Adaptive encoding speed in working memory

Joost de Jong, Hedderik van Rijn, Elkan G. Akyürek

Experimental Psychology, University of Groningen

Abstract

Humans can adapt when complex patterns unfold at a faster or slower pace, for instance when remembering a grocery list that is dictated at an increasingly fast pace. Integrating information over such timescales crucially depends on working memory, but although state-of-the-art theories assume that working memory capacity can be flexibly adapted, these theories assume that encoding speed cannot. In a series of experiments, we found that humans encoded at a faster rate when they were adapted to overall and recent rate of incoming information. Interestingly, our participants were unable to use explicit cues to speed up encoding, even though these cues were more informative than statistical information. Our findings suggest that adaptive tuning of encoding speed in working memory is a fundamental but largely implicit mechanism underlying our ability to keep up with the pace of our surroundings.

Whether it is following an increasingly rapid story told by an excited friend or listening to a piece of music that slows down its tempo; humans are experts at keeping up with the pace of our surroundings. For instance, humans rather easily comprehend speech at two-fold increases in speech rate (Foulke & Sticht, 1969). But what underlies our ability to track information efficiently over such a wide range of 'information rates'? A common bottleneck in domains like music and speech is working memory (Schulze & Koelsch, 2012). Ideally, information should be encoded sufficiently before the stimulus is no longer there, especially when interfering stimuli may follow. Intuitively, if we expect information to arrive at a fast rate, information should also be encoded at a faster rate.

Despite this intuitive idea that we can adapt working memory encoding speed to the expected information rate, such adaptations have not been tested empirically. This hiatus is surprising for three reasons. First, adaptive dynamics have been demonstrated in a wide variety of domains, such as sensory adaptation (Fairhall et al., 2001; Wark et al., 2009), decision-making (Murphy et al., 2021; Ossmy et al., 2013), timing (Remington et al., 2018), event integration (Akyürek et al., 2008), speech perception (Lerner et al., 2014), learning (Behrens et al., 2007; Piray & Daw, 2020), and motor adaptation (Gonzalez Castro et al., 2014). Overall, these lines of research show that the speed of cognitive (or neural) processes adjusts to the information rate of, or the rate of change within, the environment. For instance, the rate of evidence accumulation in decision-making speeds up when incoming evidence changes more frequently (Glaze et al., 2015).

Second, from its very inception (e.g., Sperling, 1963), working memory research has had a strong focus on encoding speed (see also, Bays et al., 2011; Bundesen,

1990; Busey & Loftus, 1994; Gegenfurtner & Sperling, 1993). However, while many factors have been found to modulate encoding speed, such as luminance/contrast (Loftus & Ruthruff, 1994), attention (Reinitz, 1990), and working memory load (Bays et al., 2011; Vogel et al., 2006), none of these qualifies as an adaptation to the temporal structure of the environment. One possible exception is the finding that encoding speed is modulated by the overall foreperiod preceding stimulus onset (Vangkilde et al., 2012, 2013). However, these adaptations are with respect to properties that facilitate preparation for memoranda, not the temporal properties of the memoranda themselves.

Third, adaptive processing in working memory has been observed with regard to its capacity. State-of-the-art computational models assume that capacity is adaptively allocated to memoranda (e.g., van den Berg & Ma, 2018). Indeed, it has been shown empirically that working memory capacity adapts to stimulus statistics (Orhan et al., 2014) and can be flexibly allocated to information that is likely to be relevant (Bays & Husain, 2008). However, none of these accounts have considered that encoding speed may adapt to the rate at which information arrives at the senses.

In this paper, we study whether and how working memory encoding speed, in a similar way as capacity, is adapted to the structure of the environment. Specifically, we hypothesize that encoding speeds up when the expected information rate is high. In a series of experiments, we leverage distribution effects, serial dependence, and cueing to systematically manipulate expectations about information rate. To preview our results, we found that humans encode information about twice as fast when they expect information to arrive at a fast rate versus a slow rate, but only when these expectations

were induced implicitly. These findings suggest that implicit encoding speed adaptations may underlie our ability to keep up with the pace of our surroundings.

Results

Our main hypothesis is that the speed of working memory encoding adapts to the expected rate of incoming information. To estimate how quickly and precisely humans encode information, we need to characterize how working memory precision varies with how long information is presented. To this end, human participants performed delayed estimation experiments, where a masked Gabor patch was presented and reproduced after a 1-second delay on each trial (see Figure 1). Presentation times (t) varied between 50 and 400 ms. From the response errors (see Figure 1 inset), we computed memory precision (inverse of circular standard deviation; σ^{-1}) for each presentation time (Figure 1; lower panel). For each participant in each condition, we fitted an exponential encoding curve (e.g., Bays et al., 2011) with two parameters: Maximum capacity (c) and encoding speed (τ^{-1}) (Figure 1 lower panel):

$$\sigma^{-1}(t) = c\,(1 - e^{-t/\tau}) \quad (1)$$

These parameters map onto the asymptotic memory precision for very long presentation times (c), and encoding speed (i.e., the speed at which precision reaches that asymptote; τ^{-1}).

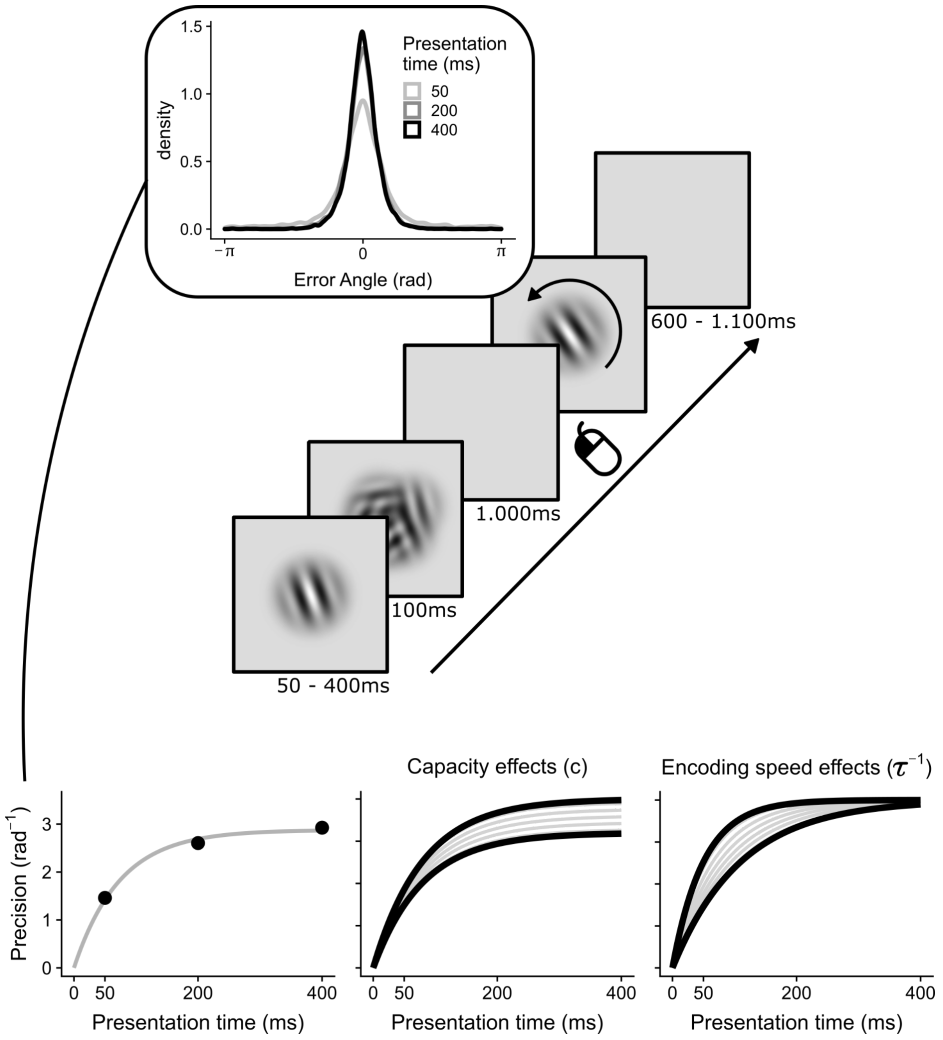

Figure 1. Experiment and analysis setup. Top: experiment setup. Participants had to remember the orientation of a Gabor that is presented for 50, 200 or 400ms and immediately followed by a pattern mask. After 1.000ms, participants had to adjust the orientation of a second Gabor to reproduce the contents of their memory. Inset: the distribution of errors (i.e., reproduced angle minus target angle). Bottom: We computed the inverse of the circular standard deviation of those errors to characterize memory precision, which smoothly increased with presentation time. An exponential encoding curve is fitted (the leftmost panel), which has two parameters: Maximum capacity and

encoding speed. The middle panel expresses differential effects of capacity, whereas the right panel depicts the effects of encoding speed.

With this definition of encoding speed, we now turn to the definition of (expected) information rate. We define information rate as the amount of information per time unit. Therefore, if we keep the amount of information (e.g., signal-to-noise ratio) constant over the presentation time, then the presentation time of information is the sole determining factor of information rate. To illustrate, when the *same* number of letters is presented in a shorter period of time, the information rate is higher. Similarly, the *encoding rate* will need to be higher to successfully encode all information compared to when this same sequence is presented over a longer period of time.

We used multiple linear regression to assess whether expected information rate specifically predicts normalized encoding speed independent of changes in normalized maximum capacity. To assess whether an individual participant showed changes in encoding parameters, we normalized parameter estimates for each participant by dividing them by the mean parameter estimate of that individual. To quantify evidence for or against the null-hypothesis, Bayes Factors for the effect of expected information rate were computed as follows (Wagenmakers, 2007):

$$BF_{10} = e^{(BIC_0 - BIC_1)/2}$$

where BIC_0 is the Bayesian-Information Criterion score for the null-model only containing c_{norm} as a predictor and BIC_1 is the Bayesian-Information Criterion score for the model that also includes expected information rate. High values for BF_{10} reflect

evidence against the null and therefore suggest an effect of information rate on encoding speed. When confounding factors were tested, they were always added and compared to a model that included both normalized capacity and information rate as predictors. Lastly, since the estimate of the circular standard deviation is systematically biased for low sample sizes, we bootstrapped (without replacement) 100 times while matching the number of observations across conditions for experiments with unbalanced designs (experiment 1 and 3). We reported the median of all test statistics (regression coefficients, *t*-values, p-values, Bayes Factors). All experimental materials, data and analysis code are available on the Open Science Framework (osf.io/d5rnu/).

Encoding speed adapts to overall information rate

In our first experiment, we tested whether humans could increase their encoding speed when the overall information rate was high. Humans are sensitive to the overall distribution of temporal intervals that they perceive (e.g., Jazayeri & Shadlen, 2010). Therefore, if participants tune their expectations to the overall stimulus duration, they should also tune their expectations about overall information rate and adapt their encoding speed accordingly. To test this hypothesis, we varied the overall information rate between blocks. Participants alternated between blocks of 160 trials where stimulus durations were either short (50ms) or long (400ms) for 70% of the trials. We hypothesized that, if humans adapt to overall information rate, they should have a higher encoding rate in the ‘short duration / high information rate’ blocks compared to the ‘long duration / low information rate’ blocks. In line with our hypothesis, encoding speed was systematically faster in blocks with a high information rate (β = 0.149, t =

3.763, $p < 0.001$, $BF_{10} = 89.82$; Figure 2). These results suggest that our participants did adapt their encoding rate to the overall information rate.

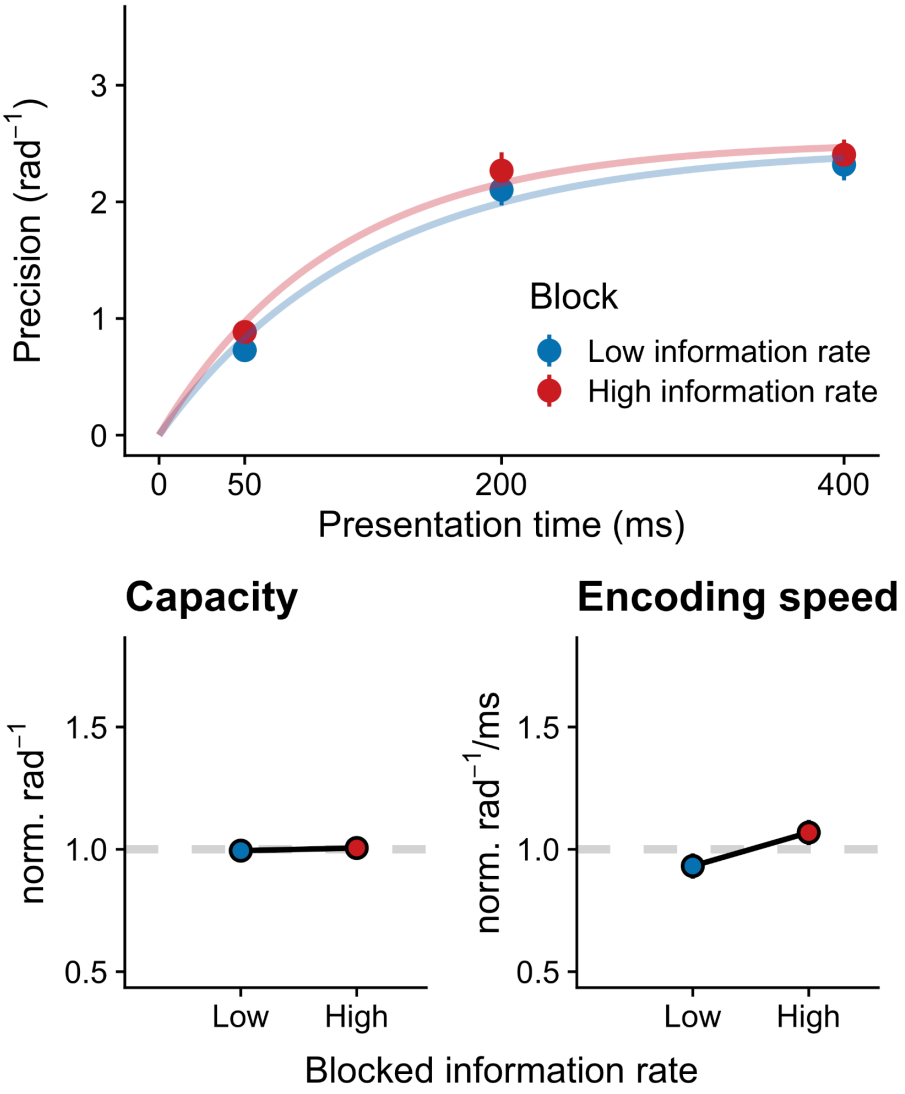

Figure 2. The effect of blocked information rate on working memory capacity and encoding speed. The top figure shows how memory precision (mean ± SE) increases with presentation time, which is accurately captured by encoding curves. Blocks either had high information rate (red; 50ms presentation time, 70%) or low information rate (blue; 400ms presentation time, 70%). The bottom figure shows how normalized

parameter values (mean ± SE, with an overall mean of 1) vary with rate. The left panel represents capacity and the right panel represents encoding speed.

Encoding speed adapts to recent information rate

In our second experiment, we tested how quickly humans can adapt their encoding rate on a trial-to-trial basis. To this end, we leveraged serial dependence observed in the perception of durations. Humans can optimally track stimulus duration on a trial-to-trial basis (de Jong et al., 2021; Glasauer & Shi, 2021). For instance, when a previous stimulus was presented briefly, the next stimulus is also expected to be relatively brief. Therefore, if the previous presentation time was brief (thus requiring a high encoding speed due to the high information rate), participants could expect a high information rate on the next trial and increase their encoding speed accordingly.

In line with our hypothesis, we found that when the previous information rate was high (i.e., the previous presentation time was brief), the encoding speed on the current trial was also high (Figure 3; $\beta = 0.013$, $t = 3.705$, $p < 0.001$, $BF_{10} = 80.96$). This finding suggests that humans speed up encoding when they expect that they have relatively little time for encoding. In line with continued adaptation, we also found faster encoding when the previous two stimuli had the same high information rate (Figure 4; $\beta = 0.037$, $t = 6.017$, $p < 0.001$, $BF_{10} > 10{,}000$). Notably, a comparison of the regression coefficients (Clogg et al., 1995) showed that the effect of the previous two stimuli was stronger than the effect of the most recent stimulus alone ($Z = 3.376$, $p < 0.001$).

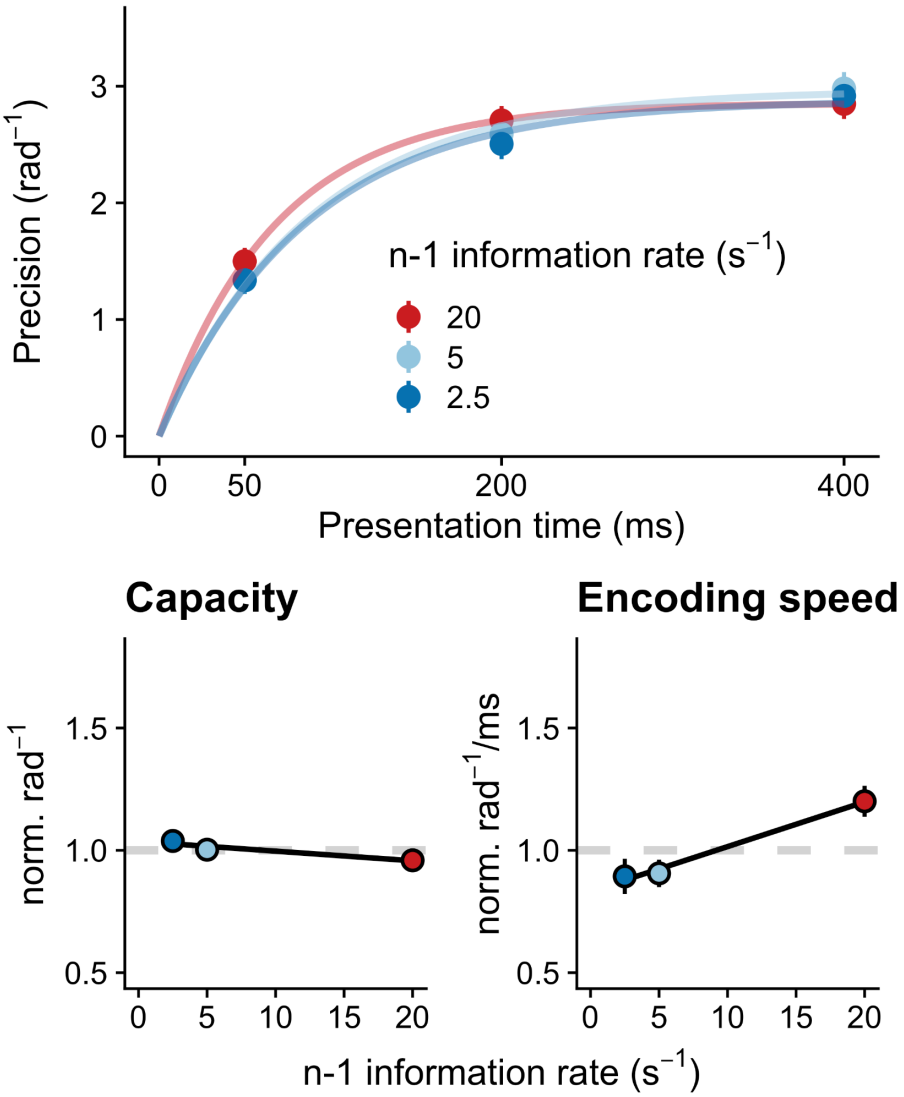

Figure 3. The effect of n-1 information rate on working memory capacity and encoding speed. The same plotting conventions are used as in Figure 2. Colors indicate whether the previous trial had a long presentation time / low information rate (blue), or brief presentation time / high information rate (red).

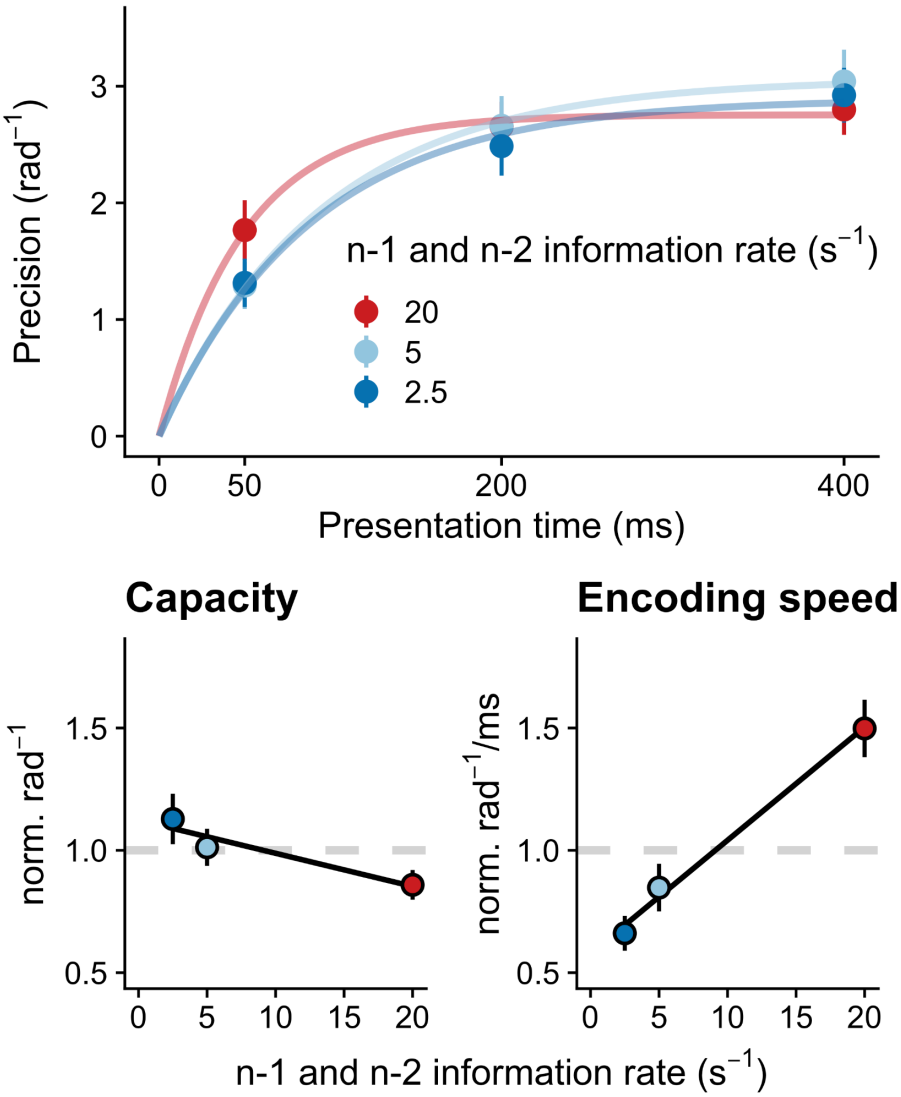

Figure 4. The effect of n-1 and n-2 information rate on working memory capacity and encoding speed. The same plotting conventions are used as in Figure 2. Colors indicate whether the previous two trials both had a long presentation time / low information rate (blue), or brief presentation time / high information rate (red).

Our hypothesis predicts that encoding speed *specifically* adapts to expected information rate, and not temporal variables that are unrelated to the memory items. Indeed, we found evidence that adaptation was specific to information rate, not the rate at which the task proceeds after a response. The previous interval between fixation onset and stimulus onset did not significantly modulate encoding speed (β = 0.098, t = 1.432, p = 0.155, BF_{10} = 0.247; Supplementary Figure 1) or interact with the effect of previous information rate (β = 0.008, t = 0.873, p = 0.384, BF_{10} = 0.129). In the same linear regression model, the effect of previous information rate was consistent (β = 0.014, t = 3.180, p = 0.002, BF_{10} = 13.104). This finding may seem at odds with reports in the literature that encoding speed in visual detection is modulated by expected foreperiod (Vangkilde et al., 2012, 2013). However, unlike these studies, we did not manipulate foreperiod in a block-wise manner and not over such a broad temporal scale (0.3 - 10s). Hence, while expected foreperiod could have some effect on encoding speed, it is unable to account for encoding speed adaptations specific to *information* rate. Overall, our findings suggest that humans specifically track information rate on a trial-to-trial basis and adapt their encoding speed accordingly.

An alternative explanation for our adaptive encoding findings is that participants adjusted their encoding speed to correct for errors on previous trials (similarly to post-error improvements of performance, Danielmeier & Ullsperger, 2011). When large errors were made on previous trials, participants may have tried to compensate by increasing their encoding speed. Hence, this account would predict a faster encoding speed following trials with large errors. However, when we did a median split on the magnitude of the errors, error magnitude on previous trials did not modulate encoding

speed on current trials. When participants made large errors on previous trials, their encoding speed on subsequent trials was the same, or if anything slightly lower (β = -0.121, t = -1.686, p = 0.094, BF_{10} = 0.374; Supplementary Figure 2). In sum, it seems that adaptations in encoding speed cannot be explained by error-related corrections.

Encoding speed is insensitive to explicit cues about information rate

So far, we have shown that encoding speed is adapted to expectations about information rate induced by distribution- and sequential-like effects. However, it is unclear whether expectations about information rate can also be induced explicitly, like with spatial cueing (e.g., Jonides, 1981). We ran an experiment where, at the start of each trial, participants were cued (‘FAST’ or ‘SLOW’) whether the upcoming stimulus would be brief (50ms) or long (400ms) with 80% cue validity. Crucially, even though these cues were more informative than our previous sequential and block-wise manipulations, they did not modulate encoding speed. If anything, ‘fast’ cues resulted in a decrease in encoding speed (Figure 5; β = -0.064, t = -0.766, p = 0.321, BF_{10} = 0.306). These findings are in line with several reports in the literature showing that explicit cues are relatively ineffective in guiding expectations about temporal features (Los et al., 2021; Maaß et al., 2019). Indeed, some have shown that temporal expectations are efficiently triggered by implicit associative cues (Salet et al., 2022), which might also be the case for expectations about information rate. However, this claim awaits further empirical support. In sum, the ineffectiveness of highly informative cues suggests that adaptation to information rate is more easily triggered by implicit processes, than by explicit cues.

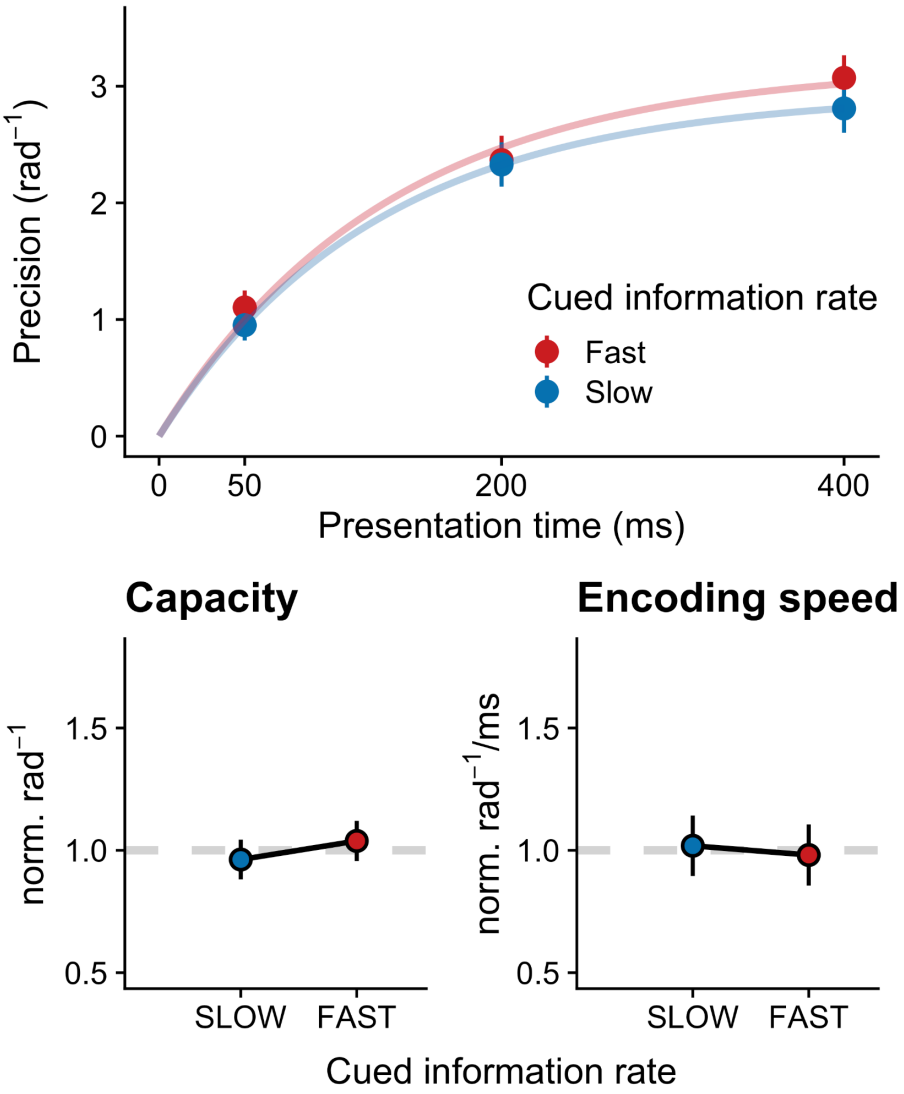

Figure 5. The effect of cued information rate on working memory capacity and encoding speed. Cues either predicted high or low information rate (50 vs 400ms presentation time, 80% validity). The same plotting conventions are used as for Figure 2.

Discussion

How do we keep up with the rate at which information reaches our senses? We hypothesized that a fundamental underlying mechanism may be the adaptation of working memory encoding *speed* to expected information rate. Indeed, we found that encoding speed was adapted to overall information rate. Moreover, encoding speed continually and specifically adapted to recent information rate on a trial-to-trial basis. We further showed that encoding speed adaptations could not be induced by highly informative cues, suggesting that these adaptations may be largely implicit. Overall, our findings suggest that the speed of working memory encoding is optimally adapted to the timescale of incoming information.

Our hypothesis was partly inspired by adaptive dynamics that are found in a wide range of perceptual, cognitive and motor domains. Indeed, we demonstrate for the first time that encoding speed in working memory may show similar adaptation to the timescale of incoming information. But what is the underlying logic behind adapting the speed of cognitive processes to match the timescale of the environment?
State-of-the-art theories of adaptive dynamics assume that task-relevant variables are tracked in an optimal way. These models propose that, in addition to tracking the mean value of some task-relevant variable, humans also track how quickly that variable is changing (Glaze et al., 2015; Ossmy et al., 2013; Piray & Daw, 2020; Wark et al., 2009). As an example (see Ossmy et al., 2013), when trying to spot a faint signal in the distance, we not only take into account its luminance, but also its duration. In other words, we keep track of how quickly luminance is changing. If we expect a brief signal, we quickly accumulate evidence before the signal is absent. However, this comes at the

expense of accumulating noise. Therefore, if we expect a long signal, we accumulate evidence at a slower rate to prevent false alarms. Indeed, in a decision-making task, Ossmy et al., (2013) found that the rate of evidence accumulation is adapted to the overall duration of the to-be-detected signal. We have shown an analogous encoding speed adaptation in working memory. Given these commonalities, it seems that these formal models of adaptive dynamics (e.g., Glaze et al., 2015) could be a powerful way to conceptualize encoding speed adaptations in working memory.

Another possible idea is that, while increments in encoding speed may improve performance, they incur a commensurate increase in some behavioral or neural cost (for a similar 'resource-rational' account of working memory capacity, see van den Berg & Ma, 2018). Whatever form these costs may take, a natural assumption is that humans attempt to maximize their net gain (i.e., performance *minus* 'encoding costs'). The second assumption is that participants weigh performance at certain presentation times by their subjective probability. That is, if participants believe that very brief presentation times are more likely (e.g., when the previous duration is short or in the 'short' blocks), they put more weight on performance at brief presentation times. These simple assumptions would naturally predict that increasing the rate of encoding is beneficial when high information rates are expected.

Given that humans can speed up working memory encoding, how might the brain achieve such a speedup? The analysis of our behavioral data provides a natural clue: The encoding curve we fit to our data (Equation 1) describes the behavior of a leaky integrator that is given a step-like input. Here, the amplitude of the input corresponds to the maximum capacity, while the speed of integration relates to encoding speed. It

would seem that, since we observed independent modulations in encoding speed, the speed of neural integration would be responsible for our effects. However, a plausible alternative is that amplified inputs do not only result in a higher measured capacity, but if capacity has a limit, some of these effects also show up in the encoding rate. To put it simply, a stronger input will drive memory precision to the limit at a faster rate without resulting in much change in that limit. Then, we would expect that increases in encoding speed would be accompanied by little to no increase in capacity. However, we found the opposite: Normalized encoding speed generally had a negative effect on normalized capacity (Supplementary Figure 3). That is, when encoding speed increases, capacity generally decreases. Indeed, this suggests that our results may be explained by a speedup in how quickly populations of neurons change their overall firing pattern (Sohn et al., 2019) or perhaps a speedup in biophysical processes in individual neurons (Durstewitz, 2003), not necessarily changes in input amplitude.

A related question pertains to the locus of encoding speed adaptations. Can our findings be explained by faster temporal summation in the visual system (Loftus & Ruthruff, 1994), as opposed to adaptations in working memory per se? The applicability of visual temporal summation to our results, however, may be limited. If temporal summation is completed around ~100ms (Gorea, 2015), it is unclear why the visual system would speed up when the previous presentation time is 200 compared to 400 ms (see Supplementary Figure 4, which shows this to be the case for 11 out of 24 participants). Similarly, one might wonder if our results could be explained by contrast adaptation: If the previous stimulus is presented for only 50ms, its contrast could be perceived as lower because temporal summation is not finished yet. Then, on the next

trial, adaptation to this lower contrast speeds up temporal summation in the visual system. However, adaptation to low contrast often entails *slower* changes in firing rates in the visual system, not faster changes (Lesica et al., 2007).

In conclusion, humans implicitly speed up encoding in working memory when they expect information to arrive at a fast pace. These adaptations may be understood in terms of optimally tracking changes in the environment by adapting the speed at which those changes are encoded. Further, additional analyses suggest that encoding speed adaptations are probably not due to amplified visual inputs. We believe that encoding speed adaptations in working memory are critical for optimal performance in environments whose pace may change suddenly and substantially.

References

Aguirre, G. K., Mattar, M. G., & Magis-Weinberg, L. (2011). De Bruijn cycles for neural decoding. *NeuroImage*, *56*(3), 1293–1300. https://doi.org/10.1016/j.neuroimage.2011.02.005

Akyürek, E. G., Toffanin, P., & Hommel, B. (2008). Adaptive control of event integration. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 569–577. https://doi.org/10.1037/0096-1523.34.3.569

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*(10), 6–6. https://doi.org/10.1167/11.10.6

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science*, *321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, *101*(3), 446–469. https://doi.org/10.1037/0033-295X.101.3.446

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients Between Models. *American Journal of Sociology*, *100*(5),

1261–1293. https://doi.org/10.1086/230638

Danielmeier, C., & Ullsperger, M. (2011). Post-Error Adjustments. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00233

de Jong, J., Akyürek, E. G., & van Rijn, H. (2021). A common dynamic prior for time in duration discrimination. *Psychonomic Bulletin & Review*, *28*(4), 1183–1190. https://doi.org/10.3758/s13423-021-01887-z

Durstewitz, D. (2003). Self-Organizing Neural Integrator Predicts Interval Times through Climbing Activity. *The Journal of Neuroscience*, *23*(12), 5342–5353. https://doi.org/10.1523/JNEUROSCI.23-12-05342.2003

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, *412*(6849), 787–792. https://doi.org/10.1038/35090500

Foulke, E., & Sticht, T. G. (1969). Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin*, *72*(1), 50–62. https://doi.org/10.1037/h0027575

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 845–866. https://doi.org/10.1037/0096-1523.19.4.845

Glasauer, S., & Shi, Z. (2021). The origin of Vierordt's law: The experimental protocol matters. *PsyCh Journal*, pchj.464. https://doi.org/10.1002/pchj.464

Glaze, C. M., Kable, J. W., & Gold, J. I. (2015). Normative evidence accumulation in unpredictable environments. *ELife*, *4*, e08825. https://doi.org/10.7554/eLife.08825

Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A., & Smith, M. A. (2014). Environmental Consistency Determines the Rate of Motor Adaptation. *Current Biology*, *24*(10), 1050–1061. https://doi.org/10.1016/j.cub.2014.03.049

Gorea, A. (2015). A Refresher of the Original Bloch's Law Paper (Bloch, July 1885). *I-Perception*, *6*(4), 204166951559304. https://doi.org/10.1177/2041669515593043

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. *Attention and Performance*, 187–203.

Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). Temporal scaling of neural responses to compressed and dilated natural speech. *Journal of Neurophysiology*, *111*(12), 2433–2444. https://doi.org/10.1152/jn.00497.2013

Lesica, N. A., Jin, J., Weng, C., Yeh, C.-I., Butts, D. A., Stanley, G. B., & Alonso, J.-M. (2007). Adaptation to Stimulus Contrast and Correlations during Natural Visual Stimulation. *Neuron*, *55*(3), 479–491. https://doi.org/10.1016/j.neuron.2007.07.013

Loftus, G. R., & Ruthruff, E. (1994). A theory of visual information acquisition and visual memory with special application to intensity-duration trade-offs. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(1), 33–49. https://doi.org/10.1037/0096-1523.20.1.33

Los, S. A., Nieuwenstein, J., Bouharab, A., Stephens, D. J., Meeter, M., & Kruijne, W. (2021). The warning stimulus as retrieval cue: The role of associative memory in

temporal preparation. *Cognitive Psychology*, *125*, 101378. https://doi.org/10.1016/j.cogpsych.2021.101378

Maaß, S. C., Schlichting, N., & van Rijn, H. (2019). Eliciting contextual temporal calibration: The effect of bottom-up and top-down information in reproduction tasks. *Acta Psychologica*, *199*, 102898. https://doi.org/10.1016/j.actpsy.2019.102898

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Murphy, P. R., Wilming, N., Hernandez-Bocanegra, D. C., Prat-Ortega, G., & Donner, T. H. (2021). Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. *Nature Neuroscience*, *24*(7), 987–997. https://doi.org/10.1038/s41593-021-00839-z

Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The Adaptive Nature of Visual Working Memory. *Current Directions in Psychological Science*, *23*(3), 164–170. https://doi.org/10.1177/0963721414529144

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013). The Timescale of Perceptual Evidence Integration Can Be Adapted to the Environment. *Current Biology*, *23*(11), 981–986. https://doi.org/10.1016/j.cub.2013.04.039

Padfield, D., & Matheson, G. (2020). *nls.multstart: Robust Non-Linear Regression using AIC Scores*. https://CRAN.R-project.org/package=nls.multstart

Piray, P., & Daw, N. D. (2020). A simple model for learning in volatile environments.

PLOS Computational Biology, *16*(7), e1007963. https://doi.org/10.1371/journal.pcbi.1007963

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reinitz, M. T. (1990). Effects of spatially directed attention on visual encoding. *Perception & Psychophysics*, *47*(5), 497–505. https://doi.org/10.3758/BF03208183

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018). A Dynamical Systems Perspective on Flexible Motor Timing. *Trends in Cognitive Sciences*, *22*(10), 938–952. https://doi.org/10.1016/j.tics.2018.07.010

Salet, J. M., Kruijne, W., van Rijn, H., Los, S. A., & Meeter, M. (2022). FMTP: A unifying computational framework of temporal preparation across time scales. *Psychological Review*. https://doi.org/10.1037/rev0000356

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, *1252*(1), 229–236. https://doi.org/10.1111/j.1749-6632.2012.06447.x

Sohn, H., Narain, D., Meirhaeghe, N., & Jazayeri, M. (2019). Bayesian Computation through Cortical Latent Dynamics. *Neuron*, *103*(5), 934-947.e5. https://doi.org/10.1016/j.neuron.2019.06.012

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *ELife*, *7*, e34963. https://doi.org/10.7554/eLife.34963

Vangkilde, S., Coull, J. T., & Bundesen, C. (2012). Great expectations: Temporal

expectation modulates perceptual processing speed. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1183–1191. https://doi.org/10.1037/a0026343

Vangkilde, S., Petersen, A., & Bundesen, C. (2013). Temporal expectancy in the context of a theory of visual attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1628), 20130054. https://doi.org/10.1098/rstb.2013.0054

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436–1451. https://doi.org/10.1037/0096-1523.32.6.1436

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wark, B., Fairhall, A., & Rieke, F. (2009). Timescales of Inference in Visual Adaptation. *Neuron*, *61*(5), 750–761. https://doi.org/10.1016/j.neuron.2009.01.019

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11. https://doi.org/10.1167/4.12.11

Methods and Materials

Participants

Participants were first-year psychology students (mean age 20.9 years; 63% female) at the University of Groningen who obtained partial course credit for participating. Informed consent was obtained before the experiment started. Participants with an average capacity estimate of higher than 50 rad^{-1} (see Analysis) were excluded from further analysis (four in experiment 1, one in experiment 2, four in experiment 3). Indeed, their average capacity far exceeded the Q3 + 3*IQR criterion for outlier detection, which was 8.0 (experiment 1), 9.3 (experiment 2) and 10.9 (experiment 3). Using this stricter exclusion criterion did not change the interpretation of our results. For experiments with bootstrapping (experiment 1 and 3), our original exclusion criterion (capacity above 50 rad^{-1}) was applied to each bootstrap. As a result, we discarded a participant within a bootstrap 5% and 15% of the time in experiment 1 and experiment 3, respectively.

Apparatus and stimuli

The experiment was programmed in OpenSesame, a free software tool for designing experiments (Mathôt et al., 2012). Stimuli were presented on a 19 inch CRT screen with a resolution of 1,280 × 1,024 pixels, running at 100 Hz. Participants were seated in a sound-attenuated room with dimmed lights approximately 60 cm from the screen. A gray background was maintained during the entire experiment. Memory items were Gabor patches with a gaussian envelope (spatial frequency: 0.05 cycles/degree, std: 12 pixels, phase: 0), presented at fixation. Its orientation was picked from a uniform distribution spanning 0 to 180 degrees. Memory items were masked with 50 overlapping

Gabor patches with random orientations (spatial frequency: 0.05 cycles/degree, std: 10 pixels, phase: 0) and their centers scattered in a 40x40 pixel square.

Procedure

After giving written consent, participants received both written and verbal instructions. Then, they performed several practice trials in a delayed estimation task (Wilken & Ma, 2004). A typical trial proceeded as follows. A fixation dot was shown for 500 to 1000ms (uniformly sampled), after which the memory item was presented at fixation. The presentation time (i.e., for how long the memory item was physically on the screen) was 50, 200 or 400ms. In order to control the time available for encoding, we attempted to eliminate any retinal or cortical after-image. To this end, a pattern mask was presented for 100ms immediately after memory item offset. Participants had to retain the memory item over the next 1000ms, after which a 'wheel probe' was presented. This was a Gabor patch that participants could turn around by moving their mouse to the remembered orientation and clicking to confirm their response. Participants only received trial-to-trial feedback in practice trials. In the real experiment, only block-wise feedback (total number of points per block) was presented. Apart from the counterbalancing of presentation times, trials per block and the inclusion of a cue at the start of each trial (Experiment 3), the procedure was identical across experiments.

Experiment 1: Blocked information rate

In our first experiment (N=57), we tested whether a higher overall information rate (i.e., 1/presentation time) increases encoding speed in working memory. To this end, each block either had a high information rate (50ms: 70%, 200ms: 15%, 400ms: 15%), or low information rate (50ms: 15%, 200ms: 15%, 400ms: 70%). Each

information rate lasted for four blocks of 40 trials (i.e., 160 trials in total), after which participants switched (from high- to low information rate, or low- to high information rate). Participants completed a total of 640 trials and nine practice trials at the start of the experiment. Participants were not informed about overall information rate or switches in information rate.

Experiment 2: Sequential information rate

In this experiment (N=24), we assessed whether encoding speed would also adapt to recent, sequential changes in information rate. We counterbalanced the order of presentation times with a de Bruijn sequence, using the software described in (Aguirre et al., 2011). A de Bruijn sequence perfectly counterbalances the order of its elements up to a certain degree using the smallest sequence length. We designed the sequence such that each possible order of three presentation times (e.g., 50 - 50 - 400) would occur with equal frequency within each block. This allowed us to fit encoding curves with an equal number of trials presentation time on trial n, each n-1 presentation time, and each n-2 presentation time. Participants performed a total of 675 trials and nine practice trials at the start of the experiment.

Experiment 3: Cued information rate

In our final experiment (N=19), we explored whether adaptations in encoding speed could be induced with explicit cues that are highly informative. At the start of each trial, we presented either the word ‘FAST’ in red, or ‘SLOW’ in blue for 400 ms. These cues were 80% valid. ‘Fast’ cues predicted 50ms presentation times 80% of the times (with 10% for both 200 and 400ms), and vice versa for ‘slow’ cues. Participants

were informed that the cues were informative and encouraged to make use of the cues. They completed a total of 500 trials and 20 practice trials at the start of the experiment.

Analysis

All analysis was performed in R (R Core Team, 2018; version 3.5.1). Error angles were computed as the reproduced angle *minus* the actual angle of the memory item multiplied by two (since Gabors have a rotational symmetry of 180 degrees instead of 360 degrees), such that errors spanned -180 to 180 degrees. We quantified memory precision as the inverse of the circular standard deviation (σ^{-1}) of the error angles at each presentation time, *t*. We subtracted the memory precision that would be expected for that sample size (*N*), which was computed by taking inverse of the mean of 1,000 circular standard deviation estimates from circular uniform distributions with sample size *N*.

To quantify encoding speed and capacity in working memory, we fitted exponential encoding curves for each participant in each condition with two parameters: Maximum capacity (c) and encoding speed (τ^{-1}):

$$\sigma^{-1}(t) = c\,(1 - e^{-t/\tau}) \quad (1)$$

These parameters map onto the asymptotic memory precision for very long presentation times (c, see Figure 1), and encoding speed (i.e., the speed at which precision reaches that asymptote; τ^{-1}, see Figure 1). These curves were fitted with the *nls.multstart* package (Padfield & Matheson, 2020; version 1.2.0), which initializes multiple starting values for each parameter when fitting the encoding curve and selects the set of parameters with the lowest AIC score. We normalized the parameters by dividing estimates by the mean estimate for each participant, which allowed us to

assess *changes* in capacity or encoding speed as induced by our experimental manipulations.

The effect of information rate on normalized encoding speed was assessed by fitting a linear regression model. The null model predicted normalized encoding speed with only normalized capacity as a predictor. The alternative model added the manipulation of expected information rate. For experiment 1, blocked information rate was dummy coded (0 for 'low' block, 1 for 'high' block). For experiment 2, n-1 and n-2 information rate (1/presentation time) were added as continuous predictors. For experiment 3, cued information rate was dummy coded (0 for 'slow', 1 for 'fast'). To quantify evidence for a modulation of encoding speed by expected information rate, we performed model comparison using Bayes Factors (Wagenmakers, 2007):

$$BF_{10} = e^{(BIC_0 - BIC_1)/2}$$

where BIC_0 is the Bayesian-Information Criterion score for the null-model only containing c_{norm} as a predictor and BIC_1 is the Bayesian-Information Criterion score for the model that also includes expected information rate. High values for BF_{10} reflect evidence against the null and therefore suggest an effect of expected information rate on encoding speed.

Experiment 1 and 3 were not balanced due to the nature of the experimental manipulation. That is, to induce expectations about information rate, some presentation times contained more observations than others in certain conditions. Estimates of circular standard deviation are biased substantially at low sample sizes, especially at high standard deviation (see Supplementary Figure 5). Such bias could have a complex

impact on the estimated capacity and encoding speed. Therefore, we bootstrapped the entire analysis for those experiments by subsampling (without replacement) for each participant and each condition to the lowest number of observations *minus* one. We sampled without replacement because estimates of circular standard deviation have higher bias and variance when duplicates can be sampled (see Supplementary Figure 6). We reported the median of all test statistics (regression coefficients, *t*-values, p-values, Bayes Factors). For plotting, we computed the standard error (SE) of the bootstrap by dividing the estimated standard deviation (SD) by $\sqrt{N}$ for each condition.