The practice session with presenting target only

Participants completed a practice session 3-6 days prior to the main task, in which only the target was present throughout the entire session. We compared this practice session (i.e., the target-only session) with the distractor present and absent conditions adopted in the mixed-search session as described in the main text. As shown in Fig. S1A right panel, a repeated-measure ANOVA showed a main effect for different conditions, $F(2,34) = 40.95$, $p < .001$, $\eta_p^2 = 0.71$. In the mixed-search session, the mean response times (RTs) were slower when a salient distractor was present (1480 ms) compared to when it was absent (1390 ms), $t(17) = 6.37$, $p < .001$, Cohen's d = 1.5 (FDR corrected). Meanwhile, the mean RTs in the target-only session (1051 ms) were the fastest, both ts > 5.48, ps < .001 (FDR corrected).

We identified 1014 responsive contacts showing High-frequency (HF) activity during this session. HF activity occurred from 126 ms to 2844 ms in the target-only session. Compared to the distractor present and absent conditions, the HF activity decreased rapidly, resulting in lower power in the 586 – 2310 ms and 610 – 1744 ms intervals respectively (see Fig. S1B).

As shown in Fig. S1C, positive target-tuning CTF slopes started at 218 ms and remained steady until 994 ms in the target-only session (cluster-based permutation test, $p < .01$). The magnitude of the target-tuning CTF slopes in the target-only session was significantly higher than that in the distractor present and absent conditions in the intervals of 224 – 348 ms and 274 – 410 ms respectively (cluster-based permutation test, $p < .01$). There was a significant difference in latency between distractor present condition and target-only session, $t(17) = 2.78$, $p = .013$, Cohen's d = 0.65, but not between distractor absent condition and target-only session, $t(17) = 1.84$, $p = .083$, Cohen's d = 0.43.

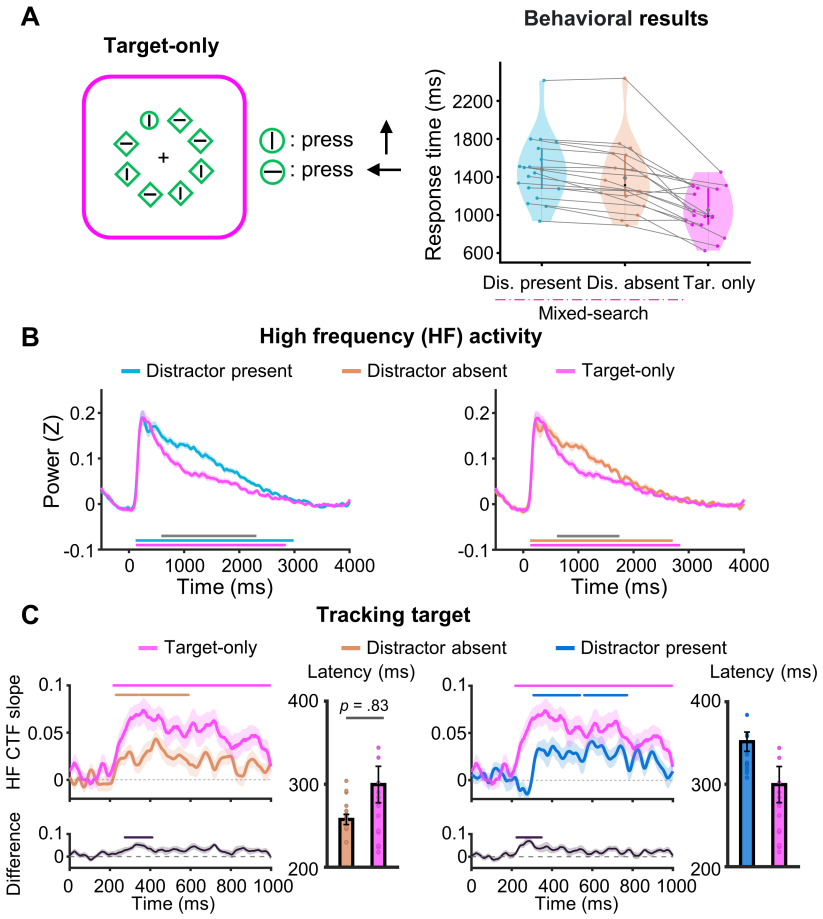

Figure S1. (A) Display setup and possible locations for search elements in the target-only session. Participants consistently searched for one unique shape (target) among search elements (e.g., circle among diamonds). The right panel shows the behavioral results (two-tailed tests), violin plots show the full distribution of n = 18 participants with dots corresponding to individual response times (RTs) respectively. (B) The observed high-frequency (HF; 60-100 Hz) power between the target-only session and distractor present/absent conditions. Significant areas are highlighted by solid lines, and the gray line marks the significant difference between conditions (cluster-based permutation test, $p < .01$). (C) The spatial selectivity (measured as CTF slopes) for the target in the target-only session, distractor present and absent conditions, including target-only-minus-distractor-absent difference and target-only-minus-distractor-present difference in the magnitude of CTF slopes and the latency of CTF slopes (n = 18 participants). Solid dots represent individual data points. Significant CTF selectivity (or in difference) areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$). The data variance is represented by ±1 SEM.

We observed differences between the target-only session and the distractor absent condition in the mixed-search session. One possible reason is that in the target-only session, participants searched consistently for a unique shape without being captured

by the salient distractor. As such, this will speed up search likely due to inter-trial priming[1]. This is different from the mixed-search session in which distractor present and absent conditions were mixed within a block. Furthermore, in the mixed-search session, it is reasonable to assume that observers anticipated the presence of a salient distractor on most trials allowing top-down control to reduce the distraction effect[2].

Time frequency results (beta and alpha bands)

The significantly negative beta activity (13-30 Hz) occurred in the intervals of 206 – 2010 ms and 202 – 1772 ms for the distractor present and absent conditions respectively. In the presence of salient distractors, the beta activity increased slowly, resulting in lower power from 1280 ms to 2286 ms relative to the distractor absent condition (see Fig. S2A left panel). Same for alpha band (8-12 Hz), the activity increased slowly, resulting in lower power from 1744 ms to 2496 ms in the distractor present condition relative to the distractor absent condition.

IEM reconstructions in beta and alpha bands

The results showed that, only in alpha band, the target-tuning CTF slopes were significant above zero from 422 ms to 998 ms (cluster-based permutation test, $p < .01$; see Fig. S2B right panel) in the distractor present condition.

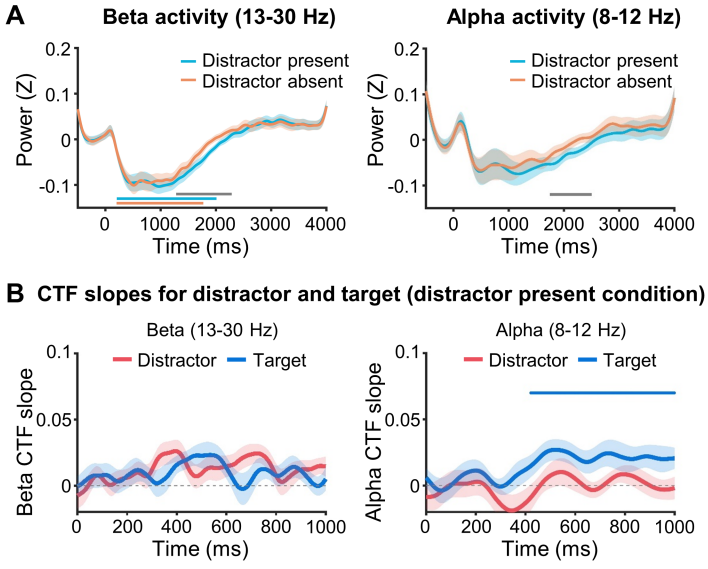

Figure S2. (A) Beta (13-30 Hz, left panel) and alpha (8-12 Hz, right panel) power for the distractor present and absent conditions. The data variance is represented by ±1 SEM. Significant areas are highlighted by the solid lines, and the gray line marks the significant difference between conditions (cluster-based permutation test, $p < .01$). (B) Spatial selectivity (measured as CTF slopes) for the target and salient distractors in the distractor present condition in beta (left panel) and alpha (right panel) band. The data variance is represented by ±1 SEM. Significant CTF selectivity areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$).

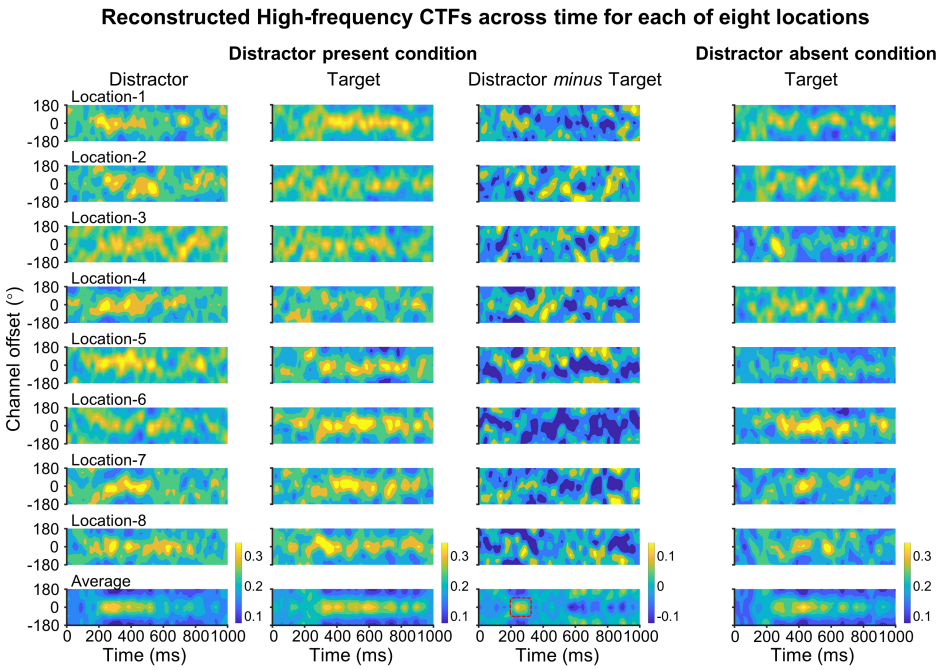

Figure S3. The reconstructed high-frequency (HF) CTFs across time for each of eight locations. A channel offset of 0° corresponds to the channel tuned for the target or distractor location. The yellow band in each subplot shows the peak channel response in different conditions.

Liner classification for the target and salient distractor locations

We also employed Support Vector Machine (SVM) to create one classifier according to HF power to classify eight target or salient distractor locations in the distractor present condition. With this method, we were able to decode he processing of target and salient distractor locations. The error-correcting output codes (ECOC) were used to classify the spatial locations for the target and salient distractors. The ECOC model addresses multiclass categorization problems by aggregating results from multiple binary classifiers. This model was implemented through the 'fitcecoc()' function in MATLAB. For each location, separate trials were allocated for training and testing purposes. Specifically, the decoding for each time point followed a threefold cross-validation procedure, where data from 2/3 of the trials (selected randomly) were used to train the classifier. Subsequently, the performance of the classifier was assessed using the remaining 1/3 of trials for each location against other seven locations. To mitigate any potential biases associated with trial assignment to groups, we iterated the entire procedure 10 times, each with new random assignments

of trials to the three groups (more details see ref. [3]).

The results confirmed those observed from the IEM model, by showing that the classification accuracy for the salient distractor exceeded chance-level (1/8) in the 216 – 470 ms interval; while that for the target exceeded chance-level during the intervals of 294 – 446 ms, 542 – 832 ms and 844 – 926 ms (see Fig. S4B).

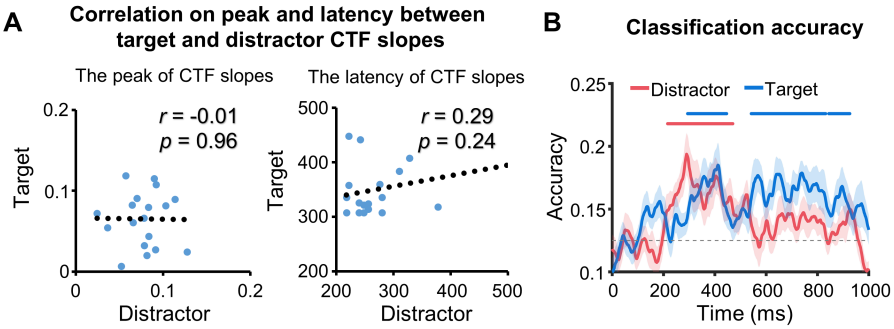

Figure S4. (A) Pearson correlation in the peak and latency of the CTF slopes between tracking the target and tracking the salient distractor. The time window for determining the peak of the CTF slopes ranged from 200 ms to 400 ms, a period during which robust distractor processing was observed. (B) Classification accuracy for the target and salient distractor locations in the distractor present condition. Chance-level performance (1/8) is marked by the dashed gray line. The data variance is represented by ±1 SEM. Accuracies above chance-level are highlighted by the solid lines (cluster-based permutation test, $p < .01$).

Random group strategy

To validate the random-group-strategy, we divided all responsive contacts into 20 groups, mimicking a group of 20 participants. Using this approach, we replicated the critical results from the raw group of 18 participants with high statistical power. Specifically, compared to the distractor absent condition, target-tuning CTF slopes were significantly lower in the distractor present condition in the early time window (226 – 288 ms; cluster-based permutation test, $p < .01$), and its latency was significantly later, $t(19) = 2.24$, $p = .037$, Cohen's d = 0.5 (see Fig. S5B left panel). Such target processing became weaker in the distractor absent condition in the late time window (694 – 786 ms; cluster-based permutation test, $p < .01$). Additionally, in the distractor present condition, the magnitude of distractor-tuning CTF slopes was stronger than that of target-tuning CTF slopes in the early time window (194 – 302 ms; cluster-based permutation test, $p < .01$), and its latency was significantly earlier, $t(19) = 7.86$, $p < .001$, Cohen's d = 1.76 (see Fig. S5B right panel). Moreover, the results confirmed that distractor-related attention declined quickly, as CTF slopes were significantly lower compared to the target ones in the late time window (580 – 768 ms; cluster-based permutation test, $p < .01$).

These findings confirm our initial hypothesis that salient distractor processing was rapid and transient. It is important to test that this cannot be attributed to arbitrary placement of contacts. To address this concern, we iterated the grouping of contacts 1000 times to generate distributions of the averaged CTF slopes (from 226 ms to 288 ms) for the target in the distractor present and absent conditions, and that of the averaged CTF slopes (from 194 ms to 302 ms) for the distractor and target in the distractor present condition (see Materials and Methods for further details). The

results revealed that the distribution of the target in the distractor absent condition was significantly different from that in the distractor present condition (chi-square (49) = 2000, $p < .001$, w = 1), with the averaged CTF slopes for target-tuning was smaller in the distractor present condition (as depicted in Fig. S5B left panel); the distribution of the distractor was significantly different from that of the target (chi-square (49) = 2000, $p < .001$, w = 1), with the averaged CTF slopes for target-tuning was smaller than those for distractor-tuning (as depicted in Fig. S5B right panel).

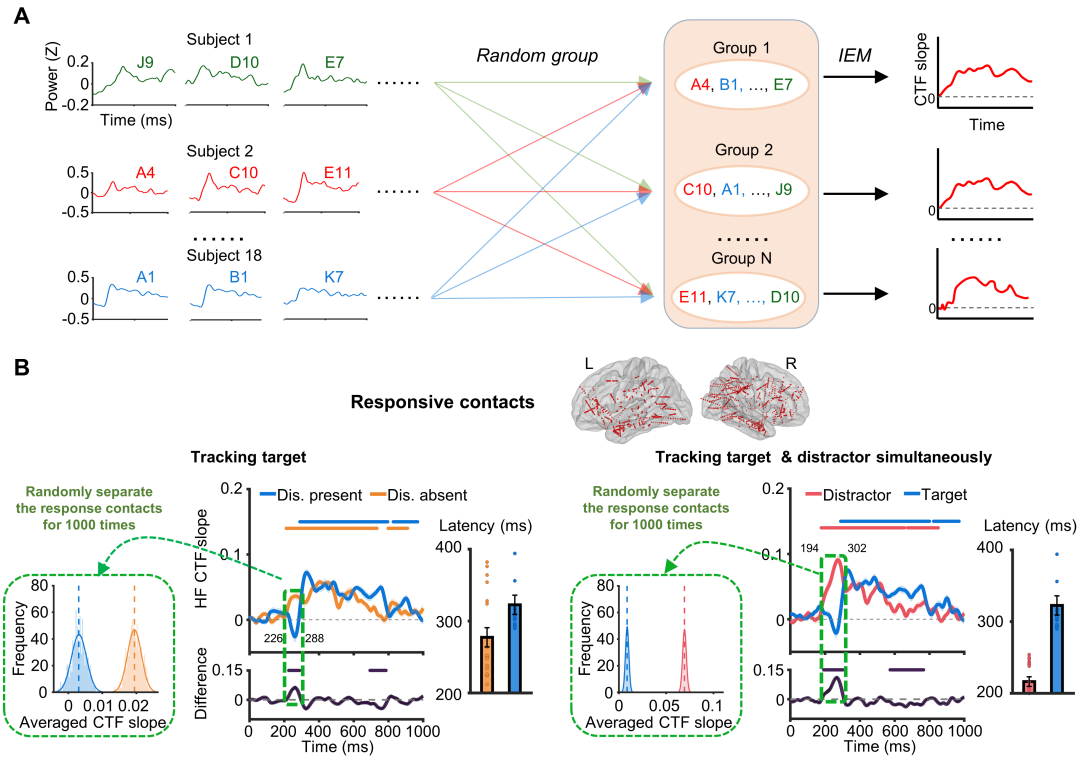

Figure S5. (A) Illustration of the random-group-strategy (see the main text for details). Because all responsive contacts across participants displayed similar patterns in HF activity, we combined these contacts and randomly grouped them into several new groups. The spatial attention was then reconstructed from these groups of contacts. (B) The results obtained using the random-group-strategy for all responsive contacts. The left panel shows the spatial selectivity (measured as CTF slopes) for the target in the distractor present and absent condition, including absent-minus-present difference in the magnitude of CTF slops and the latency of the CTF slopes (n = 20 groups, two-tailed), and the distribution of the target-tuning averaged CTF slopes for the distractor present and absent conditions from 1000 times IEM reconstructions. The right panel shows the spatial selectivity for the target and salient distractors in the distractor present condition, including distractor-minus-target difference in the

magnitude of CTF slopes, the latency of the CTF slopes (n = 20 groups, two-tailed), and the distribution for the target- and distractor-tuning averaged CTF slopes from 1000 times IEM reconstructions. The green dashed outlines indicate the time windows used to estimate averaged CTF slopes for calculating the distribution. Solid dots represent individual data points. The data variance is represented by ±1 SEM. Significant CTF selectivity (or in difference) areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$).

Tracking target-tuning attention

Target-tuning attention was reconstructed in the distractor present and absent conditions for different brain networks based on large-scale cerebral networks (see Fig. S6). For the Default network, the significantly positive target-tuning CTF slopes were observed in the intervals of 372 – 532 ms and 552 – 624 ms in the distractor absent condition, and 298 – 562 ms in the distractor present condition (cluster-based permutation test, $p < .01$). We further showed that, when a salient distractor was introduced, compared to the distractor absent condition, target-tuning CTF slopes were significantly reduced in the early time window (200 – 246 ms; cluster-based permutation test, $p < .01$).

For the Limbic network, the significantly positive target-tuning CTF slopes were observed in the intervals of 224 – 732 ms and 296 – 788 ms in the distractor absent and present conditions, respectively (cluster-based permutation test, $p < .01$). Again, compared to the distractor absent condition, target-tuning CTF slopes in the distractor present condition were significantly reduced in the early time window (236 – 288 ms; cluster-based permutation test, $p < .01$).

For the Frontoparietal network, the significantly positive target-tuning CTF slopes were observed in the intervals of 386 – 494 ms and 546 – 622 ms in the distractor absent condition, and 312 – 390 ms in the distractor present condition (cluster-based permutation test, $p < .01$). For the Ventral Attention network, only in the distractor present condition target-tuning CTF slopes were significantly above zero from 472 ms to 572 ms (cluster-based permutation test, $p < .01$).

For the Visual network, the significantly positive target-tuning CTF slopes were observed in the 480 – 564 ms interval in the distractor absent condition, and in the intervals of 404 – 824 ms and 846 – 942 ms in the distractor present condition (cluster-based permutation test, $p < .01$).

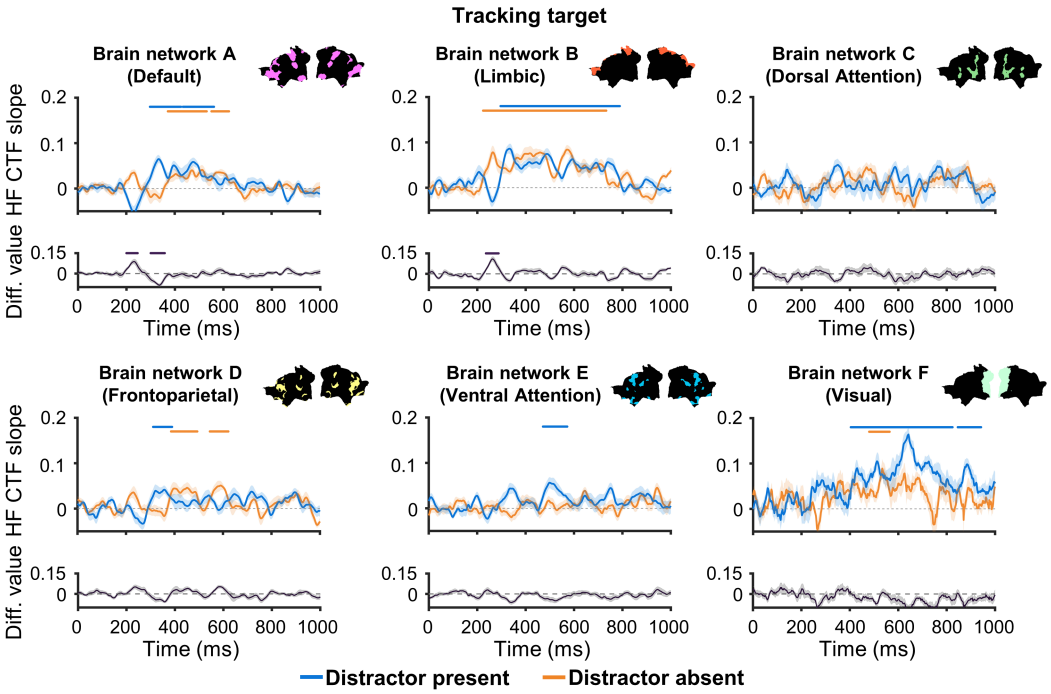

Figure S6. Spatial selectivity (measured as CTF slopes) for the target in the distractor present and absent conditions, including absent-minus-present difference in the magnitude of CTF slopes for Default (A), Limbic (B), Dorsal Attention (C), Frontoparietal (D), Ventral Attention (E) and Visual (F) networks. The data variance is represented by ±1 SEM. Significant CTF selectivity (or in difference) areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$). *$p < .05$, **$p < .01$, ***$p < .001$.

Tracking distractor- and target-tuning attention simultaneously

Distractor- and target-tuning attention were reconstructed simultaneously in the distractor present condition for different brain networks based on large-scale cerebral networks (see Fig. S7). For the Default network, distractor-tuning CTF slopes were significant above zero from 186 ms to 306 ms; while target-tuning CTF slopes were significant above zero from 298 ms to 562 ms (cluster-based permutation test, $p < .01$). The magnitude of distractor-tuning CTF slopes was stronger than that of target-tuning CTF slopes in the early time window (190 – 286 ms; cluster-based permutation test, $p < .01$).

For the Limbic network, distractor-tuning CTF slopes were significant above zero from 184 ms to 590 ms; while target-tuning CTF slopes were significant above zero from 296 ms to 788 ms (cluster-based permutation test, $p < .01$). The magnitude of

distractor-tuning CTF slopes was stronger than that of target-tuning CTF slopes in the early time window (228 – 310 ms; cluster-based permutation test, $p < .01$).

For the Frontoparietal and Ventral Attention networks, only significantly positive target-tuning CTF slopes were observed in the intervals of 312 – 390 ms and 472 – 572 ms, respectively (cluster-based permutation test, $p < .01$).

For the Visual network, distractor-tuning CTF slopes were significant above zero from 432 ms to 542 ms; while target-tuning CTF slopes were significant above zero in the intervals of 404 – 824 ms and 846 – 942 ms (cluster-based permutation test, $p < .01$). The magnitude of distractor-tuning CTF slopes was smaller than that of target-tuning CTF slopes in the late time window (570 – 694 ms; cluster-based permutation test, $p < .01$).

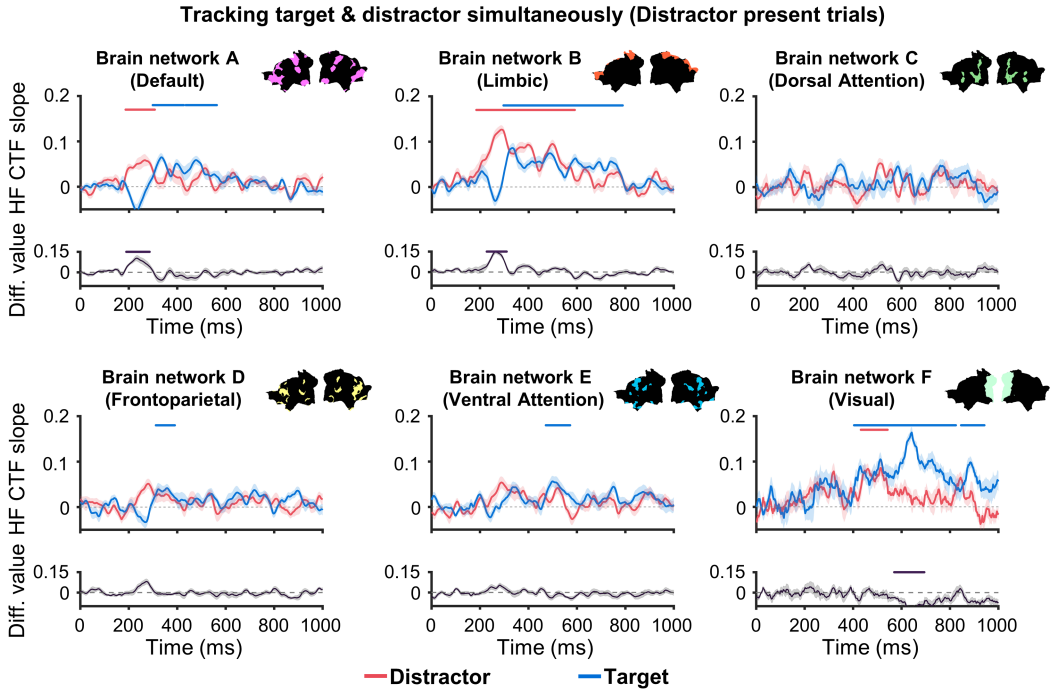

Figure S7. Spatial selectivity (measured as CTF slopes) for the target and salient distractors in the distractor present condition, including distractor-minus-target difference in the magnitude of CTF slopes for Default (A), Limbic (B), Dorsal Attention (C), Frontoparietal (D), Ventral Attention (E) and Visual (F) networks. The data variance is represented by ±1 SEM. Significant CTF selectivity (or in difference) areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$). *$p < .05$, **$p < .01$, ***$p < .001$.

Tracking distractor- and target-tuning attention simultaneously from the distractor-associated brain network (linked to the Default and Limbic networks)

The results in the distractor-associated brain network replicated those found for all responsive contacts (see Fig. 2F), showing that in the distractor present condition, compared to the distractor absent condition, target-tuning CTF slopes were significantly lower in the early time window (232 – 292 ms; cluster-based permutation test, $p < .01$), and its latency was significantly delayed, $t(9) = 2.6$, $p = .03$, Cohen's $d = 0.82$. Additionally, in the distractor present condition, the magnitude of distractor-tuning CTF slopes was stronger than that of target-tuning CTF slopes in the early time window (214 – 306 ms; cluster-based permutation test, $p < .01$). Note that the latency for distractor-tuning CTF slopes was earlier compared to that for the target ones, $t(9) = 5.65$, $p < .001$, Cohen's $d = 1.79$ (see Fig. 2F right panel).

To further validate the random-group-strategy adopted in our study, we reconstructed target- and distractor-tuning CTF slopes for the distractor-associated brain area across individuals (see Fig. S8; only 16 participants were included as two participants had less than 8 contacts as shown in Table S2). Distractor-tuning CTF slopes were significant above zero from 230 ms to 324 ms; while target-tuning CTF slopes were significant above zero from 438 ms to 612 ms (cluster-based permutation test, $p < .01$). The magnitude of distractor-tuning CTF slopes was stronger than that of target-tuning CTF slopes in the estimated early time window (230 – 324 ms), $t(15) = 2.18$, $p = .046$, Cohen's d = 0.54, and again the latency for distractor-tuning CTF slopes was earlier compared to that for the target ones, $t(15) = 19.2$, $p < .001$, Cohen's d = 4.8 (Fig. S8A). Note that, when we combined other four brain networks (except for the Default and Limbic networks) together, we did not observe any significant IEM reconstructions. This suggests that the core observations regarding distractor processing were mainly located in the distractor-associated brain network.

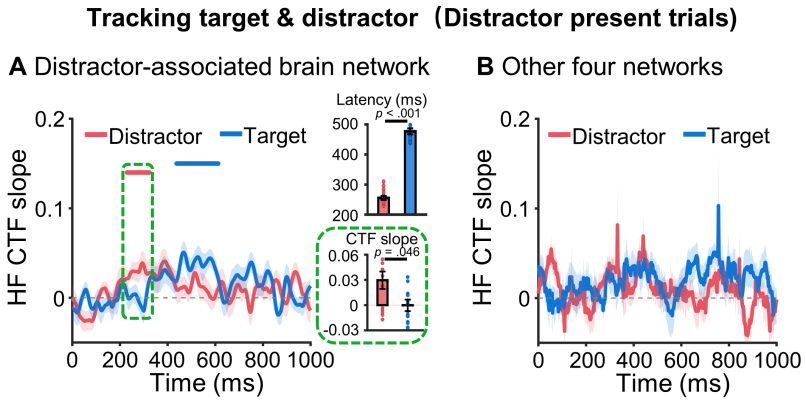

Figure S8. (A) Spatial selectivity (measured as CTF slopes) for the target and salient distractors in the distractor present condition for the distractor-associated brain network, including distractor-minus-target difference in the magnitude of CTF slopes (averaged from the early time window, 230 – 324 ms) and the latency of the CTF slopes (n = 16 participants, two-tailed). Solid dots represent individual data points. (B) Spatial selectivity for the target and salient distractors in the distractor present condition for the combination of other four networks. The data variance is represented by ±1 SEM. Significant CTF selectivity areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$).

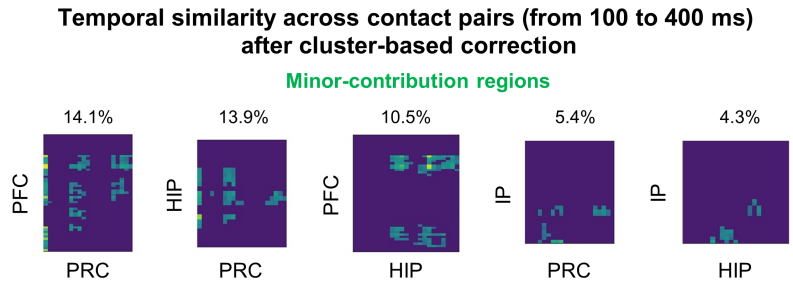

Figure S9. Temporal similarity (from 100 to 400 ms) across contact pairs between brain regions with minor contributions. The percentage presented above each sub-figure indicate how many contact pairs between brain regions have significant connections (after cluster-based correction, $p < .01$).

Responsive contacts showing seizure-onset patterns

High-frequency (HF) activity occurred in the intervals of 136 – 2730 ms, 134 – 2477 ms, and 174 – 1946 ms for all responsive contacts, the responsive contacts after excluding those that exhibited seizure-onset patterns (SOPs; identified by neurosurgeons), and the responsive contacts showing SOPs. There was no significant difference in HF activity between these conditions (see Fig. S10 left panel).

We further found that the target-tuning CTF slopes were significant above zero in the intervals of 300 – 772 ms and 796 – 954 ms for all responsive contacts, and 308 –

538 ms and 560 – 772 ms for the responsive contacts after excluding those showing SOPs (cluster-based permutation test, $p < .01$). The distractor-tuning CTF slopes were significant above zero in the intervals of 224 – 350 ms and 366 – 550 ms for all responsive contacts, and 218 – 546 ms for the responsive contacts after excluding those showing SOPs (cluster-based permutation test, $p < .01$). Notably, there was no significant difference between conditions for the target- and distractor-related CTF slopes (see Fig. S10 right panel).

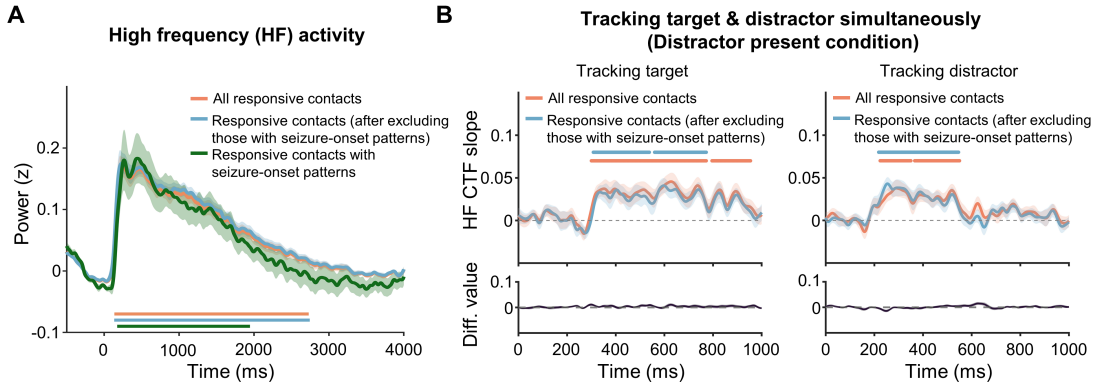

Figure S10. (A) The observed high-frequency (HF) power for all responsive contacts, responsive contacts after excluding those that exhibited seizure-onset patterns (SOPs; identified by neurosurgeons), and responsive contacts that exhibited SOPs. (B) The spatial selectivity (measured as CTF slopes) for the target (left panel) and distractor (right panel) in the distractor present condition, including their differences, for all responsive contacts and responsive contacts after excluding those that exhibited SOPs. The data variance is represented by ±1 SEM. Significant areas are highlighted by the solid lines (cluster-based permutation test, $p < .01$).

Reference

1. Pinto, Y., Olivers, C. L. & Theeuwes, J. Target uncertainty does not lead to more distraction by singletons: Intertrial priming does. *Perception & psychophysics* **67**, 1354–1361 (2005).

2. Müller, H. J., Geyer, T., Zehetleitner, M. & Krummenacher, J. Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception and Performance* **35**, 1–16 (2009).

3. Bae, G.-Y. & Luck, S. J. Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. *J. Neurosci.* **38**, 409–422 (2018).

293 **Table S1.** The information regarding the number of responsive contacts

Patients	All responsive contacts	Responsive contacts after excluding that with seizure-onset patterns
sub001	64	58
sub002	71	71
sub003	70	62
sub004	45	44
sub005	82	78
sub006	77	72
sub007	116	106
sub008	111	111
sub009	97	97
sub010	70	69
sub011	66	66
sub012	49	37
sub013	126	120
sub014	126	125
sub015	52	40
sub016	82	56
sub017	45	34
sub018	30	29
Sum	1379	1275

294

Table S2. The number of responsive contacts in different brain networks in the mixed-search condition.

Patients	Distractor-associated brain network (Default + Limbic)	Dorsal attention	Ventral attention	Frontoparietal	Visual
Sub001	32	1	8	7	0
Sub002	16	0	8	25	0
Sub003	33	0	8	15	0
Sub004	17	0	7	14	0
Sub005	17	23	10	8	12
Sub006	41	2	2	2	20
Sub007	37	12	19	12	5
Sub008	35	4	32	30	0
Sub009	20	28	7	24	14
Sub010	39	1	11	12	0
Sub011	29	5	14	13	2
Sub012	0	3	0	1	26
Sub013	62	13	8	14	0
Sub014	32	13	35	21	2
Sub015	19	2	5	7	3
Sub016	23	10	0	3	17
Sub017	16	1	5	0	9
Sub018	4	17	5	1	0
Sum	472	135	184	209	110