

Perceptions of Randomness: Why Three Heads Are Better Than Four

Ulrike Hahn and Paul A. Warren
Cardiff University

A long tradition of psychological research has lamented the systematic errors and biases in people's perception of the characteristics of sequences generated by a random mechanism such as a coin toss. It is proposed that once the likely nature of people's actual experience of such processes is taken into account, these "errors" and "biases" actually emerge as apt reflections of the probabilistic characteristics of sequences of random events. Specifically, seeming biases reflect the subjective experience of a finite data stream for an agent with a limited short-term memory capacity. Consequently, these biases seem testimony not to the limitations of people's intuitive statistics but rather to the extent to which the human cognitive system is finely attuned to the statistics of the environment.

Keywords: randomness, intuitive statistics, representativeness, gambler's fallacy, probability

In a famous set of experiments, Kahneman and Tversky (1972) demonstrated that people's perceptions of randomness were systematically biased. Asked to indicate, for example, which of the exact orders of outcomes from the toss of an unbiased coin is more likely, HTHHTT or HHHHHH (where H = heads and T = tails), people reliably choose the former. However, as the statistician knows, both orders are equally likely to occur when tossing a coin six times. Laypeople, it seems, erroneously expect that the essential characteristics of the generating process will be represented, "not only globally in the entire sequence, but also locally in each of its parts" (Kahneman & Tversky, 1972, p. 435). A random generating process such as tossing an unbiased coin will lead to global, long-run sequences that contain equal numbers of heads and tails. Our lay intuition is that this property is reflected in short sequences also. Decades of research have established three inter-related aspects of people's misperceptions of randomness:

1. People think a sequence is more likely, and hence random, if there is some irregularity in order of appearance (e.g., HHTHTH vs. HTHTHT).
2. People think a sequence is more likely, and hence random, if the equiprobable outcomes occur equally often.
3. The outcome alternation rate (i.e., how often H switches to T and vice versa) that people consider to be random is higher than that associated with chance.

(For reviews and detailed references, see, e.g., Bar-Hillel & Wagenaar, 1991; Falk & Konold, 1997; Kareev, 1992; Nickerson, 2002; Rapaport & Budescu, 1992.)

Furthermore, these misperceptions can be linked to the so-called gambler's fallacy: Some gamblers believe that after a long sequence of red in roulette, an outcome of black is more likely on the next spin. Of course, given a fair roulette wheel, this is not the case: Both outcomes are equally likely on any given trial. This misperception has likewise been attributed to the belief that local and global sequences should share the same properties: Black on the next trial would "result in a more representative sequence than the occurrence of an additional red" (Tversky, 1974, p. 151).

This set of errors and biases is one of the most well-documented phenomena in cognitive science and features in any introductory lecture on judgment and decision making. In many ways, however, these errors are baffling. People seem to be good at finding structure in the world, but how is this possible if their perceptions of randomness are so poor? Can people's undoubted sensitivity to the structure of the environment be reconciled with these seeming errors of judgment? We propose that the answer to this question is *yes*: There is a simple way in which people's supposed misperceptions reflect environmental statistics and hence a simple way in which they are actually correct.

In general, the basic tool for demonstrating the correct answer to queries about sequence probability is the combinatorial tree representing all sequences of a given length. Figure 1 shows the tree for all possible outcome combinations from four coin tosses. The exact order HHHH is simply one branch of this tree, whereas the sequence HHTT is another. Each branch represents one possible sequence from 16 and has an equal probability of 1/16 of occurring at random. This confirms the claim that people are wrong to perceive one of these exact orders as more likely than the other.

More precisely, the probability tree confirms the claim that each sequence or exact order of length 4 is equally likely given four tosses of an unbiased coin. However, one can also look at sequence probabilities in another way. Specifically, one could consider the probability of occurrence of a given sequence of length k as a *substring* in a longer, global sequence of length n (i.e., $k < n$). One

Ulrike Hahn, School of Psychology, Cardiff University, Cardiff, United Kingdom; Paul A. Warren, Wales Institute of Cognitive Neuroscience, School of Psychology, Cardiff University.

We thank Erwin Hahn, Adam Corner, Todd Bailey, Mark Johansen, Laurel Evans, Rebecca Champion, and James Gray for valuable comments on drafts of this article.

Correspondence concerning this article should be addressed to Ulrike Hahn, School of Psychology, Cardiff University, Cardiff CF10 3AT, United Kingdom. E-mail: hahn@cf.ac.uk

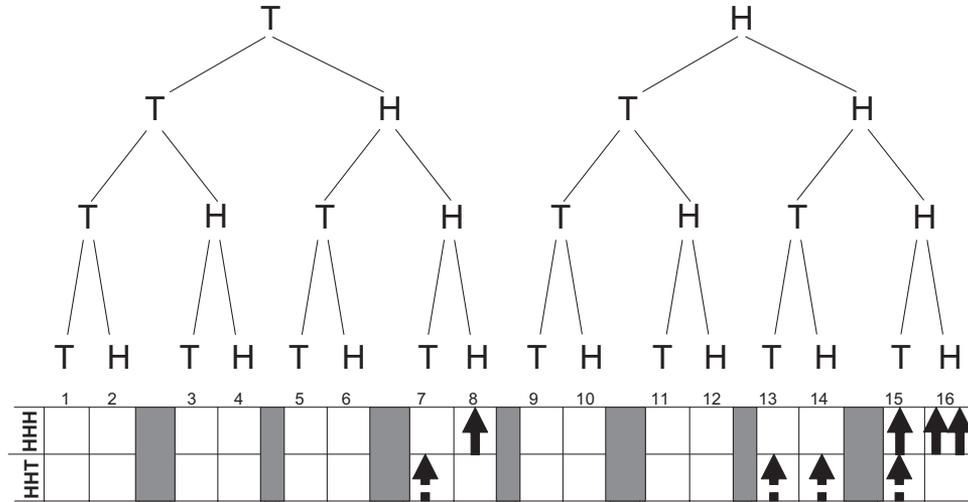


Figure 1. A probability tree indicating the 16 possible outcomes of a sequence of four coin tosses. Also marked are the outcomes on which HHH (unbroken arrows) and HHT (broken arrows) occur. Note that although there are four occurrences of both HHH and HHT in the tree, in the case of HHT, these occur in four different independent outcomes (HHHT, THHT, HHTH, HHTT), whereas for HHH they occur in only three (HHHH, THHH, HHHT) because two occurrences are in the same outcome (HHHH).

might, for example, think of this global sequence as a data stream that is unfolding over time. The probability of a sequence such as HHHH now becomes the probability of its occurrence as a substring as one moves through this global sequence, as illustrated in Figure 2.

If the global sequence is of infinite length (i.e., when $n \rightarrow \infty$), then, once again, it can be shown that all possible (substring) sequences are equiprobable. So, for example, in the $k = 4$ case, when tossing a fair coin, HHHH will appear as a substring of an infinite sequence of coin flips as often as will, say, HHTT. A Bernoulli process involving infinitely many independent trials, each with equiprobable outcomes (e.g., tossing a fair coin), will generate an infinite sequence in which not only the basic outcomes (e.g., H or T) occur equally often but also all possible sequences of a given length k will appear equally often, each with probability $1/2^k$ (by Borel’s strong law of large numbers; see, e.g., Beltrami, 1999).

All possible (substring) sequences are not, however, equiprobable if the global sequence in question is finite. Consider the probability of observing the exact order HHHH as opposed to HHTT within a longer, unfolding sequence of coin tosses of finite length, n . The probability of observing either of these two orders is not equal, and we propose that this holds the solution to the seeming puzzle of people’s misperceptions of randomness.

Once again, this result can be demonstrated using the combinatorial tree in Figure 1. For the sake of simplicity (but without loss of generality), we examine the case in which $k = 3$ and $n = 4$. Counting the number of occurrences of the exact orders HHH and HHT as local substrings within the global sequences represented by that tree, we make the following simple observation: Although the total number of occurrences of each is equal across the tree (viz., four), their distribution across branches (i.e., overall sequences of length 4) is not. Specifically, two of the four occurrences of HHH reside in a single branch of the tree, whereas the occurrences of HHT are distributed across four separate branches.

This means that given a particular global sequence of length 4, which corresponds to selecting one path through this tree, one is less likely to observe the local substring HHH within that global sequence than to observe the substring HHT. This relationship arises as a consequence of the fundamental laws of probability. A given global sequence is a particular branch of the tree, and different branches are mutually exclusive events, because one attempt at flipping a coin four times can give rise to one specific sequence only. Accordingly, the chance that one will see a particular subsequence in a series of coin tosses depends on the number of different branches in which it occurs. Examining the combinatorial tree demonstrates that in this framework, local substrings of all heads, such as HHH, will tend to accumulate through multiple occurrences in even longer strings of just heads. An intuitive metaphor for this tendency is to think of waiting for a string such as HHH as being a bit like the situation faced by the weary commuter waiting at a bus stop: For a long and frustrating period, there is no bus in sight, and then, all of a sudden, several arrive in immediate succession.

This result clearly generalizes to global sequences of finite length larger than 4 and local substrings of finite length larger than 3 (i.e. $k < n < \infty$). More formally, as a consequence of the way different substrings are distributed across the combinatorial tree, the average number of coin tosses that one has to wait before encountering a particular substring—known in the field of combinatorics as its *wait time*—is not equal across local substrings. The expected wait time for HHH is 14 tosses of a coin, whereas the



Figure 2. A finite attentional window moving through a data stream. Items within the window (square) represent a sequence held in short-term memory.

wait time for HHT is only 8 tosses of a coin. This difference, of course, is not restricted to sequences of length 3: To encounter HHTT, one must wait, on average, 16 tosses of the coin, whereas for HHHH, the wait time is 30 (there exists a simple algorithm for calculating wait times developed by John Conway; see Gardner, 1988, p. 67; for related proofs, see Guibas & Odlyzko, 1981). Figure 3A shows the average wait times for all possible sequences of length 4 ($k = 4$).

This also means that, given finite exposure, one will not be equally likely to encounter these sequences. If, for example, one stops only to observe a total, global sequence of 20 coin tosses, there is a good chance that one will not have observed HHHH, and this probability of a no show is considerably higher for HHHH than it is for, say, HHHT.

Figure 3B shows, for all possible local substrings of length 4 ($k = 4$), their probability of not occurring at all as a subse-

quence within a global, overall sequence of 20 coin tosses ($n = 20$). As can readily be seen from the shape of this graph, the probability of nonoccurrence within the global sequence is directly related to the difference in wait time.

Although interesting, these results might be less useful for explaining human behavior, if the local substring within a global sequence framework bore no resemblance to human experience. However, it seems to capture people's environmental experience very naturally: Given people's finite resources (and life span), any particular data stream that they might experience—such as someone tossing a coin several times over—is necessarily finite. Furthermore, attentional and/or memory limitations mean that people can hold in memory concurrently only a limited number of observed items. Consequently, people are also capable of monitoring concurrently only a fixed number of items within that data stream; that is,

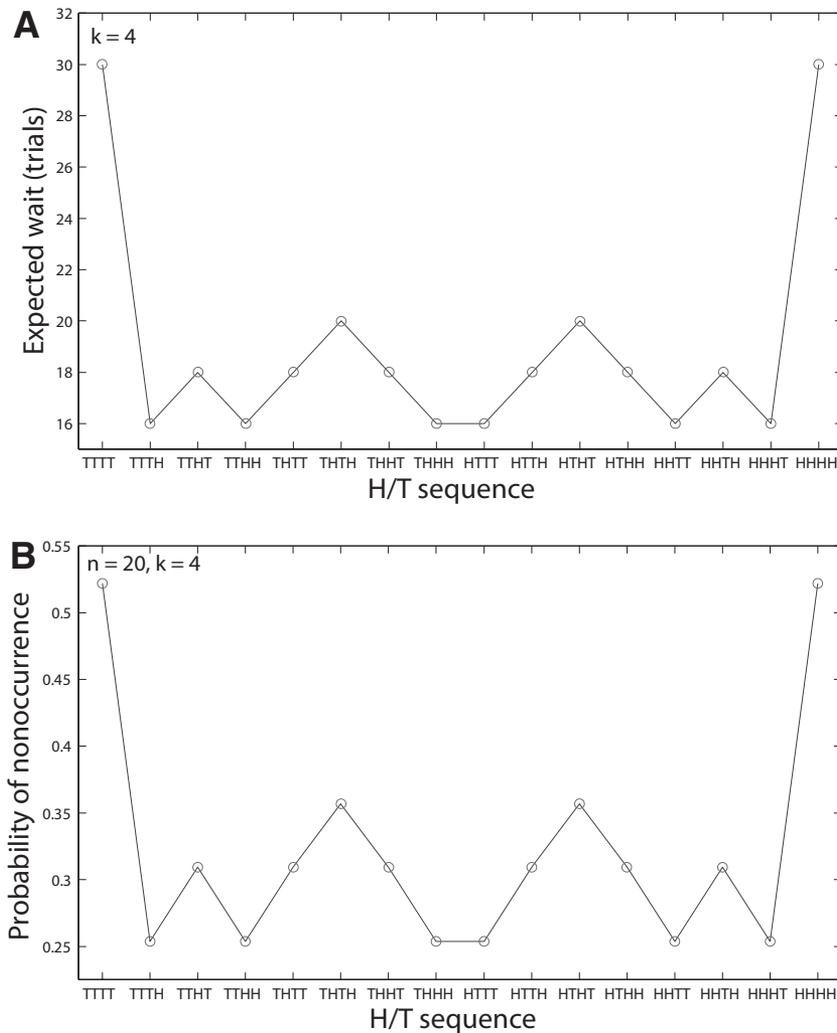


Figure 3. Average wait time (A) and probability of nonoccurrence (B) for coin toss substrings of length 4 in global sequences of length 20. Note that wait time is actually independent of global sequence length. Wait time and probability of nonoccurrence are directly related; substring HHHH has the highest average wait time and the highest probability of nonoccurrence. The alternating substrings HTHT and THTH have the next highest wait time (highest probability of nonoccurrence).

people experience something like a moving attentional window through a finite data stream (see Figure 2).

It turns out, further, that the probabilities of observing particular substrings differ substantially for global sequences of moderate length. Figure 4A shows the result of simple simulations to assess the probability of nonoccurrence of a local coin-toss substring of length 4 ($k = 4$) in global sequences of coin tosses of variable length n . For each global sequence length, the nonoccurrence probabilities are shown for the local substrings HHHH, HTHH, HHHT, HHTT, and HTHT. The graphs depicted were derived by randomly generating 10,000 global sequences and calculating the proportion for which the substring was not present, which translates directly into a stable estimate of the substring's nonoccurrence.¹

As can be seen from the figure, the relative probabilities (which are constrained to be equal in the case $k = n$) differentiate rapidly as the length, n , of the global sequence increases. At the same time, as the global sequence gets longer, the overall probabilities of nonoccurrence approach zero. Consequently, although differences persist, they become harder to detect because the absolute numbers become very small. However, the differences do seem readily detectable in shorter global sequences, such as those of length 10 to 20. This is encouraging, as it also seems to be a plausible value

for the length of data stream that people might be most likely to experience, given that only the hardened gambler or lay statistician is likely to hang around for many more tosses of a coin on any given occasion. Note, in particular, that the heads only (HHHH) and perfect alternation (HTHT) cases are different from the remainder of the possible substrings—they are significantly less likely to arise.

To dispel any concerns about generality, Figure 4B shows results for corresponding substrings of length 6. Again, the heads only and perfect alternation cases appear distinct. Moreover, they separate out more clearly from the other sequences.

Note that although the exact size of the window (i.e., the substring length k) probably has relevance in matching closely people's relative perceptions of sequences, for the general argument it matters only that there is a fixed window moving through a longer, finite data stream. The result that substring occurrence probabilities are not equal is entirely general with regard to the length of both.

In summary then, consideration of substring occurrence probability at only two extremes, that is, in the cases when global sequence and subsequence are the same length ($k = n$) or when the global sequence is infinite ($k < n \rightarrow \infty$), obscures the fact that local subsequences can be reflective of the global properties of a sequence. There is not only a sense in which laypeople are correct, given a realistic but minimal model of their experience, that different exact orders are not equiprobable, it seems that the same experience might be able to provide a useful explanation of why some sequences are perceived to be special.

One of the fundamental "misperceptions" or "biases" identified by psychological research is that people think that sequences with some irregularity in the order of appearance are more likely given an unbiased coin (Point 1 above) and that this applies even to sequences of lengths as short as 6 or 7 (Bar-Hillel & Wagenaar, 1991). Figure 4 shows that people's intuitions regarding irregularity have probabilistic support in the perceived environment, in that the most regular sequences—uniform runs and perfect alternations—are indeed less likely to be observed given a data stream from an unbiased coin.

What then of the expectation of equal numbers of heads and tails within a local sequence (i.e., the misperception mentioned in Point

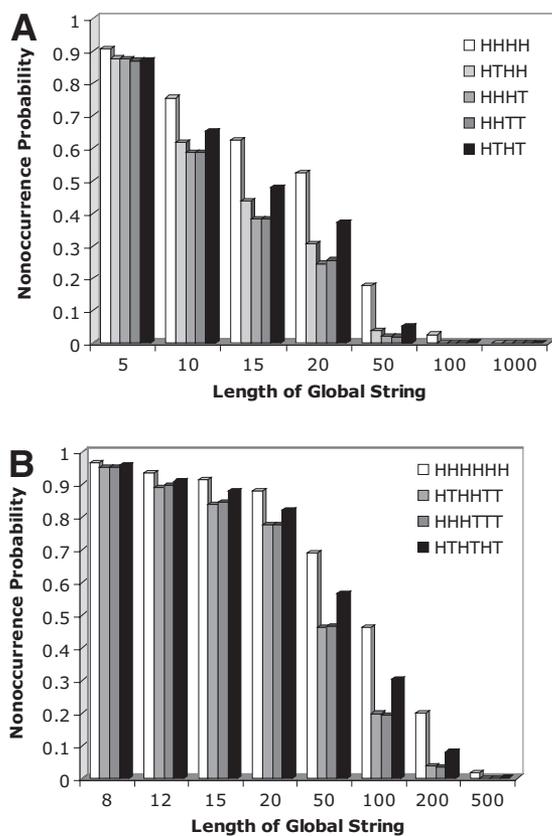


Figure 4. Results of simulations to examine the relative probability of nonoccurrence of different coin-toss substrings of length 4 (A) and length 6 (B) from among global sequences of variable length. The bars represent the proportion of global sequences within that sample of 10,000 in which the local substring in question did not occur.

¹ These simulations were conducted in the R statistics environment (R Development Core Team, 2008) and also replicated in MATLAB (Mathworks, 2007). The reader should rest assured that these results are a direct consequence of fundamental combinatorial properties and not deficiencies in the random number generators of this software. Deriving exact analytical expressions for the nonoccurrence probability of these sequences is clearly possible but nontrivial. For example, in the simplest case (HHHH), it is known that the probability of observing no string of k consecutive heads in a sequence of n coin tosses is given by $F_{n+2}^k / 2^n$, where F_{n+2}^k is the $(n + 2)$ th term of the generalized k -step Fibonacci sequence. For the $k = 4$ case simulated (see Figure 4A), the corresponding exact probabilities for the different global sequence lengths are (to 2 d.p.) {0.91, 0.75, 0.63, 0.52, 0.17, 0.03, 0.00}. These values are very close to those presented in Figure 4A. Given that the analytic solution adds nothing to the argument presented, we content ourselves with the numerical approximation. The other alternative for calculating exact probabilities would be to count substring occurrences across branches of the full factorial tree. Because this tree has 2^n branches the problem very quickly becomes intractable.

2 above)? Faced with a sequence such as HTHHTT, one might simply view this as a sequence of three heads and three tails rather than a specific order. This drastically changes the probabilities associated with the sequence. Contrasting a sequence of three heads and three tails with a sequence with five heads and one tail, such as HHTHHH, the former is three times more likely, whereas the difference between the specific orders is negligible. Indeed, researchers have noted before that the expectation of equal numbers seems commensurate with a focus on proportions rather than specific orders (Kareev, 1992).

Of course, experimental studies such as Kahneman and Tversky (1972) ask questions about exact order. Even if one assumes that people understand this question as intended by the experimenter (i.e., paying attention to order), if they do not have any direct estimates for the probability of a specific order, then this quantity must be derived from some other source.

What our sequence-subsequence model shows is that there is good reason to not distinguish carefully between the majority of sequences: Although there are some differences between the probabilities of occurrence of different sequences with the same proportion of heads and tails, these differences are small, particularly when compared with the difference between any sequence and the two special cases—uniform runs and perfect alternation—we have highlighted. For example, of the 20 possible sequences of length 6 containing equal numbers of heads and tails, all have wait times between 64 and 68, whereas the wait time for the two perfect alternations is 84, and the wait time for the uniform runs is 126 (for attendant effects on probabilities, see Figure 4). Consequently, that people might base judgments on an estimate of the proportion of outcomes rather than explicitly code each possible exact sequence seems reasonable. There is no reason to encode or monitor many exact orders with the same proportion of heads and tails if, in terms of likelihood of occurrence, they are very similar. Instead, it makes sense to combine all sequences with the same proportion of heads and tails together into a single class and reserve special status for the cases of perfect alternation and uniform runs.

Of the fundamental aspects of people's biases or misperceptions, there remains only the preference for overalternations (Point 3 above) to be explained. This also follows directly from the local substring model because the average alternation rate for short sequences is greater than the long run value of .5 (Kareev, 1992). Short sequences tend to have more alternations between heads and tails than would be expected in an infinitely long series of coin tosses. For the subset of sequences containing equal numbers of heads and tails outcomes, the percentage increase in alternations relative to the long run rate is 34% for sequences of length 4, 20% for length 6, and 12% for sequences of length 10 (calculated simply by counting and averaging the number of alternations in all relevant sequences of that length). This increase is considerable. Furthermore, the increase is not limited to sequences with equal numbers of heads and tails but arises for all relative proportions in this length range. Consequently, once again, starting from a reasonable characterization of people's experience of random sequences reveals supposed errors and biases to be accurate reflections of the statistical properties of that experience. People are not wrong to view sequences with higher than long-run alternation rates as likely outcomes of a random generating process.

Finally, the issue of overalternations leads on to the gambler's fallacy. One of the main sources of experimental evidence for

people's belief in this fallacy has come from the fact that people overalternate when asked to generate random sequences. However, it should be clear from the preceding analysis that one could easily view higher than long-run alternation rates as likely characteristics of finite-length randomly generated sequences (because they are) but not endorse the gambler's fallacy. Hence more direct evidence for belief in the gambler's fallacy is required (see also Nickerson, 2002, on other requirements for what might count as good evidence in this context).

Any such belief in the gambler's fallacy, strictly speaking, would be erroneous, given that successive tosses of a coin are independent. However, Figures 2 and 4 also indicate that the error is an understandable one. The average wait time for HHHT is almost half of that for HHHH. So under conditions that match one's experience, an occurrence of HHHT is much more likely than HHHH (it is just not more likely once one has already seen HHH). Given finite exposure, one is more likely to encounter HHHT than HHHH, and in a global sequence of length 20, the most likely outcome is not to see the substring HHHH at all (see Figure 4B).

Furthermore, as the length of the substring increases, the wait times (and hence chances of nonoccurrence) for the k heads outcome, $\text{wait}(kH)$, and the $k - 1$ heads followed by a tails outcome, $\text{wait}([k - 1]H + T)$, diverge more and more sharply. This can be seen from the general expression for the ratio of wait times for these cases:²

$$\begin{aligned} \frac{\text{wait}(kH)}{\text{wait}([k - 1]H + T)} &= \frac{2 \times (2^{k-1} + 2^{k-2} + \dots + 2^0)}{2 \times 2^{k-1}} \\ &= 1 + \sum_{i=1}^{k-1} \frac{1}{2^i} \end{aligned}$$

In other words, as the length, k , of the substring increases, the ratio of the two wait times also increases; that is, as k increases, it is necessary to wait increasingly longer for k heads relative to $k - 1$ heads followed by a tails. In a sense, there is something right about the gambler's intuition that the longer the run, the more likely, by contrast, is a sequence with a final tails.

Moreover, the differences in wait time mean that it is possible to construct a new gambling game, with a genuine winning strategy that may well appear consistent with the gambler's fallacy, even to trained statisticians. If a prior limit of 20 coin tosses were decided and one were to bet that substring HHHT would occur as opposed to HHHH, then, on average, one would win. At length 20, the probability of nonoccurrence for HHHH is roughly 0.5, whereas for HHHT it is roughly 0.25. The fact that this bet, which seems rather subtly different from the normal gambling situation, will lead to long-run gains suggests that the gambler's fallacy, even where present, is not as gross an error as it might seem.

The absolute length of the wait times also means that if you are betting only for a comparatively short time, you are quite unlikely to receive feedback that could disabuse you of the fallacy. Moreover, the difficulty of learning through exposure that the gambler's fallacy is indeed a fallacy is further compounded by the fact that it

² We derive this formula from the algorithm for calculating wait time developed by John Conway (see Gardner, 1988, p. 67).

is, of course, not a fallacy for sampling without replacement. With every removal of a red ball drawn from a finite sample of red and blue balls in an urn, the probability of drawing another red ball goes down. However, from a purely observational perspective, sampling with replacement and sampling without replacement are surprisingly hard to tell apart within the model of experience defined here. This is demonstrated in Figure 5, which shows relative probabilities of nonoccurrence of substrings (length 4) of red (R) and blue (B) balls within a global sequence of length 20 for the case of sampling with replacement (leftmost data). This case (for which subsequent draws are independent from one another) is equivalent to the coin-tossing example seen in Figure 4A. The remaining sets of data in Figure 5 show the results for the corresponding sequences in a situation of sampling without replacement. Specifically, as in Figures 4A and 4B, the bars represent the probability of nonoccurrence of each of the specified (sub)sequences of red and blue balls within a global sequence of length 20. As before, this probability is estimated by the proportion of global sequences that did not contain the substring in a simulated sample of 10,000 randomly generated global sequences. The only difference lies in the generating process, which now involves sampling without replacement. The fact that the gambler's fallacy is not a fallacy under these circumstances is reflected in the very slight increase in the probability of nonoccurrence for the uniform run RRRR, an increase that becomes more pronounced as the number of balls in the urn decreases. However, not only would this increase be hard to detect, but the overall patterns are strikingly similar. Consequently, one would be hard-pressed to decide on the basis of the outputs of the source alone whether the source involved sampling with or without replacement.

Considering all of this, the gambler's fallacy, although clearly erroneous, seems a rather more reasonable, and hence subtle, error than psychologists and statisticians seem to assume. Given the analytic tools of combinatorics and probability theory, the error of the gambler's fallacy might seem obvious (although coin tosses still hold many counterintuitive pitfalls, even for those with considerable statistical competence; e.g., see the seeming paradoxes

associated with average wait time and *pairwise precedence*, defined as the probability with which one of two specific sequences, such as HHT vs. HTH, will be encountered first; Gardner, 1988). The tools of combinatorics and probability theory themselves, however, are not at all obvious: The fact that these branches of mathematics are rather late developments (mainly 17th century) suggests caution is warranted regarding their simplicity. This, in turn, means that an error that only becomes apparent once these sophisticated tools are available might be considered rather difficult to detect. This mitigates against classification of the gambler's fallacy as a severe deficit in intuitive statistics.

Discussion

We present a very simple model of the experienced environment that is based on the undeniable assumption that human experience is finite and the uncontroversial assumption that human short-term memory is of limited capacity. Applying this model of experience to the sequential outputs of a random generating process reveals that the key aspects of laypeople's supposed misperceptions of randomness actually have probabilistic support. The supposed errors are either direct reflections of statistical properties of the experienced environment or the result of reasonable inferences from that experience.

Answering "the latter" to an experimenter's question of which exact order is more likely to arise from an unbiased coin—HHHHHH or HTHHTT—is at least as right as it is wrong. It is wrong only in two limiting cases: given the additional assumption that a coin has been tossed exactly six times, or, if the question is understood to be about the probability of occurrence of these sequences more generally, if the sample from which they are drawn is infinite. If the sample is finite, then the preference for HTHHTT is correct.

Much has been made of people's supposed errors, not just in terms of how well-calibrated people's intuitive statistics are but also in terms of the processes that underpin them. Kahneman and Tversky have used these errors to argue that likelihood judgments, in general, are based on representativeness, a heuristic based on similarity, not statistical inference (e.g., Kahneman & Frederick, 2001; Kahneman & Tversky, 1973; Tversky & Kahneman, 1982): What people are assumed to do is evaluate the resemblance or perceived similarity between local and global sequences as a way of judging them. However, these "similarities"—namely, irregularity and balanced proportions of outcomes—are equally reflections of the statistics of the experienced environment and so cannot distinguish between a heuristic or statistical account.

More generally, Rapaport and Budescu distinguished three kinds of explanation in the literature put forward to explain the "the common wisdom in this area of research," namely, "that people perform poorly when required to either recognize or generate random sequences" (Rapaport & Budescu, 1992, p. 360): (a) People have a biased notion of subjective randomness that deviates systematically from the statistical one, (b) people have no direct prior experience of such tasks and therefore lack the skills to perform well on them, and (c) people have an accurate notion of randomness but fail to reveal it in their behavior for other functional limitations (such as memory). From the present perspective, these are linked, because the observed biases reflect people's

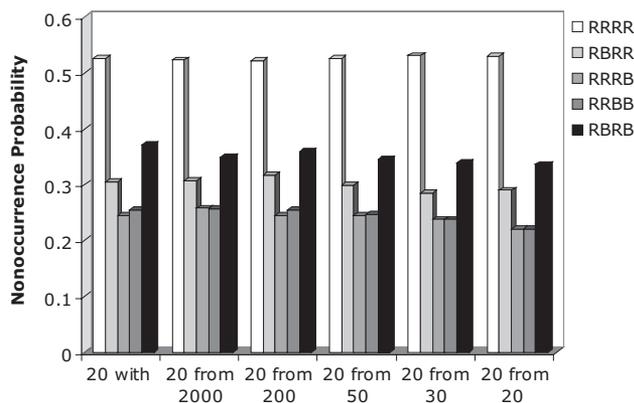


Figure 5. Results of simulations to examine the relative probability of nonoccurrence of local substrings of red (R) and blue (B) balls within a global sequence of length 20 for the case of sampling with replacement (leftmost data) and sampling without replacement from urns with decreasing initial numbers, n , of red and blue balls (urn size $n = 2,000$ balls, $n = 200$ balls, $n = 50$ balls, $n = 30$ balls, $n = 20$ balls, respectively).

experienced environment and that environment is shaped inherently by short-term memory limitations.

This is not, of course, the environment that has been assumed by experimental studies with their emphasis on long-run frequency, not just in theoretical analysis but also experimental design. For example, experimenters often ask people to generate random sequences comprising several hundred trials, that is, sequences that are far longer than any they are likely to have observed (e.g., see Falk & Konold, 1997, Table 1, for a summary). If people's biases are reflective of their experience, as we argue, it should come as no surprise that people also seem to respond well to corrective experience in the form of training and feedback in generation and prediction tasks (e.g., for generation tasks, see Neuringer, 1986; Rapaport & Budescu, 1992; for prediction, see Edwards, 1961a, 1961b; in the context of detection, see Lopes & Oden, 1987). Furthermore, there has been evidence that peoples' performance in sequence generation tasks is influenced by the engagement of attentional and memory resources; this provides corroborative support for our notion that memory and attentional windows are central to the experience of randomness (e.g., Baddeley, 1966; Kareev, 1992; Wagenaar, 1972; Weiss, 1965).

Finally, it is worth noting that closer consideration of both memory and attention on the one hand and the nature of people's actual experience on the other is gaining prominence in judgment and decision-making research more generally (e.g., Kareev, 1992, 2000; for an overview, see, e.g., Weber & Johnson, 2009). Closer consideration of everyday experience has, in several cases, led to more positive verdicts regarding seeming biases, errors, or limitations (e.g., Gigerenzer, Hell, & Blank, 1988; Hertwig & Pleskac, 2008; Oaksford & Chater, 1994). It has also become apparent that the way information is presented to participants in experiments—whether by description or by experience—can have a large effect on the patterns of behavior observed (see, e.g., in the context of judgment, Fiedler, Brinkmann, Betsch, & Wild, 2000; in the context of decision making, Barron & Erev, 2003, and Hertwig, Barron, Weber, & Erev, 2004). Finally, closer attention has been paid to the way participants subjectively interpret and experience a given experimental task. In particular, more careful consideration of the pragmatics of natural language communication has, in a number of cases, led to more positive evaluations of behavior in light of normative models (e.g., Hilton, 1995; McKenzie, 2004).

Of course, not all of the many biases and errors in human judgment and decision making have vanished after closer analysis, nor does it seem likely that all of them will. However, a good number of them have (see also, e.g., Erev, Wallsten, & Budescu, 1994; Juslin, Winman, & Olsson, 2000) and many continue to be the topic of heated debate (e.g., the *conjunction fallacy*, Hertwig, Benz, & Krauss, 2008; or, more closely related to the present paper, the *hot hand fallacy*, Bar-Eli, Avugos, & Raab, 2006). Furthermore, there is growing evidence that many seeming biases might be better characterized as the result of an unbiased statistical evaluation of an (often necessarily) biased experiential sample (e.g., Bar-Hillel, Budescu, & Amar, 2008; Fiedler, 2000; Fiedler et al., 2000; Hansson, Juslin, & Winman, 2008; Krizan & Windschitl, 2007). This possibility seems particularly pertinent to the material presented here.

Whatever the ultimate verdict on human rationality in general turns out to be, if one accepts the sequence–subsequence model presented here as a characterization of peoples' actual experiences,

then the case against lay intuitions about randomness needs to be reopened. Both how and how well people identify random sources are questions in need of reevaluation.

References

- Baddeley, A. D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.
- Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty years of “hot hand” research: Review and critique. *Psychology of Sport and Exercise*, *7*, 525–553.
- Bar-Hillel, M., Budescu, D. V., & Amar, M. (2008). Predicting World Cup results: Do goals seem more likely when they pay off? *Psychonomic Bulletin & Review*, *15*, 278–283.
- Bar-Hillel, M., & Wagenaar, W. (1991). The perception of randomness. *Advances in Applied Mathematics*, *12*, 428–454.
- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215–233.
- Beltrami, E. (1999). *What is random? Chance and order in mathematics and life*. New York: Springer.
- Edwards, W. (1961a). Probability learning in 1,000 trials. *Journal of Experimental Psychology*, *62*, 385–394.
- Edwards, W. (1961b). Unlearning the gambler's fallacy. *Journal of Experimental Psychology*, *62*, 630.
- Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, *101*, 519–527.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, *104*, 301–318.
- Fiedler, K. (2000). Beware of samples! A cognitive–ecological sampling approach to judgment biases. *Psychological Review*, *107*, 659–676.
- Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General*, *129*, 399–418.
- Gardner, M. (1988). *Time travel and other mathematical bewilderments*. New York: Freeman.
- Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 513–525.
- Guibas, L. J., & Odlyzko, A. M. (1981). String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory*, *30*, 183–208.
- Hansson, P., Juslin, P., & Winman, A. (2008). The naïve statistician: Organism–environment relations from yet another angle. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 237–260). New York: Oxford University Press.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539.
- Hertwig, R., Benz, B., & Krauss, S. (2008). The conjunction fallacy and the many meanings of *and*. *Cognition*, *108*, 740–753.
- Hertwig, R., & Pleskac, T. J. (2008). The game of life: How small samples render choice simpler. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 209–236). New York: Oxford University Press.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, *118*, 248–271.
- Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review*, *107*, 384–396.
- Kahneman, D., & Frederick, S. (2001). Representativeness revisited: At-

- tribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kareev, Y. (1992). Not that bad after all: Generation of random sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1189–1194.
- Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, 107, 397–402.
- Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin*, 133, 95–121.
- Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 392–400.
- Mathworks. (2007). MATLAB (Version 7.4). Natick, MA: Author.
- McKenzie, C. R. M. (2004). Framing effects in inference tasks—and why they are normatively defensible. *Memory & Cognition*, 32, 874–875.
- Neuringer, A. (1986). Can people behave “randomly”? The role of feedback. *Journal of Experimental Psychology: General*, 115, 62–75.
- Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review*, 109, 330–357.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608–631.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. Available from <http://www.r-project.org>
- Rapaport, A., & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General*, 121, 352–363.
- Tversky, A. (1974). Assessing uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36, 148–159.
- Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). New York: Cambridge University Press.
- Wagenaar, W. A. (1972). *Sequential response bias*. Rotterdam, the Netherlands: Bronder Offset.
- Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, 60, 53–85.
- Weiss, R. L. (1965). Variables that influence random generation: An alternative hypothesis. *Perceptual and Motor Skills*, 20, 307–310.

Received September 3, 2008

Revision received December 15, 2008

Accepted December 16, 2008 ■

Correction to Kemp and Tenenbaum (2009)

In the article “Structured Statistical Models of Inductive Reasoning” by Charles Kemp and Joshua B. Tenenbaum (*Psychological Review*, Vol. 116, No. 1, pp. 20–58), there was an error in the abstract.

The third sentence should have read “This article presents a Bayesian framework that attempts to meet both goals and describes 4 applications of the framework: a taxonomic model, a spatial model, a threshold model, and a causal model.”

DOI: 10.1037/a0015514