Metacognitive blindness in temporal selection during the deployment of spatial attention

Supplementary Material

Samuel Recht[1,2*], Vincent de Gardelle[3+] , Pascal Mamassian[2+]

[1] Department of Experimental Psychology, University of Oxford, Oxford, UK

[2] Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, PSL University, CNRS, Paris, France

[3] CNRS and Paris School of Economics, Paris, France

[+] these authors share senior authorship

Index

All our analyses were carried out using the R programming language.

1. Probing spatial attention and variability in temporal selection

1.1. Summary statistics

We assume that the distribution of errors in a given attentional condition follows a von Mises distribution, an approximation of the wrapped normal distribution:

$$f(x \mid \mu, k) = \frac{e^{k\,cos(x-\mu)}}{2\pi I_0(k)} \quad \text{for } \mu, x\ in\ (-\pi, \pi]\ \ and\ \ \ k > 0$$

where I_0 is the modified Bessel function of the first kind of order 0, μ is the location parameter (equivalent to the mean in a normal distribution), k is the concentration parameter ($1/k$ is analogous to the variance in a normal distribution) and x the angular error in a given trial.

1.2. Alternatives to the simple von Mises model

To ensure that the von Mises distribution was also preferable to an asymmetrical distribution, we considered a sine-skewed von Mises distribution (Umbach & Jammalamadaka, 2009):

$$f(x \mid \mu, k, \lambda) = \frac{e^{k\,cos(x-\mu)}}{2\pi I_0(k)}\,(1 + \lambda sin(x-\mu)) \ \text{ for } -1 \le \lambda \le 1$$

where I_0 is the modified Bessel function of the first kind of order 0, μ is the location parameter, k is the concentration parameter, λ the skewness parameter, and x the angular error in a given trial. This model with skewness as free parameter between conditions (hereinafter 'SvM', 3 parameters per condition), was also compared to a fixed parameter version across conditions ('SvM-FS').

To also rule out the existence of a bimodal distribution (which could reflect the existence of systematic early and late selection episodes), we considered a mixture of two von Mises with distinct location and concentration ('2vM', 5 parameters per condition):

$$f(x \mid \mu_1, \mu_2, k_1, k_2, w) = \mathrm{w}\ \frac{e^{k_1\,cos(x-\mu_1)}}{2\pi I_0(k_1)} + (1-w)\ \frac{e^{k_2\,cos(x-\mu_2)}}{2\pi I_0(k_2)} \quad \text{for } 0 < w < 1$$

where I_0 is the modified Bessel function of the first kind of order 0, μ_1, μ_2 are the location parameters of each von Mises, k_1, k_2 are the concentration parameters, w is the relative weight of the first mixture component, and x the angular error in a given trial. Two alternatives, with fixed concentration ('2vM-FC') or fixed location ('2vM-FL') across the two von Mises were also tested.

1.3. Trial-by-trial variability in precision level

The concentration parameter of our simple von Mises model provides a summary statistic of the overall dispersion of responses across trials. This overall precision might be distinct from the true encoding precision leading to responses. In particular, encoding precision might fluctuate from trial to trial, leading to a variability that remains undetected by our overall precision measure. To check if our measure of overall precision was distinct from the participants internal variability in memory encoding, we performed an additional model comparison.

It has been proposed that the distribution of errors following the encoding of a stimulus into working memory can be understood as a mixture of a von Mises and a uniform distribution, the later accounting for guesses, which represent no encoding at all (Zhang & Luck, 2008). However, it has been argued that such model should be interpreted with caution given the risk of inflated guess rate estimates. In particular, this risk has been shown to exist when the true generative process is a variable precision model involving zero guess rate (Ma, 2018), or when the error space is non-linearly related to the stimulus space (Schurgin, Wixted, & Brady, 2020). We nonetheless tested if the responses distribution was composed of a mixture of a von Mises and a uniform distribution. The mixture model was defined with one additional parameter g for the guess rate, as follows:

$$f(x \mid \mu, k, g) = g\,\frac{1}{2\pi} + (1-g)\,\frac{e^{k\,cos(x-\mu)}}{2\pi I_0(k)} \quad \text{for } 0 < g < 1$$

The guess rate however could be shared across attentional conditions or not, we therefore had two variants of the von Mises + guess model: one with a specific guess rate for each condition ('vM-3G', 3 parameters per condition) and a shared, fixed guess rate across conditions ('vM-FG').

A second line of thought in the working memory literature proposes that encoding from trial to trial is of variable precision rather than constant (Fougnie, Suchow, & Alvarez, 2012; Van Den Berg, Shin, Chou, George, & Ma, 2012). In such a case, errors are coming from a mixture of von Mises distributions with their concentration following a higher order distribution (often a Gamma distribution). We therefore tested a third, variable-precision model (adapted from Van Den Berg et al., 2012). Contrary to Van Den Berg and colleagues, we did not use the Fisher information (J) as the measure of precision, but we directly used the concentration parameter (k) instead. The Fisher information being monotonically related to k, we kept the latter to make it comparable to our main model.

$$f(x \mid \mu, \bar{k}, \tau) = \int_0^{\infty} \frac{e^{k\cos(x-\mu)}}{2\pi I_0(k)} Gamma(k; \frac{\bar{k}}{\tau}, \tau)\, dk \quad \text{for } \bar{k} > 0;\ \tau > 0$$

where $I_0$is the modified Bessel function of the first kind of order 0, μ is the location parameter of the von Mises distributions, $\frac{\bar{k}}{\tau}$ is the shape parameter (with $\bar{k}$ as the mean concentration) and τ the scale parameter of the gamma distribution, x the angular error in a given trial. The variable-precision model ('VP') has 3 parameters per condition. We also tested three other variants: one with a fixed shape, but variable scale parameter across conditions ('VP-FK'), one with fixed scale but variable shape parameter ('VP-FT') and finally, one with both shape and scale parameters fixed across all attentional conditions ('VP-FT-FK').

All of the tested models involved fitting a specific location parameter (μ) for each condition, in light of the strong and systematic difference in average latency observed between attentional conditions (Carlson, Hogendoorn, & Verstraten, 2006; Chakravarthi & VanRullen, 2011; Hogendoorn, Carlson, VanRullen, & Verstraten, 2010). Note that our model comparison approach was not meant to be fully exhaustive, but rather test the potential fluctuation of precision from trial-to-trial. Exp 2 being a between-subject design, the models involving fixed parameters were omitted (excepted the 2vM, where the fixed parameters are between the two von Mises), and we focused the analysis on the potential difference in model type between conditions.

Models were fitted using MLE. All analyses were carried out using R programming language. BIC and AIC were estimated for each model, and the difference between the simple vM model and the other models for each estimator is denoted ΔBIC and ΔAIC (fig. S1 & S3). A negative value suggests a better fit for the simple vM model. BIC is known to penalize more heavily the number of parameters than AIC. T-tests on BIC & AIC are left uncorrected.

2. Models fitting for Exp 1

2.1. Summary statistics

The mixture of two von Mises (2vM) was significantly worse in terms of BIC and not significantly different from the simple vM with respect to AIC (ΔBIC: $T(19) = 0$, $p<0.001$; ΔAIC: $T(19) = 111$, $p=0.840$). Importantly, the fixed precision / variable location between the two von Mises (2vM-FC) was really worse in terms of BIC ($T(19) = 0$, $p<0.001$) and not different in terms of AIC ($T(19) = 65$, $p = 0.14$), confirming that the latency was homogenous within a condition, and that the responses distribution was very likely to be unimodal. The fixed location/variable precision version (2vM-FL) was worse in terms of BIC ($T(19) = 0$, $p<0.001$), and not different

in terms of AIC (T(19) = 105, p>0.99). We also found no evidence for skewness. Both skewed von Mises models were worse than the simple vM (T(19) = 0, p<0.001 for SvM; T(19) = 5, p<0.001 for SvM-FP). The difference in AIC was not significant (T(19) = 63, p=0.123; T(19) = 81, p=0.389).

2.2. Trial-by-trial precision

When considering the von Mises + guess family models, a first important observation is that the vM-3G model, which supposes a variable guess rate between conditions, was significantly worse than the simple von Mises, according to ΔBIC (T(19) = 27, p=0.002) and not significantly different according to ΔAIC (T(19) = 146, p=0.133). Importantly, it also performed significantly worse than the model with shared guess rate across conditions (vM-FG) relative to BIC (T(19) = 201, p < 0.001), the difference in AIC between these two models was not significant (T(19) = 145, p = 0.143). It is therefore highly unlikely that a change in guess rate between attentional conditions would explain the difference in metacognition observed in our data. The benefit of adding a stable guess rate across condition (vM-FG) was unclear, with only the AIC favouring this model (T(19) = 166, p=0.021), but not the BIC (T(19) = 74, p = 0.261), and only when not correcting for multiple comparisons.

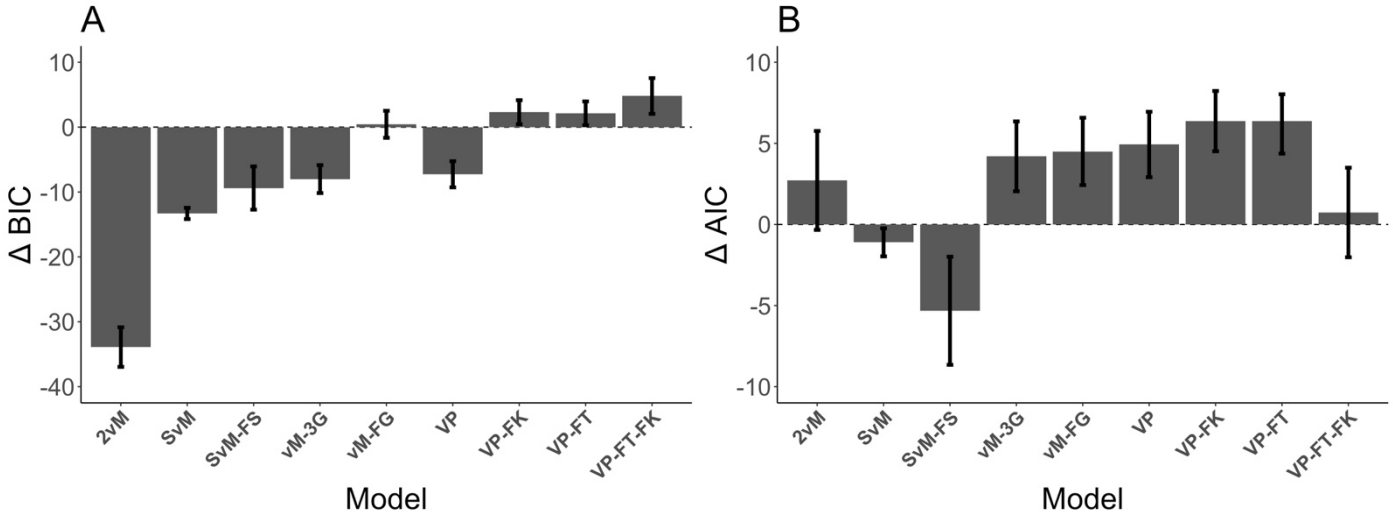


Figure S1: BIC and AIC comparison (Exp 1). (A) The difference in the Bayesian Information Criterion (BIC) between the simple von Mises and each of the other models. A negative value suggests evidence in favour of the simple von Mises model. (B) Same measure but using the Akaike Information Criterion (AIC). Bar charts represent the mean across participants. Error bars represent across participants SEM.

When considering the variable-precision models, the worst model was the full VP, which fitted a specific set of shape and scale parameters to each condition. This model's BIC was significantly worse than the simple vM (T(19)=26, p=0.002) and there was no significant

difference in AIC (T(19) = 154, p=0.069). The VP-FT-FK, which fixes the parameters across conditions, was not significantly better than the simple vM for BIC (t(19)=1.74, p=0.098) nor AIC (t(19) = 0.27, p=0.793). When fixing one parameter of the VP, we found no significant difference in BIC (for VP-FK: T(19)=118, p=0.647; for VP-FT: T(19) = 119, p=0.622), but a lower AIC for both models (for VP-FK: T(19) = 187, p=0.001; for VP-FT: T(19) = 188, p=0.001). The average ΔAIC was 6.37 for the model with fixed shape and 6.19 for the fixed scale model. Therefore, both of these two models were accounting equally well for the data, but the evidence favouring these models over the simple von Mises was fairly low, particularly when using BIC.

2.3. Testing trial-by-trial precision

Considering the guess rate family of models, only the model with fixed guess rate across conditions could be considered as a potential candidate, but adding a stable uniform across condition is mathematically equivalent to adding a probability constant, leading to no change in the precision pattern between conditions in this case.

Regarding the variable precision family of models, we tested the potential effect of a trial-by-trial variability on parameters. We selected the Variable-precision with fixed shape model (fig. S2, A and B). A repeated-measure ANOVA was first applied on latency. It confirmed the effect of condition on latency ($F(1.51, 28.74) = 203.46$, MSE=1642.46, $p<0.001$). The difference between the pre-cue and exogenous conditions ($t(19) = -6.62$, $p<0.001$), the pre-cue and endogenous conditions ($t(19) = -15.7$, $p<0.001$) and between the exogenous and endogenous conditions ($t(19) = -15.12$, $p<0.001$) were all significant after Bonferroni-correction (alpha=0.05/3). A second ANOVA was applied with average concentration as a dependent variable, and condition as an independent variable.

The effect of condition on the average concentration was significant ($F(1.98, 37.69) = 4.06$, MSE=1.00, $p=0.03$), but this effect was driven by a higher average precision in the exogenous compared to endogenous condition ($t(19) = 2.78$, $p=0.012$). The difference between the pre-cue and endogenous/exogenous conditions was not significant (all $p>0.117$, Bonferroni-corrected with alpha=0.05/3). These results are all fully consistent with what was observed using the simple vM model.

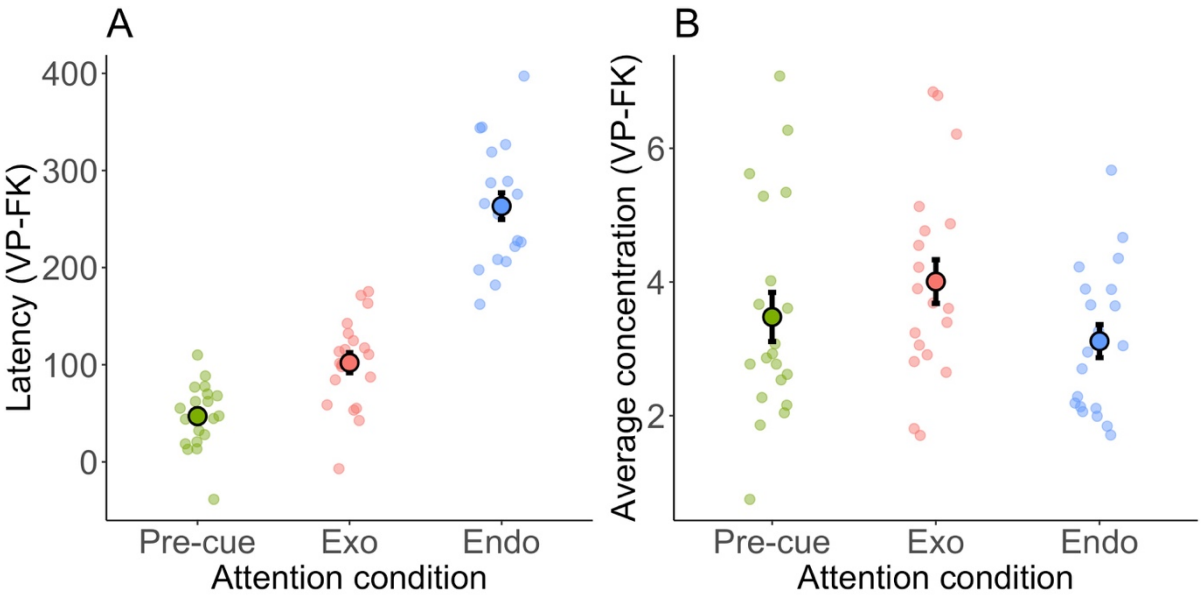


Figure S2: Effects of variable encoding precision. (A) The latency for each attentional orienting condition in the Variable-precision with fixed shape parameter model, we do not expect a change in the latency profile given the model. (B) The average concentration for each attentional orienting condition in the Variable-precision with fixed shape parameter model. There is no significant difference between the pre-cue and exogenous/endogenous conditions. Coloured dots correspond to individual participants in the given condition. Black-outlined dots represent the mean across participants. Error bars represent across participants SEM.

Together these results suggest that our attentional manipulation strongly affected average latency (μ), but not the overall precision of responses. It also suggests that our measure of overall precision (k) of the response distribution was likely to mirror trial-level encoding precision for this experiment. Moreover, adding a guess rate parameter was only very weakly beneficial when the guess rate was fixed across conditions.

3. Models fitting for Exp. 2

3.1. Summary statistics

As for Exp 1, we found no compelling evidence for a bimodality or asymmetry of the empirical distributions, confirming the appropriateness of the simple von Mises model (see the 2vM/2vM-FC and the SvM models in fig. S3 and table S4).

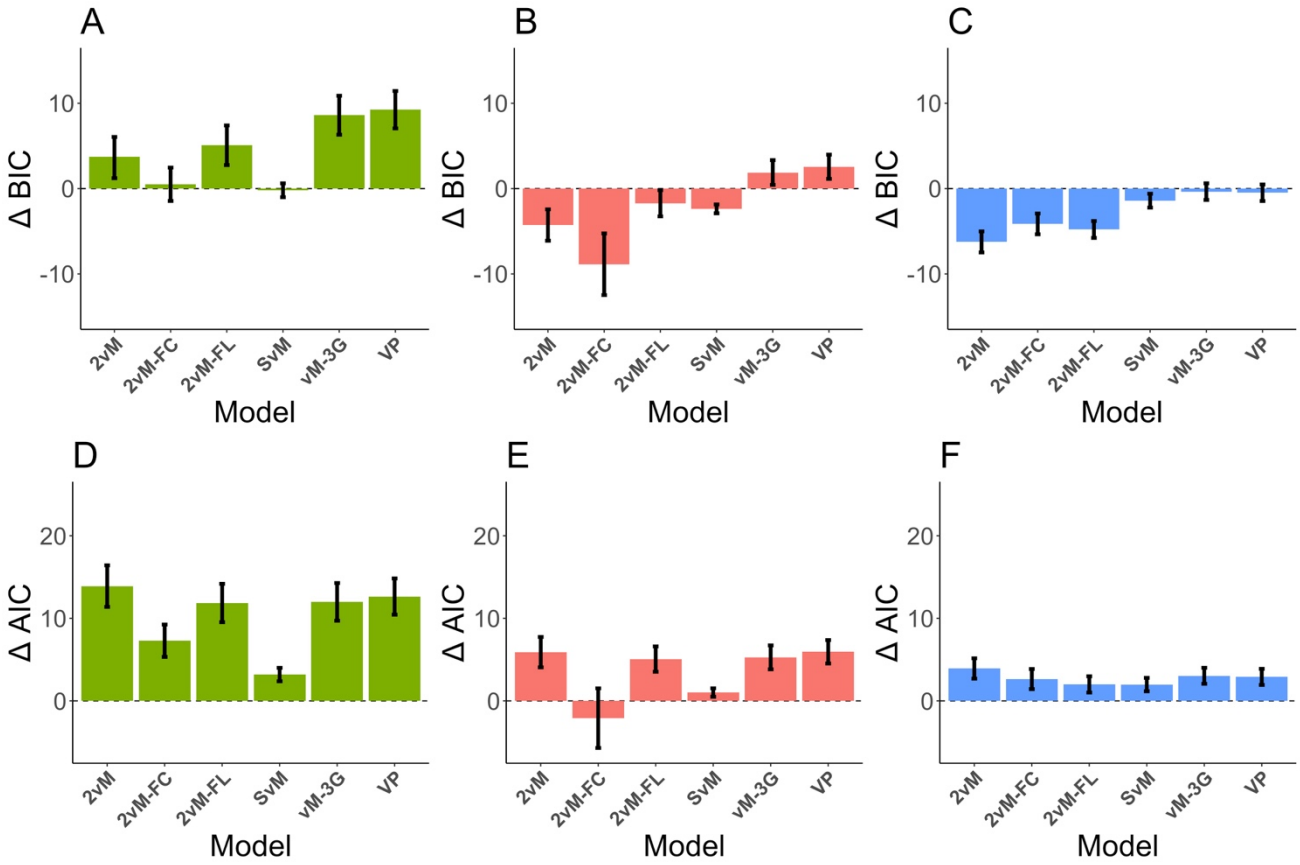


Figure S3: BIC and AIC comparison (Exp 2). (A, B, C) The difference in the Bayesian Information Criterion (BIC) between the simple von Mises and each of the other models for pre-cue (A), exogenous (B) and endogenous (C) conditions. A negative value suggests evidence in favour of the simple von Mises model. (D,E,F) Same measure but using the Akaike Information Criterion (AIC). Bar charts represent the mean across participants. Error bars represent across participants SEM.

3.2. Trial-by-trial precision

Because Exp. 2 has a between-subjects design, we oriented our analyses on the potential differences in encoding model between conditions (fig. S3 and table S4). Interestingly, we found strong evidence for the family of mixture models in the pre-cue group, and moderate evidence for it in the exogenous group, in sharp contrast to the endogenous group. In particular, the model with a mixture of a von Mises and a uniform (vM-3G) and the model with a mixture of von Mises following a Gamma distribution (VP) were significantly outperforming the simple von Mises model. This suggests that despite overall response variability being seemingly matched between conditions, the trial-by-trial fluctuation of internal precision level might be differentially distributed between pre-cue/exogenous and endogenous conditions. While the evidence against mixture models in the endogenous condition was unconclusive (fig. S3, C & E), it was markedly lower than for the pre-cue and endogenous conditions.

Group	Model	ΔBIC	Statistical test (ΔBIC)	ΔAIC	Statistical test (ΔAIC)
Pre-cue	2vM	3.73	T(31) = 320, p = 0.31	13.90	T(31) = 500, p < 0.001
	2vM-FL	5.07	T(31) = 354, p = 0.09	11.85	T(31) = 494, p < 0.001
	2vM-FC	0.51	T(31) = 238, p = 0.64	7.29	T(31) = 445, p < 0.001
	SvM	-0.20	T(31) = 227, p = 0.50	3.19	T(31) = 441, p < 0.001
	vM-3G	8.60	T(31) = 435, p < 0.001	11.99	T(31) = 498, p < 0.001
	VP	9.24	T(31) = 468, p < 0.001	12.64	T(31) = 509, p < 0.001
Exogenous	2vM	-4.27	T(35) = 178, p = 0.01	5.90	T(35) = 506, p = 0.005
	2vM-FL	-1.72	T(35) = 211, p = 0.06	5.06	T(35) = 526, p = 0.002
	2vM-FC	-8.87	T(35) = 83, p < 0.001	-2.09	T(35) = 315, p = 0.79
	SvM	-2.37	T(35) = 103, p < 0.001	1.01	T(35) = 405, p = 0.26
	vM-3G	1.88	T(35) = 355, p = 0.74	5.28	T(35) = 519, p = 0.002
	VP	2.56	T(35) = 410, p = 0.23	5.95	T(35) = 557, p < 0.001
Endogenous	2vM	-6.25	T(37) = 89, p < 0.001	3.92	T(37) = 543, p = 0.01
	2vM-FL	-4.79	T(37) = 98, p < 0.001	1.99	T(37) = 451, p = 0.25
	2vM-FC	-4.14	T(37) = 164, p = 0.002	2.64	T(37) = 450, p = 0.26
	SvM	-1.41	T(37) = 208, p = 0.02	1.98	T(37) = 450, p = 0.25
	vM-3G	-0.35	T(37) = 290, p = 0.25	3.04	T(37) = 506, p = 0.05
	VP	-0.48	T(37) = 256, p = 0.10	2.91	T(37) = 543, p = 0.01

Table S4: Statistics. The average difference (ΔBIC, ΔAIC) between the simple vM and the considered model. Negative values are evidence for the simple von Mises.

3.3.Testing trial-by-trial precision

To better understand the nature of the trial-by-trial fluctuation in precision, we estimated the parameters of the Variable precision model for each condition. For the reasons to prefer this model over the guess rate model, please see Section 1.3. of the present document and Section 2 of the paper. We did not fit the VP model per confidence group, because given the limited samples per group (approx. 90) and the complexity of the model, drawing conclusions about estimated parameters in this case would have been relatively meaningless.

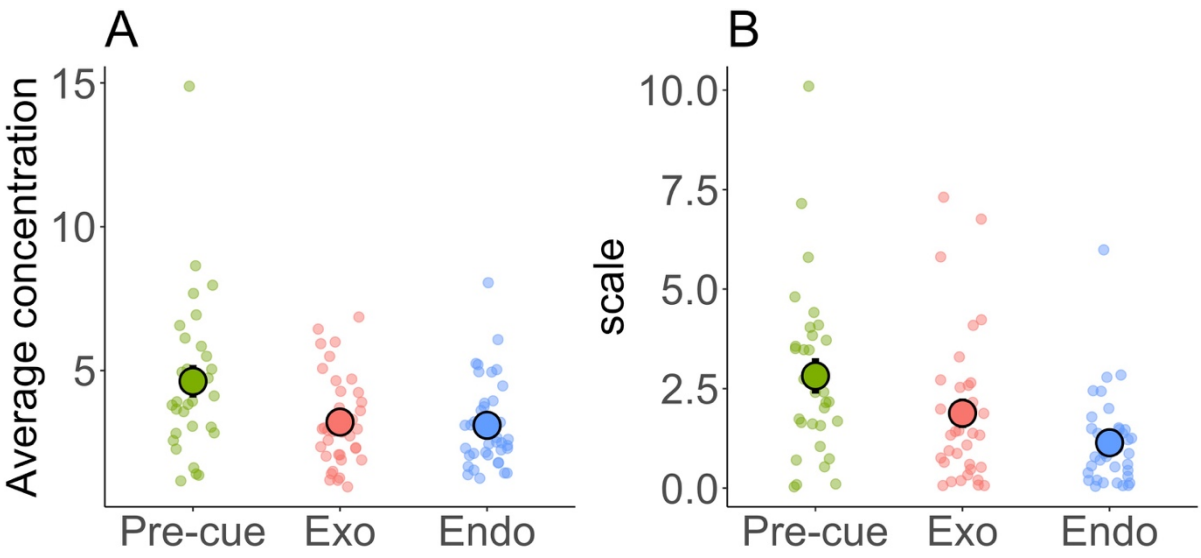


Figure S5: Parameters estimate for the VP model (Exp 2). The significant effect of condition on the average concentration (A) and scale (B) parameters of the higher order Gamma distribution suggest that internal precision is fluctuating from trial to trial in a more pronounced way for the pre-cue condition.

We did not test the latency, since this parameter is the same in both the simple von Mises and VP model. We applied two distinct ANOVA on the average concentration parameter and the scale (fig. S5). The effect of condition on both average concentration ($F(2, 103) = 6.33$, $MSE=3.86$, $p=0.003$) and scale ($F(2, 103) = 8.28$, $MSE=2.96$, $p<0.001$) was significant. Corrected t-test confirmed a difference in average concentration between pre-cue and exogenous ($T(66) = 781$, $p = 0.01$), pre-cue and endogenous ($T(68) = 846$, $p = 0.005$), but not for exogenous vs. endogenous ($T(72) = 706$, $p = 0.82$). Regarding the scale parameter, we found a significant difference between pre-cue and endogenous ($T(68) = 944$, $p < 0.001$), but not between pre-cue and exogenous ($T(66) = 394$, $p = 0.02$) and not between exogenous and endogenous conditions ($T(72) = 847$, $p = 0.08$).

One could therefore speculate that orienting endogenous attention temporary shrinks trial-by-trial precision variability (fig. S6). This shrinking could be partly responsible for the difference in metacognitive ability found in the present study: with lower trial-by-trial fluctuation in evidence signal, error discrimination might prove more difficult, despite average selection variability remaining fairly stable. It has been recently proposed that variable precision models can be related to the stochastic sampling interpretation of neural population coding, where multiple samples of a stimulus are taken depending on resource allocation (Schneegans, Taylor, & Bays, 2020). The quality of representation depends on both

the number of samples and the encoding precision, the latter being weighted by a Gamma-like distribution over samples. Within this framework, greater average precision is also synonymous of broader variability in precision level (fig. S6). One could speculate that such mechanism also facilitates metacognition.

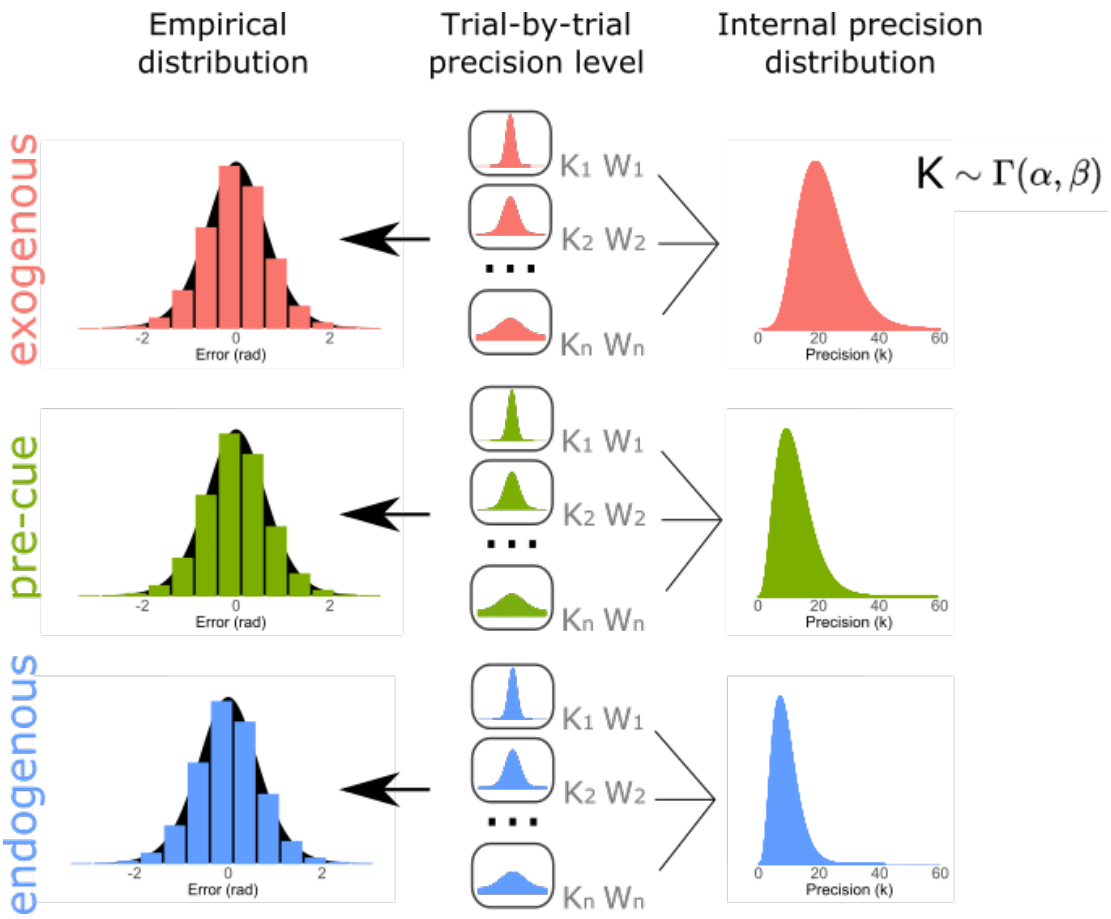


Figure S6: Schematic diagram of the difference between overall and trial-by-trial precision (Exp 2). While the overall dispersion of the empirical distribution of error could remain relatively similar between conditions (left panels, from top to bottom), the different components (centre panels) of the distribution can be produced by different levels of internal precision. Each particular precision level (K_n) is weighted (W_n) via a higher-order distribution (right panels). The distribution of internal precision has been shown to often follow a Gamma distribution. The two parameters (α,β) of the internal distribution determine trial-to-trial variability in precision (K). We used the group-averaged estimated parameters of the fits to produce the different internal precision distributions. Lower average precision level correlates with lower fluctuation in precision (e.g. endogenous vs. pre-cue).

4. Clocks' eccentricity in Exp. 1

Our experimental paradigm in Exp 1 involved two distinct eccentricities: 4 of the clocks were located at 4° eccentricity and 2 clocks at 6° eccentricity. The eccentricity here was also depending on the position relative to the horizontal meridian (the 6° eccentricity landing on the meridian). We nonetheless checked if eccentricity has an effect on our results. We added the eccentricity factor to the ANOVAs reported in section 3.1.2. of the main manuscript, where latency (or concentration) was predicted by confidence (higher vs. lower) and attention condition (pre-cue vs. exogenous vs. endogenous).

We found no evidence for an effect of eccentricity on latency ($F(1,19) = 0.08$, $MSE = 721.78$, $p = 0.78$), no eccentricity x confidence interaction ($F(1,19) = 0.51$, $MSE = 224.13$, $p = 0.48$) and no eccentricity x condition interaction ($F(1.74, 33.09) = 0.40$, $MSE = 609.91$, $p = 0.64$). The triple interaction confidence x condition x eccentricity was also not significant ($F(1.71, 32.57) = 0.43$, $MSE = 528.87$, $p = 0.62$).

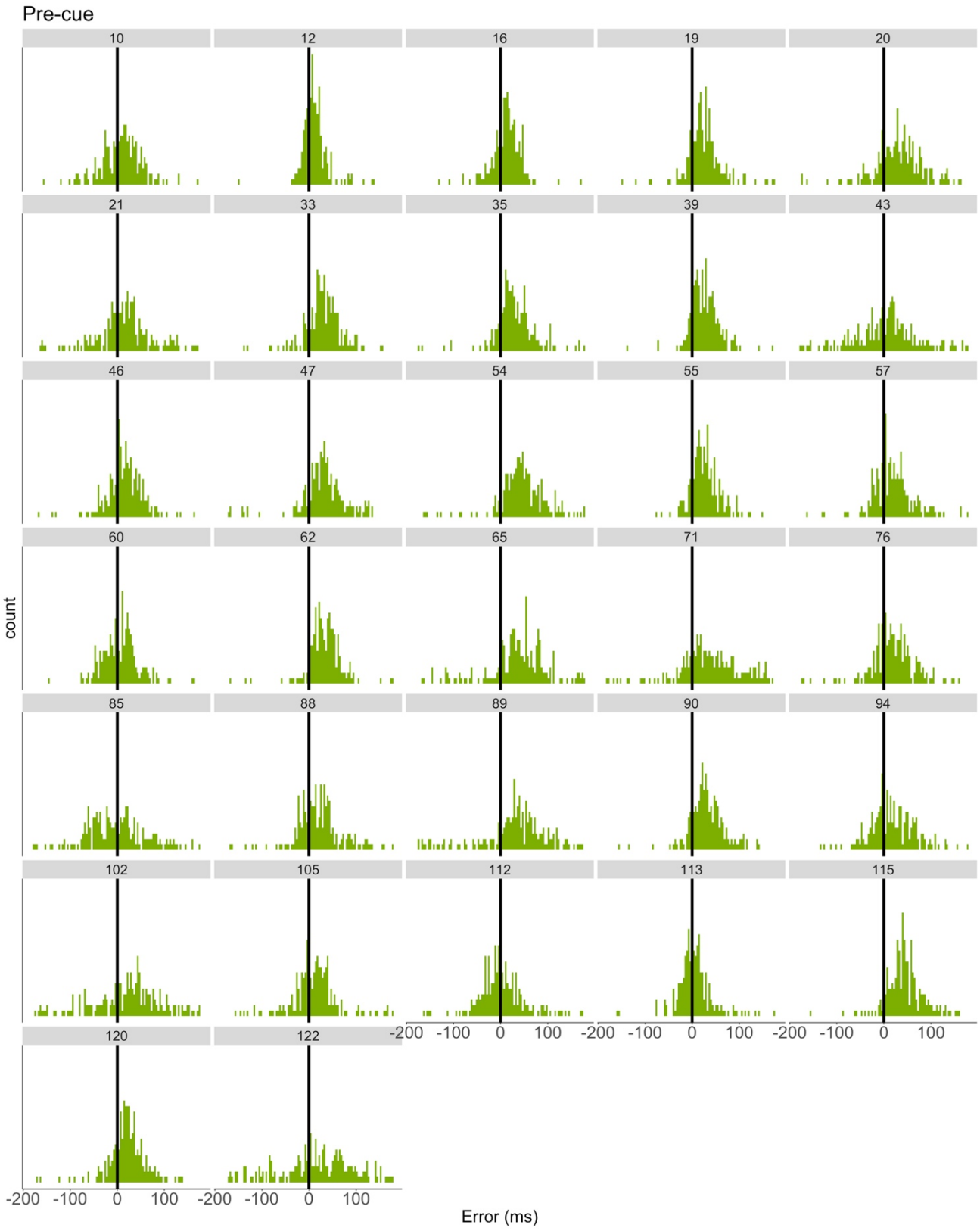
Regarding concentration, we found a main effect of eccentricity ($F(1, 19) = 13.56$, $MSE = 1.86$, $p = 0.002$), but no confidence x eccentricity interaction ($F(1, 19) = 2.98$, $MSE = 0.72$, $p = 0.10$) or condition x eccentricity interaction ($F(1.99, 37.72) = 0.73$, $MSE = 1.27$, $p = 0.49$). The triple interaction confidence x condition x eccentricity was also not significant ($F(1.46, 27.78) = 0.72$, $MSE = 1.97$, $p = 0.45$).

Therefore, we can conclude that eccentricity (or the clock's position relative to the horizontal meridian) affected perceptual performance via encoding precision, as reflected by the main effect on the concentration parameter. However, it did not affect metacognition, since the effect of eccentricity on precision did not interact with confidence level. The effect of eccentricity on performance also did not seem to depend on attention. To rule out any effect of eccentricity on our conclusions, in Experiment 2, all clocks were placed at the same eccentricity along a virtual circle centred on the fixation point. However, it would be interesting to consider the effect of peripheral presentations distance in future studies.

5. Raw histograms of Exp. 2

The excluded outliers are presented at the end.

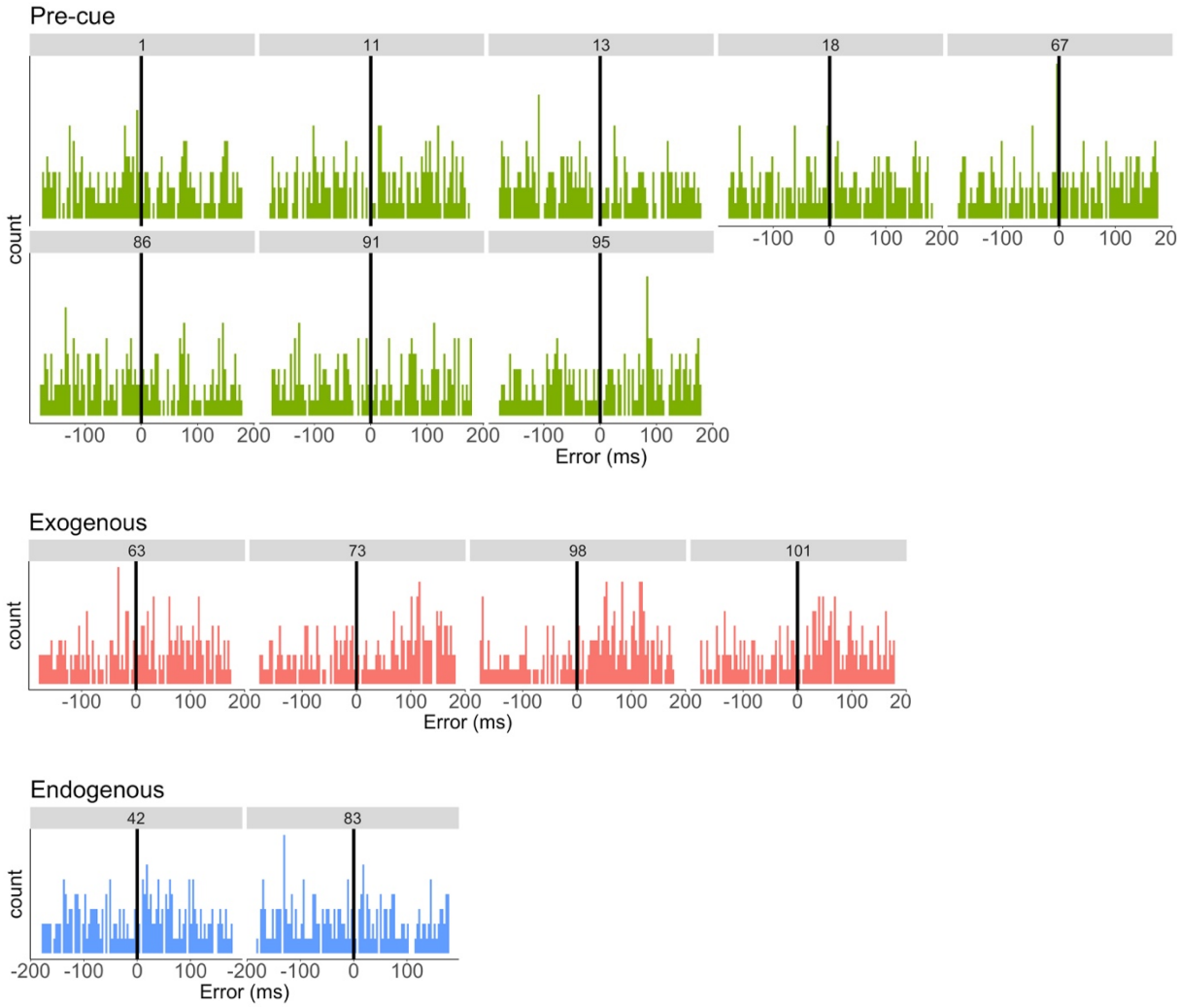
5.1.Included participants




Exogenous
2
3
4
6
9
22
24
25
27
30
34
37
38
40
44
45
50
53
56
59
61
66
68
70
74
77
79
81
92
93
99
100
104
107
117
119
count
-100
0
100
200
Error (ms)


Endogenous
5
8
14
15
17
23
26
28
29
31
32
36
41
48
49
51
52
58
64
69
72
75
78
80
82
84
87
96
97
103
106
108
110
111
114
116
121
123
count
-100
0
100
200
Error (ms)

5.2.Excluded participants (outliers)



REFERENCES

Carlson, T. A., Hogendoorn, H., & Verstraten, F. A. J. (2006). The speed of visual attention: What time is it? Journal of Vision, 6, 1406–1411. https://doi.org/10.1167/6.12.6

Chakravarthi, R., & VanRullen, R. (2011). Bullet trains and steam engines: Exogenous attention zips but endogenous attention chugs along. Journal of Vision. https://doi.org/10.1167/11.4.12

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. Nature Communications. https://doi.org/10.1038/ncomms2237

Hogendoorn, H., Carlson, T. A., VanRullen, R., & Verstraten, F. A. J. (2010). Timing divided attention. Attention, Perception, and Psychophysics, 72(8), 2059–2068. https://doi.org/10.3758/APP.72.8.2059

Ma, W. J. (2018). Problematic usage of the Zhang and Luck mixture model. BioRxiv. https://doi.org/10.1101/268961

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. Nature Human Behaviour, 1-17. https://doi.org/10.1038/s41562-020-00938-0

Van Den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.1117465109

Umbach, D., & Jammalamadaka, S. R. (2009). Building asymmetry into circular distributions. Statistics & probability letters, 79(5), 659-663. https://doi.org/10.1016/j.spl.2008.10.022

Schneegans, S., Taylor, R., & Bays, P. M. (2020). Stochastic sampling provides a unifying account of visual working memory limits. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(34), 20959–20968. https://doi.org/10.1073/pnas.2004306117

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature. https://doi.org/10.1038/nature06860