Title

The speed of statistical perception

Permalink

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Sherman, Brynn E
Yousif, Sami R
Reiner, Caroline
et al.

Publication Date

2022

The speed of statistical perception

Brynn E. Sherman (brynn.sherman@yale.edu)

Sami R. Yousif (sami.yousif@yale.edu)

Caroline L. Reiner (caroline.reiner@yale.edu)

Frank C. Keil (frank.keil@yale.edu)
Department of Psychology, Yale University
New Haven, CT 06520 USA

Abstract

In virtually every activity we engage in — from analyzing economic trends, to predicting which of two football teams is more likely to win a game — our minds are tasked with separating signal from noise. Such computations benefit from the fact that our minds are highly attuned to the statistical structure of the world. But how *quickly* do we detect statistical structure — and to what extent is our sensitivity to structure rooted in perceptual processes? To address this, we asked observers to judge whether briefly presented visual stimuli were generated randomly or non-randomly. In as little as a tenth of a second, people exhibited the same stable biases of statistical perception that they exhibit in classic cognitive tasks (i.e., without time constraints). These results suggest that certain biases of subjective probability may arise not from how we *think* about randomness, but from how we *perceive* statistical information in the first place.

Keywords: randomness; statistical learning; over-alternation bias; perception; subjective probability

Introduction

Imagine flipping a coin 10 times and receiving 10 'tails' in a row. You would likely find that outcome surprising. You may even question whether that outcome was truly random. We are constantly faced with problems like this: From discerning whether traffic is going to stop given one car slowing down, to interpreting graphs of stock values, we are tasked with parsing out what is *random* from what is *structured*. We know that our minds are attuned to structure in the external environment (e.g., via *statistical learning;* Saffran et al., 1996; see Sherman et al., 2020 for review), but how do we detect the presence (or lack) of structure in the first place?

Prior work studying randomness has emphasized the ways in which judgments of randomness are intrinsically and systematically biased (e.g., Kahneman & Tversky, 1972; Gilovich et al., 1985; Nickerson, 2002; Zhao et al., 2014; Reimers et al., 2018). This body of work has shown our impressions of randomness are often objectively incorrect. Take, for example, the two sequences of coin-flips represented in Fig 1A. Which sequence do you think was most likely to have been generated randomly? Most people indicate that Sequence #1 seems more random (see, e.g., Bar-Hillel & Wagenaar, 1991), yet Sequence #2 is in fact more likely to occur by random chance (in the sense that the *proportion of transitions* from heads-to-tails or vice versa is more likely to be 50% than 70%). This phenomenon, often referred to as an 'over-alternation bias', has been generalized across a range of tasks; it is present both when participants judge sequences and when they generate sequences themselves (for review, see Bar-Hillel & Wagenaar, 1991), and is consistent across visual, auditory, spatial, and temporal presentations (Yu et al., 2018a).

Although some work has attempted to explain the underlying perceptual and cognitive processes that cause biases in randomness judgments (Falk & Konold, 1997; Hahn & Warren, 2009; Griffiths et al., 2018; Warren et al., 2018; Yu et al., 2018b), there is still a fundamental open question about the *speed* of these processes: Are subjective impressions of randomness (and the biases associated with them) truly *cognitive* biases, or might they be rooted in perceptual systems (see, e.g., Yu et al., 2018b)?

One possibility is that these biases are indeed cognitive in nature: that our impressions of the randomness of, e.g., coin flips, arise from explicit reasoning about the probability of those events. However, given the stability of these same basic patterns (e.g., the 'over-alternation bias') across many modalities and paradigms, one may wonder whether certain rules governing impressions of randomness are wired into perceptual systems themselves.

If these biases are built into the visual system, how would we know? Prior work has sought to characterize the properties that distinguish 'visual' from higher-level cognitive features (see Scholl & Gao, 2013; Hafri & Firestone, 2021) and has identified *speed* as one relevant factor. Put simply, we would predict that truly visual features are likely to be perceived very quickly, and that non-visual features are likely to be perceived less quickly (although latency alone is not sufficient to determine whether a property counts as a true visual feature). Here, we aim to apply this logic to randomness, providing an initial investigation into the perceptual roots of randomness judgments.

Current Study

We conducted two experiments that addressed how quickly impressions of randomness are formed. We presented

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

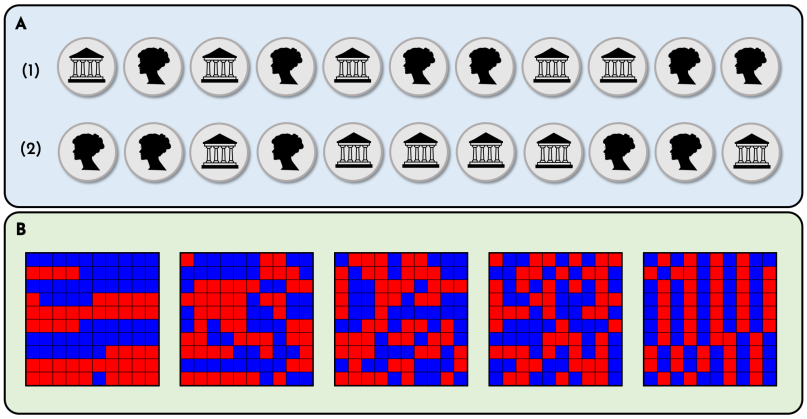

Figure 1: Examples of stimuli used in experiments on randomness. A) Studies often ask participants to judge whether a sequence of coin flips is random or not. In reality, Sequence 2 is "more random" in the sense that the transition probability among heads and tails is 0.5, but participants tend to judge Sequence 1 (which as a transition probability of 0.7) as more random. This effect is referred to as the 'over-alternation bias'. B) Example stimuli used in the current study. Each grid was generated with a transition probability (the probability of a switch between red and blue; going left to right, top to bottom) of 0.1, 0.3, 0.5, 0.7, and 0.9, respectively.

observers with binarily-colored 10x10 grids and asked them to judge whether or not each grid was generated randomly.

These grids were designed to be visual analogs of the coin flips which have been used in previous studies (see Fig 1B). To manipulate randomness, we generated sets of grids with different transition probabilities (i.e., the likelihood that a cell will switch from blue to red, or vice versa).

Our key question is how quickly observers will reliably make judgments of randomness. Inspired by similar paradigms used to measure the speed of face perception (e.g., Willis & Todorov, 2006; Colombatto et al., 2021), we ask how much exposure time is required for participants to make consistent judgments. Specifically, each stimulus was shown at five different time intervals (from 16ms to 500ms, followed by a mask). Assuming observers make reliable responses with 500 ms of exposure time (Reiner et al., 2021), we can then ask how quickly observers come to make that same judgment; for example, do their judgments made with 16ms of exposure agree with their judgments made with 500ms of exposure?

Experiment 1

Method

To assess the speed at which observers form stable randomness judgments, we designed a task in which participants were briefly presented with binarily-colored matrices (followed by a backwards mask) and were asked to judge whether or not each grid was generated randomly.

Participants Twenty observers participated in this task in exchange for course credit. One additional observer was excluded from analyses based on preregistered exclusion criteria (in this case, extreme response times). All observers gave informed consent prior to participation. Sample size, design, and analysis choices were pre-registered. Pre-registrations for both experiments, as well as raw data and stimuli, are available at the following OSF page: https://osf.io/6ze2b/

Stimuli The experimental stimuli consisted of 10x10 grids generated in Python, using PsychoPy libraries (Peirce, 2007). Each cell in a grid was either red or blue, and each grid was generated via a transition probability (0.1, 0.3, 0.5, 0.7, 0.9; see Fig 1B), governing the probability that cell $n+1$ in the grid was different from cell n. We adopted transition probability as a method to manipulate randomness, in the tradition of many prior studies of randomness (Bar-Hillel & Wagenaar, 1991); to make the stimuli a visual analog of these tasks, we used grids, as has been used in prior studies as well (e.g., Zhao et al., 2014; Yu et al., 2018b).

The grids were generated from left to right, top to bottom, meaning, e.g., that the leftmost cell of the second row of the grid was dependent on the rightmost cell of the first row of the grid. Eight grids were generated at each transition probability; half started with a red cell, and the other half started with a blue cell. To ensure that the number of blue vs. red cells could not be used as a cue, each color was represented approximately equally, regardless of transition probability; we constrained the grids such that there could be no more than 55 (out of 100) of one color. Further, we verified that each grid's true transition probability approximately reflected the desired transition probability. We also computed transition probability in the vertical direction (i.e., the likelihood of transitioning from blue to red or vice versa, if moving top-to-bottom, left-to-right) to ensure that our transition probability manipulation did not create spurious effects in the vertical dimension. On average, the transition probability in the vertical dimension did not differ

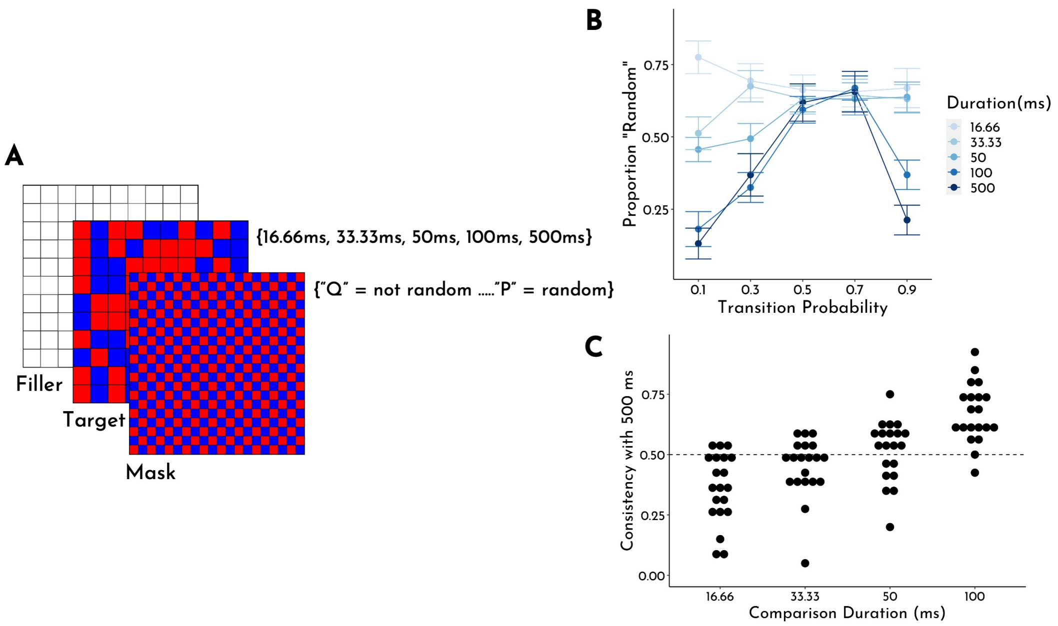

Figure 2: Experiment 1 design and results. A) On each trial, observers first viewed a "filler" grid stimulus, followed by the briefly presented target grid stimulus. Observers were then presented with a 20x20 checkerboard stimulus, which served as a backwards mask and remained on the screen until observers made their response. B) Randomness judgments as a function of transition probability and exposure duration. Error bars indicate +/- 1 SEM. C) Within-observer consistency results, separated out by each exposure duration compared to 500 ms. Each dot is a single observer.

from 0.5, and was unrelated to transition probability in the horizontal direction.

We additionally generated a single 20x20 grid, in which each cell simply alternated between red and blue. This served as a backwards mask on all trials.

Design and Procedure Observers performed the experiment while seated in front of a monitor with a 60 Hz refresh rate. The experiment was presented using PsychoPy libraries in Python (Peirce, 2007). Observers were instructed that they would be viewing a series of grids and be asked to determine whether each grid was generated randomly or not. No explicit instructions or definitions of randomness were given.

The trial sequence is illustrated in Fig 2A. Each trial began with a filler stimulus (a blank 10x10 grid) for 3 seconds, followed by the brief presentation of a grid stimulus. This target grid was presented for either 16.66 ms, 33.33 ms, 50 ms, 100 ms, or 500 ms and was followed immediately by the presentation of a mask stimulus. This backwards mask was critically important for limiting the post-perceptual or iconic memory processing that may influence observers' judgments (Breitmeyer & Ogmen, 2000). The mask remained on the screen until observers made their judgment. Observers responded with one of four button presses, indicating that the stimulus was 'definitely not random', 'maybe not random', 'maybe random', or 'definitely random'. For the purposes of analyses, we collapsed across 'definitely' and 'maybe' responses.

Observers completed 200 trials total (8 grids x 5 transition probabilities x 5 viewing durations). The task sequence was randomized for each observer, and observers were not informed that each grid was presented multiple times throughout the experiment. Observers completed five representative practice trials before beginning the task.

Results & Discussion

First, we assessed how judgments of randomness were influenced by transition probability and exposure duration (Fig 2B). An ANOVA revealed a significant main effect of transition probability, $F(4, 76) = 24.81$, $p < .001$, $\eta_p^2 = .57$. indicating that observers were indeed sensitive to the randomness manipulation. Specifically, especially at the longer durations (100 and 500 ms), observers' judgments of randomness generally tracked the true transition probability (i.e., they were least likely to indicate the stimuli with transition probabilities of 0.1 and 0.9 to be random), but exhibited the characteristic over-alternation bias, in which the stimuli with a transition probability of 0.7 were judged to be maximally random. There was also a significant main effect of exposure duration $F(4, 76) = 9.11$, $p < .001$, $\eta_p^2 = .32$, as well as a significant interaction between transition probability and duration $F(16, 304) = 11.20$, $p < .001$, $\eta_p^2 = .37$, indicating that observers' randomness judgments were indeed dependent on exposure duration.

To address our critical question of the minimal exposure time required to make stable judgments of randomness, we examined the consistency in randomness judgments between each of the four shorter durations (16.66 ms, 33.33 ms, 50 ms, 100 ms) relative to the longest duration (500 ms). We took both an across-observer and within-observer approach to answer this question. First, to address this in an across-observer manner, we computed the mean randomness rating across observers for each of the 40 stimuli; we did this separately for each exposure duration, and then correlated the stimulus-level randomness ratings between judgments of the stimuli at 500 ms and each of the shorter intervals. Comparing the 100 ms trials to the 500 ms trials, we found a high degree of correlation in randomness judgments, $r = .84$, $t(38) = 9.41$, $p < .001$. We also found a reliable correlation between judgments at 50 ms and 500 ms, $r = .52$, $t(38) = 3.79$, $p < .001$. There was not a reliable correlation between judgments at 33.33 ms and 500 ms, $r = .27$, $t(38) = 1.75$, $p = .09$. Surprisingly, there was a reliable *negative* correlation between judgments at 16.66 ms and 500 ms, $r = -.34$, $t(38) = -2.26$, $p = .03$. These results suggest that — at the group level — reliable judgments of randomness may be realized with as little as 50 ms.

However, in order to best address the stability in judgments of randomness, we would want to examine the minimum exposure time for which *individual* observers form reliable judgments. To ask this question, we computed a within-observer consistency measure. Specifically, for each stimulus, we compared the response between the trial in which that stimulus was presented at 500 ms and the trial in which that stimulus was presented at one of the shorter durations; if the response was the same across those two trials, it was coded as a '1', and if it was different it was coded as a '0'. We then averaged across all stimuli within a given observer to obtain a mean consistency value (with 0 meaning no two matched trials had a consistent response, and 1 meaning all trial pairs had a consistent response) for each observer. Consistency values greater than 0.5 indicate that an observer reliably agrees with themselves more than what would be expected by chance.

The results of the within-observer consistency analysis are shown in Fig 2C. There was a significant main effect of duration on consistency judgments, $F(3, 57) = 27.88$, $p < .001$, $\eta_p^2 = .59$. To evaluate the durations which showed reliable consistency with 500 ms, we ran one-sample t-tests comparing each duration to chance level of 0.5. We found that 100 ms trials exhibited reliable consistency with 500 ms, $M = .67$, $t(19) = 6.34$, $p < .001$, $d = 1.42$. Neither 50 ms nor 33.33 ms trials exhibited reliable consistency; 50 ms: $M = .52$, $t(19) = .62$, $p = .54$, $d = .14$; 33.33 ms: $M = .45$, $t(19) = -1.75$, $p = .097$, $d = -.39$. Surprisingly (though consistent with the across-observer analysis), there was below chance consistency between 16.66 ms trials and 500 ms trials, $M = .36$, $t(19) = -4.22$, $p < .001$, $d = -.94$.

Together, these data shed light on the visual nature of randomness judgments. Both transition probability and duration affected judgments of randomness, indicating both that participants are sensitive to transition probability, and that exposure duration influences judgments. As is apparent from Fig 2B, exposure durations 100 ms and 500 ms (and to a lesser extent, 50 ms) produced similar results — both to each other and to previous findings in the literature. Specifically, judgments made at these two exposure durations roughly tracked the true transition probability, but exhibited the characteristic 'over-alternation bias', with a peak in randomness judgments for stimuli with a transition probability of 0.7. Our within-observer consistency analysis formalized this observation, demonstrating that observers were reliably consistent with themselves across 100 ms and 500 ms exposure durations. The results of the across-observer correlation analysis slightly diverged, suggesting that there may also be reliable consistency across 50 ms and 500 ms.

Experiment 2

Experiment 1 suggested that stable judgments of randomness may be realized with as little as 50-100 ms. However, one limitation of the previous study concerned how the grids were generated: Specifically, the grids were generated from left to right, top to bottom, providing a sort of "reading direction" (although participants were unaware of this). Although we are interested in the stability of judgments at an individual stimulus level, which should not be affected by the reading direction, we aimed to rule out the possibility that our results rely on the directionality of the grids. Therefore, in Experiment 2, we replicated our results, but rotated all of our stimuli by 90 degrees (see Fig 3A).

Method

An independent set of 20 observers completed this task. Two additional observers were excluded from analyses based on preregistered exclusion criteria (in this case, extreme response times). The design and procedure for Experiment 2 were identical to that of Experiment 1, except that all of the presented stimuli were rotated 90 degrees, relative to how they were presented in Experiment 1 (Fig 3A).

Results & Discussion

We again first assessed how judgments of randomness were influenced by transition probability and exposure duration (Fig 3B). An ANOVA revealed a significant main effect of transition probability, $F(4, 76) = 12.93$, $p < .001$, $\eta_p^2 = .40$, a significant main effect of exposure duration, $F(4,76) = 20.13$, $p < .001$, $\eta_p^2 = .51$, and a significant interaction between the transition probability and exposure duration, $F(16, 304) = 4.60$, $p < .001$, $\eta_p^2 = .19$. As is evident from the figure, the results are quite convergent across Experiments 1 and 2, indicating that observers' judgments of randomness were relatively stable despite the rotation of the stimuli.

Next, we assessed whether there was reliable consistency in randomness judgments across-observers. Correlating group-level randomness judgments across stimuli at 100 ms and 500 ms, we again found a reliably strong correlation, $r =$

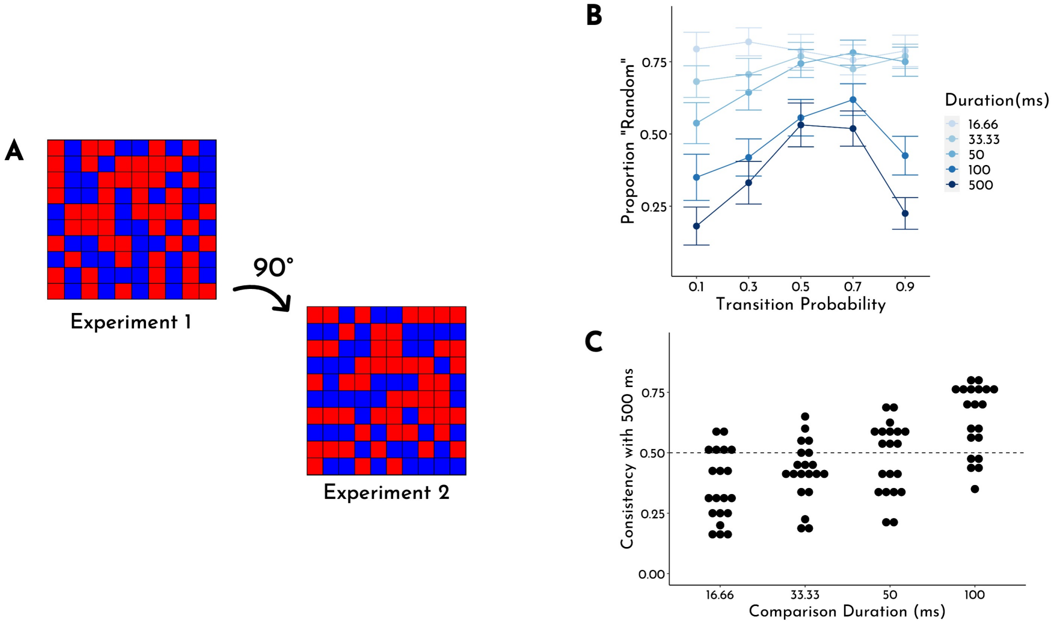

Figure 3: Experiment 2 design and results. A) The key difference between Experiment 1 and Experiment 2 is that all of the target grid stimuli in Experiment 2 were rotated 90 degrees relative to their presentation in Experiment 1. B) Randomness judgments as a function of transition probability and exposure duration. Error bars indicate +/- 1 SEM. C) Within-observer consistency results, separated out by each exposure duration compared to 500 ms. Each dot is a single observer.

.70, $t(38) = 5.98$, $p < .001$. Similarly, there was also a reliable correlation in randomness judgments made at 50 and 500 ms, $r = .51$, $t(38) = 3.68$, $p < .001$. However, judgments made at both 33.33 and 16.66 ms were uncorrelated with judgments made at 500 ms; 33.33: $r = .21$, $t(38) = 1.31$, $p = .20$; 16.66: $r = -.09$, $t(38) = -.58$, $p = .57$. These results converge with the findings of Experiment 1, demonstrating consistency in stimulus-level randomness judgments with as little as 50 ms.

In assessing the consistency in individual observers' responses across the different durations, we again found a converging pattern of results to Experiment 1 (Fig 3C). There was a significant main effect of duration on consistency judgments, $F(3, 57) = 23.27$, $p < .001$, $\eta_p^2 = .55$. Evaluating the consistency of each interval with 500 ms, we again found that 100 ms trials exhibited reliable consistency with 500 ms, $M = .64$, $t(19) = 4.33$, $p < .001$, $d = .97$. The 50 ms trials did not exhibit reliable consistency, $M = .48$, $t(19) = -.64$, $p = .53$, $d = -.14$. Similar to Experiment 1, there was also below chance consistency between 16.66 ms trials and 500 ms trials, $M = .36$, $t(19) = -4.32$, $p < .001$, $d = -.97$, as well as between the 33.33 ms trials and 500 ms trials, $M = .42$, $t(19) = -2.86$, $p = .01$, $d = -.64$.

Taken together, the results across Experiments 1 and 2 converge, with randomness judgments tracking transition probability similarly in the two studies, and stable judgments emerging with less than 100 ms.

The key difference between Experiments 1 and 2 was the rotation of the stimuli in Experiment 2. This manipulation addressed potential effects of the "reading direction" of the stimuli. If the results of these two experiments had diverged, observers might well have been "reading" the stimuli from left-to-right. However, the convergence of results across both experiments suggests that judgments of randomness arise from holistic properties of the visual stimuli themselves, independent of their orientation.

General Discussion

Here, we examined the minimal exposure time required to make stable judgments of randomness. Across two experiments, we found converging evidence that observer's judgments of randomness given only 100 ms of exposure reliably agreed with their ratings given 500 ms. Further, we replicated previous observations of an 'over-alternation bias,' (e.g., Bar-Hillel & Wagenaar, 1991) in which stimuli with more switches than would be expected by chance (i.e., a transition probability of .7) are judged to be maximally random; indeed, this bias was also evident with only 100 ms, or perhaps as little as 50ms, of exposure. Together, these data suggest that observers can form stable — and biased — impressions of randomness in under a tenth of a second.

The speed of impressions

This finding contributes to a growing body of work examining the *speed* of processing of various visual attributes. Although there is not a single processing speed that defines the boundary between 'vision' and 'high-level cognition', our finding that judgments become reliable at approximately 100 ms is broadly consistent with timing of face trait judgments (Willis & Todorov, 2006), physical stability judgments (Firestone & Scholl, 2017), and various forms of categorization (e.g., Greene & Oliva, 2009; Mack &

Palmeri, 2015). Thus, our finding points toward the interpretation that impressions of randomness can be formed very quickly and may be wired into the visual system itself.

Our findings were somewhat equivocal with respect to the precise timing; the across-observer analyses suggested that reliable judgments may be formed with 50 ms, but the within-observer analyses suggested that 100 ms is required. However, the within-observer analysis is a stronger test of our question, assessing the extent to which individual observers form stable judgments of randomness.

Notably, we found evidence of reliable *in*consistency between judgments made at the shortest durations (e.g., 16.66 ms) and 500 ms. This result most likely reflects a response bias, in which participants' guesses — given the extremely brief, masked viewing duration — are biased toward the response "random"; this can be evidenced by Figs 2B and 3B, in which it is apparent that judgments at those shortest durations have high proportion of "random" responses (and are insensitive to transition probability).

One notable difference between the stimuli in the current study, relative to other studies which have examined the speed of perception, is the relative visual complexity; to form an impression of randomness, observers must judge the relations among a set of 100 discrete objects displayed in two dimensions. This leaves open a question of whether the randomness of shorter sequences (say, of 10 items, displayed in a line) may be perceived even more quickly.

Relation to prior studies of randomness

One important feature of the current study design is that we examine *consistency* in observers' judgments of randomness (relative to themselves), rather than comparing observers' judgments relative to an objective truth. This approach enables us to draw conclusions about the extent to which well-known biases in randomness judgments, such as the over-alternation bias, emerge from biases in the visual system, as opposed to higher-level cognition or reasoning. One could imagine, for example, that judgments made with limited exposure better reflect objective reality, whereas judgments made with extended time give rise to such biases. Or, one could imagine different kinds of systematic biases emerging at different viewing durations. However, we find that the exact same biases that are present at long viewing durations (and even in high-level cognitive tasks) are also the dominant biases present after very limited exposure times, suggesting that these biases are present in early perception.

Our results converge with other lines of work examining the perceptual roots of impressions of randomness. Specifically, some work has argued that biases in judgments of randomness emerge from low-level perceptual and cognitive constraints, such as attention and short-term memory capacity (Hahn & Warren, 2009; Warren et al., 2018; Yu et al., 2018b). Our work takes a different approach, examining the speed at which these impressions are formed as a proxy for their dependence on perceptual systems.

In doing so, our work builds on existing work demonstrating the rapidity at which other kinds of statistical information (e.g., information about the *mean* size of a set of shapes; e.g., Whitney & Leib, 2018; Luo & Zhao, 2018). Here, we extend this body of work, demonstrating that not only properties such as averages are relatively automatically extracted and represented; higher-order properties like the statistical contingencies among items may also be realized within the visual system itself.

Here, we took advantage of a well-known method to manipulate randomness (i.e., transition probability) and a well-known bias in randomness perception (i.e., over-alternation-bias) to examine the reliability in observers' judgments of the likelihood of binary outcomes. However, this raises several open questions about how our findings generalize. First, although we operationalized randomness in this study with transition probability in the horizontal direction, there are other ways to quantify randomness (e.g., with measurements of formal complexity) and to manipulate transition probability in multiple dimensions at once (e.g., Brady & Tenenbaum, 2013). Generalizing our findings to other randomness metrics could not only help to establish robustness of our findings, but also shed light on what *kind* of randomness observers are extracting and representing from these displays.

Second, we experience many other forms of randomness in the world that are not merely limited to binary outcomes. For example, we observe structure and randomness in things like the clustering of items in space or even in looking at economic trends. The randomness in these kinds of information cannot be as easily quantified by simple properties like transition probability, but rather may result from a more complex statistical inference process (e.g., Griffiths et al., 2018). It thus may be fruitful for future work to examine the time course of randomness judgments using other kinds of randomly generated stimuli; such work may reveal the boundary conditions or circumstances under which certain kinds of randomness judgments rely on visual versus higher-level cognitive processes. In other words, this may help us to understand what kinds of structure are and are not directly perceivable in the first place. Replicating these experiments with other kinds of random stimuli will also help to create a deeper understanding of how the visual system parses spatially structured information and translates it into judgments of randomness.

Conclusion

The ability to perceive structure in the environment — to separate signal from noise — is foundational to all cognitive organisms. Indeed, all forms of learning rely on the ability to detect meaningful statistical structure in the environment. While most work on statistical impressions in humans has focused on higher-level cognition, here we shed light on the perceptual bases of these impressions. We show that stable biases of subjective randomness manifest with perhaps as little as a twentieth of a second of exposure time, providing a strong signal that biases of statistical processing may be deeply rooted — perhaps even wired into perceptual systems themselves.

References

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics, 12*(4), 428-454.

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*, *120*(1), 85-109.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62*(8), 1572-1595.

Colombatto, C., Uddenberg, S., & Scholl, B. J. (2021). The Efficiency of Demography in Face Perception. *Attention, Perception, & Psychophysics, 83*(8), 3104-3117.

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*(2), 301-318.

Firestone, C., & Scholl, B. (2017). Seeing physics in the blink of an eye. *Journal of Vision, 17*(10), 203-203.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*(3), 295-314.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20*(4), 464-472.

Griffiths, T. L., Daniels, D., Austerweil, J. L., & Tenenbaum, J. B. (2018). Subjective randomness as statistical inference. *Cognitive Psychology, 103*, 85-109.

Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences, 21*(6), 475-492.

Hahn, U., & Warren, P. A. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review, 116*(2), 454-461.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430-454.

Luo, A. X., & Zhao, J. (2018). Capacity limit of ensemble perception of multiple spatially intermixed sets. *Attention, Perception, & Psychophysics*, *80*(8), 2033-2047.

Mack, M. L., & Palmeri, T. J. (2015). The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General, 144*(3), 551-569.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 109*(2), 330-357.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Reimers, S., Donkin, C., & Le Pelley, M. E. (2018). Perceptions of randomness in binary sequences: Normative, heuristic, or both? *Cognition, 172*, 11-25.

Reiner, C., Yousif, S., Sherman, B., & Keil, F. (2021). Common structure underlying visual and non-visual judgments of randomness. *Journal of Vision, 21*(9), 2254-2254.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention, 4629*, 197-229.

Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15-20.

Warren, P. A., Gostoli, U., Farmer, G. D., El-Deredy, W., & Hahn, U. (2018). A re-examination of “bias” in human randomness perception. *Journal of Experimental Psychology: Human Perception and Performance, 44*(5), 663-680.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105-129.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592-598.

Yu, R. Q., Gunn, J., Osherson, D., & Zhao, J. (2018). The consistency of the subjective concept of randomness. *Quarterly Journal of Experimental Psychology, 71*(4), 906-916.

Yu, R. Q., Osherson, D., & Zhao, J. (2018). Alternation blindness in the representation of binary sequences. *Journal of Experimental Psychology: Human Perception and Performance, 44*(3), 493-502.

Zhao, J., Hahn, U., & Osherson, D. (2014). Perception and identification of random events. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1358-1371.