

The smart potential behind probability matching

Wolfgang Gaissmaier^{a,b,*}, Lael J. Schooler^a

^aMax Planck Institute for Human Development, Berlin, Germany

^bHarding Center for Risk Literacy, Lentzeallee 94, 14195 Berlin, Germany

A B S T R A C T

Probability matching is a classic choice anomaly that has been studied extensively. While many approaches assume that it is a cognitive shortcut driven by cognitive limitations, recent literature suggests that it is not a strategy per se, but rather another outcome of people's well documented misperception of randomness. People search for patterns even in random sequences, which results in probability matching at the outcome level. Previous studies have supported this by the finding that distracting people with a secondary verbal working memory task presumably prevents the pattern search, resulting in more maximizing behavior that is considered more rational. The current paper demonstrates with two experiments that there is actually truth in both accounts. For some participants, probability matching indeed seems to be the result of a cognitive shortcut, a simple "win stay, lose shift" strategy, and in one experiment identified these as participants low in working memory capacity. For others, however, a potentially smart pattern search strategy underlies probability matching. These probability matchers (who still look irrational in the absence of patterns) actually have a higher chance of finding a pattern if one exists. Contrary to the almost uniformly negative perception of probability matching, we therefore conclude that there can be a potentially smart strategy behind probability matching.

Keywords:

Probability matching
Working memory capacity
Pattern search

1. Introduction

"Clarice, does this *random* scattering of sites seem overdone to you? Doesn't it seem *desperately* random? Random past all possible convenience? Does it suggest to you the elaborations of a bad liar?" (Harris, 1988/1991, p. 293).

In Thomas Harris's "The Silence of the Lambs", the psychiatrist Dr. Hannibal Lecter, an imprisoned murderous cannibal, helps FBI agent Clarice Starling find a serial killer. His comment on a map of the sites where the victims were found is crucial: The serial killer tries to hide his location among seemingly random sites, but by being too random to be credibly random, he unintentionally helps agent Starling discover an important pattern that leads to him.

This example illustrates both that people are not overly successful in producing randomness and that they are very

good in detecting patterns. Both may be two sides of the same coin – the well documented misperception of randomness (e.g., Bar Hillel & Wagenaar, 1991) may facilitate the early detection of patterns (Lopes, 1982). Sometimes, however, people even detect patterns where there are none, and it is easier to convince them that a sequence is structured than that it is random (Hyman & Jenkin, 1956).

Probability matching, a classic choice anomaly, could be a further consequence of misperceiving randomness. In a typical experiment, people have to predict which of two events that have different probabilities of occurring will take place. For example, event E1 could occur with a probability of $p(E1) = .67$, while event E2 occurs with $p(E2) = 1 - p(E1) = .33$. Given that the sequence of events is random, the best strategy would be always to predict the more frequent event E1. This strategy is called *maximizing* and would yield an average accuracy of 67%. However, most often *probability matching* is observed, that is, predicting the events in proportion to their probability of

* Corresponding author. Tel.: +49 30 82406691; fax: +49 30 82406394.
E-mail address: gaissmaier@mpib-berlin.mpg.de (W. Gaissmaier).

occurrence. In the example this would mean predicting E1 in 67% of the trials and E2 in 33% of the trials. Probability matching is considered suboptimal, because it would yield an accuracy of only 55.78% on average ($.67 \times .67 + .33 \times .33$). There are numerous attempts to explain this choice anomaly (see Vulkan (2000), for a review). A commonality behind many of these approaches is that they tie probability matching to cognitive limitations, based on the idea that “the default processing strategy of most participants is a nonnormative cognitive shortcut” (West & Stanovich, 2003, p. 249).

Yet there is also a quite different view on probability matching, which we would like to elaborate on here: Probability matching occurs because people do not see or believe the randomness of the sequence and try to outdo the optimal maximizing strategy (and inevitably fail). For instance, people maximize more strongly if they believe in the randomness of the sequence. Goodnow (1955) showed that people were more likely to maximize if the task was framed as gambling rather than problem solving. Wolford, Newman, Miller, and Wig (2004) manipulated the perceived randomness of a sequence of events more directly by increasing the alternation rate of the binary event. If events alternate (slightly) more often than would be expected by chance, then the sequence is actually less random, but people perceive it to be more random (cf. Ayton & Fischer, 2004; Falk & Konold, 1997; Gilovich, Vallone, & Tversky, 1985). Wolford et al. found that people maximized more strongly if the alternation rate was higher than chance than when it was actually random.

In line with Wolford et al. (2004), we believe that probability matching is not a strategy per se, but rather the outcome of a more complex, cognitively more demanding strategy—the search for patterns. Wolford, Miller, and Gazzaniga (2000) pointed out that any plausible pattern a person might try has to match the probabilities. As long as people do not believe that the sequence is random, they will try to improve their predictive accuracy by looking for regularities in the sequence. In the last block of a binary choice experiment conducted by Yellott (1969), participants always received feedback indicating that their predictions were correct, irrespective of what they predicted. They continued to match probabilities, and when asked for their impressions afterwards, most responded that they finally found the pattern in the sequence. Congruently, Unruh and Corominas (2007) showed that participants who reported having found complex rules in a random sequence of binary events were closer to probability matching than those who did not report such rules.

The pattern search idea is further supported by Wolford et al.'s (2004) finding that distracting people with a secondary verbal working memory task presumably prevents the pattern search, resulting in more maximizing behavior (and thus in behavior that is considered more rational). Similarly, Kareev, Lieberman, and Lev (1997) showed that people with low short term memory capacity maximized more consistently than did people with a higher short term memory capacity in a very similar task. Gaissmaier, Schooler, and Rieskamp (2006; cf. Gaissmaier, Mata, & Schooler, 2008) replicated this result and interpreted it along the lines of Wolford et al. by hypothesizing that peo-

ple with low short term memory capacity lack the cognitive capacity needed to search much for patterns. Moreover, Gaissmaier et al.'s modeling efforts suggest that people with higher short term memory capacity tended to explore more. Although this explorative behavior put them at a disadvantage as long as the environment was stable, it actually helped them track changes in the environment.

In this paper we report 2 experiments that further investigate what underlies probability matching and test two competing hypotheses. If (i) probability matching indeed reflects the tendency to search for patterns, then people who probability match in the typical binary choice task should be more likely to find patterns if they exist. Indeed, those people who have been traditionally interpreted as being irrational (because they probability match in the absence of patterns) would actually look smart when patterns are present. Alternatively, if (ii) probability matching results from a lazy cognitive shortcut, as proposed by West and Stanovich (2003), then it should not be related to finding patterns. We studied a well defined shortcut that leads to probability matching, “win stay, lose shift”. This strategy consists of repeating predictions if they were successful in the last trial, but changing them if they failed.

Moreover, we investigated how strategy use is related to working memory capacity, which itself to some degree reflects differences in strategic processing of information (Cokely, Kelley, & Gilchrist, 2006). While pattern search should go along with higher working memory capacity (e.g., DeCaro, Thomas, & Beilock, 2008), relying on a cognitive shortcut should go along with lower working memory capacity, such as West and Stanovich (2003) propose.

2. Experiment 1

2.1. Method

2.1.1. Participants

Eighty people (44 female), mostly students, with an average age of 24 years ($SD = 4.06$) participated in the study.

2.1.2. Design and procedure

The main task was to predict repeatedly whether a red square would appear in the upper half of the screen or a green square in the lower half, similar to Wolford et al. (2004). After 10 training trials, there were 576 trials altogether, divided into halves of 288 trials. Participants earned € .01 for each correct trial in addition to a € 5 show up fee. Every participant saw both a no pattern and a pattern half, the order of which was counterbalanced.

The no pattern half consisted of 288 trials of the binary choice task without any pattern. That is, the sequence of events was serially independent. The probability of the more common event was $p = .67$, while the probability of the less common event was $1 - p = .33$.

The pattern half consisted of a pattern of length 12, in the sequence 001010001100. The probability of the more common event was thus also $p = .67$ (i.e., 8/12), while the probability of the less common event was $p = .33$ (i.e., 4/12). The pattern was repeated for 288 trials.

Between the halves, there was a short break and participants were reminded that the second half could potentially be different. To make it easier to detect that something was indeed different, the more common event was reversed between the two halves.¹

An automated operation span task (Unsworth, Heitz, Schrock, & Engle, 2005; cf. Conway et al., 2005) was administered before the main task to assess working memory capacity.

2.2. Results and discussion

For all participants, we counted how often they predicted the more frequent event in the no pattern half. We will subsequently call this variable *maximizing*, as it reflects the degree to which someone is applying a maximizing strategy. A value of 1 would reflect perfect maximization (i.e., always predicting the more frequent event), a value of .67 would reflect probability matching, and a value of .5 would reflect a guessing strategy.

2.2.1. Median split analyses

For each participant, we computed how close this participant's behavior was to probability matching by calculating the absolute difference between their degree of maximizing and probability matching (i.e., .67) in the no pattern half. We divided participants into probability matchers and non matchers at their median proximity to probability matching. This median split was done separately for each of the counterbalancing conditions, so that there were 20 probability matchers and 20 non matchers each.

Based on this classification in the no pattern half, we predicted that probability matchers would be more successful in picking up the pattern than non matchers in the pattern half. To investigate this, we looked at the proportion of trials where they made correct predictions in the pattern half (*pattern accuracy* subsequently). The pattern accuracy of probability matchers and non matchers was similar early in the pattern half. But over time, probability matchers improved their accuracy more than non matchers, suggesting that they were better able to pick up the pattern (Fig. 1A). In a repeated measures analysis of variance, we included the within subject factor block (consisting of the accuracy in three pattern blocks of 96 trials each) and two between subjects factors, probability matcher vs. non matcher and order of presentation. The growing separation between probability matchers and non matchers was reflected in the linear contrast between the factors probability matcher vs. non matcher and block, $F(1, 76) = 7.86, p = .006, \eta_p^2 = .09$. None of the interactions including the factor order of presentation was substantial, nor was its between subjects effect.

2.2.2. Curvilinearity analyses

Median splits as used here can decrease statistical power and introduce error (Irwin & McClelland, 2003). Fur

¹ Which event was more common in which half was additionally counterbalanced, but will be disregarded subsequently since it did not affect the main dependent variables.

thermore, by using the absolute distance to probability matching, we have treated deviations from probability matching in both directions (i.e., toward chance performance and toward maximizing) identically. That is, we have looked at a linear relation between proximity to probability matching in the no pattern half and the ability to pick up patterns in the pattern half.

However, if one does not use proximity to probability matching but the proportion of maximizing responses in the no pattern half as independent variable, then one would predict a curvilinear relation to the ability to pick up patterns. Namely, pattern finding should peak at probability matching level and drop both toward chance performance and toward maximizing.

Such a curvilinear trend was reflected in a quadratic regression with the proportion of maximizing responses in the no pattern half as independent variable and accuracy in the last pattern block as dependent variable, $F(2, 77) = 2.63, p = .08, R^2 = .06$. The linear regression only accounted for less than 1% of the variance, $F(1, 78) = 0.06, p = .80, R^2 = .001$. The linear contrast effect in the median split analysis suggests that the improvement in pattern accuracy differs between probability matchers and non matchers, as probability matchers started with about the same accuracy as non matchers, but ended up higher. Similarly, the explained variance in the quadratic regression is greater if instead of accuracy on the last pattern block the improvement between the first and the last pattern block is predicted, $F(2, 77) = 7.64, p = .001, R^2 = .17$ (Fig. 2A). Again, the linear regression only accounted for 1% of the variance, $F(1, 78) = 0.84, p = .36, R^2 = .01$.²

2.2.3. Win stay, lose shift

As pointed out in the introduction, probability matching can also result from a cognitive shortcut, namely a simple "win stay, lose shift" strategy. This strategy predicts that participants repeat predictions after successful trials, but change them after failures.

For each participant, we calculated the proportion of trials in the no pattern half that could be described by "win stay, lose shift". Participants using that strategy would also end up probability matching, but we would not assume that they are looking for patterns. Congruently, "win stay, lose shift" negatively correlated with pattern accuracy in the last pattern block, $r(38) = .37, p = .02$, and improvement across pattern blocks, $r(38) = .40, p = .01$ among probability matchers (as defined by the median split).

We then dichotomously classified participants as using "win stay, lose shift" or not. We did this by comparing the fit of "win stay, lose shift" to the participants' responses to

² In addition to the curvilinearity analyses on decision outcomes, we also ran similar analyses to predict response times. In Experiment 1, there was a curvilinear relation between maximizing and response times, indicating that probability matchers were both slower than maximizers and than guessers, supporting the idea that they were applying a more complex strategy. The results from Experiment 2, however, were inconsistent with these results. First of all, the relation between maximizing and response times was linear rather than curvilinear – the higher the degree of maximizing, the faster. More importantly even, working memory capacity explained more variance in speed, with higher working memory going along with faster response times, which was not the case in Experiment 1.

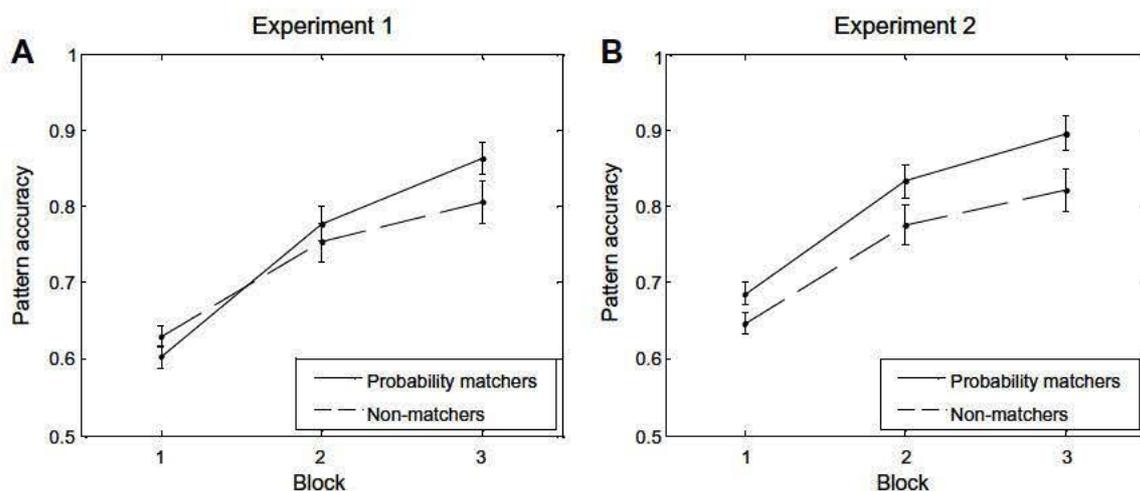


Fig. 1. Mean accuracy in the pattern half of the experiment (\pm standard error of the mean) for probability matchers and non-matchers as classified in the no-pattern half. The data is depicted for three blocks of 96 trials each, separately for (A) Experiment 1 and (B) Experiment 2.

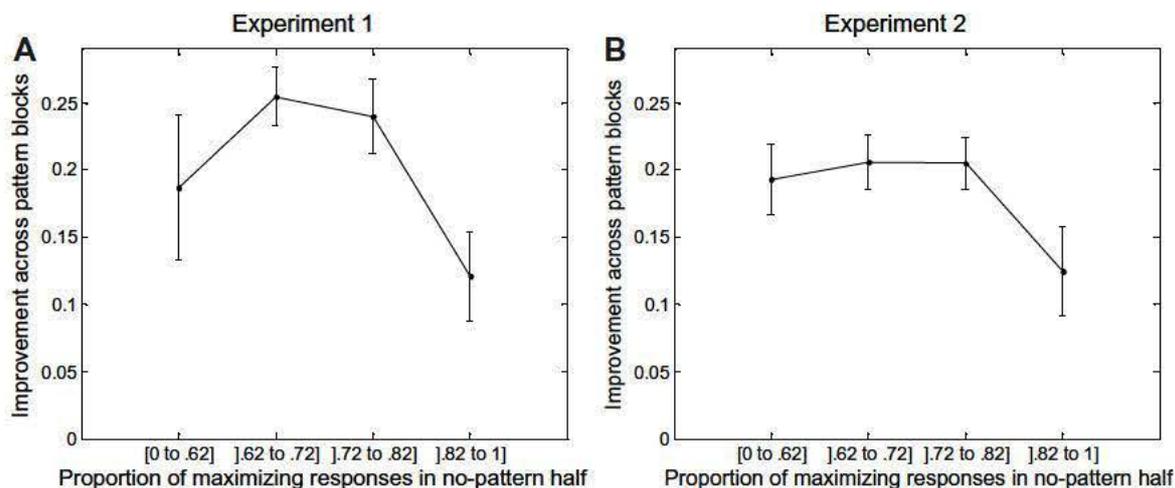


Fig. 2. Curvilinear relation between the mean proportion of maximizing responses in the no-pattern half and mean improvement in accuracy in the pattern half from block 1 to block 3 (\pm standard error of the mean) for (A) Experiment 1 and (B) Experiment 2. To define the bins, we have started with the bin including probability matching (i.e., .67) and have widened it by adding and subtracting .05 (i.e., it goes from .62 to .72). Since there are only few people below .62, all of them are included in one bin. The number of participants in the bins is 10, 34, 22, and 14, respectively, from left to right, in Experiment 1, and 29, 42, 54, and 14, respectively, in Experiment 2.

the fit that would be expected, given that someone did not use “win stay, lose shift”, but simply responded according to her base rates of predicting one option over the other.³ Among probability matchers, 40.0% were classified as using “win stay, lose shift”, compared to 27.5% of non matchers.

³ The expected fit of “win-stay, lose-shift” depends on the base rates of predicting one event over the other. Here, it is $.67 \times p + .33 \times (1 - p)$ on average, where p is the proportion of trials that a participant predicts the more frequent event. We simulated 10,000 runs for each participant. For each run, we used the stimuli the participant saw, and calculated the proportion of trials that would be fit by “win-stay, lose-shift”, given that participants made predictions according to their base rate, p . Participants were classified as using “win-stay, lose-shift” if the fit of this strategy to their actual predictions was higher than the best fit in the 10,000 runs of the simulation.

Those probability matchers who used “win stay, lose shift” only achieved a final pattern accuracy of 81.3%, compared to 89.5% of those who did not use “win stay, lose shift”, $t(21.9) = 1.75$, $p = .09$, $d = 0.59$.

2.2.4. Working memory capacity

The use of more complex strategies such as pattern search should be related to working memory. Congruently, working memory capacity was related to pattern accuracy on the last pattern block, $r(78) = .22$, $p = .05$. However, it did not predict proximity to probability matching, $r(78) = .02$, $p = .89$.

We showed that probability matching is the result of pattern search for some, but the result of “win stay, lose shift” for others. Supporting the idea that “win stay,

lose shift” is a cognitive shortcut, this strategy describes probability matchers with a lower working memory capacity better, $r(38) = .36, p = 0.02$.

3. Experiment 2

Experiment 1 only employed one particular pattern. To test whether the results would generalize to other patterns, we replicated the experiment but instead each participant received a randomly generated pattern. We expected, however, that random patterns could introduce another source of variability that could wash out the systematic relation between the strategy a person employs and their ability to find patterns. Whether or not people find a pattern is not only a function of their strategy but also of the difficulty of the pattern. For example, some patterns could be so transparent that someone wedded to maximizing could not miss them. Others could be so difficult that even a dedicated pattern searcher would not find them. To deal with this, we increased the number of participants. Furthermore, as we did not predict any differences between the two counterbalancing conditions, and as we did not find any differences in Experiment 1, we eliminated one of the counterbalancing conditions; all participants now started with the no pattern half followed by the pattern half.

3.1. Method

3.1.1. Participants

One hundred and thirty nine people (83 female), mostly students, with an average age of 24 years ($SD = 3.11$) participated in the study.

3.1.2. Design and procedure

Design and procedure were identical to Experiment 1 with the following exceptions. Now all participants started with the no pattern half followed by the pattern half. The pattern half consisted of a pattern of length 12. This time, the pattern was not identical for each participant, but instead each participant received a randomly generated pattern. The only restriction of the pattern was that the probability of the more common event was again $p = .67$ (i.e., 8/12), while the probability of the less common event was $p = .33$ (i.e., 4/12). The pattern was repeated for 288 trials.

3.2. Results and discussion

The following analyses are identical to the major analyses in Experiment 1, which are motivated and described in detail in Experiment 1.

3.2.1. Median split analyses

Again, we divided participants into probability matchers and non matchers at their median absolute distance to probability matching (i.e., .67) in the no pattern half. Based on this classification in the no pattern half, we then tried to predict how good people were in finding the randomly generated pattern in the pattern half.

In a repeated measures analysis of variance, we included the within subject factor block (consisting of the accuracy in three pattern blocks of 96 trials each) and probability matcher vs. non matcher as between subjects factor. This time, there was no growing separation between probability matchers and non matchers, as there was basically no linear contrast effect between the factors probability matcher vs. non matcher and block, $F(1,137) = 2.47, p = .119, \eta_p^2 = .02$. Instead, probability matchers were better than non matchers overall, reflected in a between subjects effect for the factor probability matcher vs. non matcher, $F(1,137) = 7.08, p = .009, \eta_p^2 = .05$ (Fig. 1B). Individual t tests revealed that this difference held on all three blocks, all $t_s > 2.10$, all $p_s < .04$, Cohen's d ranging from 0.36 to 0.49.

3.2.2. Curvilinearity analyses

In Experiment 1, pattern accuracy was highest for people closest to probability matching and dropped for people whose predictive behavior in the no pattern half was closer to guessing or to maximizing. This shows up again: A curvilinear trend was reflected in a quadratic regression with the proportion of maximizing responses in the no pattern half as independent variable and accuracy in the last pattern block as dependent variable, $F(2,136) = 3.23, p = .04, R^2 = .05$. The linear regression only accounted for less than 1% of the variance, $F(1,137) = 1.21, p = .27, R^2 = .009$. That there was no linear contrast effect in the median split analysis in Experiment 2, contrary to Experiment 1, suggests that it is not so much the improvement in pattern accuracy that differs between probability matchers and non matchers this time. Congruently, although there was a curvilinear trend reflected in a quadratic regression when predicting the improvement between the first and the last pattern block (such as in Experiment 1), it was not stronger than predicting final pattern accuracy, $F(2,136) = 3.86, p = .02, R^2 = .05$ (Fig. 2B). Again, the linear regression was not substantial, $F(1,137) = 0.22, p = .64, R^2 = .002$.

3.2.3. Win stay, lose shift

In Experiment 1, we found that if participants are probability matching because they are relying on a “win stay, lose shift” strategy, this made it less likely that they found the pattern. As in Experiment 1, the accordance of a participant's behavior with “win stay, lose shift” among probability matchers (as defined by the median split) negatively correlated with pattern accuracy in the last pattern block, $r(73) = .28, p = .01$, but not with improvement across pattern blocks this time, $r(73) = .14, p = .22$.

Again, we dichotomously classified participants as using “win stay, lose shift” or not, such as in Experiment 1. Among probability matchers, 50.7% were classified as using “win stay, lose shift”, compared to 51.6% of non matchers. Those probability matchers who used “win stay, lose shift” only achieved a final pattern accuracy of 85.3%, compared to 94.0% of those who did not use “win stay, lose shift”, $t(61.6) = 2.81, p = .007, d = 0.65$.

3.2.4. Working memory capacity

In Experiment 1, working memory capacity positively correlated with pattern accuracy and negatively correlated

with accordance to “win stay, lose shift”. This time, working memory capacity was not related to pattern accuracy on the last pattern block, $r(137) = .09$, $p = .28$, but was weakly correlated to overall pattern accuracy, $r(137) = .15$, $p = .09$. As in Experiment 1, it was not related to proximity to probability matching per se, $r(137) = .06$, $p = .52$. This time, working memory capacity was not related to “win stay, lose shift” among probability matchers, $r(73) = .08$, $p = 0.49$.

4. General discussion

We showed differences between participants in their tendencies to search for patterns. Predictive behavior more closely in line with probability matching in the absence of a pattern was related to pattern accuracy when a pattern was introduced in both experiments. At least for some people, probability matching thus results from relentless search for patterns. For them, probability matching is actually not a strategy per se, but rather the outcome of a more complex search for patterns.

However, there were also probability matchers who did not pick up regularities. We identified some of these people if their probability matching could be better described by a “win stay, lose shift” strategy. At least in Experiment 1, “win stay, lose shift” went along with lower working memory capacity. Although this relation to working memory capacity did not replicate in Experiment 2, we believe it is justified to say “win stay, lose shift” can be seen a cognitive shortcut. It is simple but at the cost of hindering pattern detection in this setting. Still, this does not mean that “win stay, lose shift” is generally a bad strategy, and it is known to be very successful in competitive situations such as the prisoner’s dilemma game (Nowak & Sigmund, 1993).

Taken together, these results demonstrate the importance of looking beyond the outcome of decisions and to the underlying strategies. Here, the same outcome, probability matching, is reached by some participants by a simple “win stay, lose shift” strategy. For others, however, rather the opposite seems to be true. They can be described as probability matching, because they are looking for patterns rather than relying on cognitive shortcuts.

Lopes (1982) discussed the predisposition towards pattern detection in a signal detection framework, arguing that it can be seen as a liberal criterion to classify some thing as signal (i.e., generated by a nonrandom process) in stead of noise (i.e., generated by a random process). Thereby, people limit the number of misses but increase the number of false alarms (cf. Kareev, 1995, for a similar argument regarding the detection of correlation). To conclusively evaluate the predisposition towards pattern detection, one needs to consider the structure of the real environment (Todd & Gigerenzer, 2007).

And regarding the natural environment, Pinker (1997, p. 346) speculated that patterns are ubiquitous: “Many events work like that. They have a characteristic life history, a changing probability of occurring over time.” Ayton and Fischer (2004), cf. Ayton, Hunt, & Wright, (1989) have pointed out that there are not many natural environments where one can safely assume conditional independence of

a succession of events, except for casinos and psychology laboratories. However, one can easily come up with examples where incorrectly assuming patterns although there are none can cause great harm. For example, people look for trends and patterns in investment decisions and often pay too much attention to recent outcomes (Rieskamp, 2006). Additionally, one could argue that in the current experimental setting, participants could have employed a maximizing strategy and at the same time looked for patterns, as feedback was independent of choice. The results indicate that this is not what people were doing, and this is in line with typically observed differences between *acting* and *observing*, which have been reported for maximizing in binary choice (Newell & Rakow, 2007) and for recency effects (Rakow, Demes, & Newell, 2008). And although maximizing and looking for patterns at the same time would be possible in this setting, in many other environments it would not be, because often feedback will depend on choice.

Therefore, we refrain from finally evaluating the misperception of randomness and the resulting epiphenomenon probability matching as rational or irrational. Still, the results reported here illustrate the importance of considering the structure of the environment when evaluating behavior. What works poorly in one environment (here: searching for patterns if there are none and ending up with suboptimal probability matching) may work well in another (here: actually finding patterns if they do exist). Contrary to the almost uniformly negative perception of probability matching, we therefore conclude that there can be a potentially smart strategy behind probability matching.

Acknowledgements

This research was funded by the Max Planck Institute for Human Development, Berlin, Germany. We thank Bettina von Helversen, Julian N. Marewski, Valerie F. Reyna and George Wolford for insightful comments on the project and Anita Todd for editing the manuscript.

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.
- Ayton, P., Hunt, A. J., & Wright, G. (1989). Psychological conceptions of randomness. *Journal of Behavioral Decision Making*, 221–238.
- Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics*, 12, 428–454.
- Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin and Review*, 13, 991–997.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin and Review*, 12, 769–786.
- DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, 107, 284–294.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104, 301–318.
- Gaissmaier, W., Schooler, L. J., & Mata, R. (2008). An ecological perspective to cognitive limits: Modeling environment–mind interactions with ACT-R. *Judgment and Decision Making*, 3, 278–291.

- Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 966–982.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295–314.
- Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *American Journal of Psychology*, 68, 106–116.
- Harris, T. (1991). *The silence of the lambs*. New York: St. Martin's Press (first edition in 1988).
- Hyman, R., & Jenkin, N. S. (1956). Involvement and set as determinants of behavioral stereotypy. *Psychological Reports*, 2, 131–146.
- Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, 40, 366–371.
- Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, 56, 263–269.
- Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, 126, 278–287.
- 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.
- Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin and Review*, 14, 1133–1139.
- Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in Prisoner's Dilemma. *Nature*, 364, 56–58.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes*, 106, 168–179.
- Rieskamp, J. (2006). Positive and negative recency effects in retirement savings decisions. *Journal of Experimental Psychology: Applied*, 121, 233–250.
- Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science*, 16, 167–171.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498–505.
- Unturbe, J., & Corominas, J. (2007). Probability matching involves rule-generating ability: A neuropsychological mechanism dealing with probabilities. *Neuropsychology*, 21, 621–630.
- Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, 14, 101–118.
- West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory and Cognition*, 31, 243–251.
- Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *The Journal of Neuroscience*, 20(RC64), 1–4.
- Wolford, G., Newman, S., Miller, M. B., & Wig, G. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology*, 58, 221–228.
- Yellott, J. I. Jr. (1969). Probability learning with noncontingent success. *Journal of Mathematical Psychology*, 6, 541–575.