

Perception of Randomness: Subjective Probability of Alternation

Yanlong Sun (Yanlong.Sun@uth.tmc.edu)

Hongbin Wang (Hongbin.Wang@uth.tmc.edu)

School of Biomedical Informatics, University of Texas Health Science Center
Houston, TX 77030 USA

Abstract

We present a statistical account for the subjective probability of alternation in people’s perception of randomness. By examining the *spatio-temporal distances* between pattern events, specifically, the *frequency* and *delay* of binary patterns in a Markov chain, we obtain some normative measures to calibrate people’s expectation of randomness. We suggest that it can be fruitful to study subjective randomness in the context of human object representation and perception of time and space.

Keywords: subjective randomness; probability of alternation; waiting time; perception of time and space.

Introduction

Much is known that subjective randomness—people’s intuitive judgment on how an event or a series of events appears random—systematically deviates from the stochastic randomness described by normative probability theories. Among many statistics describing such discrepancy, the *probability of alternation* (p_A) has been the most extensively studied in psychology literature (e.g., Budescu, 1987; Falk & Konold, 1997; Kahneman & Tversky, 1972; Kareev, 1992; Lopes & Oden, 1987; Nickerson, 2002; Sanderson, 2011). In a binary sequence generated by independent and stationary Bernoulli trials (for example, repeatedly tossing a fair coin), p_A can be defined as the probability that the outcome of any single event is different from the preceding one. If the process is truly random (e.g., the same fair coin is being tossed independently), the probability of alternation has the expected value $p_A = .5$. However, reviewed by Falk and Konold (1997, Table 1, p.304), in almost all of the studies with the tasks of recognizing or generating randomness, the modal *subjective probability of alternation* was approximately .60. That is, people tend to perceive sequences with $p_A \approx .60$ as the most random and generate random sequences with $p_A \approx .60$.

One particular reason that the probability of alternation p_A receives a great deal of attention in the studies on subjective randomness is that it is highly correlated with many other sequential statistics, such as the runs test and serial correlation. Together, these statistics cover a variety of empirical phenomena, for example, the perception of streaks in basketball shooting (Burns, 2004; Gilovich, Vallone, & Tversky, 1985; Sun & Wang, 2010b), the recency effect (Ayton & Fischer, 2004), the working memory capacity and detection of covariances in short sequences (Kareev, 1992), and the encoding of subjective complexity (Falk & Konold, 1997; Falk, Falk, & Ayton, 2009). (For a recent review, see Oskarsson, Van Boven, McClelland, & Hastie, 2009).

To explain the biased probability of alternation in subjective randomness, Falk and Konold (1997) developed a “difficulty predictor” (DP) as a measure of “subjective complexity”.

Based on the concept that random sequences are irreducibly complex (i.e., algorithmic complexity, Kolmogorov, 1965), Falk and Konold propose that people’s sense of randomness is not based on the deviations from the equiprobability of patterns of the same length (i.e., “ n -grams”), rather, it may be based on the difficulty level when people attempt to memorize or copy a sequence by its minimal description. Given any binary sequence, the difficulty predictor is defined by adding twice the number of alternating runs to the number of pure runs. For example, the following sequence is partitioned into 5 segments, where pure runs (streaks) are underlined and alternating runs are double underlined:

H H T H T H T H T H T H T H T H T H T H

Thus, the DP score for this particular sequence is $1 + 2 + 1 + 1 + 2 = 7$. By this measure, a perfect streak would be perceived as the most nonrandom because of its lowest DP score (i.e., the easiest to remember). In contrast, sequences with more alternating runs—hence greater p_A —are more difficult to encode therefore tend to be perceived as more random.

Overall, it has been demonstrated that DP correlates remarkably well with participants’ ratings of randomness, memorization time, assessed difficulty of memorization, and copying difficulty. And, the mean ratings of randomness show the classic preference for over-alternating sequences (Falk et al., 2009; Falk & Konold, 1997). However, Griffiths and Tenenbaum (2003) point out that DP remains a subjective measure since its objective counterpart, algorithmic complexity, is not computable. Instead, Griffiths and colleagues propose to use Bayesian inferences to account for subjective randomness (Griffiths & Tenenbaum, 2001, 2003; Hsu, Griffiths, & Schreiber, 2010). By this account, the subjective randomness of a particular sequence X is defined as

$$\text{random}(X) = \log \frac{p(X|\text{random})}{p(X|\text{regular})} \quad (1)$$

Then, the problem of judging randomness can be reduced to comparing two probabilities—whether the sequence is produced by a random process (e.g., independent and stationary Bernoulli trials), or, by a process with some regularities. To specify $p(X|\text{regular})$, Griffiths and Tenenbaum (2003) develop a hidden Markov model that makes transitions between hidden states depending on whether a motif is to be repeated or altered, where a motif is a short pattern such as H, T, HT, or, TH. By maximizing $p(X|\text{regular})$, they obtain a set of parameters that provide a better fit to the mean randomness ratings reported by Falk and Konold (1997).

The difficulty predictor and the Bayesian account have the

advantage to test against specific encoding strategies. However, DP has some counterintuitive properties. For example, in Figure 1, sequence (a) may appear to be more random than sequence (b), but the former actually has a lower DP score ($DP = 5$) than the latter ($DP = 6$). The Bayesian approach can fix this problem by adding more motifs of various length, but at the cost of computational complexity—to include all motifs of length 4, the hidden Markov model will have 22 motifs and 72 states (Griffiths & Tenenbaum, 2003).

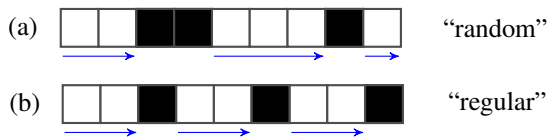


Figure 1: “Fast detection” of regularities. Without exactly counting alternating runs or calculating probabilities, it would be easily discernible that sequence (a) appears more “random” than sequence (b). The regularity in sequence (b) can be detected by the equal distances between patterns (for example, between the filled squares or the interruptions of unfilled squares). In addition, this example shows some of the counterintuitive properties of the difficulty predictor in that sequence (a) has a lower DP score than sequence (b).

Perhaps more interestingly, Figure 1 also prompts a speculation: whether the judgment of randomness can be reached at *before* any effort of encoding or estimating the probabilities of the observed sequences. For example, the regularity in Figure 1(b) might be quickly spotted by the equal distances between patterns. Such a speculation has actually led us to consider the recent advances in the investigations on *perception of time and space*. For instance, it has been posited that complex achievements such as mathematics and geometry, which are uniquely human in their full linguistic and symbolic realization, rest nevertheless on a set of core knowledge systems driven by the representations of object, space, time, and number, and these representations may have an early developmental origin shared by human infants as well as animals (e.g., Dehaene & Brannon, 2010; Spelke & Kinzler, 2007; Spelke, Lee, & Izard, 2010). Applied to the research on subjective randomness, it would be plausible to hypothesize that when people attempt to judge randomness (or detect regularities), the processing of spatio-temporal distances between observations and patterns is the primitive driving force, *before* any encoding effort of memorizing, copying the observed sequences or comparing the probabilities of specific processes.

Spatio-Temporal Distances between Patterns

In the present paper, we propose to utilize the spatio-temporal principles in object representation and human perception of time and space (e.g., Spelke & Kinzler, 2007) to study subjective randomness. Our approach is to first examine the spatial and temporal distributions of pattern events produced by random processes then match them to the psychological

spatio-temporal distances in people’s perception of randomness. To study the spatio-temporal distances between events, we focus on two sets of statistics, namely, given a random or a regular process, *how often* or *how likely* an event or a series of events would occur; And, from the start of an observation, *when* or *where* the events of interest would be encountered.

Apparently, *how often*, *when*, and *where* are different statistical properties and may bear different psychological relevancies. To set ideas, consider a simple case of coin tossing. If we tossed a coin three times and got three heads in a row as HHH (H = heads and T = tails), many of us might start getting suspicious about the fairness of the coin. But we would think it not at all noteworthy if the three tosses resulted in the pattern THH. The apparent randomness (or nonrandomness) cannot be explained by the frequency of encounters since the probability of obtaining either pattern in their exact orders is precisely the same, $(\frac{1}{2})^3 = \frac{1}{8}$ (i.e., the equiprobability of the “ n -grams”, Falk & Konold, 1997). However, less is known that there is a set of statistical properties that may very well explain why people consider a streak pattern rare and remarkable. When a fair coin is tossed repeatedly, it takes on average 8 tosses to observe the first occurrence of THH, but it takes on average 14 tosses to observe the first occurrence of HHH. Moreover, when both patterns are monitored simultaneously in one global sequence, the odds are 7 to 1 that one is more likely to first encounter THH than to first encounter HHH. That is, despite equal probabilities of occurrences, the time it takes to first encounter HHH is significantly “delayed” than that of THH.

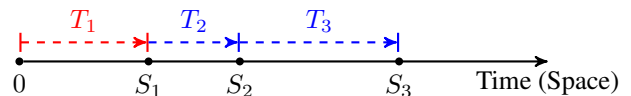


Figure 2: Spatio-temporal intervals between encounters of random events. An observation starts *from scratch* at Time = 0. T_1 is the first arrival time and its expected value $E[T^*]$ is called *waiting time*. T_2, T_3, \dots , are the interarrival times between successive occurrences of the events, and their expected value $E[T]$ is called *mean time*.

For formal definitions, we record the time (or location) when an event occurs at S_1, S_2, S_3, \dots , counting from the very beginning of the process (Figure 2). Then, $T_1 = S_1$ is the first arrival time with an expected value $E[T_1] = E[T^*]$ called *waiting time*. $T_2 = S_2 - S_1, T_3 = S_3 - S_2, \dots$, are the interarrival times between successive occurrences of the events, and their expected value $E[T]$ is called *mean time*.¹ It can be shown that the mean time of a pattern is in effect a measure of *frequency* as the inverse of probability of occurrence, and the waiting time is a measure of *delay* in that a pattern’s expected first arrival time may be longer but not shorter than

¹Note that since the first arrival time may have a different distribution than later interarrival times, we use T^* to denote the first arrival time and T to denote interarrival times.

its mean time (Ross, 2007).

Consider an irreducible Markov chain $\{X_n, n \geq 0\}$ with transition probabilities $P_{i,j}$ and stationary probabilities π_i , $i \geq 0$. Then, for pattern (i_1, i_2, \dots, i_k) , its mean time (i.e., inverse of the pattern's frequency) is the mean number of transitions between successive visits to the pattern,

$$E[T_k] = \frac{1}{\pi_{i_1} P_{i_1, i_2} \dots P_{i_{k-1}, i_k}} \quad (2)$$

For the pattern's waiting time, we first consider whether a successive arrival of the pattern can "reuse" any of the elements from its previous arrival. For example, in sequence THHH, pattern HH has occurred twice and its second arrival has reused the last element from the first arrival. An overlap index s is defined as the maximum number of elements at the end of the pattern that can be used as the beginning part of the next arrival,

$$s = \max \{j < k : (i_{k-j+1}, \dots, i_k) = (i_1, \dots, i_j)\} \quad (3)$$

For example, $s_{HT} = 0$ and $s_{HH} = 1$. Let $\mu(a, b)$ denote the mean number of transitions for the Markov chain to enter state b from state a . If the pattern has no overlap, $s = 0$, its waiting time is,

$$E[T_k^*] = \mu(a, i_1) - \mu(i_k, i_1) + E[T_k] \quad (4)$$

If the pattern has an overlap $s > 0$, we first consider $E[T_s^*]$, the waiting time for a shorter sub-pattern (i_1, i_2, \dots, i_s) , which is consisted of the first or the last s elements in pattern (i_1, i_2, \dots, i_k) . Then,

$$E[T_k^*] = E[T_s^*] + E[T_k] \quad (5)$$

By recursively applying Equation (5) until we reach the shortest sub-pattern with no overlap, we can obtain the waiting time for the original pattern. Comparing Equations (4) and (5), we can see that when looking for the first arrival of a pattern, if the pattern has an overlap $s > 0$, anything that goes wrong after the first s elements will make the counting process start from scratch. In other words, a pattern's *waiting time* can be *delayed* by the pattern's overlapping property. In contrast, Equation (2) shows that a pattern's *mean time* or *frequency* is not affected by the overlapping property.

Frequency and Delay by p_A

To generate binary patterns, we can use a Markov chain with two states H and T, where $P_{H,T} = P_{T,H} = p_A$, and $P_{H,H} = P_{T,T} = 1 - p_A$. This Markov chain is equivalent to the models used by Lopes and Oden (1987), where $p_A < .5$ represents the tendency of repetition, and $p_A > .5$ represents the tendency of alternation.

Assuming that the initial state is equally likely to be in either H or T, from equation (2), we have

$$E[T_{HT}] = E[T_{TH}] = \frac{2}{p_A} \quad (6)$$

$$E[T_{HH}] = E[T_{TT}] = \frac{2}{1 - p_A} \quad (7)$$

From Equations (4) and (5), we have

$$E[T_{HT}^*] = E[T_{TH}^*] = 1 + \frac{3}{2p_A} \quad (8)$$

$$E[T_{HH}^*] = E[T_{TT}^*] = \frac{-2p_A^2 + 5p_A + 1}{2p_A(1 - p_A)} \quad (9)$$

Figure 3 plots the mean time and waiting time for patterns of length 2 as the functions of probability of alternation p_A . When $p_A = .5$, we have a case of independent and stationary Bernoulli trials. We first note that the mean time is the same for all patterns of the same length. For example, $E[T_{HT}] = E[T_{HH}] = 4$. However, the waiting time can be different depending on the pattern's overlapping property. For example, $s_{HT} = 0$, $E[T_{HT}^*] = 4$, and, $s_{HH} = 1$, $E[T_{HH}^*] = 6$. That is, the waiting time is longer for the shortest streak patterns HH or TT than for the shortest alternating pattern HT or TH. Solving the equality between Equations (8) and (9), we obtain $p_A = \frac{1}{3}$. Thus, as long as $p_A > \frac{1}{3}$, we have $E[T_{HH}^*] > E[T_{TH}^*]$.

Moreover, let $\text{Var}(T)$ denote the variance of the interarrival times, it can be shown that patterns may differ substantially in how evenly they are distributed over time (or space), for

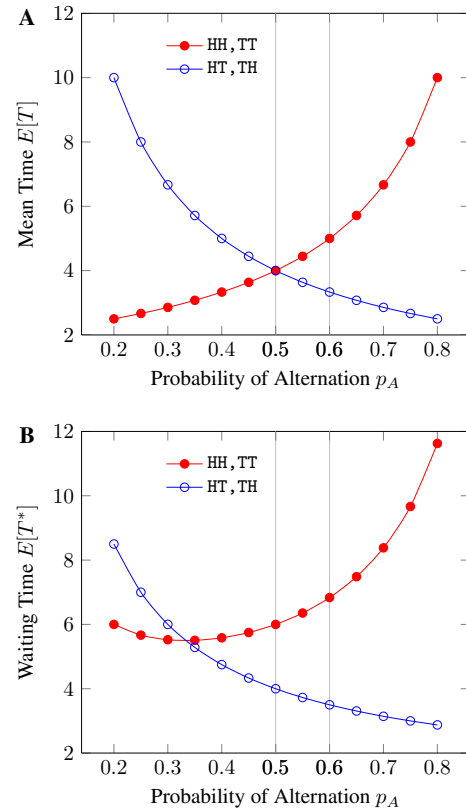


Figure 3: Mean time (A) and waiting time (B) as the functions of probability of alternation p_A . When $p_A = .5$, the process is equivalent to independent and stationary Bernoulli trials.

example, $\text{Var}(T_{\text{TH}}) = 4$, and $\text{Var}(T_{\text{HH}}) = 20$. In addition, the waiting time is highly correlated with the variance of interarrival times since both values are extended by a pattern’s overlap tendency (see Table 1). (For the calculation of variances, see Sun & Wang, 2010a.)

Table 1: Mean and variance of the first arrival time (T^*) and interarrival times (T) for binary patterns in independent Bernoulli trials ($p_A = 0.5$) when tossing a fair coin. Reciprocal patterns are listed only once, for example, HH is equivalent to TT, and, HT is equivalent to TH.

Patterns	s	$E[T]$	$\text{Var}(T)$	$E[T^*]$	$\text{Var}(T^*)$
H	0	2	2	2	2
HT	0	4	4	4	4
HH	1	4	20	6	22
HHT	0	8	24	8	24
HTT	0	8	24	8	24
HTH	1	8	56	10	58
HHH	2	8	120	14	142

Psychological Implications

The Markov model described above may provide quantitative measures to calibrate subjective randomness, particularly regarding the seemingly miscalibrated subjective probability of alternation p_A . Different from previous studies that focus on the frequency of patterns, here we examine the spatio-temporal distances between pattern events that cover both frequency and delay (including variances), either from the very beginning of the process (waiting time), or between successive occurrences of the pattern given that the pattern has occurred before (mean time).

It should be noted that so far our analyses are limited to short patterns. This consideration is based on the empirical findings that people are sensitive to even the shortest patterns of length 2. For example, it has been reported that participants would report “a streak is occurring” beginning at the third repeating event (Carlson & Shu, 2007). In an fMRI study, Huettel, Mack, and McCarthy (2002) show that a distributed set of regions in prefrontal cortex are exquisitely sensitive to the presence and the termination of streak patterns even when pattern length was only 2, despite the fact that participants were informed of the random order of the sequences.

It appears that the waiting time statistics fit well to one of the most influential accounts for subjective randomness, the “representativeness heuristic” (e.g., Gilovich et al., 1985; Kahneman & Tversky, 1972). By this account, people expect smaller sequences to resemble the balanced distributions in the long run, such that a streak of heads would seem to be rare and remarkable as it would not be representative of the process of random coin tossing. This would have explained the biased subjective probability of alternation in that people tend to expect fewer and shorter streaks than would be mathematically probable when observing sequences produced by a random

process, and, they avoid repetitions of the same elements when instructed to generate such sequences (e.g., Falk & Konold, 1997; Wagenaar, 1972).

Despite its plausibility, the representativeness account has been criticized for the lack of definition (Ayton & Fischer, 2004; Falk & Konold, 1997; Gigerenzer, 1996). Nevertheless, it has been proposed that the waiting time statistics may provide quantitative explanations to this account (Hahn & Warren, 2009; Sun, Tweney, & Wang, 2010a; Sun & Wang, 2010a, 2010b). Specifically, people judge the frequency of an event on the basis of how it is representative of the underlying population or process (representativeness), and how easily an example can be brought to mind (availability). When people think of a truly random process (by actually tossing a coin or conducting a mental experiment), a streak pattern—even at its shortest length of 2 (e.g., HH in tossing a coin)—may be perceived as the most nonrepresentative and the most unavailable. Compared with other patterns of the same length, a streak is the most delayed in its first arrival and has the largest variance of interarrival times thus the most uneven or clustered distribution over time (see Table 1). Note that these particular properties are not limited to independent Bernoulli trials where $p_A = .5$. For example, Figure 3B shows that as long as $p_A > \frac{1}{3}$, the waiting time for the streak pattern HH will be longer than that of non-streak pattern HT.

Moreover, there has been direct evidence suggesting that people may at least have an approximate sense of the waiting time, namely, the delayed occurrence of streak patterns. Oppenheimer and Monin (2009) report that when participants were asked to estimate the number of coin flips before the occurrence of a pattern of length 5, they believed that a sequence of coin flips was nearly twice as long before a streak (mean estimate = 16.2) than when there was no streak (mean estimate = 8.7). Applying Equations (4) and (5), we can show that for patterns of length 5, the waiting time for a streak is 62 tosses, and the average waiting time for non-streak patterns is approximately 34.3 tosses: the former is nearly twice as long as the latter.

In the light of the delayed first arrival for streak patterns, we speculate that people’s expectation of the probability of alternation as $p_A \approx 0.6$ might in effect have been driven by their experiences of pattern events as random patterns unfold in time and space. For example, we can reconstruct the task of generating randomness with the Markov chain described above. Since at any given moment, participants face the choice of either repeating or reversing the current outcome (an H or a T), the generation process is equivalent to the process of choosing patterns of length 2—either a streak pattern (HH or TT) or a non-streak pattern (HT or TH). Then, $p_A = .6$ means that 60% of the time participants choose a non-streak pattern, indicating a false belief that in tossing a fair coin independently (i.e., $p_A = .5$), streak patterns should occur less frequently than non-streak patterns. Such belief can be formulated as a ratio of pattern mean times. From Equations (6) and (7), when

$p_A = .6$,

$$E [T_{HH,TT}] : E [T_{HT,TH}] = p_A : (1 - p_A) = 3 : 2$$

Then, comparing the waiting time in the process of independent coin tossing yields exactly the same ratio, where $p_A = .5$,

$$E [T_{HH,TT}^*] : E [T_{HT,TH}^*] = 6 : 4 = 3 : 2$$

That is, measured by the mean time, participants have failed the task of producing randomness (i.e., the independence property where $p_A = .5$) as if they have falsely believed that patterns HH and TT would occur less frequently than patterns HT and TH. Quantitatively, this comparison indicates that the observed bias might have stemmed from participants' expectation of waiting time from a truly random process, since the mean time does not distinguish any patterns when $p_A = .5$ (e.g., see Figure 3A).

Discussion

The development of waiting time statistics appears to be promising and may have potential significance in explaining a range of human cognitive functions (e.g., Oppenheimer & Monin, 2009; Sun et al., 2010a; Sun, Tweney, & Wang, 2010b; Sun & Wang, 2010b, 2010a, 2011). Specifically, we argue that it can be fruitful to study subjective randomness in the context of *human perception of time and space*. And, the *frequency* and *delay* of pattern events, rather than individual events (e.g., a single coin toss), may be the key statistics and theoretical constructs that underlie human perception of randomness.

It has been posited that by exposing to the various environmental statistics, human mind may have evolved an accurate sense of randomness but may fail to reveal it by the standard of a particular measuring device (e.g., Pinker, 1997). Given that the waiting time and the variance of interarrival times can be substantially different for patterns with the same mean time (e.g., Table 1 and Figure 3), one may logically assume that these statistics may play a critical role in shaping people's perception and judgment of randomness. Unfortunately, in the long lasting investigations on subjective randomness, the mean time of patterns serves as the sole normative measure of randomness. Despite various forms of experimental tasks (e.g., randomness generation or recognition, probabilistic predictions) and statistical methods (e.g., runs test, serial correlation, Bayesian inferences), existing studies have been focusing on the discrepancies between subjective responses and the probabilities of the occurrences of random patterns. Nevertheless, the absence of waiting time statistics in the investigations on subjective randomness may be due to its late and still ongoing development in statistical research (e.g., Pozdnyakov, 2008; Ross, 2007), and, remain fairly novel to the audience in psychology (c.f., Hahn & Warren, 2009; Konold, 1995; Nickerson, 2007; Sun et al., 2010a, 2010b).

More significantly, recent advances in the behavioral and neurological sciences on human cognitive achievements all point to the role of the perception of time and space. It has

been proposed that complex achievements such as mathematics and geometry, which are uniquely human in their full linguistic and symbolic realization, rest nevertheless on a set of core knowledge systems that are driven by the representations of object, space, time and number (Dehaene & Brannon, 2010; Spelke & Kinzler, 2007; Spelke et al., 2010). And, these representations may have a common perceptual metric in the form of a *mental number line* (Burr & Morrone, 2011; Dehaene, Piazza, Pinel, & Cohen, 2003) and have an early developmental origin shared by human infants as well as other animals (de Hevia & Spelke, 2010; Haun, Jordan, Vallortigara, & Clayton, 2010; Hubbard, Piazza, Pinel, & Dehaene, 2005). Applied to the research on subjective randomness, it would be plausible to hypothesize that when people attempt to judge randomness (or detect regularities), the processing of spatio-temporal distances between observations and patterns is the primitive driving force, *before* any encoding effort for memorizing, copying, or assessing the probability of pattern occurrences (e.g., see Figure 1).

Nonetheless, we need to collect more empirical evidence to investigate whether and how human cognition is sensitive to the statistics of random patterns, for example, via experiments that manipulate the overlapping and delay properties of pattern events then measure the psychological responses. Moreover, a theoretical breakthrough would also require us to firmly demonstrate the psychological relevance of the pattern time statistics and the spatio-temporal principles in object representation and human perception of time and space, in order to develop a mental calculus of how these constructs work together.

Acknowledgments

Supported by the Office of Naval Research (ONR) grant number N00014-08-1-0042, and Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract number D10PC20021. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained hereon are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government.

References

- Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, 32(8), 1369–1378.
- Budescu, D. V. (1987). A Markov model for generation of random binary sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 25–39.
- Burns, B. D. (2004). Heuristics as beliefs and as behaviors: The adaptiveness of the "hot hand". *Cognitive Psychology*, 48(3), 295–331.
- Burr, D. C., & Morrone, M. C. (2011). Spatiotopic coding and remapping in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564), 504–515.

- Carlson, K. A., & Shu, S. B. (2007). The rule of three: How the third event signals the emergence of a streak. *Organizational Behavior and Human Decision Processes*, 104(1), 113–121.
- de Hevia, M. D., & Spelke, E. S. (2010). Number-space mapping in human infants. *Psychological Science*, 21(5), 653–660.
- Dehaene, S., & Brannon, E. M. (2010). Space, time, and number: A Kantian research program. *Trends in Cognitive Sciences*, 14(12), 517–519.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3-6), 487–506.
- Falk, R., Falk, R., & Ayton, P. (2009). Subjective patterns of randomness and choice: Some consequences of collective responses. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 203–224.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104(2), 301–318.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592–596.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.
- Griffiths, T. L., & Tenenbaum, J. B. (2001). Randomness and coincidences: Reconciling intuition and probability theory. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the cognitive science society* (pp. 398–403). Mahwah, NJ: Lawrence Erlbaum Associates.
- Griffiths, T. L., & Tenenbaum, J. B. (2003). Probability, algorithmic complexity, and subjective randomness. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual conference of the cognitive science society* (pp. 480–485). Cognitive Science Society.
- Hahn, U., & Warren, P. A. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review*, 116(2), 454–461.
- Haun, D. B. M., Jordan, F. M., Vallortigara, G., & Clayton, N. S. (2010). Origins of spatial, temporal and numerical cognition: Insights from comparative psychology. *Trends in Cognitive Sciences*, 14(12), 552–560.
- Hsu, A., Griffiths, T. L., & Schreiber, E. (2010). Subjective randomness and natural scene statistics. *Psychonomic Bulletin & Review*, 17(5), 624–629.
- Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, 6(6), 435–448.
- Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, 5(5), 485–490.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430–454.
- Kareev, Y. (1992). Not that bad after all: Generation of random sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1189–1194.
- Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1, 1–7.
- Konold, C. (1995). Confessions of a coin flipper and would-be instructor. *The American Statistician*, 49(2), 203–209.
- Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning Memory and Cognition*, 13(3), 392–400.
- Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review*, 109(2), 330–357.
- Nickerson, R. S. (2007). Penney Ante: Counterintuitive probabilities in coin tossing. *The UMAP Journal*, 28(4), 503–532.
- Oppenheimer, D. M., & Monin, B. (2009). The retrospective gambler's fallacy: Unlikely events, constructing the past, and multiple universes. *Judgment and Decision Making*, 4(5), 326–334.
- Oskarsson, A. T., Van Boven, L., McClelland, G. H., & Hastie, R. (2009). What's next? Judging sequences of binary events. *Psychological Bulletin*, 135(2), 262–285.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pozdnyakov, V. (2008). On occurrence of patterns in Markov chains: Method of gambling teams. *Statistics & Probability Letters*, 78(16), 2762–2767.
- Ross, S. M. (2007). *Introduction of probability models* (9th ed.). San Diego, CA: Academic Press.
- Sanderson, Y. B. (2011). Color charts, esthetics, and subjective randomness. *Cognitive Science*.
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, 10(1), 89–96.
- Spelke, E. S., Lee, S. A., & Izard, V. (2010). Beyond core knowledge: Natural geometry. *Cognitive Science*, 34(5), 863–884.
- Sun, Y., Tweney, R. D., & Wang, H. (2010a). Occurrence and nonoccurrence of random sequences: Comment on Hahn and Warren (2009). *Psychological Review*, 117(2), 697–703.
- Sun, Y., Tweney, R. D., & Wang, H. (2010b). Postscript: Untangling the gambler's fallacy. *Psychological Review*, 117(2), 704–705.
- Sun, Y., & Wang, H. (2010a). Gambler's fallacy, hot hand belief, and time of patterns. *Judgment and Decision Making*, 5(2), 124–132.
- Sun, Y., & Wang, H. (2010b). Perception of randomness: On the time of streaks. *Cognitive Psychology*, 61(4), 333–342.
- Sun, Y., & Wang, H. (2011). Probability theory and perception of randomness: Bridging “ought” and “is”. *Behavioral and Brain Sciences*, 34(5), 271–272.
- Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, 77(1), 65–72.