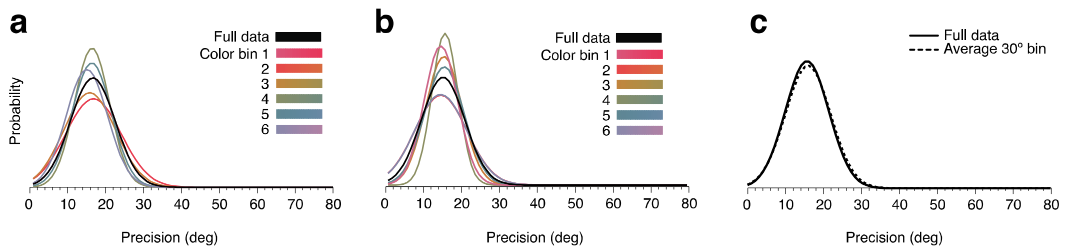
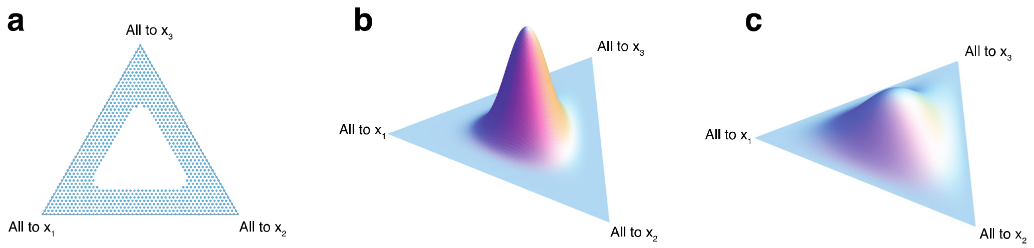
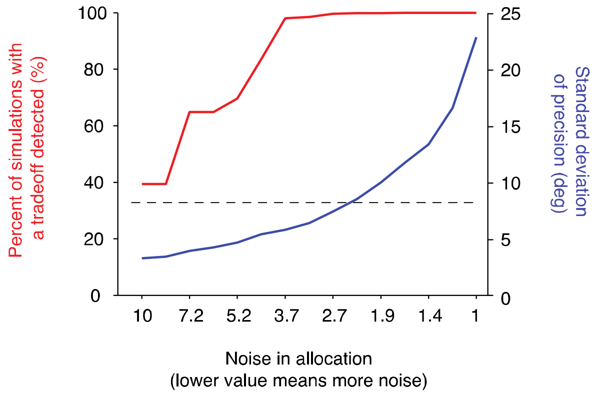


Supplementary Figure S1. Memorability differences between colors does not drive variation in precision. (a) Best-fit precision distributions for six color bins, each of 30°. Each distribution is labeled with the color at the bin's center, and the black curve is all of the data together. The legend shows the range of colors included in each bin. The curves were fit simultaneously to all subjects' data. **(b)** How much variability is to be expected by chance? We simulated six data sets by sampling precision values from the black curve. Even though each simulation is based on the same distribution, the estimated model parameters display a range of values. This range is comparable to that seen in the observed binned data in panel **(a)**. **(c)** The variability in the full data set (solid black line) is the same as that for the average bin (dashed black line).



Supplementary Figure S2. Simulating uneven allocation policies. (a) A ternary plot that shows the noiseless uneven allocation policies used for simulations. At each point, the allocation to x_1, x_2, and x_3 sums to 1. The corners correspond to the three policies for which one item is given all of the commodity. The center corresponds to a policy of even allocation across the items. Each filled circle corresponds to a fixed, uneven allocation policy that could produce at least as much variability in precision as we observed in our data. **(b & c)**. Noisy allocation distributions are formalized as a Dirichlet distribution over the space of allocation policies, centered on even allocation. The distribution in panel **(b)** has lower noise levels (concentration parameters of 9, 9, and 9) than the one in panel **(c)** (concentration parameters of 4, 4, and 4).



Supplementary Figure S3. Variability from noisy even allocation produces detectable tradeoffs. Noisy even allocation was formalized as a symmetric Dirichlet distribution on the simplex, centered at even allocation, with concentration parameters (α,α,α) that determine the noise level. Monte Carlo simulations were performed for 15 levels of noise, evenly-spaced in log units from low noise where $\alpha = (10,10,10)$ to high noise where $\alpha = (1,1,1)$. Simulated data was fit with the variable-precision model in order to estimate the standard deviation of precision (blue line). We also performed a trade-off analysis to determine the percentage of simulations at each noise level (red line) in which at least one of the six possible sortings showed a difference in estimated precision. The dashed line is drawn at the variability estimated in the data. Critically, at the level of allocation noise necessary to produce this amount of variability our tradeoff detection procedure is nearly perfect.

Supplementary Table S1. AIC$_c$ values for Experiment 1.

Subject	Fixed (n=1)	Variable (1)	Fixed (3)	Variable (3)	Fixed (5)	Variable (5)
1	6371	6367	7876	7870	7550	7545
2	6463	6457	7785	7780	7326	7312
3	6171	6160	7207	7184	7390	7373

Lower AIC$_c$ values are better fits.

Supplementary Table S2. AIC$_c$ values for perceptual control study.

	No noise		Noise	
Subject	Fixed	Variable	Fixed	Variable
1	4528	4528	7141	7140
2	4734	4744	7116	7116
3	4423	4435	7866	7882

Lower AIC$_c$ values are better fits.

Supplementary Notes

The power of tradeoff detection

In Experiment 3, we asked participants to report multiple items per display and found independence of precision across items. From this we concluded that variability in precision is not due to tradeoffs in the allocation of a commodity to remembered items. In the Supplementary Methods, we describe our method for showing that the tradeoff-detection method is powerful enough to reliably detect tradeoffs if they were to exist. We considered two sources of tradeoffs: (1) noiseless uneven allocation, where the participant knowingly attends or gives preferential processing to a chosen item, and (2) noisy even allocation, where the participant tries to distribute her commodity evenly, but fails, unwittingly giving preference to some of the items.

Each allocation (i.e., allotment of resources) can be described as a vector of proportions of the commodity dedicated to each item (e.g., [1/3, 1/3, 1/3] for even allocation across three items, or [0.1, 0.1, 0.1, 0.2, 0.5] for uneven allocation across five). The set of all such possible allocations defines a simplex (Supplementary Figure 2A).

Guessing

The present data was not as well described by models that did not include a uniform error component. Indeed, the AIC_C values indicated that models with a uniform component provided a better fit to the data (all AIC_C values were decisively,

>100, better). Furthermore, estimates of this guess state do not differ to a large extent between the variable- and fixed-precision models. Estimates of the rate of guessing in Experiment 1 were 1.1%, 14.1%, and 21.6% for the variable-precision model for set sizes 1, 3, and 5 respectively). This was comparable to the guess rates for the fixed-precision model (1.4%, 15.9%, and 22.0%). In Experiment 2 there was a large drop in guess rate between the 'random probe' and 'choose the best' condition (variable-precision: 18.7% to 7.5%; fixed-precision: 20% to 8.2%). The data do suggest that fixed-precision models may slightly over-estimate guess rates, presumably because these models misclassify some highly imprecise responses as guess responses.

Supplementary Methods

Variability in perception

We eliminated the memory demands of the task by keeping the stimulus present while the participant responded. The stimulus was a single colorful circle, 0.75° in radius, placed in the center of the display. The participant was asked to select the color of the dot from a color wheel containing all of the possible colors. Three participants completed 750 trials of this perceptual task.

Naturally, participants were better at the perceptual task than at the memory task, and so in a variant of the perceptual task, we added noise to the stimulus so that the error matched that of the working memory task. This was accomplished by setting eighty percent of the stimulus' pixels to either black or white. The pixels were selected randomly for each participant but held constant across trials to minimize variability caused by differences in the noise. Three participants completed 750 trials of this perceptual task with noise.

Variability in memory for specific colors

To measure any contribution of differences in memorability to the variability in precision observed in the first experiment, we replicated the study and compared the variability measured in the full data set to that measured in subsets of trials binned by color. Variability found within one of these subsets cannot be attributed to differences in the memorability of colors because all of the trials within a bin were similar in color. Two participants performed 3,120 trials of a color working memory task with three items.

Variability in memory is not due to eye movements

We performed an experiment on three participants using an EyeLink eye tracker, in which the stimulus was displayed conditional on 1000 ms of continuous, uninterrupted fixation. The fixation region was defined by a square (1.35° from center to edge) centered on the fixation mark. The stimulus display consisted of the presentation of three colorful circles for 100 ms (see Methods for information on the visual angle and screen position of the stimuli).

Noiseless uneven allocation

To detect whether our analysis is sensitive to trade-offs due to noiseless uneven allocation (perhaps as a byproduct of a quantized storage capacity that can not be evenly divided among items, e.g., 4 slots among 3 objects), we performed Monte Carlo simulations for each point on the simplex capable of producing sufficient variability to explain the present data. The possible allocation policies of the commodity to n items are defined by the set of points (x_1, x_2, …, x_n), where x_i is the proportion of the commodity dedicated to the ith item and where the x_i's sum to 1. In the case of $n = 3$ items, these policies can be visualized in a ternary plot whose vertices correspond to the three policies where one item is given all of the commodity, and whose center corresponds to even allocation across the items (Supplementary Fig. 2a). To convert from the proportion of allocated commodity to the precision of report error, we used a sample-size model in which the variance of expected report error is inversely proportional to an item's allocation [1,38,39]. The proportionality constant was determined by setting the precision for an item under even allocation to 21.5°, the median precision across participants in Experiment 3. The sample size model makes it possible to estimate the expected precision under different allocation policies. Note that since we have defined precision as the square root of the report error variance, changing the allocation to one item will lead to a change in precision equivalent to dividing 21.5° by the square root of the change in allocation. For example, doubling the allocation to an item (66.7%) will lead to an expected precision of $21.5°/\sqrt{2} = 15.2°$.

Rather than make assumptions about the particular allocation policy chosen by the participant, instead we simulated data for each of the possible allocation strategies, in increments of 2%, that are capable of producing the variability observed in our experiments. (For example, there's no need to consider the case of even allocation because it is incapable of producing any variability at all.) For each of the 963 allocation policies that met these requirements, we simulated response error data for 16 participants, where each participant makes three responses on each of 450 trials. We then asked whether sorting by the first response led to an observable difference on the precision of the second response, and *vice versa*. If either result was statistically significant at $\alpha = 0.05$, we considered it as evidence of non-independence between the two responses. We repeated this procedure 1,000 times. For the worst-case allocation policy (i.e., the one that was hardest to detect), we found evidence of non-independence 96.6% of the time. Averaged over all tested allocation policies, we found evidence of non-independence 99.9% of the time. Thus, comparing responses within a trial is a powerful test for detecting tradeoffs in allocation between items.

Noisy even allocation

We performed Monte Carlo simulations to determine whether our analysis method is sensitive to trade-offs induced by noisy even allocation. We formalize noisy even allocation through a symmetric Dirichlet distribution on the simplex, centered at perfect evenness, and whose concentration parameter corresponds to the amount of noise. We simulated 15 levels of noise ranging from little noise (concentration parameter = 10, 10, 10) to very high levels of noise (concentration parameter = 1, 1, 1) (Supplementary Figure 2b). (The simulated noise levels were evenly spaced intervals in log space between these extremes). For each noise level, 50 separate simulations were performed. For each simulation, 500 trials (consisting of 3 items) were independently generated for 16 participants. Precision values for each item were generated by applying a sample size model [1,38,39] (as in the noiseless uneven allocation simulations) to allocation vectors drawn from a Dirichlet distribution. Measurement error for each simulation was generated by drawing values from a von Mises distribution according to the stimulated precision of each item. We then asked whether sorting responses by whether another response was an accurate or inaccurate judgment led to an observable difference in precision. We determined the percentage of simulations at each noise level that showed evidence of a trade-off for at least one of the six possible sortings. At low noise there is no measurable variability and consequently trade-offs are rarely detected. With increased noise, variability increases and trade-offs are easier to detect. Importantly, the likelihood of detecting tradeoffs rises above 95% well before the mean estimate of variability reaches 8.5° (the estimate of the variability in precision in the observed data) (Supplementary Figure 3).

These simulations demonstrate that our method of sorting error distributions by performance on another item in the same trial is a highly sensitive method test for detecting trade-offs between items regardless of whether these trade-offs emerge from a noisy allocation policy or a fixed, but uneven allocation process.

Supplementary References

38. Bonnel, A.M., and Miller, J. (1994). Attentional effects on concurrent psychophysical discriminations: Investigations of a sample-size model. *Attention, Perception, & Psychophysics 55*, 162-179.

39. Palmer, J. (1990). Attentional limits on the perception and memory of visual information. Journal of Experimental Psychology: Human Perception and Performance *16*, 332-350.