Current Biology *24*, 1918–1922, August 18, 2014 http://dx.doi.org/10.1016/j.cub.2014.07.013

Report

Asymmetric Predictability and Cognitive Competition in Football Penalty Shootouts

Erman Misirlisoy[1,*] and Patrick Haggard[1,*]
[1]Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

Summary

Sports provide powerful demonstrations of cognitive strategies underlying competitive behavior [1]. Penalty shootouts in football (soccer) involve direct competition between elite players and absorb the attention of millions. The penalty shootout between Germany and England in the 1990 World Cup semifinal was viewed by an estimated 46.49% of the UK population [2]. In a penalty shootout, a goalkeeper must defend their goal without teammate assistance while an opposing series of kickers aim to kick the ball past them into the net. As in many sports [3], the ball during a penalty kick often approaches too quickly for the goalkeeper to react to its direction of motion; instead, the goalkeeper must guess the likely direction of the kick, and dive in anticipation, if they are to have a chance of saving the shot [4–6]. We examined all 361 kicks from the 37 penalty shootouts that occurred in World Cup and Euro Cup matches over a 36-year period from 1976 to 2012 and show that goalkeepers displayed a clear sequential bias. Following repeated kicks in the same direction, goalkeepers became increasingly likely to dive in the opposite direction on the next kick. Surprisingly, kickers failed to exploit these goalkeeper biases. Our findings highlight the importance of monitoring and predicting sequential behavior in real-world competition. Penalty shootouts pit one goalkeeper against several kickers in rapid succession. Asymmetries in the cognitive capacities of an individual versus a group could produce significant advantages over opponents.

Results and Discussion

Previous game-theoretic models have viewed football penalty kicks as a zero-sum, simultaneous-move game. Analyses of penalty kicks during match play have suggested that goalkeepers and kickers both use a mixed strategy, randomly choosing between left and right dives or kicks [7–9]. In this case, the penalty kick situation resembles a mixed-strategy Nash equilibrium. Crucially, this equilibrium can hold only if both goalkeepers' and kickers' choices are serially independent and uncorrelated, because any predictable pattern of sequential behavior could be exploited by the other party. Remarkably, choices in match-play penalty kicks were indeed found to be random and serially independent [7, 8], even though human attempts to "act randomly" are generally poor [10, 11]. However, penalty kicks in match play are infrequent and are generally separated by long periods of time, so serial independence may be unsurprising.

*Correspondence: e.misirlisoy.11@ucl.ac.uk (E.M.), p.haggard@ucl.ac.uk (P.H.)

In contrast, the penalty shootout is based on a series of multiple kicks taken in rapid succession, after normal match play has failed to produce a clear winner. Each party could easily monitor for sequential regularities in the other's behavior and could exploit any departure from randomness. The goalkeeper might anticipate the direction of the kick, for example by second-guessing the kicker's behavioral choices [12, 13]. In particular, sequences of human choices exhibit a "gambler's fallacy" [10, 14]. This involves the belief that in a series of independent random binary events, such as a coin toss, the alternative event becomes increasingly likely to occur after progressively longer runs of one outcome. Following repeated kicks of the ball in the same direction, a goalkeeper subject to the gambler's fallacy will anticipate the ball going in the opposite direction on the next kick and will therefore dive accordingly. Importantly, the gambler's fallacy is not based on the value of any single previous outcome, but on the run length over which an outcome has already occurred. If goalkeepers exhibit this regularity, kickers could potentially exploit it to gain advantage. Previous game-theoretic studies mostly considered match-play penalty kicks [7]. Although some examined penalty shootouts [15], they focused primarily on motivational factors predicting performance, rather than on runs of decisions by individual players.

We examined online videos and statistics of all penalty shootouts from FIFA World Cup and UEFA Euro Cup finals tournaments from 1976 to 2012, comprising 361 penalty kicks in 37 shootouts. We recorded the direction each ball was kicked (left/right/center) and also the direction in which the goalkeeper moved. These correspond to action selection decisions by kickers and goalkeepers, respectively. Goalkeepers remained in the center only very rarely (2.49%), and kicks to the center were also rare (9.14%). Moreover, we had no strong prior assumptions about sequential effects on center choices, so we removed them, leaving a data set of 321 penalty kicks involving clear left/right choices.

Goalkeepers were approximately equally likely to dive left or right, showing no significant difference from 50% (46.73% left; $p = 0.27$, binomial test). Similarly, kickers were equally likely to shoot left or right (46.42% left; $p = 0.22$, binomial test). A chi-square test showed no significant difference between these proportions ($p = 0.94$). Importantly, goalkeepers were no more likely to dive in the direction of the kick than would be expected by chance (53.58%; $p = 0.22$, binomial distribution). This suggests that goalkeepers are indeed unable to react to kicker behavior and endogenously choose each dive direction.

We next investigated possible exploitable regularities, by analyzing effects of repeated ball direction across consecutive kicks on goalkeeper behavior and determining whether goalkeepers showed a gambler's fallacy. We also examined whether sequences of kick direction did indeed depend on previous kick history. We thus measured the direction of the kick, and of the goalkeeper's dive, after runs of one ($n = 159$), two ($n = 66$), or three ($n = 16$) consecutive kicks in the same direction during a single penalty shootout. We also considered the possibility that goalkeepers' previous dive directions, rather than previous kick directions, could predict future behavior (see Figures S1 and S2 available online). However,

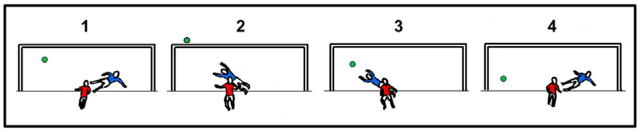

Figure 1. Four Consecutive Shots by Portuguese Kickers against the English Goalkeeper during the Euro 2004 Semifinal Penalty Shootout

The Portuguese kickers are shown in red and the English goalkeeper in blue; the green circle is the ball. All four kicks were to the left of the goal. The goalkeeper made a rightward dive following three consecutive kicks to the left.

we found no evidence that dive history was relevant for kickers or goalkeepers: kick direction was more important than dive direction. Therefore, we present here the results based on runs of kick directions. We used bootstrap resampling [16, 17] to compare the actual data to the same data randomly shuffled 10,000 times. We tested whether goalkeepers and kickers were more likely than chance to switch direction in their behavior following long repetitions of kicks in the same direction.

When kickers repeatedly kicked in the same direction, goalkeepers became progressively more likely to dive in the opposite direction on the next kick, confirming a gambler's fallacy. Figure 1 shows an example in which the goalkeeper dived to the right after three successive kicks to the left, thus failing to save the ball that the fourth kicker again directed to the left.

Figure 2 shows overall goalkeeper data (red line) for percentage of dives in the opposite direction to the last kick. This percentage increased monotonically as run length increased and then crossed the bootstrapped 95% confidence interval after three kicks in the same direction. Goalkeepers' action choices therefore show a pattern similar to the gambler's fallacy [10, 14].

Figure 3 indicates that kickers showed a rather different sequential pattern. Kickers showed a trend to switch to a new direction after a run of two kicks in the same direction ($p = 0.07$). Importantly, this trend did not escalate further as run lengths became longer: kickers were no more likely to switch directions after a run of three than would be expected by chance. Overall, kickers seem to show less predictable behavioral sequences than goalkeepers, with a less obvious gambler's fallacy pattern. One important reason for this difference could be that the kicks in penalty shootouts are generated by multiple agents, who act relatively independently. In contrast, the behavioral sequence of dives is generated by the goalkeeper alone, so that cognitive limitations on an individual's random generation figure prominently in his action decisions.

The goalkeeper's choices could arise for two reasons. They may believe that the sequence of kick directions is genuinely random and simply be subject to the pervasive gambler's fallacy [10, 14]. Alternatively, the goalkeeper's behavior could reflect a "cognitive hierarchy": the goalkeeper may believe that the *kickers* will display a gambler's fallacy [18], though in fact kickers do not. In any case, goalkeepers appear to behave nonoptimally, since they show a sequential bias not aligned with kicker behavior. The goalkeepers' sequential bias was not functional, since the dive direction that they chose after runs of three repeated kick directions was no more likely than chance to coincide with the actual direction of the kick on that attempt (same direction 56.25%; $p = 0.79$, binomial distribution).

Analyses of gambler's fallacy assume that incrementing runs of repeated events affect behavioral choices [19, 20]. However, other forms of serial dependence could also occur. Lagged regression offers a general framework for considering independent contributions of multiple previous events in predicting the current state. We therefore also applied discrete-choice regression for unbalanced panel data to our data set. This analysis approach has been popular in econometrics (e.g., [21]) and also in behavioral economics, including game-theoretic analyses of tournaments [8, 15]. We investigated whether current dive direction depended on kicking direction, and on dive direction one to three kicks previously (see Supplemental Experimental Procedures for data, analysis, and summary output). First, we used a random-effects model, as recommended for short panels [8, 15]. The overall model did not provide significant evidence for serial dependence ($F_{6,95} = 1.63$, $p = 0.15$), consistent with a previous report [7]. Interestingly, the partial coefficients nevertheless showed a significant negative relation between current dive direction and the kick direction three kicks previously ($p = 0.04$), after the contribution of other events was taken into account. We also fitted an additional model using first-difference estimators; these have been preferred because they can be shown to provide unbiased and consistent modeling that can account for unobserved heterogeneity [22]. The overall first-difference model was significant ($p < 0.001$). There was again a significant negative effect of the kick direction three kicks before the current dive ($p = 0.03$). However the goalkeeper's behavior now showed a negative relation to their own previous dive direction at lags one to three (all $p < 0.01$). Recent studies have pointed to additional factors that may affect penalty shootout scores as a result of their motivational significance, notably kicking order (first versus second kicking team in the shootout) and difference between the two teams' scores immediately before each kick [15]. Adding these motivational predictors did not change the pattern of significance for other terms in either regression model.

Such lagged regressions do not consider effects of repetitions within sequences and therefore provide a different perspective from our runs analysis. However, the results are in broad agreement with the runs analysis, in showing a serial dependence on kick direction three kicks before the current dive. They also suggest that goalkeepers may have a general tendency to avoid repetitions in their sequential diving direction behavior, consistent with human behavior during random sequence generation [23].

Kickers tend to use the inside of their dominant foot to make contact with the ball during a penalty kick. This means that right-footed and left-footed players find it easier to kick to their left and right, respectively [7, 8]. A goalkeeper may use information about the kicker's "natural side" in deciding which way to dive. In our data set, goalkeepers dived to the kicker's natural side 178 times, against 143 dives to the nonnatural side ($p = 0.06$, binomial test), suggesting that this information may be important. We therefore repeated our analyses after recoding all kick directions as "natural" or "nonnatural" for that kicker. Interestingly, this analysis showed that goalkeepers were significantly more likely than chance to dive in the opposite direction to the last kick after a run of one repeated kick direction ($p = 0.03$) (i.e., they more likely to dive to the nonnatural side of kicker 2 if kicker 1 had just kicked to their natural side, and vice versa). However, goalkeepers did not use information about natural kicking direction for longer runs. In fact, the probability of switching dive direction relative to current kicker preference (i.e., natural side) did not increase with run length (see Figure S3). Goalkeepers therefore may use information about the kicker's preferences at the start of a run, but they

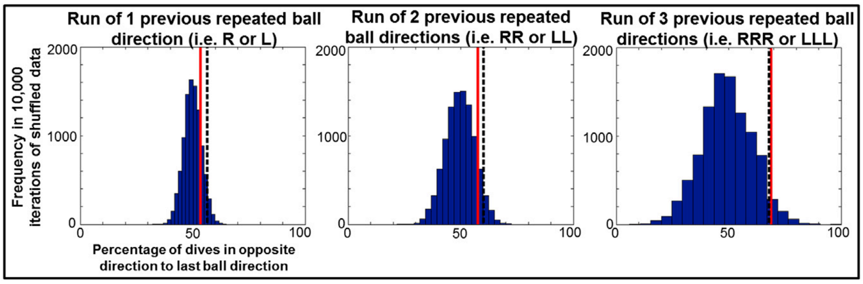

Figure 2. Percentage of Goalkeeper Dives in the Opposite Direction to the Last Ball Direction following Runs of One, Two, or Three Repeated Kicks in the Same Direction

Blue bars indicate 10,000 iterations of random data across 20 bins. Dashed black lines are one-tailed 95% confidence intervals of the shuffled distribution. Red lines indicate actual goalkeeper data.

do not appear to keep track of kicker-specific preferences over extended sequences. To do so, the goalkeeper would need reliable information about each preceding kicker's preferred foot and the corresponding actual ball direction. This information would rapidly approach memory capacity. Instead, goalkeepers initially take kicker preference into account in choosing dive direction but thereafter simply track sequences of kicks to left and right. The gambler's fallacy effect identified in our main analysis is therefore based on accumulating information about left and right ball directions without comprehensive modeling of kicker preference.

The gambler's fallacy is well known in both game theory and cognitive psychology. Perruchet and colleagues link the gambler's fallacy to conscious expectancy about the likely pattern of events [19], while others suggest that the gambler's fallacy arises because of a cognitive default based on sampling without replacement [20]. Paradoxically, in simple laboratory reaction tasks, the conscious expectancy that a long run will shortly end seems to coexist with increasingly fast reaction times [19]. Extending this idea to penalty shootouts, the goalkeeper shown in Figure 1 might have dived very rapidly, and thus had a particularly high chance of saving the fourth kick, if the kick had in fact gone in the direction of his dive.

Previous game-theoretic studies of elite sports have not specifically considered the gambler's fallacy but have investigated serial dependence of choices in general. In-match penalty kicks did not show serial dependence [7, 8]. Kovash and Levitt [24] found clear negative serial correlations in baseball pitches and NFL football passes, though their data sets were much larger than that used here. The gambler's fallacy is conventionally measured as increasing perceptual expectancy of a novel event with increasing run length [19]. Kovash and Levitt instead reported lag-one correlations in *production* of events. Nevertheless, their data are consistent with the broad view that the gambler's fallacy may be an important factor limiting minimax play (i.e. an optimum strategy that minimizes worst possible losses) in elite sports [7].

Since goalkeepers are carefully selected and highly trained, one may ask why their vulnerability of regular behavioral sequences persists. Kickers may simply fail to detect, or otherwise fail to exploit, the fundamental limitation in goalkeepers' random generation, which means that the vulnerability can continue without being penalized. Goalkeepers may not be aware of the vulnerability, perhaps because it has not been prioritized in selection and training. In fact, penalty shootouts are relatively rare—our sample contained 168 matches that could potentially have produced a penalty shootout, but only 37 were eventually settled in this way. Moreover, in-match penalty kicks do *not* show regular patterns of kicking direction [7, 8]. Therefore, the occurrence and implications of the gambler's fallacy in penalty shootouts may have gone unnoticed.

The penalty shootout pits the will of one goalkeeper against the will of many kickers. The cognitive functions of monitoring past performance and generating novel actions are highly relevant to both parties. These cognitive functions have been extensively studied in individuals, but little is known about how they are coordinated across individuals within a group [25]. We suggest that the goalkeeper's and kickers' performance can be understood by considering how asymmetric competition between a single individual (the goalkeeper) and a group of individuals (the kickers) depends on behavior-monitoring and behavior-generating processes. For example, goalkeepers may have complete autobiographical memory of the sequence of kick directions, and of their own diving behavior, because they face every kick. In contrast, kickers have autobiographical memory for their own kick only and may rely on less-direct experience for their teammates' kicks. Kickers may therefore fail to monitor the goalkeeper's diving behavior. They may then fail to detect and exploit regular patterns of goalkeeper dives.

The goalkeeper should have a monitoring advantage over kickers, since self-related actions and material are better retained than other classes of material [26, 27]. However, goalkeepers have the disadvantage of producing predictable sequential behavior, due to limitations in generating random sequences. In contrast, the group of kickers may collectively produce a more random sequence than any single individual, because each kicker chooses their kick direction only once. Distributed cognition across the group of kickers may increase randomness. Similarly, other animals can produce stochastic or random behavior at the group level [28, 29]. Thus, in summary, the individual goalkeeper has an advantage in the monitoring function but a disadvantage in random generation. The group of kickers, in contrast, has a disadvantage in monitoring, but an advantage in random generation.

Game-theoretic analyses have shown the value of unpredictable, mixed strategies in many games [30], and specifically in football penalty kicks during match play [7]. Experimental studies confirm that departures from randomness can be readily exploited [31, 32]. Penalty shootouts involve a unique combination of a simple, competitive decision-making game and a cognitive asymmetry between the individual goalkeeper and the group of kickers. There has been relatively little research on games in which the same individual competes iteratively against a set of several individuals, each of whom takes a turn, since most studies focus on two-player games. However, cognitive hierarchy models [18] suggest that players try to anticipate opponents' decisions but fail to take into account the possibility that their opponents may also be doing this as well as they are, or even better. Importantly, this strategic anticipation is known to be constrained by working memory capacity: the typical number of strategic thinking steps

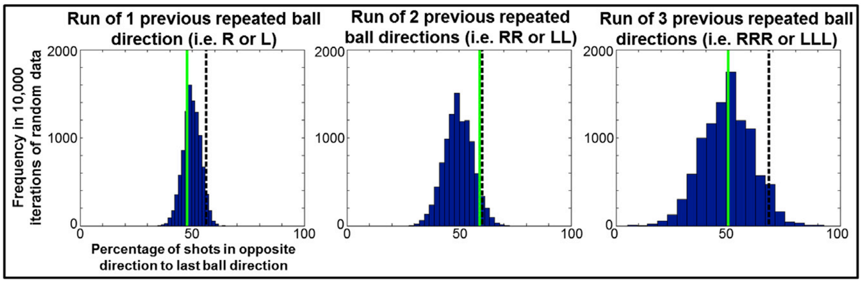

Figure 3. Percentage of Kicker Shots in the Opposite Direction to the Last Ball Direction following Runs of One, Two, or Three Repeated Kicks in the Same Direction

Blue bars indicate 10,000 iterations of random data across 20 bins. Dashed black lines are one-tailed 95% confidence intervals of the shuffled distribution. Green lines indicate actual kicker data.

about opponents' likely decisions is around 1.5 [18]. Our data show that the kickers collectively lack the degree of strategic thinking required to exploit goalkeeper behavior. Groups that communicate freely generally reach better perceptual decisions than isolated individuals [33], though it remains unclear whether the same group advantage applies in strategic, competitive tasks involving random generation. Interestingly, kickers in penalty shootouts typically act as a series of individuals, with little intercommunication among the group during the shootout. Better coordination of the kickers' collective cognitive capacities might in principle give them a strategic advantage over the goalkeeper. Paradoxically, our results suggest that the kickers should not monitor the goalkeeper's dive direction sequences. Rather, they should monitor previous kick directions and use these to simulate the goalkeeper's predictions, along the lines of the cognitive hierarchy model. This is because the goalkeeper's gamblers' fallacy is driven by previous kick directions. Thus, groups of kickers should deploy a triad of advanced cognitive functions to exploit goalkeeper bias: working memory for monitoring kick directions, theory of mind for estimating the goalkeeper's likely dive direction, and communication to pool the group's cognitive capacities.

In the penalty shootout, both fairness and success may depend on the balance between the individual goalkeeper's ability in random generation and kickers' collective ability to predict the goalkeeper's intention by monitoring sequential behavioral patterns. Our research might motivate cognitive training in random generation for goalkeepers, and in intercommunicative monitoring and prediction for kickers. These could help teams better prepare for nail-biting penalty shootouts in future football tournaments.

Supplemental Information

Supplemental Information includes four figures, Supplemental Experimental Procedures, and a data set used for lagged regression analyses and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2014.07.013.

Acknowledgments

This work was supported by the Economic and Social Research Council (ESRC, grant RES-062-23-2183). P.H. was additionally supported by an ESRC Professorial Fellowship and by European Research Council Advanced Grant HUMVOL (Human Volition, Agency and Responsibility). We would additionally like to thank Professor Nick Chater, Dr. Nobuhiro Hagura, Dr. Max-Philipp Stenner, and three anonymous reviewers for critical comments and revisions to the manuscript.

Received: March 18, 2014
Revised: June 5, 2014
Accepted: July 4, 2014
Published: July 31, 2014

References

1. Yarrow, K., Brown, P., and Krakauer, J.W. (2009). Inside the brain of an elite athlete: the neural processes that support high achievement in sports. Nat. Rev. Neurosci. *10*, 585–596.
2. BBC News (2004). Record audience for England match. June 25, 2004. http://news.bbc.co.uk/2/hi/entertainment/3839067.stm.
3. Land, M.F., and McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. Nat. Neurosci. *3*, 1340–1345.
4. Savelsbergh, G.J.P., Williams, A.M., Van der Kamp, J., and Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. J. Sports Sci. *20*, 279–287.
5. Savelsbergh, G.J.P., Van der Kamp, J., Williams, A.M., and Ward, P. (2005). Anticipation and visual search behaviour in expert soccer goalkeepers. Ergonomics *48*, 1686–1697.
6. McMorris, T., and Hauxwell, B. (1997). Improving anticipation of goalkeepers using video observation. In Science and Football III, T. Reilly, J. Bangsbo, and M. Hughes, eds. (London: Taylor & Francis), pp. 290–294.
7. Palacios-Huerta, I. (2003). Professionals play minimax. Rev. Econ. Stud. *70*, 395–415.
8. Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. Am. Econ. Rev. *92*, 1138–1151.
9. Azar, O.H., and Bar-Eli, M. (2011). Do soccer players play the mixed-strategy Nash equilibrium? Appl. Econ. *43*, 3591–3601.
10. Tune, G.S. (1964). Response preferences: a review of some relevant literature. Psychol. Bull. *61*, 286–302.
11. Wagenaar, W.A. (1972). Generation of random sequences by human subjects: A critical survey of literature. Psychol. Bull. *77*, 65–72.
12. Gallagher, H.L., and Frith, C.D. (2003). Functional imaging of "theory of mind". Trends Cogn. Sci. *7*, 77–83.
13. Siegal, M., and Varley, R. (2002). Neural systems involved in "theory of mind". Nat. Rev. Neurosci. *3*, 463–471.
14. Tversky, A., and Kahneman, D. (1971). Belief in the law of small numbers. Psychol. Bull. *76*, 105–110.
15. Apesteguia, J., and Palacios-Huerta, I. (2010). Psychological pressure in competitive environments: Evidence from a randomized natural experiment. Am. Econ. Rev. *100*, 2548–2564.
16. Efron, B., and Tibshirani, R.J. (1994). An Introduction to the Bootstrap (Boca Raton: CRC Press).
17. Wasserman, S., and Bockenholt, U. (1989). Bootstrapping: applications to psychophysiology. Psychophysiology *26*, 208–221.
18. Camerer, C.F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. Q. J. Econ. *119*, 861–898.
19. Perruchet, P., Cleeremans, A., and Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. J. Exp. Psychol. Learn. Mem. Cogn. *32*, 955–965.
20. Ayton, P., and Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? Mem. Cognit. *32*, 1369–1378.
21. Arellano, M., and Honoré, B. (2001). Chapter 53. Panel data models: some recent developments. In Handbook of Econometrics Volume 5, J.J. Heckman and E. Leamer, eds. (Amsterdam: Elsevier), pp. 3229–3296. http://www.sciencedirect.com/science/article/pii/S1573441201050061.
22. Wooldridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data (Cambridge: MIT Press).
23. Spatt, J., and Goldenberg, G. (1993). Components of random generation by normal subjects and patients with dysexecutive syndrome. Brain Cogn. *23*, 231–242.

24. Kovash, K., and Levitt, S. (2009). Professionals Do Not Play Minimax: Evidence from Major League Baseball and the National Football League, NBER Working Paper 15347 (Cambridge: National Bureau of Economic Research). http://www.nber.org/papers/w15347.
25. Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. Nat. Rev. Neurosci. *4*, 165–178.
26. Symons, C.S., and Johnson, B.T. (1997). The self-reference effect in memory: a meta-analysis. Psychol. Bull. *121*, 371–394.
27. Macrae, C.N., Moran, J.M., Heatherton, T.F., Banfield, J.F., and Kelley, W.M. (2004). Medial prefrontal activity predicts memory for self. Cereb. Cortex *14*, 647–654.
28. Bednekoff, P.A., and Lima, S.L. (1998). Randomness, chaos and confusion in the study of antipredator vigilance. Trends Ecol. Evol. *13*, 284–287.
29. Perony, N., Tessone, C.J., König, B., and Schweitzer, F. (2012). How random is social behaviour? Disentangling social complexity through the study of a wild house mouse population. PLoS Comput. Biol. *8*, e1002786.
30. Sato, Y., Akiyama, E., and Farmer, J.D. (2002). Chaos in learning a simple two-person game. Proc. Natl. Acad. Sci. USA *99*, 4748–4751.
31. Lee, D., Conroy, M.L., McGreevy, B.P., and Barraclough, D.J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. Brain Res. Cogn. Brain Res. *22*, 45–58.
32. Vickery, T.J., Chun, M.M., and Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron *72*, 166–177.
33. Bahrami, B., Olsen, K., Latham, P.E., Roepstorff, A., Rees, G., and Frith, C.D. (2010). Optimally interacting minds. Science *329*, 1081–1085.