Supplemental Information

As supplemental information, we provide details of several additional analyses. Specifically, we show the data supporting the conclusion that the accuracy was a function of a lag of both more and less recent/imminent probe item. For JOR we show that incorrect responses deepened only on the lag of the selected probe. For JOI the results of the analysis of incorrect responses did not lead to strong evidence. Here we also provide broken out RT plots for each of the two experiments (figures in the main text included data from both experiments put together). We also show details of the analysis that compares performance in JOR and JOI in Experiment 2 (in Experiment 2 the same subjects did both tasks). We did not find a significant correlation for either RT, accuracy, or scanning rate.

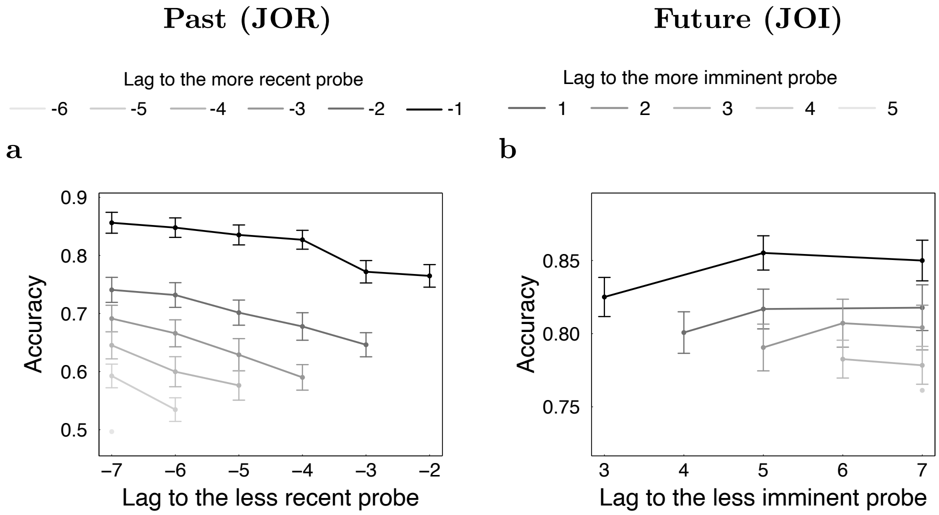

Figure S1. : **Accuracy in JOR and JOI. a.** Accuracy in JOR showed a reliable recency effect and a reliable distance effect. **b.** Accuracy in JOI was higher for probes that were closer in time to the present but did not show a reliable distance effect. See text for details.

JOR accuracy showed a recency effect and a distance effect

The probability that participants selected the more recent probe was $.70 \pm 0.01$ in Experiment 1 and $.69 \pm 0.01$ in Experiment 2 (mean accuracy across participants, taking into account only trials when participants responded by selecting one of the probes - excluding "UP" responses and trials that ended with no response). The accuracy was $.82 \pm .01$ when the lag of the more recent probe was -1 and dropped to $.49 \pm .02$ when the lag was -6 in Experiment 1. Similarly, in Experiment 2, the accuracy was $.81 \pm .02$ when the lag of the more recent probe was -1 and dropped to $.51 \pm .03$ when the lag was -6. In Experiment 1 at lag -6 the probability of choosing the more recent probe was not different from chance (Chi-squared prop test, $\chi^2(96) = 89.1$, p-value not significant). Lag -5 had an accuracy of 0.56 ± 0.01 and was significantly higher than chance (Chi-squared prop test, $\chi^2(96) = 142.6$, $p < 0.01$). In Experiment 2 at lag -6 and -5 the probability of choosing the more recent probe was not different from chance (for lag -6: Chi-squared prop test, $\chi^2(31) = 29.34$,

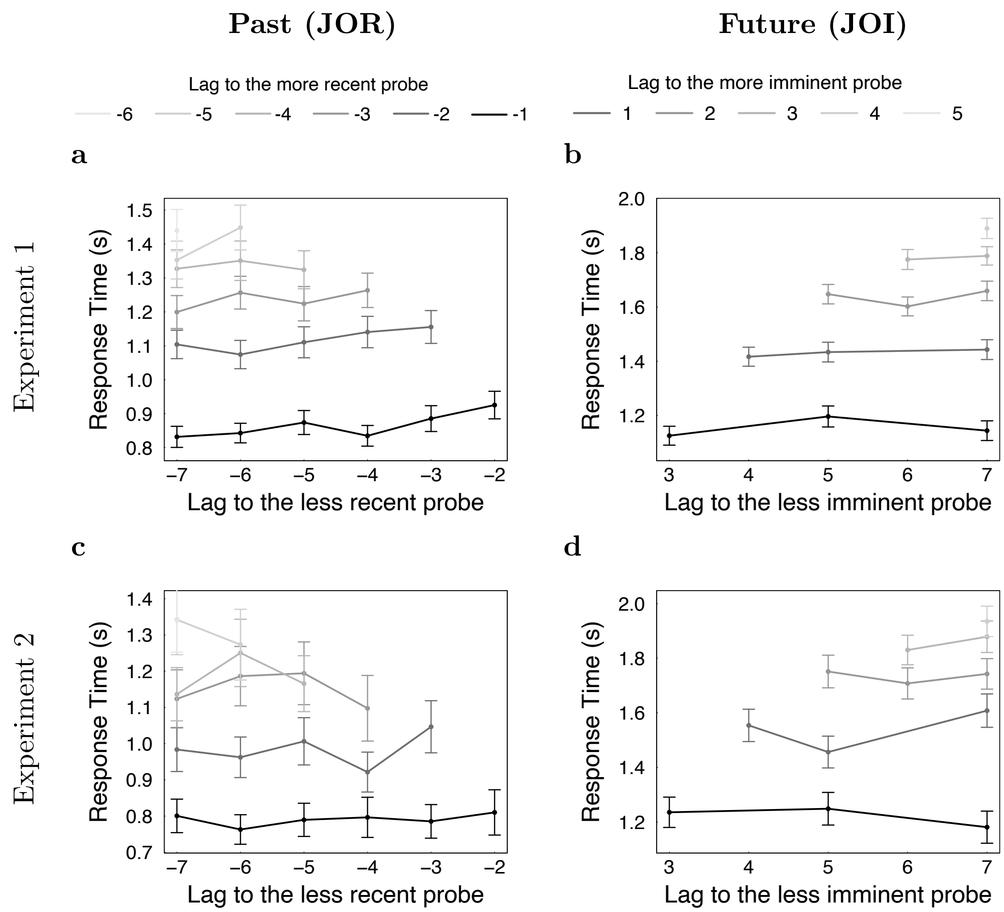

Figure S2. : **Analogous to Figures 3a and b but with results for each experiment shown separately.** Median correct RTs depend on the lag of the more recent/imminent probe but not on the lag of the less recent/imminent probe. This is consistent with a serial self-terminating search model.

p-value not significant; for lag -5: Chi-squared prop test, $\chi^2(31) = 44.28$, p-value not significant). Lag -4 had an accuracy of 0.60 ± 0.02 and was significantly higher than chance (Chi-squared prop test, $\chi^2(31) = 90.52$, $p < 0.01$).

Accuracy also depended on the temporal distance between the more recent probe and the less recent probe. For a lag of the more recent probe, the accuracy improved as the lag of the less recent probe increased (distance effect). The upward-sloping lines in Figures S4a and c indicate the presence of this distance effect. To quantify the distance effect for each participant, we calculated the slope of each line in Figure S4a and c. A Bayesian t-test (Rouder et al., 2009) on the obtained slopes revealed "decisive evidence" (Wetzels & Wagenmakers, 2012; Kass & Raftery, 1995; Jeffreys, 1998) favoring the hypothesis that the slopes are different from 0 (JZS Bayes Factor $> 10^2$ in both Experiment 1 and Experiment 2).

The effects of lags of the two probes on accuracy was quantified using a linear mixed

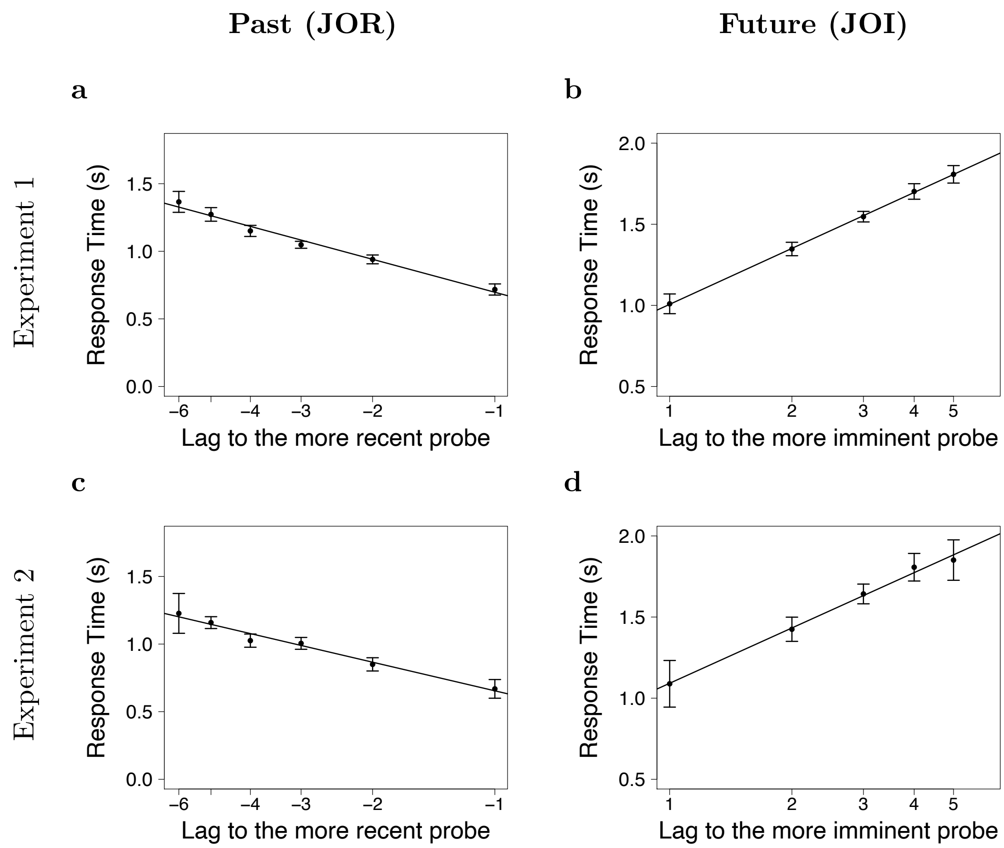

Figure S3. : **In each of the two experiments median correct RT varied sub-linearly with lag.** This figure is analogous to Figure 3c,d, but with results for each experiment shown separately.

effect analysis with independent intercepts for each participant. The accuracy increased with a decrease in the lag of the more recent probe by $.077 \pm .002$, $t(1918) = -31.9$, $p < 0.01$ per unit change in lag in Experiment 1 and by $.070 \pm .004$, $t(618) = -16.49$, $p < 0.01$ per unit change in Experiment 2. Accuracy also increased with the lag of the less recent probe by $.023 \pm .002$, $t(1918) = 9.73$, $p < 0.01$ per unit in Experiment 1 and $.028 \pm .004$, $t(618) = 6.56$, $p < 0.01$ per unit in Experiment 2. These findings are consistent with the findings from prior studies.

In JOR we observed that less recent item had an impact on accuracy, but not on RT, causing the distance effect only on accuracy. To explain this in the context of serial scanning mechanisms we discuss two possible causes for participants to make an incorrect judgment according to scanning models. First, participants might not properly encode the more recent probe and simply scan past it until reaching the less recent probe. In that case, the lag of a less recent probe does not impact the accuracy. Second, the noise level in the evidence accumulation process might be large enough such that participants do not accumulate enough evidence for the more recent probe when they reach the less recent probe

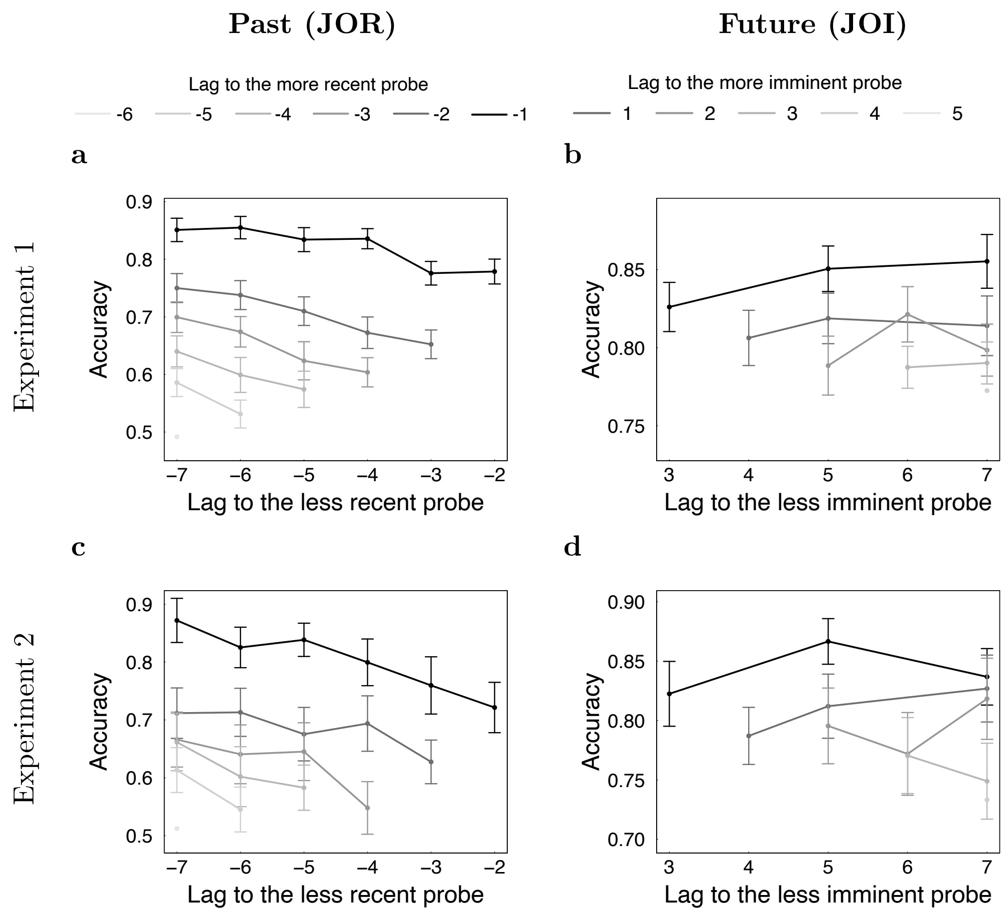

Figure S4. : **Accuracy in JOR and JOI across the two experiments.** This figure is analogous to Figure S1, but with results for each experiment shown separately. In both experiments, accuracy in JOR showed a reliable recency effect and a reliable distance effect. Also, in both experiments accuracy in JOI was higher for probes that were closer in time to the present but did not show a reliable distance effect.

during the scanning process. Consistent with the second explanation, the results reported here support a hypothesis that items on the timeline are represented with broad overlapping fields that have sequential peaks.

Incorrect RT depended only on the lag of the selected probe in the JOR task

In a self-terminating backward scanning model, if the scan misses the more recent probe, it would then terminate on the less recent probe. These responses would be errors and the scanning time for these errors would depend on the lag to the less recent probe.

Given the overall error rate of .30 $\pm$ 0.01 in Experiment 1 and .31 $\pm$ 0.01 in Experiment 2, there was less than half the number of observations for incorrect RTs as there was for correct RTs. Also note that the number of errors was not evenly distributed over lags, so that some points have many fewer observations than others (Figure S5). Nonetheless, error

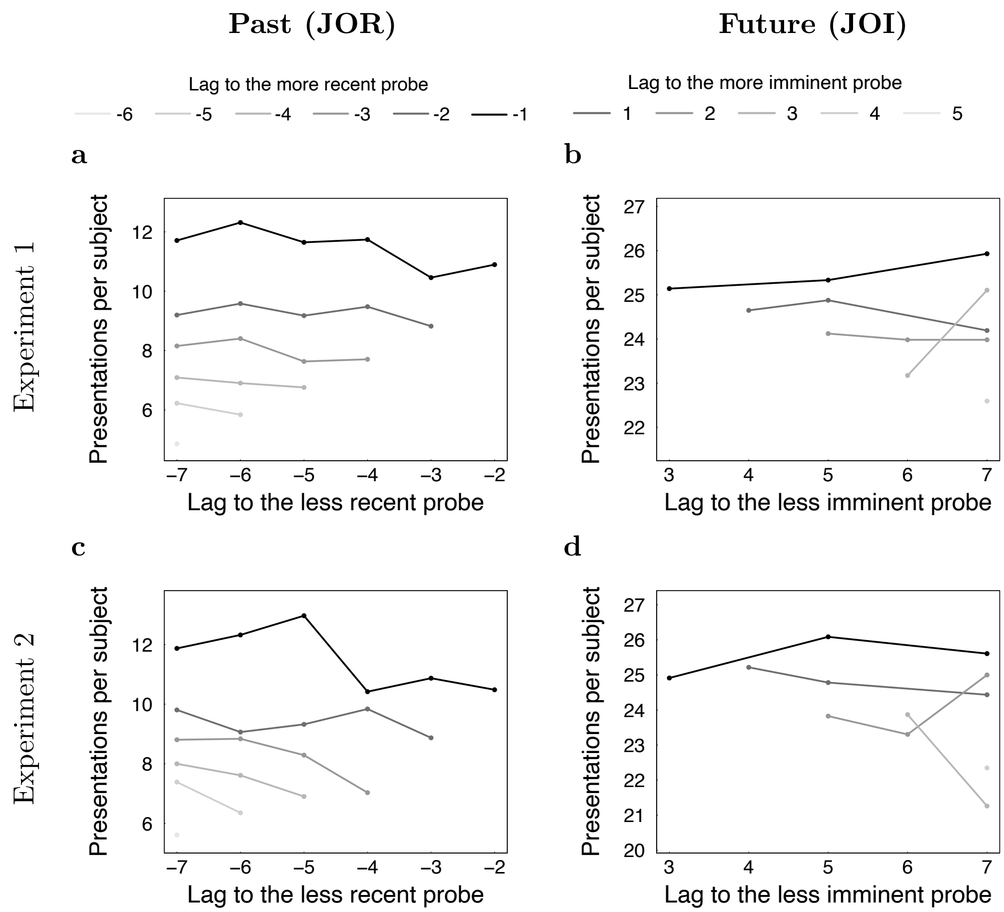

Figure S5. : **Number of probe presentations for every available combination of more and less recent/imminent lag in correct trials.** Since in JOR participants had an option to press "UP" button to indicate that they do not remember seeing either of the probes, the number of correct presentations was lower in JOR compared to JOI.

RTs appear to depend reliably on the lag to the less recent item. There does not appear to be a strong effect of the lag to the more recent probe.

To evaluate whether there was an effect of the lag to the more recent probe we calculated the slope of the distance effect for each value of the lag to the less recent probe (see also Figures S8a and c). This is analogous to the distance effect calculation for correct RTs except the distance effect is calculated separately for the lag to the less recent probe rather than the more recent probe. A Bayesian t-test showed "strong evidence" favoring the hypothesis that the slopes of the median RTs as a function of the more recent lag are not different from 0 (JZS Bayes Factor = 14.5) in Experiment 1 and "substantial evidence" (JZS Bayes Factor = 5.79) in Experiment 2.

A linear mixed effects analysis, allowing each participant to have an independent intercept, and the less recent lag as regressor showed a significant effect of the lag to the less recent probe on the median RT of incorrect responses in Experiment 1 ($.033 \pm .007$ s,

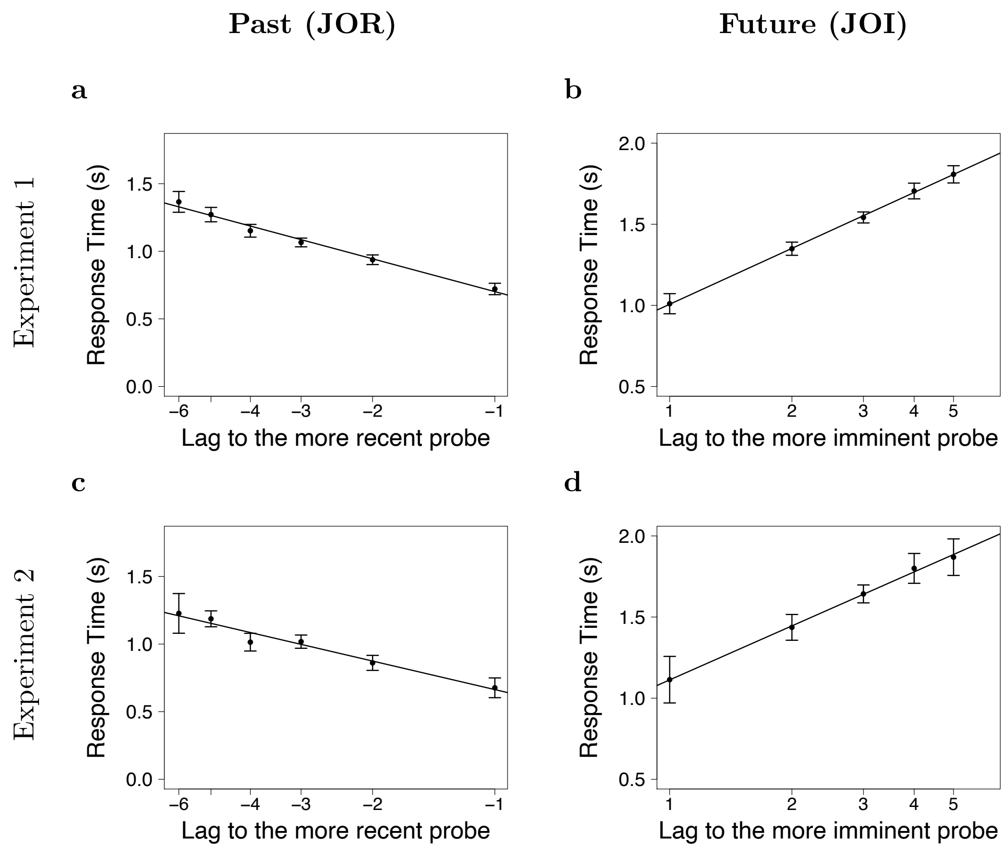

Figure S6. : Sublinear relationship between lag and RT was preserved when the dataset was modified such that each pair of more and less recent/imminent lag had the same number of correct trials (the number of observations for each pair was reduced to the number of observations of a least frequent pair by performing random sampling).

$t(473) = 4.9$, $p < 0.001$), but not in Experiment 2 ($.019 \pm .012$ s, $t(153) = 1.57$).

JOI accuracy depended only on the lag of the more imminent probe

The overall accuracy in JOI was $.81 \pm .02$ in Experiment 1 and $.80 \pm .03$ in Experiment 2. In Experiment 1 the accuracy varied from $.84 \pm .02$ when the most imminent lag was 1 to $.77 \pm .02$ when the most imminent lag was 5. In Experiment 2 the accuracy varied from $.84 \pm .03$ when the most imminent lag was 1 to $.73 \pm .02$ when the most imminent lag was 5.

In order to check if the somewhat upward-sloping lines in Figures S4b and d indicate the presence of a reliable distance effect, the slopes of each of the lines were calculated for every participant. A Bayesian t-test (Rouder et al., 2009) on the obtained slopes revealed “substantial evidence” favoring the hypothesis that the slopes are not different from 0 (JZS Bayes Factor 7.2 in Experiment 1 and 8.53 in Experiment 2).

Thus the accuracy depended on the lag to the more imminent probe alone. The effects of the more imminent probe on accuracy were quantified using a linear mixed analysis with independent intercepts for each participant. Accuracy decreased with an increase in the

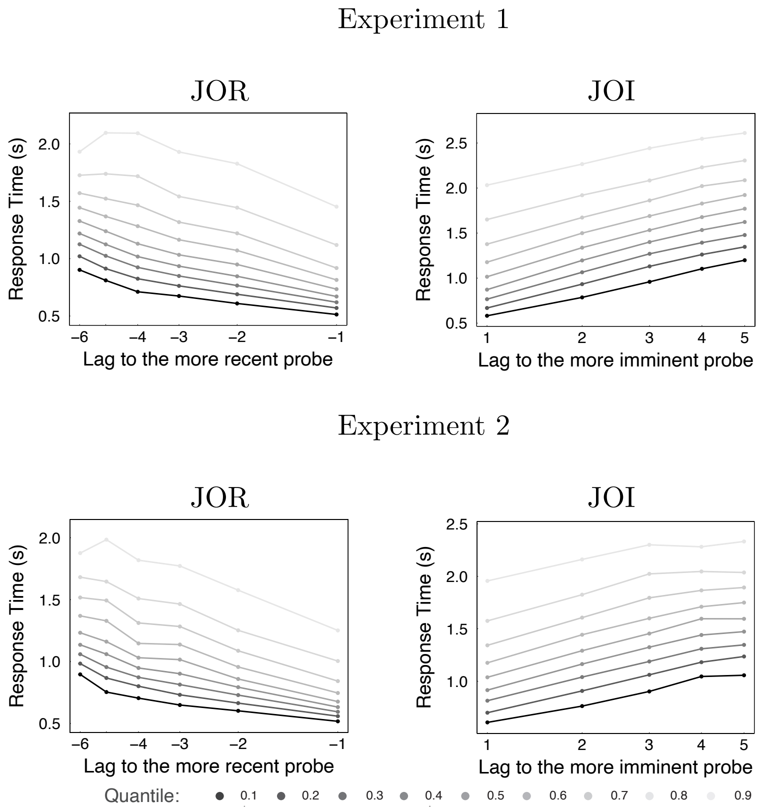

Figure S7. : **Response time as a function of lag for every decile across all the participants separated per experiment.** The results were consistent across the two experiments. See Figure 7 and the main text for more details.

lag to the more imminent probe by $.017 \pm .002$, $t(227) = -6.9$, $p < 0.01$ per unit change in lag in Experiment 1 and by $.026 \pm .004$, $t(91) = -6.9$, $p < 0.01$ in Experiment 2. The finding that participants were more accurate based on the lag of the more imminent probe is different from the accuracy results seen in JOR (Hacker, 1980) where accuracy depends on both the more recent and the less recent probes. Keeping in mind that JOI had fewer pairs of more and less imminent lags than JOR, an additional reason for this finding could be in the fact that in JOR participants could use "UP" button to indicate that they do not remember seeing the probe items in a given trial. Availability of "UP" button in JOR and not JOI was motivated by the fact that in JOI participants have seen all the items many times. On the other hand, if participants did not pay much attention during a particular JOR trial, they might not remember seeing the probe items. However, this also resulted in participants making more guesses in JOI than in JOR, making it harder to find the distance effect.

Incorrect RT showed mixed results in the JOI task

When participants make errors, the forward scanning model predicts that the RT should depend on the lag of the less imminent probe. However this was not observed

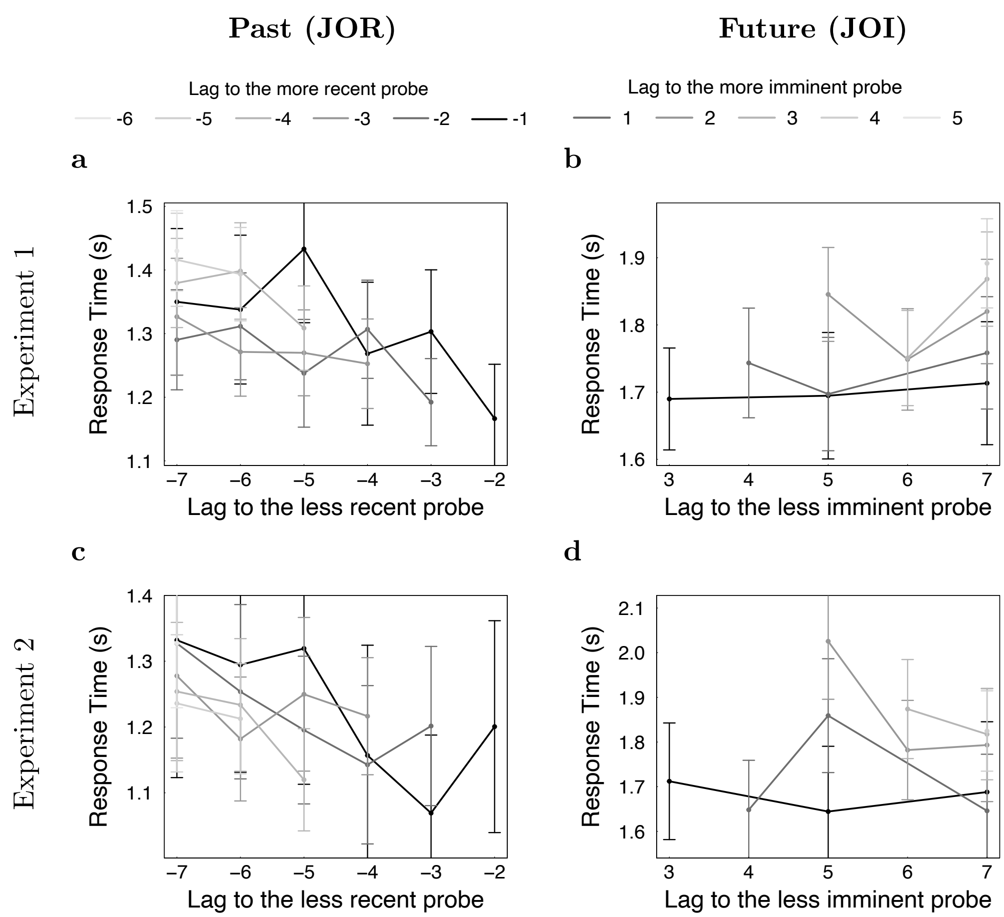

Figure S8. : **RT as function of the lag of the more recent/imminent probe for trials where participants made an incorrect response.**

in the error responses here, perhaps because accuracy was much higher than in JOR. In Experiment 1 the overall error rate was $.19 \pm .02$ and for each participant on average 6 error responses per condition were obtained. In Experiment 2 the overall error rate was $.20 \pm .03$ and for each participant on average 3 error responses per condition were obtained.

To evaluate whether the more imminent probe contributed to the change in median RTs of a particular less imminent probe (distance effect), the slope across the more imminent lag for each less imminent lag was calculated (see also Figures S8b and d). This is analogous to the distance effect calculation for correct RTs except the distance effect is calculated separately for the lag of the less imminent probe rather than the more imminent probe. The less imminent lags of 3 and 4 only have one possible combination so a slope could not be calculated for these lags. Across the remaining slopes, a Bayesian t-test showed that the evidence for a distance effect was "barely worth a mention" in both Experiment 1 and Experiment 2 (JZS Bayes Factor = 1.1 in both experiments).

Given the overall error rates of .19 and .20, there was less than one fourth the number of observations for incorrect RTs as there was for correct RTs. While a linear mixed effects analysis, allowing each participant to have an independent intercept, and both lag as regressors, showed a significant effect of the lag of the more imminent probe on the median RT of incorrect responses, $.05 \pm .05$ s, $t(586) = 2.9$, $p < 0.01$ in Experiment 1 and $.05 \pm .03$ s, $t(228) = 2.1$, $p < 0.05$ in Experiment 2, a linear model with both the more and the less imminent lag did not show any significant effect of the more or the less imminent probe, $.04 \pm .02$ s, $t(641) = 1.7$ and $.01 \pm .02$ s, $t(641) = .4$ respectively in Experiment 1 and $.06 \pm .04$ s, $t(250) = 1.4$ and $.00 \pm .04$ s, $t(250) = .0$ in Experiment 2.

Comparison of cross-participant performance in JOR and JOI

Out of the 32 participants that performed JOR and JOI in Experiment 2, 22 were above chance in both tasks. We investigated the correlation of different measures related to response time and accuracy in JOR and JOI across those 22 participants. Briefly, we did not find reliable correlations in any of the dependent measures we considered. RT in JOR was not correlated with RT in JOI, $r(20) = .41$, $p = .050$. Similarly, accuracy was not correlated across participants in JOR and JOI, $r(20) = .18$. We also computed the slope of RT as a function of log lag of a more recent/imminent probe for each participant and found no correlation between JOR and JOI, $r(20) = .01$.

Individual participant results

Additional supplementary material shows the results for each participant in each of the two experiments. For participants in Experiment 2, results from JOR and JOI are shown side by side for easy comparison.

Data availability

The entire dataset is available on Open Science Framework website: “https://osf.io/ur8yw/?view˙only=2d279249c1464de096e121dcb49bd89f”.