Do Causal Beliefs Influence the Hot-Hand and the Gambler's Fallacy?

Giorgio Gronchi (giorgio.gronchi@gmail.com)
Università di Firenze, Dipartimento di Psicologia, via della Cittadella 7
50144 Firenze, Italy

Steven A. Sloman (steven_sloman whoisat brown.edu)
Brown University, Department of Cognitive and Linguistic Sciences, Box 1978
Providence, RI 02912 USA

Abstract

An open problem in the study of randomness perception is why, given the same sequence of independent and equiprobable events, sometimes people exhibit a positive recency effect (hot hand fallacy) and sometimes a negative recency effect (gambler's fallacy). In this paper we provide evidence of the role of causal belief about the source generating the sequence in predicting the next outcome. People that faced a streak with a nonrandom generating source tended to commit the hot hand fallacy, while people that faced a streak with a random source tended to commit the gambler's fallacy. No differences between the generating source were found in sequences that looked random.

Keywords: psychology, causal reasoning, decision making.

Introduction

In a sequence of independent and equiprobable binary outcomes like coin flips, when both elements have the same probability of occurring, the likelihood of the next item being one or the other value (heads or tails) is identical. However, two opposing errors are commonly committed by people who are asked to predict the next outcome. Some exhibit the negative recency effect (or gambler's fallacy) claiming that tails (T) is more likely to come up than heads (H) when the former is less frequent and recent than the latter. In contrast, others overestimate the positive contingency among the elements, believing that Heads will be more likely than Tails (positive recency effect or hot-hand fallacy). Positive recency has the effect of leading people to expect sequences of outcomes to exhibit streaks while negative recency leads to the expectation of alternation. Indeed, these biases are particularly common when evaluating sequences. Both biases are robust phenomena, yet the factors that give rise to the gambler's fallacy versus the hot hand fallacy remain unclear.

Starting with the seminal paper of Tversky and Kahneman (1971), evidence for the belief in the *law of small numbers* has been reported suggesting that people believe that a random process will correct itself in order to ensure equiprobability and irregularity in each subsequence generated. Kahneman and Tversky (1972) proposed an explanation based on local representativeness: People have beliefs about random process that mediate the evaluation of binary series and the prediction of outcomes. These beliefs lead people to expect equal frequencies of both outcomes and an irregular pattern and that small local sequences of outcomes should have these global characteristics. Thus a streak of one type of outcome should be followed by a different type. In contrast, Gilovich, Vallone and Tversky (1985) observed that basketball fans believed that if a basketball player hit a series of shots, then the next shot is more likely to be a hit (he has the "hot-hand"). However, a statistical analysis of actual shooting records provided no evidence for a positive correlation beteween successive shots. Again, Gilovich, Vallone and Tversky (1985) attributed this result to the representativeness heuristic: a streak of succesful hits is representative of and therefore induces belief that the sequence is not random.

The limits of the account from local representativeness have been discussed by various authors (Gigerenzer, 1991; Kubovy & Gilden, 1991). As Boynton (2003) says "it is unclear how a single heuristic could provide a complete explanation for positive and negative recency" (p. 119).

Hastie and Dawes (2001) made a proposal that complements local representativeness. They claim that people construct causal models of the generating source of the sequence. On one hand, if people believe that the source will produce a nonrandom output, the positive bias should occur. On the other hand, if one believes that the source is random, the gambler's fallacy should be observed.

A little empircal evidence supports this point. Burns (2002) showed that the mechanism by which a sequence was generated is critical in determining how people predict. He showed different sequences of 8 elements of Heads and Tails varying the proportion of elements, the presence of a final streak, and the generating source (random vs nonrandom). When the mechanism was nonrandom, participants were more likely to continue the streaks while when the mechanism was random, streaks and frequency of the previous outcomes had no impact. In fact, the gambler's fallacy was not observed in this experiment. He concluded that how the mechanism generating a sequence is understood is critical to how the sequence is interpreted. Attributing the sequence to a causal mechanism leads to the expectation of streaks, but attributions to a random process leads people to ignore information about streaks and the frequency of elements. However, Burns may have failed to observe a gambler's fallacy because of the charateristics of the sequence; it contained high frequency elements that were not very recent and vice versa.

Burns and Corpus (2004) asked participants to make a prediction after 100 equiprobable outcomes with varying scenarios (random vs nonrandom) including basketball shots, gambling with the roulette, and selling cars. They told participants only that there was a final streak of 4 elements and that the entire sequence was composed 50% of one outcome and 50% the other. They observed that with scenarios suggesting nonrandomness (like basketball), participants were more likely to continue a streak.

Recently, causal beliefs have been investigated in a degree of randomness evaluation task used previously by Falk and Konold (1997). A method to quantify the charateristics of a sequence is the probability of alternation (P(A)). Given a sequence with n elements, there are n-1 transitions between successive outcome and r-1 actual changes of symbol. The probability of alternation is the ratio between the actual changes and the total number of transitions. For example, having 8 elements, TTTTHHHH has 7 transitions but only a single actual change between elements (the fourth T followed by the fifth H), thus P(A) = 1/7. Asking participants to classify sequences as chance, streaky or alternating, Gilovich, Vallone and Tversky (1985) found that low and medium P(A) values were associated with streaky series. However, high values of P(A) were classified as chance series. Very high level of P(A) were classified as alternating (e.g., HTHTHTHTHT). This finding has been replicated many times (cf. Falk & Konold, 1997). Generally, sequences with a P(A) of about .65 are judged the most random. However, from a normative point of view, sequences with the highest values of objective randomness (computed as second-order entropy) are those with P(A) = .5. Variyng the probability of alternation and scenarios, McDonald and Newell (2007) found that highly alternating sequences were perceived as more random in a coin scenario compared to a basketball scenario. Thus, their study also suggests that causal belief has the potential to influence judgments of randomness.

However, some evidence suggests that causal models of the generating source do not have a role in predicting the next outcome. Boynton (2003) asked participants to predict a series of 100 outcomes and gave feedback after each event. He told participants that the generating source was sometimes random and sometimes not. Before each prediction, he assessed if the source was believed to be generating random or nonrandom output. He found that intuitions about the randomness or nonrandomness of the source did not influence the response.

In this paper we ask whether causal beliefs affect the prediction of a random binary series. We worry that previous experiments (Burns, 2002; Burns & Corpus, 2004) employing written scenarios did not succeed in inducing different causal beliefs about the generating source. To strengthen the causal belief manipulation, we employed videos of a random generating source (a die being rolled) and of a game between two basketball players making clear the competitive nature of the event.

On this basis, we expect that predictions will be mediated by both causal beliefs and the P(A) of the sequence. For streaky sequences (low and medium P(A) values), in the case of a random generating source people should commit the gambler's fallacy because they will think that the nonstreaky outcome is due (the law of small numbers; Tversky & Kahneman, 1971). However, for a nonrandom generating source people should exhibit a positve recency effect because they expect streaks to continue (especially when the player faces a streak of hits) as the work of Gilovich, Vallone and Tversky (1985) suggests.

Sequences with high values of P(A) that appear to be generated by chance should not elicit predictions of a particular outcome regardless of beliefs about the generating source. In sequences that appears to be random there is not a particular outcome that is too frequent or too recent, so in the case of the random generating source the law of small numbers does not apply (Tversky & Kahneman, 1971). At the same time, a chance looking sequence of hits or misses does not induce the belief that the player is hot. The player actually hits and misses the basket in a random way so no information is provided in order to make a prediction (Gilovich, Vallone & Tversky, 1985; Hastie & Dawes, 2001). Thus, people should not prefer a particular outcome.

Finally, causal beliefs about the generating source should not make a difference for alternating sequences (very high P(A) values). Also in this case, the recency and frequency of outcomes are similar and so there is no reason to apply the law of small numbers or to overestimate the contingency of a particular outcome. In sum, in the case of alternating sequences, people should not prefer a particular outcome but should respond at chance or continue the alternating pattern of the sequence.

Experiment

To assess the question of whether causal beliefs about the generating source influence the prediction of the next element in a random binary sequence, we conducted an experiment varying the scenarios and the probability of alternation of the sequence. We predict that both factors should mediate the prediction of the next element of a string.

Method

Participants Participants were 60 undergraduate and graduate students from Brown University, University of Pisa, and the University of Florence. Brown University students were recruited by an internet advertisement and paid 3 dollars for participations, while Italian students participated on a voluntary basis.

Stimuli Stimuli were videos of a die being rolled (Random condition) and two basketball players (an attacker and a defender) playing against each other (Nonrandom condition). Fourteen videos per condition were employed. The videos of the die showed a blackboard in which a die with three red dots and three blue dots was rolled 8 times.

The videos of the basketball match showed two players (A and B) playing one-on-one against each other. In each video the attacker shot 8 times at the basket while the defender tried to prevent him from scoring. The matches were as realistic as possible, however the distance between the basket and the position of each shot was kept constant. Each video, in both conditions, had a length of about 1 minute. Sequences were prepared in order to have two series of both conditions for each different value of P(A). Table 1 shows the different sequences and their associated predictions.

Table 1: Sequences employed in the experiment together with prediction

P(A)	Sequence	Prediction (Random)	Prediction (Nonrandom)
7/7	Series A: THTHTHTH	indifferent*	indifferent*
	Series B: HTHTHTHT	indifferent*	indifferent*
6/7	Series A: THTHHTHT	indifferent	indifferent
	Series B: HTHTTHTH	indifferent	indifferent
5/7	Series A: TTHTHHTH	indifferent	indifferent
	Series B: HHTHTTHT	indifferent	indifferent
4/7	Series A: TTHTHHHT	indifferent	indifferent
	Series B: HHTHTTTH	indifferent	indifferent
3/7	Series A: TTTHTHHH	alternation	repetition
	Series B: HHHTHTTT	alternation	repetition
2/7	Series A: THHHHTTT	alternation	repetition
	Series B: HTTTTHHH	alternation	repetition
1/7	Series A: HHHHTTTT	alternation	repetition
	Series B: TTTTHHHH	alternation	repetition

*it also possible to predict that the pattern will continue (alternating).

For each value of probability of alternation there were two equivalent sequences substituting Tails with Heads and vice versa. In the Nonrandom condition the sequences were the same (Heads corresponded to hitting the shot and Tails to missing the shot). For each probability of alternation, series A showed player A against player B and vice versa for series B. For P(A) = 1/7 and P(A) = 2/7, series A was associated with a final streak of misses (and series B a final streak of hits). In contrast, for P(A) = 3/7, series A was associated with a final streak of hits and series B to a final streak of misses. The last two columns shows the predictions for each sequence and condition in terms of repeating the last outcome of the sequence (repetition), predicting a different outcome (alternation) or no preference between the two alternatives (indifferent). A special case is represented by the sequence with P(A) = 7/7 as people could predict a continuation of the pattern (alternation in both cases).

Procedure Participants were randomly assigned to one of the two randomness conditions. Half the participants read the following instructions:

> You are about to see a series of scenes of a fair die being rolled. The die has 3 sides with a blue dot and 3 sides with a red dot so the probabilities of rolling a blue or a red are the same and both equal to 50%. You will see the die rolled 8 times and then you will be asked to predict the outcome of the 9th roll.

In the basketball condition the task was:

> You are about to see two basketball players (an attacker and a defender) playing one-on-one against each other. Each player has a probability of 50% of hitting the basket. For each scene, I'd like you to predict the most probable outcome of the final shot. You we will see one of the player shoot 8 times and you will be asked to predict the 9th shot.

Both groups performed the task in front of a personal computer writing their responses in a questionnaire. Each participant saw the fourteen sequences in a random order with the only constraint being that sequences of series A were always followed by sequences of series B and vice versa. After observing a single sequence they were requested to predict if the next outcome was 'blue' or 'red' (Random condition) or if the attacker will hit or miss the basket on the next shot (Nonrandom condition). After each judgment they wrote how confident they were in their prediction on a 7-point Likert scale. Each participant performed the task alone in a quiet place. Both tasks together required about 20 minutes to complete.

Results

Table 2 shows the percentages of participants predicting a repetition of the last outcome for each sequence in the two conditions.

Table 2: Percentages predicting a repetition of the last outcome along with results of Chi^2 tests comparing conditions.

P(A)	Series	Random vs Nonrandom	$Chi^2(1)$	*p*
7/7	A	20.0 vs 53.3	5.81	p<.01
7/7	B	20.0 vs 50.0	4.69	p<.05
6/7	A	50.0 vs 43.3	0.07	-
6/7	B	53.3 vs 50.0	0.00	-
5/7	A	30.0 vs 53.3	2.14	-
5/7	B	50.0 vs 50.0	0.00	-
4/7	A	60.0 vs 50.0	0.27	-
4/7	B	56.7 vs 76.7	1.88	-
3/7	A	26.7 vs 56.7	4.39	p<.05
3/7	B	30.0 vs 56.7	3.33	p<.05
2/7	A	23.3 vs 56.7	5.63	p<.01
2/7	B	30.0 vs 70.0	8.07	p<.01
1/7	A	30.0 vs 46.7	2.46	-
1/7	B	30.0 vs 70.0	11.32	p<.001

With regard to the random condition, similar percentages were observed for each different P(A) with the exception of series B with P(A) = 5/7.

For the Nonrandom condition, differences between series were observed only for P(A) = 1/7 and P(A) = 2/7. Series B was more frequently predicted to continue.

With regard to streaky sequences (P(A) = 1/7, P(A) = 2/7 and P(A) = 3/7) significant differences were observed between the two conditions (with the exception of series A with P(A) = 1/7). As predicted by the causal belief hypothesis, percentages continuing the streak were lower in the Random condition. In the basketball scenario, streaks were most likely to be expected to continue for series with P(A) = 1/7 and P(A) = 2/7 that included a final streak of hits.

Again, as predicted by the hypothesis, no differences were found between the two conditions for chance sequences (P(A) = 4/7, P(A) = 5/7, P(A) = 6/7). With some exceptions, percentages for both scenarios were around 50%.

Finally, in the alternating series (P(A) = 7/7), a statistically significant difference between the two condition was observed. In the Random condition participants tended to not repeat the last outcome in order to continue the alternating sequence (HTHTHT...). In the Nonrandom condition, participants showed no preference.

With regard to confidence judgments, no significant differences were observed across conditions except that the mean confidence rating in the Nonrandom condition (M = 3.73, sd = 1.05) was higher than the mean confidence rating in the Random condition (M = 3.10, sd = 1.28) (t_{58}=2.10, $p < .05$).

Discussion

In summary, this experiment confirms the role of causal beliefs about the generating source in determining whether people commit the gambler's fallacy or the hot hand fallacy. First, given the same sequences people were less confident in their prediction with a random than a nonrandom generating source. A possible explanation is that people were unsure about their judgments when the generating source was random but felt they had grounds for a prediction when they had causal beliefs about the operation of the source. Second, the probability of alternation of the sequences mediated the influence of causal belief.

In the case of the dice, participants that faced a streak tended to commit the gambler's fallacy. In contrast, observing two basketball players playing against each other, participants that faced a streak tended to have no preference or, in the case of a streak of hits, to commit the hot-hand fallacy. It turned out that participants chose to continue the streak only about 50% of the time in four sequences of the nonrandom condition over the six streaky strings (P(A) = 1/7, 2/7 and 3/7). By contrast, in the random condition people were consistently below 30% in all six streaky strings. The hot-hand fallacy was apparently more likely in the face of a streak of hits than a streak of misses. When facing a streak of misses, the probability of repeating the same outcome was consistently around 50%, while when facing a streak of hits in two sequence out of three probability was 70%.

No difference in the predictions was observed when sequences appeared to be random and people didn't show any preferences for the two outcomes. Finally, participants were more likely to continue an alternating pattern in the case of the die but not for the basketball scenario.

The perceived randomness of the string may have modulated the influence of the causal belief on predictions. If the sequence looked random, participants responded at chance ignoring information about the generating source. If the sequence presented a streaky pattern, people overestimated or underestimated the contingency between the events in a way consistent with their causal beliefs. This type of strategy led them to the gambler's fallacy (due to the randomness of the generating source, the sequence should correct itself) or the hot-hand fallacy (e.g., because the player is hot, it is more likely a hit than a miss).

In the case of an alternating sequence, causal beliefs seem to also have a role. On one hand, in the basketball scenario the attacker was unlikely to be willingly hitting and missing the basket as the regularities in the pattern suggested, i.e. any basketball player would presumably to make every basket. This may explain why predictions were at chance. On the other hand, a fair die that produced a clearly nonrandom looking output (such as an alternating pattern) is likely to induce people to think that the die or the way it was rolled was not fair. Thus, participants were more likely to follow the observed regularities in making predictions.

In sum, the gambler's fallacy and the hot-hand fallacy depend on whether the sequence is streaky or not, i.e. whether the recency and frequency of the elements are extreme enough to lead to overestimations of the contingency among outcomes.

In this study participants were forced to choose the next event in the sequence and they did not have the option of saying that the two events were equally likely. Maybe giving three options would magnify our small differences among conditions. Furthermore, it is possible that previous studies didn't find a clear role of causal beliefs because they employed only a written task without showing the actual generating process. One of the strongest pieces of evidence against the role of causal belief comes from Boynton (2003) who didn't find an influence of the belief about the generating source on prediction. An explanation of these results might be that the task didn't describe in a clear way the characteristics of the generating source (more specifically, a computer that sometimes produces a random and sometimes a nonrandom output). Moreover, Boynton's (2003) task had participants deciding if the computer was generating a random or nonrandom output. As Nickerson (2002) points out, speaking of a random product can be misleading. In our opinion, in order to determine if the causal belief about the generating source has an influence, it is fundamental to stress the role of the process and not of the product.

Our results suggest that belief about the generating source, together with the characteristics of the sequence (streaky, chance or alternating) do indeed mediate predictions. They confirm the role that causal schema have in how we perceive events and how we make judgments about them. A broad literature (reviewed in Sloman, 2005) shows that causal models represent an effective tool for understanding and predicting events. However, they can also affect our judgment inducing particular biases such as the positive/negative recency effects.

Acknowledgments

This work was supported by a University of Florence Fellowship for graduate students. Many thanks to Dave Buchanan, Adam Darlow, Phil Fernbach, Rachel Ostrand, Ju-Hwa Park, Caterina Primi, John Santini, Qun Yang and 4 anonymous reviewers for valuable input.

References

Boynton, D. M. (2003). Superstitious responding and frequency matching in the positive bias and the gambler's fallacy effects. *Organizational Behavior and Human Decision Processe*s, 91(2), 119-127.

Burns, B. D. (2002). Does the law of small numbers explain the gambler's fallacy? *48th Annual Meeting of the Psychonomic Society*, Kansas City, MO, November 2002.

Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: "Gambler's fallacy" versus "hot hand". *Psychonomic Bulletin & Review, 11*(1), 179-184.

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

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond 'Heuristics and Biases'. *European Journal of Social Psychology, 17*, 295-314.

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

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world*. London: Sage.

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

Kubovy, M., & Gilden, D. (1991). Apparent randomness is not always the complement of apparent order. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of the structure.* Washington DC: APA.

McDonald, F. E. J., & Newell, B.R. (2007). Judgments of randomness: is there a bias towards high alternations? *48th Annual Meeting of the Psychonomic Society*, Long Beach, CA, November 2007.

Nickerson, R. S. (2002). The production and the perception of randomness. *Psychological Review, 109*(2), 230-357.

Sloman, S. A. (2005). *Causal models: How we think about the world and its alternatives*. New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1971). Belief in the 'Law of small number'. *Psychological Bulletin, 76*, 105-110.