lambda = 45.7$, $d_{0.05} = 0.064$ and $d_{0.01} = 0.101$ priority units. Thus,
 860 relatively small differences in effective-priority height can move a candidate from
 861 competitive to effectively unavailable. Nonlinear shaping prevents raw differences in
 862 salience or target-feature overlap from producing such disproportionate effects by placing
 863 both sources within an appropriate competitive range before selection.

864 **Targeted Model Variants**

865 Fixed-parameter generalization shows that the full neural dynamic model can
 866 reproduce the benchmark structure, but not which theoretical commitments are responsible
 867 for that success. We therefore used the reduced diagnostic surrogate to construct targeted
 868 variants, each removing or simplifying one commitment while leaving the remaining
 869 structure fixed. None of the variants was refitted. Broad failure indicates a commitment
 870 that constrains performance across benchmark families; more selective failure identifies
 871 conditions in which that commitment becomes especially consequential.

872 **No delayed guidance.** This variant removes the temporal separation between
 873 early salience and later target guidance by setting

$$t_G = 0. \quad (77)$$

874 It tests whether an initial interval dominated by salience is required for the
 875 additional-distractor effects.

876 **Linear salience.** This variant removes nonlinear salience shaping:

$$Z_S^{\text{red}}(x, m) = a_{S,m}^{\text{red}} w_S u_S^{\text{red}}(x). \quad (78)$$

It tests whether normalized local feature contrast can influence selection directly or must first be transformed into the operating range of competition.

Linear guidance. This variant removes nonlinear guidance shaping:

$$Z_G^{\text{red}}(x, m) = a_{G,m}^{\text{red}} w_G \frac{u_G^{\text{red}}(x)}{\max_{y \in \Omega} u_G^{\text{red}}(y)}. \quad (79)$$

It tests whether normalized target-feature overlap is sufficient as a linear guidance signal or must be transformed before influencing selection.

Fixed guidance operating point. This variant retains nonlinear guidance but removes its dependence on the number of target-defining features:

$$Z_G^{\text{red}}(x, m) = a_{G,m}^{\text{red}} w_G \frac{1}{1 + \exp[-\beta_G (u_G^{\text{red}}(x) - 0.5)]}. \quad (80)$$

It tests whether a single operating point is sufficient across target definitions.

No selection failures. This variant removes episodes in which no usable selection stabilizes:

$$p_{\text{fail},m} = 0 \quad \text{for all } m. \quad (81)$$

It tests whether search costs can be explained entirely by selecting and rejecting nontargets or also depend on unresolved competition that consumes time without producing a selection.

Variant results

Variant outputs were evaluated relative to the fixed surrogate baseline. Search benchmarks were assessed from their set-size functions and slopes; additional-distractor benchmarks were assessed from their condition means and distractor-present contrasts. An outcome was classified as *preserved* when its defining benchmark pattern remained intact relative to baseline and *collapsed* otherwise.

Removing nonlinear shaping produced the broadest failures (Extended Data

Figure 4). Linear salience increased balanced NRMSE from 0.0673 to 0.3263 and collapsed 14 of 21 benchmark outcomes. Linear guidance increased balanced NRMSE to 0.3335 and collapsed 19 of 21. Thus, the benchmark structure cannot be reproduced by a simple weighted combination of normalized local contrast and target-feature overlap; both inputs must be shaped before entering competition. The remaining variants produced more selective impairments. Removing delayed guidance caused all seven additional-distractor outcomes to fail, while preserving all seven task-irrelevant feature outcomes and six of seven triple-conjunction outcomes. The only triple-conjunction failure was 3D(0), consistent with its dependence on the early salience-dominated phase. Fixing the guidance operating point primarily impaired triple-conjunction search: five of seven outcomes failed, whereas all task-irrelevant feature and additional-distractor outcomes were preserved. This indicates that triple-conjunction performance depends strongly on guidance adapting to the number of target-defining features. Removing selection failures had the narrowest effect: all additional-distractor outcomes were preserved, whereas three task-irrelevant feature outcomes and the high-overlap 3D(2) triple-conjunction condition failed.

Extended Data Figure 5 shows the corresponding condition-level changes. These patterns reveal that the benchmark families do not isolate independent components of the model. Rather, all depend on the same coupled dynamics, while particular conditions expose different vulnerabilities within that shared process: additional-distractor effects are especially sensitive to temporal asymmetry, triple-conjunction search to target-dependent guidance, and high-competition search conditions to selection failures. The variant analysis therefore strengthens the fixed-parameter generalization result. Cross-paradigm performance depends on specific, interpretable commitments rather than unrestricted flexibility: weakening a core component either produces broad failure across benchmarks or selective failure where that component becomes most consequential. These variants are diagnostic interventions within the present architecture, not an exhaustive comparison with all possible models.

Sensitivity Analysis and Local Stress Testing

The targeted model variants asked whether particular theoretical components were necessary. The analyses here ask a complementary question: once those components are retained, how far can the reduced model’s parameters move before benchmark performance deteriorates? We used three diagnostics: Morris screening within a local $\pm 20\%$ neighborhood of the calibrated baseline, one-at-a-time perturbations of the most influential parameters, and a joint nominal $\pm 10\%$ stress test in which all 23 reduced-model parameters varied simultaneously. Together, these analyses test whether performance depends on a small, interpretable subset of parameters or on diffuse compensation across many parameters.

Morris Screening

The Morris analysis (Campolongo et al., 2007; Morris, 1991) screened all 23 reduced-model parameters around the fixed baseline. For most parameters, the sampled range extended $\pm 20\%$ from the calibrated value, with structurally bounded ranges truncated accordingly. The screen used four levels, 15 trajectories and a fixed random seed; timing parameters were treated as integer-valued quantities. These results describe local sensitivity within the sampled neighborhood and should not be interpreted as posterior uncertainty estimates or as characterizing the full feasible parameter space.

Extended Data Figure 6A plots each parameter’s mean absolute elementary effect, μ^* , against the variability of those effects, σ . The guidance sigmoid slope β_G showed the largest combination of average influence and variability, followed by the guidance bias θ_G , salience excitation spread σ_+ , and guidance feature-count shift κ_G . The ordered ranking in panel B similarly identified β_G , θ_G , σ_+ , κ_G , and guidance weight w_G as the five most influential parameters. Local dependence was therefore concentrated primarily around guidance shaping and feature-count sensitivity, with additional contributions from salience structure.

The one-at-a-time follow-up in panel C isolated the directional consequences of

perturbing the highest-ranked parameters while holding all others fixed. The largest penalties followed a lower β_G , a higher θ_G , and a lower κ_G , whereas perturbations of equal magnitude in the opposite direction were substantially less damaging. The elementary-effect distributions in panel D showed corresponding variation across Morris trajectories. Together, these results indicate both directional nonlinearity and dependence on the surrounding parameter configuration rather than independent additive parameter effects.

The family-specific rankings in panels E–G showed that the same parameters recur across benchmark families but with different emphasis. Triple-conjunction outcomes were especially sensitive to salience excitation spread σ_+ , guidance weight w_G , and feature-count shift κ_G , linking performance to both salience structure and target-dependent guidance. Task-irrelevant feature outcomes were dominated by the guidance transform, particularly β_G , θ_G , and κ_G , showing that these effects cannot be explained by bottom-up salience alone. Additional-distractor outcomes depended on both salience and guidance, with strong contributions from σ_+ , β_G , and θ_G , and a further contribution from guidance onset t_G . Thus, the benchmark families jointly constrain a common set of mechanisms rather than mapping independently onto separate parameters.

Joint Local Stress Test

The Morris screen perturbs parameters along sampled local directions. The joint stress test instead asks what happens when all parameters vary simultaneously. We generated 500 joint random perturbations of the fixed reduced-model baseline by independently drawing uniform coordinates on $[0, 1)$ for all 23 parameters and mapping them to local parameter bounds. Bounds were nominally $\pm 10\%$, clipped to the allowed parameter ranges. Integer timing bounds were rounded outward and sampled timing values were rounded to integer milliseconds, yielding ranges of 27–33 ms for fast-cycle duration and 45–56 ms for slow-cycle duration. This analysis is a local stress test of the calibrated mechanism, not an estimate of posterior uncertainty.

Extended Data Figure 7A shows that the fixed baseline lies near the low-error edge of the perturbation distribution. Some draws remained close to baseline performance, but most increased error, and a distinct subset produced substantially broader benchmark failure. The families did not deteriorate in lockstep (panel B): task-irrelevant feature outcomes were the most frequently preserved, whereas additional-distractor outcomes failed more readily and triple-conjunction outcomes showed an intermediate failure burden.

Panels C–E show that these failures were structured rather than simply noisier versions of the baseline predictions. As triple-conjunction failures accumulated, predicted search slopes increased disproportionately in several conditions, including *3D1*, *12D1*, and *3D2*. In the task-irrelevant feature family, high-failure draws flattened the condition-specific slope pattern toward zero. In the additional-distractor family, increasing failure attenuated the negative contrasts. Different benchmark families therefore failed in qualitatively different ways when the shared parameter configuration was perturbed.

Descriptive associations between sampled parameter values and failure counts identified κ_G as the strongest overall correlate of joint failure ($\rho = -0.53$; panel F): lower values were associated with more failed benchmark criteria. Other associations were smaller and differed by family. Triple-conjunction failures were most strongly associated with κ_G , with additional contributions from salience structure, weighting, and guidance timing. Task-irrelevant feature failures were associated with feature-count sensitivity, guidance bias and weight, guidance slope, and fast-cycle timing. Additional-distractor failures were associated with feature-count sensitivity, guidance onset, slow-cycle timing, and guidance bias. Because all parameters varied simultaneously, these associations are descriptive and should not be interpreted as isolated causal effects.

Panel G makes the dominant overall relationship with κ_G explicit: draws below the calibrated value generally produced more failed criteria, whereas failure decreased and then leveled off at higher values. The remaining parameter trends were substantially weaker. Joint performance was therefore not maintained by diffuse compensation among many

equally influential parameters; local viability remained organized around feature-count-sensitive guidance, with benchmark-specific contributions from guidance shaping, timing, and salience structure.

Pairwise Local-Viability Geometry

Sensitivity rankings identify influential parameters but not how their values interact. Extended Data Figure 8 therefore shows pairwise projections of the joint perturbation sample for the leading parameters in each benchmark scope.

Across the overall and family-specific panels, low-error solutions form curved or sloped regions rather than broad axis-aligned plateaus. The overall criterion shows particularly clear interactions between κ_G and θ_G , slow-cycle timing, and salience structure. Triple-conjunction viability depends on combinations of feature-count-sensitive guidance, salience structure, salience weighting, and guidance timing. Task-irrelevant feature viability is tightly structured by the guidance transform and fast-cycle timing, whereas additional-distractor viability depends strongly on compatible combinations of guidance strength, guidance onset, and slow-cycle timing.

The calibrated baseline is therefore not surrounded by a broad region in which parameters can vary independently with little cost. Nearby parameter sets that preserve benchmark performance do exist, but they occupy structured combinations: a change in one parameter can be tolerated for some values of another parameter and damaging for others. The calibrated solution thus lies within a constrained local viability region shaped by interactions among the same mechanisms identified by the targeted variants and sensitivity analyses.

Summary of Diagnostic Analyses

The diagnostic analyses converged on a common picture. Nonlinear shaping of both salience and guidance was required broadly across the benchmark set, whereas temporal asymmetry, target-dependent guidance and selection failures produced more selective impairments when removed. The benchmark families therefore did not map onto

independent model components; rather, they placed different constraints on the same coupled dynamics.

The sensitivity analyses supported this interpretation from a complementary direction. Local performance was concentrated around parameters governing guidance shaping and feature-count sensitivity, with additional contributions from salience structure, timing and competitive selection. Joint perturbations produced structured, family-specific failures rather than uniform degradation, and low-error solutions occupied curved local regions in parameter space rather than a broad plateau of independently tolerated values.

Together, these results indicate that the fixed-parameter generalization is not sustained by diffuse parameter flexibility. It depends on a constrained set of interacting dynamical commitments whose importance is expressed across multiple benchmark families, with particular conditions exposing different vulnerabilities of the shared mechanism.

Generative AI Acknowledgment

The reduced diagnostic surrogate model, reduced model variants, reduced model robustness analyses and plotting helper functions were implemented in Julia. ChatGPT Codex was used as an AI-assisted coding tool for code drafting, refactoring and debugging. All reduced model equations, analysis choices, parameter settings and interpretations were specified, reviewed and validated by the authors.

Data availability

All reconstructed benchmark displays, empirical benchmark values used in the analyses, model-generated benchmark outputs and analysis-ready data underlying the figures are available at OSF.

Code availability

Code, model configuration files and analysis scripts required to reproduce the simulations and analyses are available at OSF. This includes the full neural model configuration file for the Cedar framework, the reduced diagnostic surrogate, targeted model variants, sensitivity and stress-test analyses, and scripts for reproducing the

1059 reported figures and summary statistics. Cedar is available separately at GitHub.

1060 **Author contributions**

1061 R.G. conceived and designed the study, developed and implemented the models and
1062 software, curated the data, performed and validated the analyses, prepared the
1063 visualizations, and wrote the original manuscript. J.P.S. acquired funding, supervised the
1064 research, and contributed to reviewing and editing the manuscript. G.S. supervised the
1065 research and contributed to reviewing and editing the manuscript.

1066 **Competing interests**

1067 The authors declare no competing interests.

1068 **Funding**

1069 This work was funded by RPG-2021-350 from the Leverhulme Trust awarded to
1070 J.P.S.

References

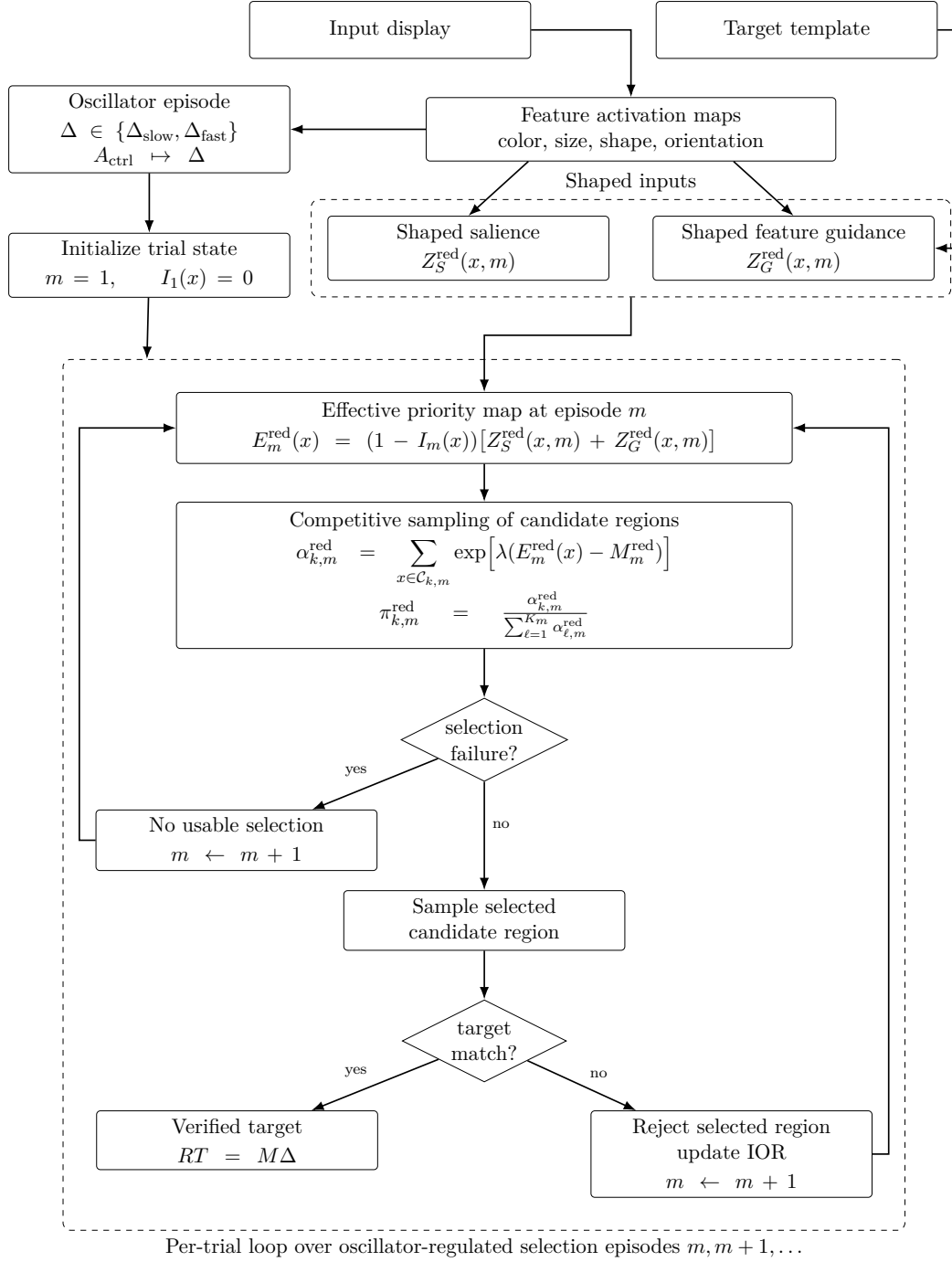
- Anderson, B. A., Kim, H., Kim, A. J., Liao, M.-R., Mrkonja, L., Clement, A., & Grégoire, L. (2021). The past, present, and future of selection history. *Neuroscience & Biobehavioral Reviews*, *130*, 326–350.
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443.
- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485–496.
- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, *303*(5919), 696–698.
- Bisley, J. W., & Mirpour, K. (2019). The neural instantiation of a priority map. *Current Opinion in Psychology*, *29*, 108–112.
- Buschman, T. J., & Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron*, *63*(3), 386–396.
- Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, *22*(10), 1509–1518.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature reviews neuroscience*, *13*(1), 51–62.
- Deco, G., & Rolls, E. T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, *44*, 621–642.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.
- Donk, M., & Van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, *19*(7), 733–739.

- 1098 Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity.
1099 *Psychological Review*, 96(3), 433–458.
- 1100 Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is
1101 contingent on attentional control settings. *Journal of Experimental Psychology:*
1102 *Human perception and performance*, 18(4), 1030.
- 1103 Found, A. (1998). Parallel coding of conjunctions in visual search. *Perception &*
1104 *Psychophysics*, 60(7), 1117–1127.
- 1105 Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of
1106 salient-but-irrelevant sensory inputs. *Psychological Science*, 26(11), 1740–1750.
- 1107 Gaspelin, N., & Luck, S. J. (2018). Combined electrophysiological and behavioral evidence
1108 for the suppression of salient distractors. *Journal of Cognitive Neuroscience*, 30(9),
1109 1265–1280.
- 1110 Gaspelin, N., Ma, X., & Luck, S. J. (2025). Signal suppression 2.0: An updated account of
1111 attentional capture and suppression. *Psychonomic bulletin & review*, 32(6),
1112 2648–2668.
- 1113 Grieben, R., Tekülve, J., Zibner, S. K., Lins, J., Schneegans, S., & Schöner, G. (2020).
1114 Scene memory and spatial inhibition in visual search: A neural dynamic process
1115 model and new experimental evidence. *Attention, Perception, & Psychophysics*.
- 1116 Hamker, F. H. (2005). The emergence of attention by population-based inference and its
1117 role in distributed processing and cognitive control of vision. *Computer Vision and*
1118 *Image Understanding*, 100(1-2), 64–106.
- 1119 Hamker, F. H. (2006). Modeling feature-based attention as an active top-down inference
1120 process. *BioSystems*, 86(1-3), 91–99.
- 1121 Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts
1122 of visual attention. *Vision Research*, 40(10-12), 1489–1506.
- 1123 Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature*, 334(6181),
1124 430–431.

- 1125 Kröse, B. J., & Julesz, B. (1989). The control and speed of shifts of attention. *Vision*
 1126 *Research*, 29(11), 1607–1619.
- 1127 Lien, M.-C., Ruthruff, E., & Hauck, C. (2022). On preventing attention capture: Is
 1128 singleton suppression actually singleton suppression? *Psychological Research*.
- 1129 Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting.
 1130 *Current Opinion in Psychology*, 29, 160–167.
- 1131 Liesefeld, H. R., & Müller, H. J. (2020). A theoretical attempt to revive the
 1132 serial/parallel-search dichotomy. *Attention, Perception, & Psychophysics*, 82,
 1133 228–245.
- 1134 Lomp, O., Richter, M., Zibner, S. K. U., & Schöner, G. (2016). Developing dynamic field
 1135 theory architectures for embodied cognitive systems with cedar. *Frontiers in*
 1136 *Neurorobotics*, 10.
- 1137 Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress
 1138 toward resolving the attentional capture debate. *Visual Cognition*, 29(1), 1–21.
- 1139 Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments.
 1140 *technometrics*, 33(2), 161–174.
- 1141 Niu, Z., Mordkoff, J. T., & Hollingworth, A. (2025). Template-based and saliency-driven
 1142 attentional control converge to coactivate on a common, spatially organized priority
 1143 map. *Journal of Experimental Psychology: Human Perception and Performance*.
- 1144 Nordfang, M., & Wolfe, J. M. (2014). Guided search for triple conjunctions. *Attention,*
 1145 *Perception, & Psychophysics*, 76(6), 1535–1559.
- 1146 Powell, M. J. (2009). The bobyqa algorithm for bound constrained optimization without
 1147 derivatives. *Cambridge NA Report NA2009/06, University of Cambridge,*
 1148 *Cambridge*, 26(26-46), 1.
- 1149 Proulx, M. J. (2007). Bottom-up guidance in visual search for conjunctions. *Journal of*
 1150 *Experimental Psychology: Human Perception and Performance*, 33(1).

- 1151 Ptak, R. (2012). The frontoparietal attention network of the human brain: Action, saliency,
1152 and a priority map of the environment. *The Neuroscientist*, 18(5), 502–515.
- 1153 Sandamirskaya, Y., & Schöner, G. (2010). An embodied account of serial order: How
1154 instabilities drive sequence generation. *Neural Networks*, 23(10), 1164–1179.
- 1155 Schöner, G., Spencer, J. P., & DFT Research Group, T. (2016). *Dynamic thinking: A*
1156 *primer on dynamic field theory*. Oxford University Press.
- 1157 Tam, J., Callahan-Flintoft, C., & Wyble, B. (2022). What the flip? what the pn flip can tell
1158 us about proactive suppression. *Journal of Cognitive Neuroscience*, 34(11),
1159 2100–2112.
- 1160 Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*,
1161 51(6), 599–606.
- 1162 Theeuwes, J. (2023). The attentional capture debate: When can we avoid salient
1163 distractors and when not? *Journal of Cognition*, 6(1).
- 1164 Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive*
1165 *Psychology*, 12(1), 97–136.
- 1166 Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic*
1167 *Bulletin & Review*, 1(2), 202–238.
- 1168 Wolfe, J. M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic*
1169 *Bulletin & Review*, 1–33.
- 1170 Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the
1171 feature integration model for visual search. *Journal of Experimental Psychology:*
1172 *Human perception and performance*, 15(3), 419.
- 1173 Wyble, B., Callahan-Flintoft, C., Chen, H., Marinov, T., Sarkar, A., & Bowman, H. (2020).
1174 Understanding visual attention with ragnaroc: A reflexive attention gradient
1175 through neural attractor competition. *Psychological Review*, 127(6).

1176 Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and
1177 their interactions with visual working memory. *Annals of the new York Academy of*
1178 *Sciences*, 1339(1), 154–164.

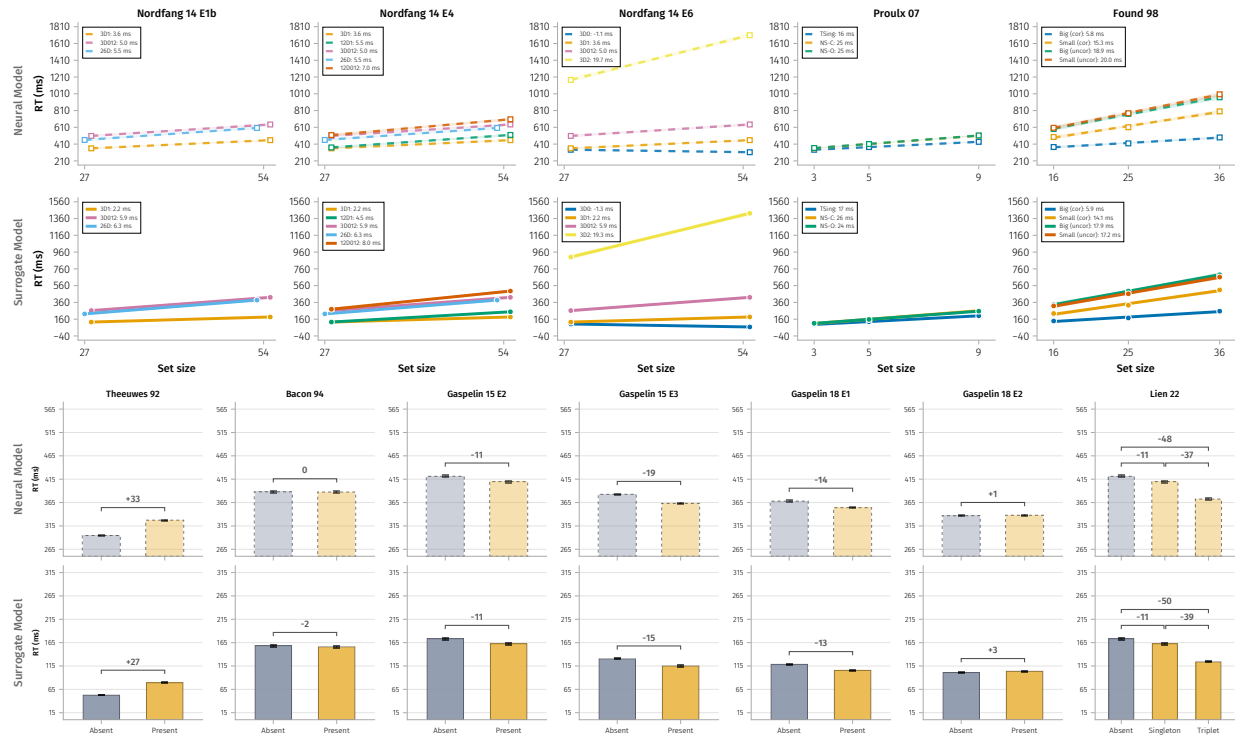


Extended Data Figure 1

Workflow of the reduced diagnostic surrogate. Displays provide spatial feature maps, and a display-level control signal selects the slow or fast oscillator regime. Shaped salience and delayed shaped guidance jointly define an episode-level approximation of the full model's attentional selection landscape. Candidate regions are then sampled competitively to approximate the distribution of stabilized attentional peaks, with some episodes ending in selection failure. Selected nontargets update inhibition of return, whereas verification of the target terminates the trial.

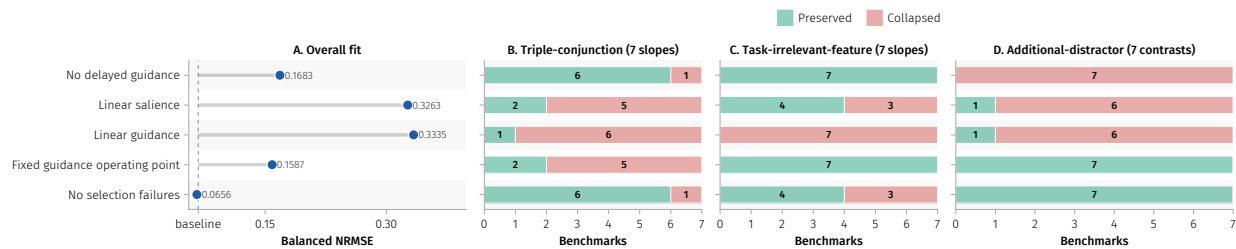
Selection geometry and effective competitive range in the reduced priority landscape. Top: in the surrogate approximation, differences in candidate peak height influence selection odds through the exponential term, whereas differences in spatial support act through the multiplicative support term. Shaped guidance therefore acts predominantly through candidate height, while shaped salience acts predominantly through spatial support. Bottom: as selection decisiveness λ increases, progressively smaller differences in effective priority are sufficient to render a lower-priority candidate effectively noncompetitive.

Selection geometry and effective competitive range in the reduced priority landscape. Top: in the surrogate approximation, differences in candidate peak height influence selection odds through the exponential term, whereas differences in spatial support act through the multiplicative support term. Shaped guidance therefore acts predominantly through candidate height, while shaped salience acts predominantly through spatial support. Bottom: as selection decisiveness λ increases, progressively smaller differences in effective priority are sufficient to render a lower-priority candidate effectively noncompetitive.



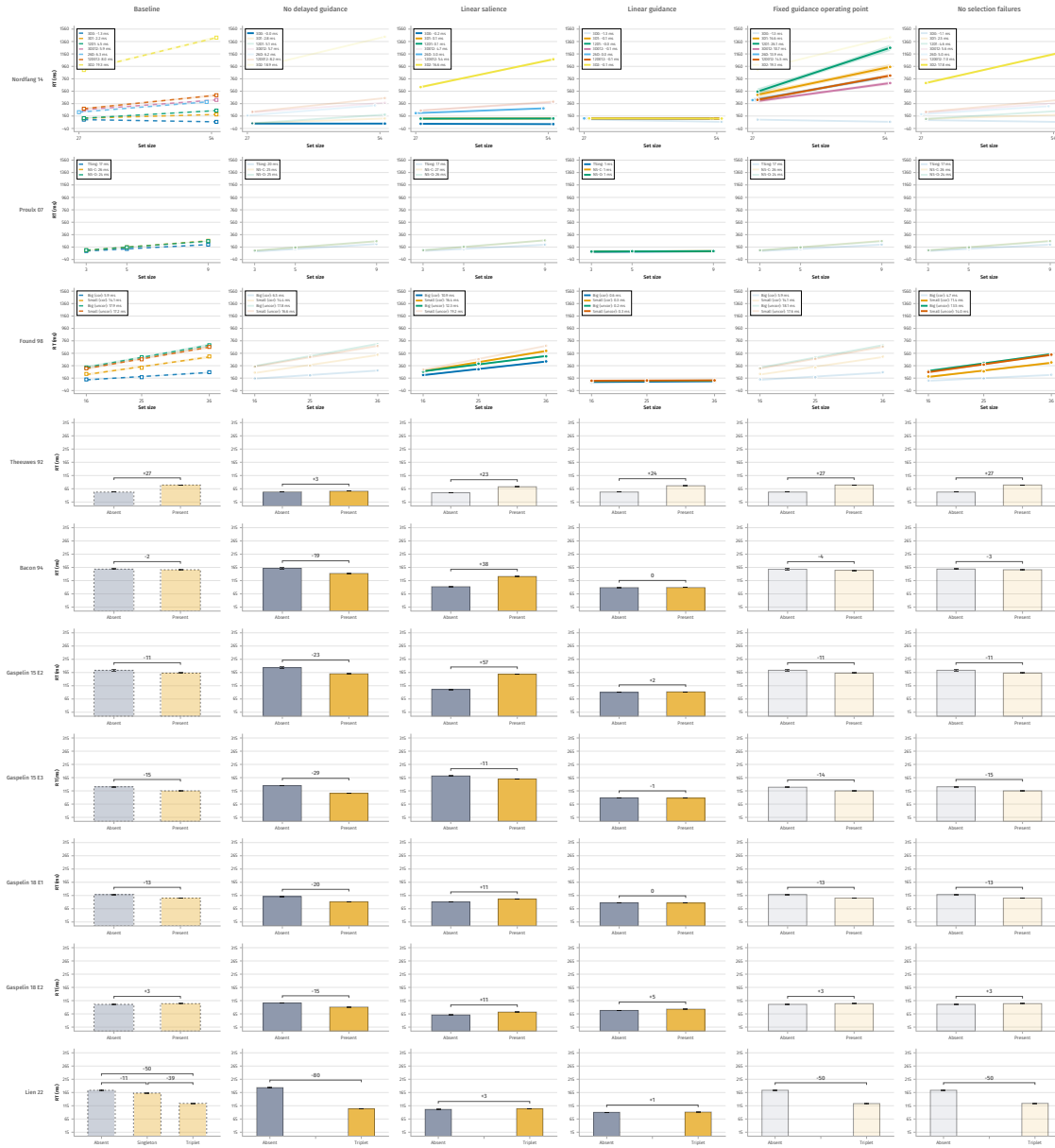
Extended Data Figure 3

Validation of the reduced diagnostic surrogate against the full neural dynamic model. The figure compares the complete benchmark functions and condition means generated by both models. The surrogate closely preserves the benchmark-level geometry of the full neural model across search slopes and distractor-present contrasts, supporting its use for targeted ablations and sensitivity analyses.



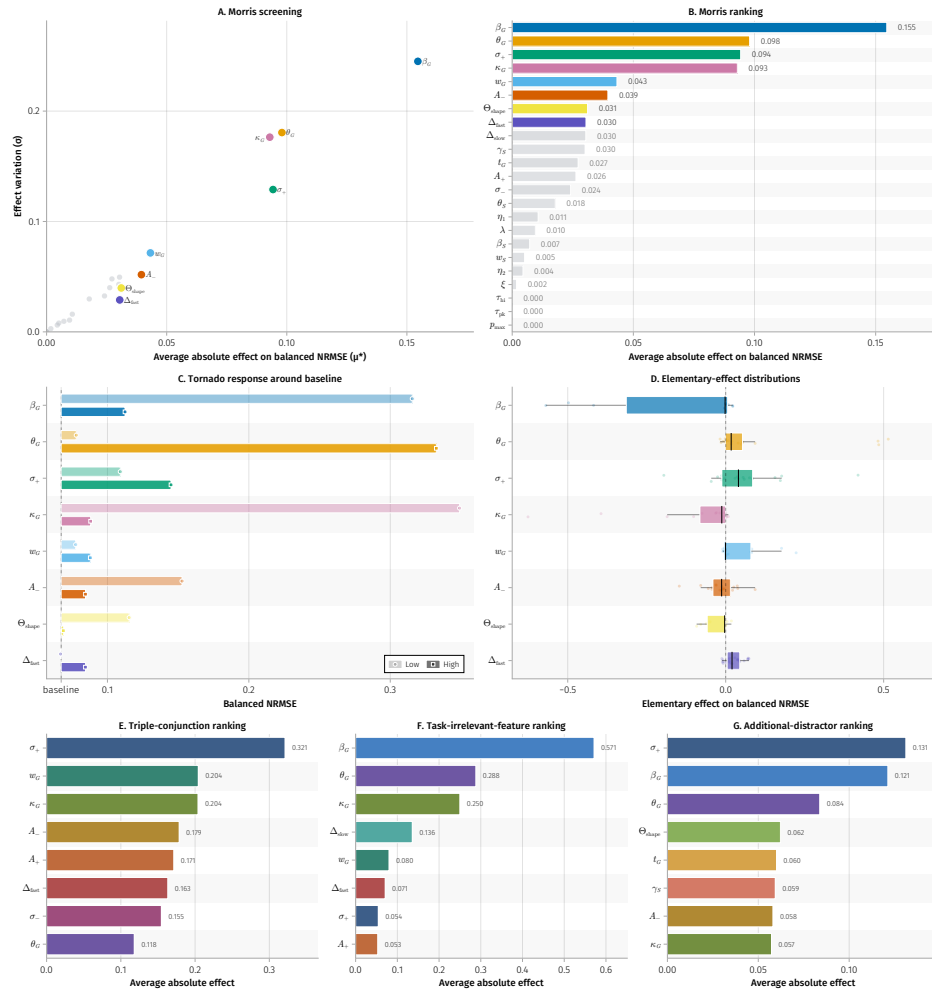
Extended Data Figure 4

Theory-diagnostic variants of the reduced surrogate. Panel A shows balanced NRMSE for each targeted variant; the dashed vertical line marks the fixed-surrogate baseline, and lower values indicate better agreement with the empirical benchmarks. Panels B–D show the number of preserved and collapsed outcomes for the seven triple-conjunction slopes, seven task-irrelevant-feature slopes, and seven additional-distractor contrasts, respectively. Green denotes preserved outcomes and pink denotes collapsed outcomes.



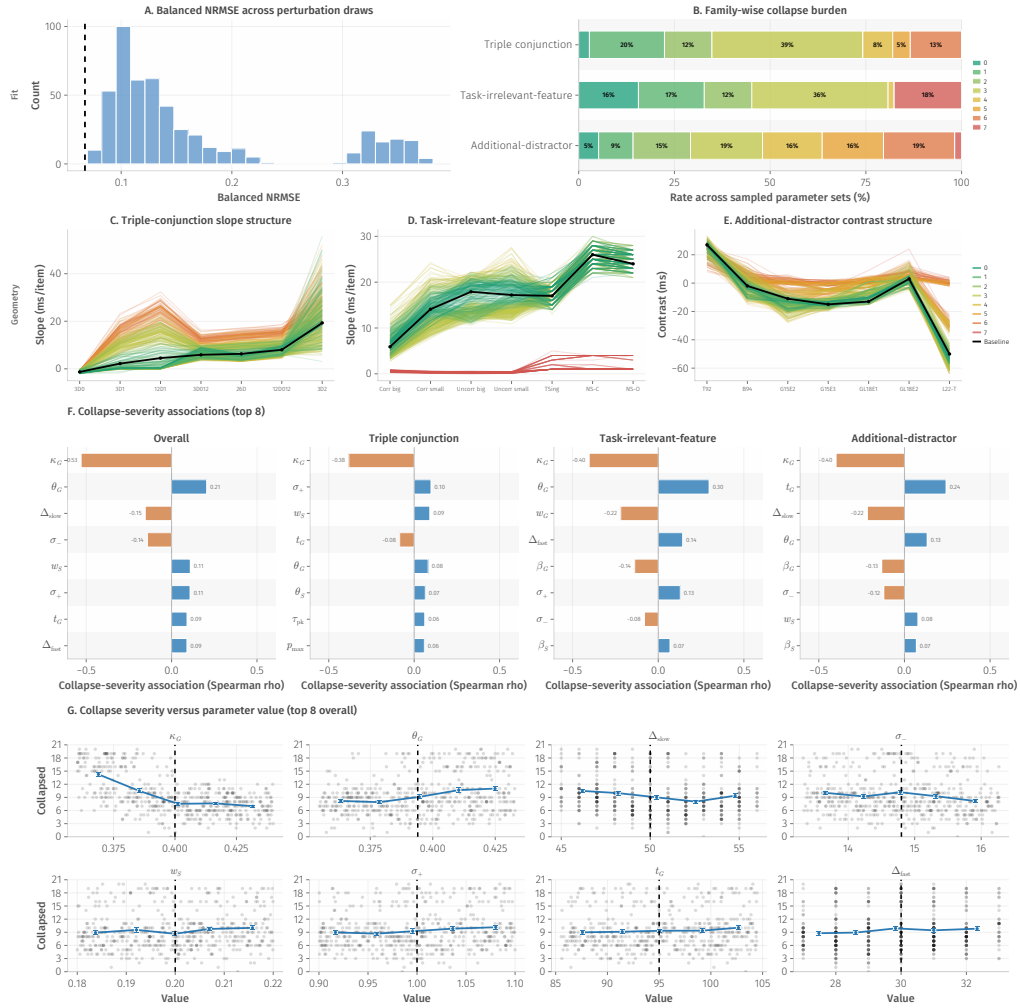
Extended Data Figure 5

Condition-level consequences of targeted model variants. Columns show the fixed reduced-surrogate baseline and five targeted variants; rows show benchmark conditions grouped by family. Each family contains seven unique, non-replication benchmark conditions used for the diagnostic analyses. Search benchmarks are shown as set-size functions, whereas additional-distractor benchmarks are shown as condition means and resulting distractor-present contrasts. Outcomes classified as preserved are plotted with high transparency, whereas collapsed outcomes are emphasized. In the Lien et al. (2022) benchmark, the singleton condition is omitted from the diagnostic grid because it replicates the corresponding singleton configuration from Gaspelin et al. (2015) Experiment 2 and therefore does not constitute an additional unique benchmark condition.



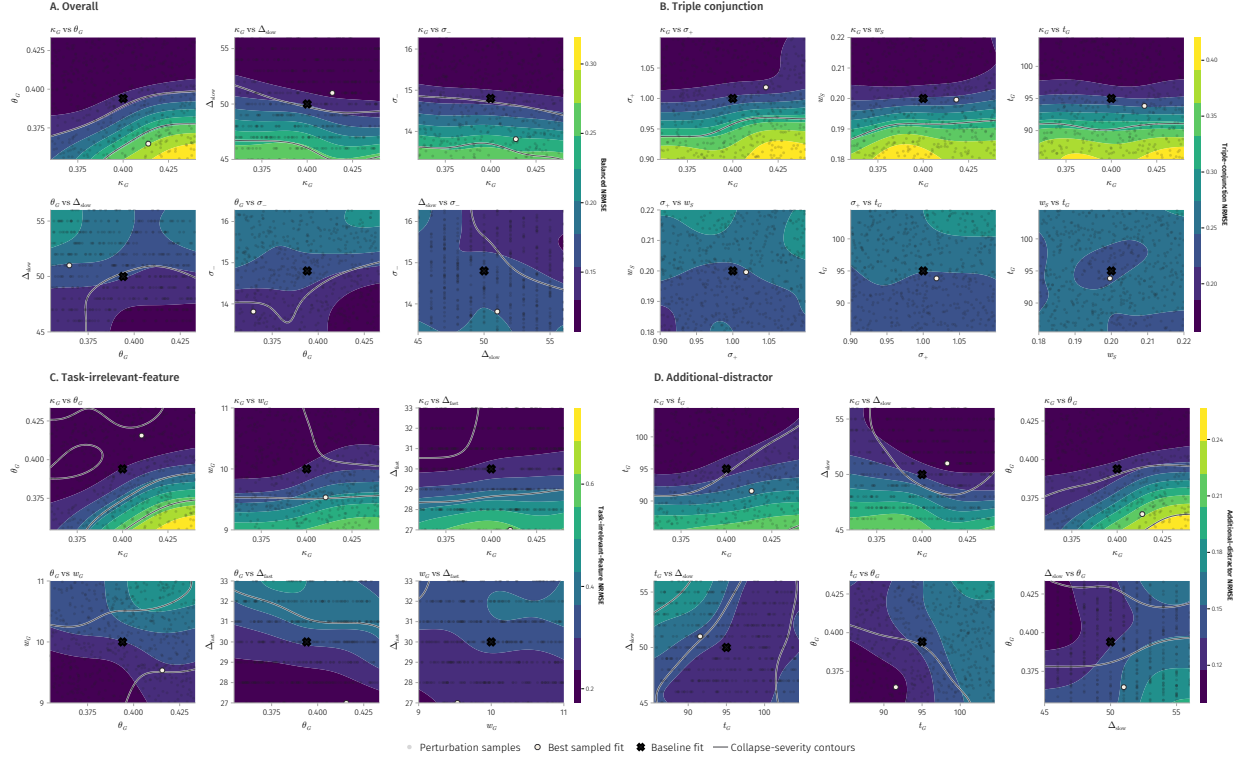
Extended Data Figure 6

Local sensitivity of the reduced shaped-priority model. Panel A plots the mean absolute Morris elementary effect (μ^ , horizontal axis) against the variation of elementary effects across trajectories (σ , vertical axis) for all 23 parameters; farther right indicates greater average influence on balanced NRMSE, whereas higher values indicate greater context dependence across the sampled local parameter space. Panel B ranks parameters by μ^* . Panel C shows the one-at-a-time follow-up, in which each selected parameter was moved separately to a low or high local value while all others remained at the calibrated baseline; the resulting balanced NRMSE is shown relative to baseline. Panel D shows the distributions of signed elementary effects across Morris trajectories. Panels E–G show family-specific Morris rankings for the triple-conjunction, task-irrelevant-feature, and additional-distractor benchmarks, respectively. All Morris analyses were conducted within the screened $\pm 20\%$ neighborhood of the fixed baseline.*



Extended Data Figure 7

Overview of the joint local stress test around the fixed reduced-model baseline. Panel A shows the distribution of balanced NRMSE across 500 independent random perturbation draws; the dashed line marks the fixed baseline. Panel B shows, for each benchmark family, the percentage of sampled parameter sets producing 0–7 failed benchmark criteria. In panels C–E, each thin line shows the benchmark profile from one perturbation draw, line color denotes the family-specific failure count, and the black line denotes the baseline profile. Panel C shows triple-conjunction slopes, panel D task-irrelevant-feature slopes, and panel E additional-distractor contrasts. Panel F shows the eight largest descriptive Spearman associations between sampled parameter values and failure count overall and within each family; negative values indicate that lower parameter values are associated with more failures, whereas positive values indicate that higher values are associated with more failures. Panel G shows the corresponding failure counts and descriptive trends for the eight leading overall parameters; vertical dashed lines mark their baseline values. All 23 reduced-model parameters were varied independently within a nominal $\pm 10\%$ neighborhood of the fixed baseline, subject to the parameter bounds and integer rounding described in Methods.



Extended Data Figure 8

Pairwise projections of local fit quality by benchmark scope. Panels A–D show the overall, triple-conjunction, task-irrelevant-feature, and additional-distractor criteria, respectively. Each panel contains six pairwise projections for leading parameters identified in the joint stress test. Darker colors indicate lower kernel-smoothed NRMSE within the stated scope. Gray points show the 500 perturbation draws, the black marker denotes the calibrated baseline, and the white marker denotes the best sampled fit within that scope. Light contours summarize failure severity. The curved and sloped low-error regions indicate that local viability depends on combinations of parameter values rather than broad independent tolerance along individual parameter axes. These maps are descriptive projections of the full 23-parameter perturbation sample within the nominal $\pm 10\%$ neighborhood described in Methods; they are not posterior uncertainty surfaces or estimates of isolated causal effects.