Signal Power as the Limited Resource of Working Memory

S. Thomas Christie (tchristie@umn.edu)
Cognitive Science, 75 E River Rd
Minneapolis, MN 55455 USA

Paul Schrater (schrater@umn.edu)
Psychology, 75 E River Road
Minneapolis, MN 55455 USA

Abstract

Working memory is famously capacity-limited, but the nature of the capacity is still a matter of debate. While recent research supports a continuous-resource model of working memory, these models do not account for the phenomenology of cognitive effort or the capacity's apparent sensitivity to task demands and incentives. We suggest that describing working memory as a noisy information channel accounts for these phenomena. In this view, capacity is a function of continuous signal power that can be allocated across multiple signals. We used a computational simulation to infer signal power for subjects performing the N-back task, for $N = 1, 2, 3$. We found that as task difficulty increases, signal power per response is lower, but overall signal power increases, suggesting that subjective feelings of cognitive effort may be related to signal power. Our approach provides a parsimonious explanation for working memory effects that grounds the continuous resource view in an established theory while providing a plausible account of subjective effort and cognitive control.

Keywords: working memory; information theory; n-back

Introduction

While human long-term memory storage capacity is virtually unlimited, short-term or working memory is severely constrained. Early theorists modeled capacity as a set of discrete slots (Miller, 1956), but recent research supports models of memory as a variously allocated continuous resource (Ma, Husain, & Bays, 2014). Despite growing acceptance of a continuous resource model of working memory, the nature of the resource remains unclear, and the severity of its limitation an object of puzzlement. It appears that working memory does not have a fixed capacity, but rather a one that varies (while remaining limited) in the presence of incentives (Heitz, Schrock, Payne, & Engle, 2008). In addition, explanations of cognitive control identify the close relationship demands on working memory and the phenomenology of mental effort (see Shenhav et al. (2017) for a recent review), suggesting that capacity itself is an actuator for effortful cognitive control.

We suggest that viewing working memory as a noisy information channel provides a parsimonious account of the capacity limitations of working memory, the flexibility of those limitations, and the ability to adapt to incentives. In this view, the limited continuous resource supporting working memory is simply signal power. A signal is the neural representation of information in working memory. This can be represented in terms of the magnitude vector of activity across a neural population. Signal power refers to the overall variance, or power, of this vector.

Considering working memory as a controllable, variable-power channel can explain capacity increases in the presence of incentives. Information transmission rate through a channel is a function of both inherent channel noise power and signal power (Cover & Thomas, 2012). Increased signal power allows for faster information transmission, which could manifest as increased precision memory of a single item, or as similar precision memory of more items, depending on task demands.

Signal power must be paid for with energy, a fact that immediately grounds our proposal in terms of available energetic resources. If noise power is fixed, increased information transmission rate scales with (roughly) the logarithm of the signal power, providing diminishing returns. If signal power is identified with neural gain, as seems reasonable (Van den Berg, Shin, Chou, George, & Ma, 2012), the rate of available energy via blood glucose or astrocytic glycogen becomes both a cost and a limitation. Incentives could induce a temporary boost in signal power, but only to the extent that surplus energy is available (see Christie and Schrater (2015) for further discussion).

In the current work, we instantiate a proof-of-concept model of working memory as signal through an information channel, and use this model to estimate comparative signal power of subjects performing the N-back task. We show that as N increases, accuracy falls while overall signal power increases, a finding not accounted for by current models. Signal power is thus a candidate for connecting for the 'resource' underpinning working memory, and connects the phenomenological experience of mental exertion, the fidelity constraints of memory, and the energetic limits of neural gain.

Model

In the N-back task (Kirchner, 1958), subjects are asked to view a series of letters (or other images or cues), and indicate whether the letter most recently seen is the same as the letter N trials ago. Each observed letter is encoded into a neural representation, transmitted through time via working memory, and finally compared with the representation of a more recent letter. The N-back task can be viewed as involving the transmission of cue information from the past to the present through a noisy information channel.

Model specification

The model consists of three distinct information channels, abstracted as conditional distributions which represent the critical transformations in most cognitive tasks. Each of these distributions has a parameter which acts to adjust the amount of information transmitted. Updating forms a simple model of bandwidth limitations corresponding to constructs of working memory and/or updating. In the following channels, s_* represents signal power and λ_* represents errors and noise.

Encoding: Sensing to Latent State Representation

$$p(x_t|y_t; s_y, \lambda_y) \tag{1}$$

Update: Updating State Representation

$$p(z_t|z_{t-1}, x_t; s_z, \lambda_z) \tag{2}$$

Ouput: State to Response

$$p(r_t|z_t; s_r, \lambda_r) \tag{3}$$

Distinct input symbols y are embedded as vectors x into a latent memory representation z which is updated by noisy shift and add operations. A response r is generated through a readout mapping on z.

The encoding transformation is a simple vector embedding (e.g. 1-hot). The state update equation is given by:

$$z_t = (\mathbf{S} + \mathbf{R}(\varepsilon))z_{t-1} + \begin{bmatrix} 0 \\ s_z \end{bmatrix} x_t + \eta \tag{4}$$

$$\eta \sim \text{Poisson}(\lambda_z) \tag{5}$$

$$\varepsilon \sim \text{Normal}^+(\lambda_{z2}, \sigma_z) \tag{6}$$

Letters y_t are encoded as 1-hot vectors x_t of dimension $(n \times 1)$ (equation 1), which are subsequently multiplied by signal amplitude s_z. We interpret the outcome as a Poisson code. Letter encodings are then embedded in a single memory vector z_t of length $((N+1)n \times 1)$ (equation 2). As new letters are observed, the memory vector z_t is updated via a shift matrix $\mathbf{S}$ (equation 3) with the possibility of some interference from other letters, indicated by $\mathbf{R}$. $\mathbf{S}$ and $\mathbf{R}$ are dimension $((N+1)n \times (N+1)n)$. A shifted update requires that a set of recently observed letters is represented as concatenated vector, akin to memory 'slots' but with the difference that each incorporates signal power, noise, and interference. As an image representation is transmitted through memory, it is corrupted by i.i.d. noise η with fixed power λ_z, which we treated as Poisson noise (equations 3 and 5). Figure 1 shows schematic of an encoding of a single letter.

Finally, the response r_t is calculated by comparing argmax of the most recent, and latest, sections of z_t.

As an example, for $N = 2$ and n letter possibilities, the update distribution is:

$$\vec{z}_t = \begin{bmatrix} \bar{x}_{t-2} \\ \bar{x}_{t-1} \\ \bar{x}_t \end{bmatrix} = \left(\begin{bmatrix} 0 & I_n & 0 \\ 0 & 0 & I_n \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} I_n\varepsilon & 0 & I_n\varepsilon \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \begin{bmatrix} \vec{x}_{t-3} \\ \vec{x}_{t-2} \\ \vec{x}_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ s_z \end{bmatrix} \vec{x}_t + \eta$$

In the above equations, a new encoded image $\vec{x}_t$ is observed and incorporated into the working memory representation. I_n is the identity matrix of size n. Barred vectors on the left side represent noisy updated estimates of the corresponding variables on the right side. A subject response of 'same letter' or 'different letter' is generated by comparing $\arg\max \bar{x}_t$ and $\arg\max \bar{x}_{t-2}$, which was observed $N = 2$ time steps previously.

The encoding and update scheme we chose is one of the simplest of many possible models, easy to visualize and allows for explicit representations of signal and noise power. Our conclusions are unchanged for additive Gaussian White-Noise signal corruption, and for more complex embedding schemes.

Results

Simulations To illustrate the relationship between signal power and subject response accuracy in the N-back task, we used the model above to simulate subjects performing the task, varying levels of interference λ_z and signal power s_z. We set a fixed i.i.d. noise power of $\lambda_z = 1000$ (arbitrary units), specified as the variance of the generating Poisson distribution. Accuracy curves for a fixed $\varepsilon = 0.3$ are shown in Figure 2. For a given power budget, accuracy drops as N increases, modeling the ubiquitous set-size effect.

Task description Subjects performed the N-back task and were shown 6 blocks of 70 images (letters) each via Amazon Mechanical Turk; data from additional blocks with a different manipulation were recorded but are not included in this analysis. We calculated overall accuracy on the 1-back, 2-back, and 3-back tasks for 26, 26, and 31 subjects respectively. Using power/accuracy curves generated by computational simulation (see Figure 2), we estimated signal power for each subject. Estimated total signal power values are plotted in Figure 3.

Results We found that expended signal power budgets increase as a function of N. Increase from $N = 2$ to $N = 3$ is small but still significant ($p < 0.05$). Total power outlay was calculated as N times the power estimated in a single signal. The scale of increase resembles observed fMRI data (see figure 2B in Loughead et al. (2009)), though within-subject N-back conditions would allow for more precise estimates.

From this, we conclude that (1) overall signal power budget is not fixed, but adapts to task conditions, (2) signal power scales by much less than a direct multiple of N, indicating an intrinsic cost (though not a *fixed* capacity) associated with
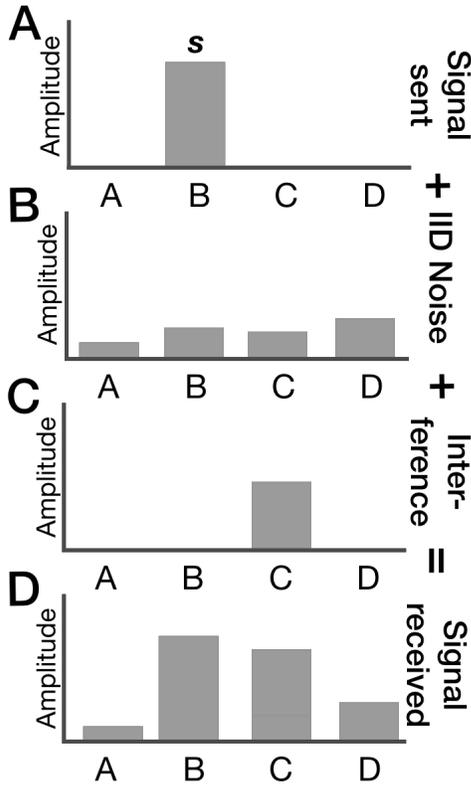
Figure 1: Model of letter encoding and corruption by noise. (A) The 'sent' signal is a 1-hot encoding of observed letter with some signal power. (B) i.i.d. Poisson noise is added to each bin. (C) Interference from other letters is represented as *decayed* signal. (D) The 'received' signal is a noisy version of the 'sent' signal. The received letter is assumed to be the letter with the highest count.
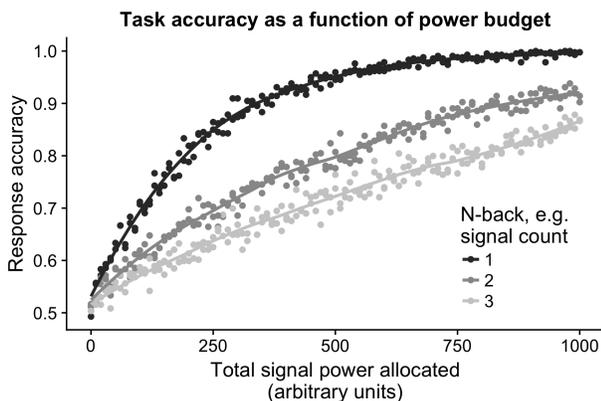
Figure 2: Simulated accuracy increases as a function of input signal power s and noise power $\lambda_z = 1000$ on the N-back task with interference proportion $\varepsilon = 0.3$. We assumed that a total power budget is allocated evenly across N signals.
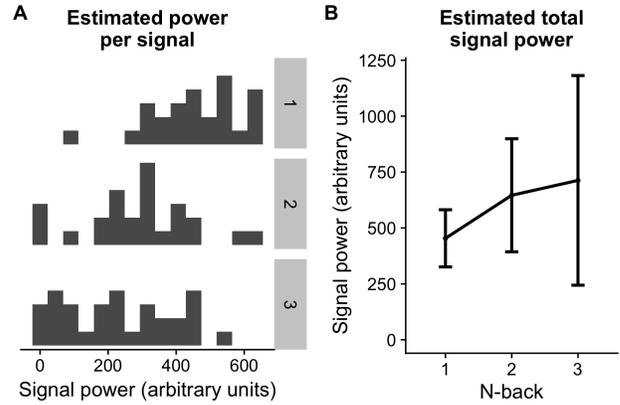
Figure 3: (A) Histogram of average per-signal power by subject. Row labels indicate N. (B) Estimated total signal power outlay per subject, obtained by multiplying estiated signal power by N. Bars represent one standard deviation.

additional resources. By requiring parameterized noise, a signal-power interpretation also accounts for subject guessing, a necessary requirement for any working memory model.

Discussion

The application of information theory to experimental psychology, and in particular the treatment of individuals as an information channels, is nearly as old as information theory itself (Hyman, 1953). However, concerns regarding signal independence and efficient coding mechanisms have led to a near abandonment of its use in psychology (see Luce (2003), but for an exception see Sims (2016)). We feel that this reaction throws the baby out with the bathwater, abandoning information theory's computational-level conceptual framework involving signal, noise, and transmission due to concerns about implementations at a lower level of analysis. With working memory research in particular, research suggests a continuous resource that is allocatable, flexible, and sensitive to both incentives and metabolic manipulations. Information theory offers a well-established quantitative framework that can account for these findings while providing straightforward integration with neural data.

In the current work we re-introduce salient concepts from information theory and provide a proof-of-concept plausibility argument for its application to a working memory task. Future work is needed in order to make concrete, quantitative predictions. In particular, a full Bayesian inference of signal and interference noise parameters will allow model comparison with leading variable resource theories. A follow-up study using a within-subject design with higher values of N and simultaneous brain imaging would provide better sensitivity to individual differences and insight into the possible connections between modeled signal power and neural gain.

References

Christie, S. T., & Schrater, P. (2015). Cognitive cost as dynamic allocation of energetic resources. *Frontiers in neuroscience*, *9*, 289.

Cover, T. M., & Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *45*(1), 119–129.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of experimental psychology*, *45*(3), 188.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, *55*(4), 352.

Loughead, J., Wileyto, E., Valdez, J., Sanborn, P., Tang, K., Strasser, A., . . . Lerman, C. (2009). Effect of abstinence challenge on brain function and cognition in smokers differs by comt genotype. *Molecular psychiatry*, *14*(8), 820.

Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of general psychology*, *7*(2), 183.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, *17*(3), 347.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, *40*, 99–124.

Sims, C. R. (2016). Rate-distortion theory and human perception. *Cognition*, *152*, 181–198.

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*, *109*(22), 8780–8785.