

When a coin toss does not appear random: Causal belief and judgments of randomness.

Fiona E. J. McDonald (FMcDonald@psy.unsw.edu.au)

School of Psychology, UNSW
Sydney, NSW 2052 Australia

Ben R. Newell (Ben.Newell@unsw.edu.au)

School of Psychology, UNSW
Sydney, NSW 2052 Australia

Abstract

Distinguishing between random and non-random data is important for inductive reasoning. Prior research has found a bias towards judging binary sequences with alternation rates above 0.5 as most random. In most of this research the concept of randomness was explained to participants via the example of a coin being tossed. The current experiment examined the influence of such example explanations on the perception of the randomness of binary sequences. Participants were told that sequences had been generated by a coin toss, a basketball player taking free throws, or were given no prior belief about the generating process (control). In the control condition there was no bias towards rating high alternation rate sequences as most random, however the bias persisted when a causal mechanism (coin, basketball player) was provided. Previously found correlations between perceived memorisability and perceived randomness were only found when a belief about the generating mechanism was provided.

Keywords: randomness judgment; binary sequences; causal belief; coin toss; gambler's fallacy; hot hand belief.

A coin being tossed is the classic example used to illustrate randomness, as it satisfies many of the elusive features of randomness. It has often been stated that randomness is hard to define and requires consideration of the randomness of the source as well as the output (e.g. Bar-Hillel & Wagenaar, 1993; Falk & Konold, 1997; Nickerson, 2002). Using a coin toss to explain what is meant by randomness avoids complex and potentially biasing instructions in experiments. Wagenaar (1972) comments that the coin toss example is well accepted by participants as an "ideal randomizer" (p. 112). There is a long history of research into our understanding of randomness, much of it concluding we are poor at both generating and judging randomness (e.g. Falk & Konold, 1997; Tune, 1964; Wagenaar, 1970a, 1970b; Weiss, 1965). In this paper we will explore whether using examples like a coin toss have actually impeded measurement of our understanding of randomness.

When making predictions about future events we tend to observe the past, look for patterns and then make inductions about what is likely to occur in the future. Thus the ability to distinguish between random and non-random sequences is a key human behavior (Lopes, 1982). Statistical analysis of sequence structure can provide details on how representative each sequence or set of sequences is of a

stochastic process. A prototypically random sequence has an alternation rate of 0.5 (Falk & Konold, 1997). Sequence alternation rate describes the proportion of times an outcome is the opposite of the previous one. Sequences with long runs of one outcome have low alternation rates while those where the outcome changes frequently have high alternation rates.

Judging Randomness

While early research into our understanding of randomness focused on generation of random sequences, Wagenaar (1970a) suggested that judgment tasks better assess our understanding as they avoid attentional and memory limitations which could impact generation performance. In judgment tasks people select random sequences from sets of sequences with varying properties. As generation studies had found that people tend to produce sequences with too many alternations between outcomes, Wagenaar studied the effect of manipulating sequence alternation rate on perceptions of randomness. Wagenaar (1970a) found that, consistent with the generation studies, people have a tendency to select sequences with alternation rates around 0.6 as the most random. This indicates a bias towards judging sequences with too many alternations as ideally random. The preference for alternation in both judgment and generation tasks is consistent with the gambler's fallacy, that is, the belief that following a run of one outcome, the other outcome is now due.

Similar studies by Falk (1981), and Falk and Konold (1997) also found that people tend to rate sequences with alternation rates around 0.6 to 0.7 as the most random. Falk and Konold propose perceived sequence complexity as the mechanism underlying the bias in randomness ratings. That is, people decide how complex a sequence appears and use this to rate its randomness. To study this, they equated perceived sequence complexity with judgments of sequence encoding difficulty (operationalised as actual and perceived difficulty to memorise or copy the sequences). As they found strong correlations between measures of encoding difficulty and judgments of randomness across sequence alternation rate they concluded that people use a subjective measure of sequence complexity as a proxy for determining sequence randomness.

Is judgment always poor?

The bias towards perceiving high alternation rate sequences as random suggests a consistent but flawed human understanding of randomness. However, other research has found that certain experimental conditions can alter this bias, indicating that there is not a consistent biased view of randomness. Wiegiersma (1982) found that when he asked participants to produce a random sequence using music tones there was less bias than in a conventional generation task. He also found that altering the physical presentation of the symbols, such as changing the font or changing the spacing, impacted on the degree of bias (1987). Additionally, Lopes and Oden (1987) found that if they gave people information about the properties of the hypothetical machines causing the sequences that people were better able to identify the random sequences.

Furthermore, research into the role of causal belief on gambler's fallacy behaviour suggests that knowledge of the mechanism producing the sequence can influence our judgments about its randomness. Studies by Ayton and Fischer (2004), and Burns and Corpus (2004) looking at the gambler's fallacy found that people's belief about the cause of the sequence affects when they are likely to follow the gambler's fallacy and when they are likely to follow its converse, the hot hand belief. The hot hand belief originated in basketball and describes the belief that a run of one outcome (e.g. making shots in basketball) will continue. Burns and Corpus found that beliefs about the randomness of the generating mechanism influenced whether people predict a short run of the same outcome would continue or end. When the causal mechanism was perceived to be random, such as a roulette wheel, people were more likely to end a run than when they perceived the cause to be non-random, such as a basketball player (see also Ayton & Fischer).

The inference from these studies is that knowledge of the cause of a sequence can influence our interpretation of its randomness. Thus the instructions given to participants in perception of randomness tasks might also influence the judgments participants make.

The role of instruction

In their studies on perceived randomness, Wagenaar (1970a) and Falk and Konold (1997), gave participants the example of a coin toss to aid explanation of the concept of randomness. In her earlier study Falk (1981) used colored cards and asked participants how "well shuffled" (p. 225) the cards were. Wiegiersma (1982) also used the example of a coin toss for his conventional perception of randomness task, but used the example of a roulette wheel for his random music tone generation task. In their judgment task Lopes and Oden (1987) refer to sequences produced by various machines with different process properties. Thus it appears that in studies of perceived randomness, some type of causal 'agent' or mechanism is mentioned and in many cases it is a coin being tossed.

The coin toss example was introduced to make the task clearer for participants and presumably because, as a coin toss is known to be random, it should not influence people's judgments about randomness. Is it possible however that people do not believe coin tossing is really random? Roney and Trick (2003) found that people believe there is interconnectedness between successive coin tosses. In their study, people predicted a series of coin toss outcomes while getting feedback. They found that participants committed the gambler's fallacy except when they were told a new block of tosses was starting. Gambler's fallacy behavior is only sensible when you believe there is interconnectedness between outcomes, which mention of the new block broke. Clearly, whenever people follow the gambler's fallacy, they are behaving as if a coin toss is not random. As apparently all previous perception of randomness studies have used an example such as a coin toss, it is possible that the bias in randomness judgments is due to the coin example and not due to an innate bias in our understanding of randomness.

A well cited study by Gilovich, Vallone and Tversky (1985) on the hot hand belief had participants judge the randomness of sequences that were nominally created by basketball players. They asked participants to rate sequences with different alternation rates as chance, streak or alternate shooting and found that sequences with alternation rates of around 0.7-0.8 were judged as typical of a chance sequence. It has been suggested by Gilovich and colleagues that people believe in the hot hand effect because they do not believe that basketball shots are independent events. This is similar to believing that successive coin tosses are interconnected and thus it is interesting that a similar bias is found when the sequence cause is a basketball player. It is possible that introducing a causal agent, regardless of the type of cause, reduces beliefs about outcome independence, thus resulting in a biased assessment of randomness.

Current experiment

The current experiment directly assesses the role of causal agent on judgments of randomness by comparing the perceived randomness (PR) ratings of people given a causal belief with those given none. Falk and Konold (1997) proposed that randomness judgments are influenced by perceptions of sequence complexity, operationalised as encoding or memorisability difficulty. They presumed that the bias in people's judgments of randomness is due to an innate misunderstanding of randomness and sought to find the underlying mechanism. They took the strong correlations they found between various measures of encoding difficulty and measures of perceived randomness as evidence for their theory that judgments of randomness are influenced by perceptions of sequence complexity. If judgments of randomness are biased due to belief examples such as a coin toss and not encoding difficulty, then difficulty of encoding may not correlate with perceptions of randomness when no causal belief is given. So in addition, one measure of encoding difficulty, perceived difficulty to memorise (PM) the sequences, was also explored.

In this experiment, sets of sequences with various alternation rates were presented to participants who rated the perceived randomness or perceived memorisability of each sequence. To investigate the role of causal belief on these judgments participants were given either no causal belief or one of two causal beliefs. The two causal beliefs used were a coin being tossed and a basketball player taking free throws. It was expected that as causal beliefs reduce beliefs about the independence of outcomes, they would influence perceptions of randomness. In particular, when no belief is given, the highest randomness ratings would be given for sequences with alternation rates close to the ideal 0.5, whereas when a causal belief is given the highest randomness ratings would be for sequences with higher alternation rates.

It was expected that when no causal belief was given there would be no correlation between PR and PM ratings. In addition, perceived memorisability scores would not be influenced by causal belief as the ratings should be influenced only by the sequence structure. In keeping with the findings of Falk and Konold (1997) it was expected that sequences judged to be most difficult to memorise would be those with alternation rates around 0.6 to 0.7.

Method

Participants

Second year students ($N = 241$, mean age = 20.7 years, $SD = 2.7$) from UNSW took part in the experiment as part of a class activity. Different tutorial groups were assigned to each of the six experimental conditions ($n = 35$ to 46).

Stimuli

The sequences shown to participants were generated using the random number generator in Microsoft Excel and sequences with the required alternation rates were selected. As well as controlling for alternation rate, sequences were selected with approximately even numbers of each outcome. These sequences were then converted to @ and # symbols. Figure 1 provides examples of sequences used for three of the alternation rates.

```
0.2:  @ @ @ @ @ # # @ @ @ @ # # # # # # # # @ @
0.5:  # @ # # @ # @ # @ @ # # @ @ @ @ @ @ # # # #
0.8:  # @ @ # @ @ # # @ # @ # @ # @ # @ # @ # #
```

Figure 1: Examples of sequences used with alternation rates of 0.2, 0.5, and 0.8.

There were six sequences for each of the seven alternation rates (0.2 to 0.8) resulting in a total of 42 sequences. Sequences were presented in blocks of seven with one example of each alternation rate in each block. The sequences were chosen and ordered randomly. Each sequence was used once for each participant.

Procedure

Participants read instructions appropriate for their condition and then rated six sets of seven sequences. The instructions were kept as similar as possible but the middle section was modified for the three belief conditions. The instructions for the memorisability task are consistent with those used by Falk and Konold (1997). The section below highlights the differences in instructions.

Perceived randomness (PR) conditions:

Everyone:

The aim of this experiment is to find out how random sequences appear to people.

Control condition:

You are about to see some sequences of @ and # symbols. You need to rate how random you think each sequence is on a scale from 1 to 7. A random sequence is one governed by chance in which each outcome is independent of the one before it.

Coin toss condition:

When a coin is tossed you get a sequence of head and tail outcomes. You are about to see some sequences of @ and # symbols that were created by tossing a coin. You need to rate how random you think each sequence is on a scale from 1 to 7. A random sequence is one governed by chance in which each outcome is independent of the one before it, thus all of these coin toss sequences can be considered to be equally random; however we are interested in how random they appear to you.

Basketball condition:

In the game of basketball people make free throws, which are attempts to shoot a goal. This creates a sequence of hits and misses. You are about to see some sequences of @ and # symbols that were created by people making free throws in basketball. You need to rate how random you think each sequence is on a scale from 1 to 7. A random sequence is one governed by chance in which each outcome is independent of the one before it.

Everyone:

Please look at all the sequences first then assign a 7 to the sequence or sequences that appear most random to you and a 1 to the sequence or sequences that appear least random to you. You need to use the rating 1 and 7 at least once per page. Use the remaining numbers (2-6) to rate the rest of the sequences.

Perceived memorisation (PM) condition:

Everyone:

The aim of this experiment is to find out how people rate their ability to memorise information.

Control condition:

You are about to see some sequences of @ and # symbols.

Coin toss condition:

When a coin is tossed you get a sequence of head and tail outcomes. You are about to see some sequences of @ and # symbols that were created by tossing a coin.

Basketball condition:

In the game of basketball people make free throws, which are attempts to shoot a goal. This creates a sequence of hits and misses. You are about to see some sequences of @ and # symbols that were created by people making free throws in basketball.

Everyone:

You need to rate how difficult you think it would be to memorise each sequence on a scale from 1 (not at all difficult) to 7 (extremely difficult). Imagine that you have to memorise each sequence and then reproduce it on a piece of paper without looking at the computer screen. The difficulty of memorising might be affected by things like how easy it would be to divide the sequence in to memorable 'chunks' or the presence of patterns in the sequence. Please look at all the sequences first then assign a 7 to the sequence or sequences that appear most difficult to memorise and a 1 to the sequence or sequences that appear least difficult to memorise. You need to use the rating 1 and 7 at least once per page. Use the remaining numbers (2-6) to rate the rest of the sequences.

Participants then rated the sequences as per the instructions. While participants were forced to use the extremes of the scale they were free in their use of the other ratings. Thus the ratings are not strictly ordinal.

Results

Each participant's average rating for each alternation rate was calculated. Consistent with Falk and Konold (1997) average group ratings were calculated and normalized (0 to 1). These group ratings are presented separately for the two tasks in Figures 2 and 3. (Note: No bar is visible when the rating is 0.)

The results for both tasks were analysed using ANOVAs. In both analyses there are several significant main effects for alternation rate, so for each analysis only the largest effect is reported.

As expected, both tasks had significant main effects for alternation rate, PR (quadratic): $F(1,115) = 81.32, p < .05, \eta^2 = .41$; PM (linear): $F(1,120) = 749.84, p < .05, \eta^2 = .86$. There were no main effects for belief for either task which was expected as the belief manipulation was between participants and the same range of ratings was used for each belief. For the PM task there was no significant interaction between belief and alternation rate, $F(2, 120) = 1.81, p = .17$, indicating that the belief manipulation had no effect on ratings of perceived randomness. For the PR task a difference was expected between the two belief conditions and the control condition. Therefore the results from the two belief conditions were combined and compared against the control condition. A small but significant interaction was found between belief and alternation rate, $F(1, 116) = 4.33,$

$p < .05, \eta^2 = .04$. This interaction indicates that people given a causal belief rate sequences with higher alternation rates as more random compared to those not given a causal belief (control).

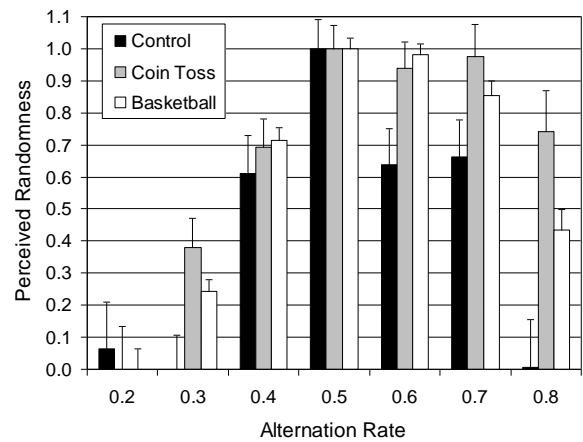


Figure 2: Average perceived randomness ratings (1 is most random) for sequences with different alternation rates for groups with different causal beliefs (error bars = SEM).

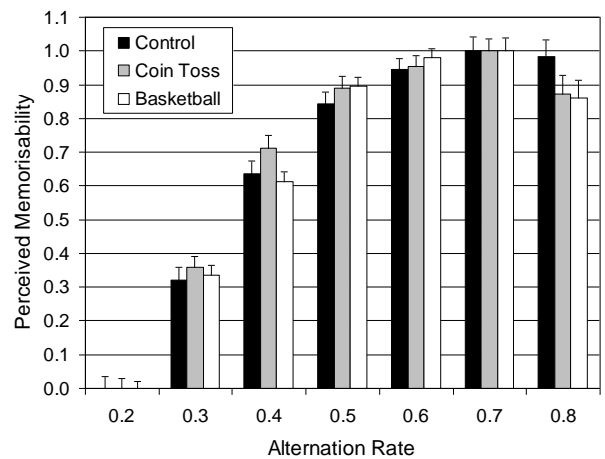


Figure 3: Average perceived memorisability ratings (1 is hardest to memorise) for sequences with different alternation rates for groups with different causal beliefs (error bars = SEM).

Taken together these results indicate that the pattern of results is different for the two tasks and that there are small effects for belief only in the PR task. Earlier research into perceived randomness concluded that people are biased towards seeing higher alternation rate sequences as more random than the mathematically random 0.5 alternation rate sequences. That is, earlier research has found that the peak randomness rating by participants was for sequences with alternation rates above 0.5. To explore the role of belief within each task further, peak PR and peak PM values were calculated.

For each participant a quadratic curve was fitted to their rating data allowing the alternation rate with the peak rating

to be calculated. Thirteen participants from the PR task were excluded as it was not possible to fit a curve to their data and thus no peak perceived randomness could be calculated. For the PM task only one participant was excluded. For each belief and task the peak ratings were compared against a value of 0.5 using a single sample t-test. The value of 0.5 was selected as it should be the peak if ratings were based on the mathematically correct value. Results from this analysis are presented in Table 1 below. These results show that for all conditions except the control PR group, the peak rating was above 0.5.

Table 1. Peak PR and PM values for each belief, t-test results and 95% confidence intervals (CI).

	Peak	t-test comparison with 0.5	95% CI
Control	0.52	$t_{32}=.33, p=.74$	0.45-0.57
PR			
Coin Toss	0.61	$t_{39}=3.36, p<.01$	0.54-0.66
Basketball	0.59	$t_{30}=2.278, p<.01$	0.52-0.65
PM			
Control	0.70	$t_{34}=14.03, p<.01$	0.67-0.73
Coin Toss	0.67	$t_{41}=10.76, p<.01$	0.64-0.71
Basketball	0.68	$t_{44}=11.34, p<.01$	0.65-0.71

Falk and Konold (1997) found correlations between various encoding measures and the randomness ratings, suggesting that when judging randomness people rate the sequences based on encoding difficulty. In the current study correlations between PR and PM were calculated using the group averages. Strong and significant correlations were found between PR and PM for the conditions with causal beliefs (coin toss: $r^2=.981, p=.000$; basketball: $r^2=.885, p=.008$) but no significant correlation was found for the control condition ($r^2=.534, p=.217$). The differences in the correlations are due to the effect of causal belief on PR. Causal belief influences PR ratings but not PM ratings, so it is not possible for PR and PM to correlate for all beliefs.

Discussion

Peoples' inability to accurately select the appropriate sequence as the most random has been used as evidence of our poor understanding of randomness (Falk, 1981; Falk & Konold, 1997; Wagenaar, 1972). The results from the PR task reported here suggest that when not influenced by causal information people are able to accurately select the most random sequence. This contrasts with instances where a causal belief is given, where perceptions of randomness are biased towards judging sequences with higher alternation rates as more random. The results are generally consistent with previous research, as studies that found randomness ratings biased towards higher alternation rates have involved causal agents (Falk, 1981; Falk & Konold, 1997; Wagenaar, 1970a). However, the results from the control condition, finding no bias, lead to a different interpretation, namely that under certain conditions people are able to accurately judge the randomness of a sequence.

The results show that causal belief does influence perceptions of randomness by shifting the peak randomness rating to higher alternation rate sequences.

The results of the present study contrast with those of Ayton and Fischer (2004), and Burns and Corpus (2004) who found that the type of causal belief influences judgments. That is the animacy (Ayton & Fischer) or the perceived randomness (Burns & Corpus) of the belief was crucial in determining people's assessment of the sequence. The current experiment compared belief with no belief and found that belief, regardless of the type, results in a bias in judgments of randomness.

While causal belief influences perceptions of randomness it appears to have no effect on perceptions of memorisability. Falk and Konold (1997) suggested that encoding difficulty may be used as a proxy in rating the randomness of a sequence. In this experiment we have operationalised encoding difficulty as perceived difficulty to memorise the sequence. When no belief was given there is no correlation between PR and PM, suggesting that ratings of memorisability or encoding difficulty are not directly influencing ratings of randomness. In the two experiments where Falk and Konold measured perceived memorisability they do not appear to have mentioned a causal agent as part of the instructions. Despite manipulating causal belief, the PM task results reported here are consistent with those found by Falk and Konold, showing a bias towards judging higher alternation rate sequences as harder to memorise. As belief has no effect on PM ratings, it appears that participants are focused on the sequence structure.

The high correlation between randomness rating and difficulty of encoding task found by Falk and Konold (1997) are replicated in the coin toss condition, while the correlation in the basketball condition exists but is slightly weaker. This occurs because the coin toss and basketball PR results are biased towards high alternation rates. For the control condition there is no bias in the PR task but still a bias in the PM task so the correlation is weak and not significant.

The PR task appears susceptible to manipulation of experimental variables. For example, Wieggersma (1987) was able to reduce the bias in a perception task by varying the font and symbol spacing. In the current study the causal beliefs seem to have created the impression that the individual outcomes are not independent. While this is not unusual for the basketball belief it is unexpected for the coin toss belief. However, Roney and Trick (2003) have previously demonstrated that people do hold faulty beliefs about the interconnectedness of consecutive outcomes of coin tosses. A similar design as that used by Roney and Trick, where interconnectedness is manipulated, could be used to test whether reducing the apparent interconnectedness of outcomes would reduce the effect of a causal belief on judgments of randomness.

The finding from the current study, that people are not biased in their understanding of randomness, appears at odds with the overwhelming prevalence of the gambler's

fallacy in everyday life. It is important to note however, that when the gambler's fallacy is observed there is always a known cause for the sequence. It is possible that the gambler's fallacy could be reduced if people made predictions of future outcomes without knowing the cause of the sequence.

This study has not explored the role of causal belief in generation of random sequences. At least some previous studies have included the coin toss example to guide participants in their concept of randomness (e.g. Kareev, 1992; Rapoport & Budescu, 1992). Budescu and Rapoport (1992; 1994) found that under game conditions, and when no causal belief was given, people produced less biased sequences than when asked to produce random sequences with a coin toss (1992) or a die (1994) as examples. It would be worthwhile to separate the belief instruction from the task type to see whether just asking participants to produce a random sequence without mentioning a causal agent is sufficient to reduce the bias in generation of sequences. However, generation and judgment may still be distinct tasks involving different processes and perhaps the process of generation encourages people to construct interconnectedness between outcomes. Alternatively, people may be limited in their ability due to attention or capacity restrictions as suggested elsewhere (Falk & Konold, 1997; Wagenaar, 1970a).

The results presented provide a new mindset from which to view previous findings and the oft stated comment that people have a poor understanding of randomness. It appears that when not influenced by causal information we are able to accurately perceive randomness. However, when a causal agent is present, as is the case in most everyday situations, interconnectedness between outcomes is perceived, resulting in biased judgments. No support was found for the suggestion by Falk & Konold (1997) that judgments of randomness are directed by perceptions of sequence complexity.

References

- Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory and Cognition*, *32*, 1369-1378.
- Bar-Hillel, M. & Wagenaar, W. A. (1993). The perception of randomness. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues*. New Jersey: Lawrence Erlbaum Associates.
- Budescu, D. V., & Rapoport, A. (1994). Subjective randomization in one- and two-person games. *Journal of Behavioral Decision Making*, *7*, 261-278.
- Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: "Gambler's fallacy" versus "hot hand". *Psychonomic Bulletin and Review*, *11*, 179-184.
- Falk, R. (1981). The perception of randomness. In *Proceedings of the Fifth International Conference for the Psychology of Mathematics Education*, *1*, 222-229.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, *104*, 301-318.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295-314.
- Kareev, Y. (1992). Not that bad after all: Generation of random sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1189-1194.
- Lopes, L. L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 626-636.
- Lopes, L.L., & Oden, G.C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 392-400.
- Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review*, *109*, 330-357.
- Rapoport, A., & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General*, *121*, 352-363.
- Roney, C. J. R., & Trick, L. M. (2003). Grouping and gambling: A Gestalt approach to understanding the gambler's fallacy. *Canadian Journal of Experimental Psychology*, *57*, 69-75.
- Tune, G.S. (1964). A brief survey of variables that influence random-generation. *Perceptual and Motor Skills*, *18*, 705-710.
- Wagenaar, W. A. (1970a). Appreciation of conditional probabilities in binary sequences. *Acta Psychologica*, *34*, 348-356.
- Wagenaar, W. A. (1970b). Subjective randomness and the capacity to generate information. *Acta Psychologica*, *33*, 233-242.
- Wagenaar, W. A. (1972). *Sequential response bias: A study on choice and chance*. Unpublished doctoral dissertation. Utrecht University.
- Weiss, R.L. (1965). "Variables that influence random-generation": An alternative hypothesis. *Perceptual and Motor Skills*, *20*, 307-310.
- Wiegiersma, S. (1982). Can repetition avoidance in randomization be explained by randomness concepts? *Psychological Research*, *44*, 189-198.
- Wiegiersma, S. (1987). The effects of visual conspicuousness and the concept of randomness on the recognition of randomness in sequences. *Journal of General Psychology*, *114*, 157-165.