3. Møller, A. P. & Pomiankowski, A. Why have birds got multiple sexual ornaments. *Behav. Ecol. Sociobiol.* **32,** 167–176 (1993).

4. Johnstone, R. A. Multiple displays in animal communication: 'backup signals' and 'multiple messages'. *Phil. Trans. R. Soc. Lond. B* **351,** 329–338 (1996).

5. Brooks, R. & Endler, J. A. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution* **55,** 1644–1655 (2001).

6. Kodric-Brown, A. & Nicoletto, P. F. Age and experience affect female choice in the guppy (*Poecilia reticulata*). *Am. Nat.* **157,** 316–323 (2001).

7. Moore, P. J. & Moore, A. J. Reproductive aging and mating: The ticking biological clock in female cockroaches. *Proc. Natl Acad. Sci. USA* **98,** 9171–9176 (2001).

8. Borgia, G. Bower quality, number of decorations and male mating success of male satin bowerbirds (*Ptilonorhynchus violaceus*): an experimental analysis. *Anim. Behav.* **33,** 266–271 (1985).

9. Patricelli, G. L., Uy, J. A. C., Walsh, G. & Borgia, G. Sexual selection: male displays adjusted to female's response. *Nature* **415,** 279–280 (2002).

10. Patricelli, G. L., Uy, J. A. C. & Borgia, G. Female signals enhance the efficiency of mate assessment in satin bowerbirds (*Ptilonorhynchus violaceus*). *Behav. Ecol.* **15,** 297–304 (2004).

11. Uy, J. A. C., Patricelli, G. L. & Borgia, G. Complex mate searching in the satin bowerbird (*Ptilonorhynchus violaceus*). *Am. Nat.* **158,** 530–542 (2001).

12. Uy, J. A. C., Patricelli, G. L. & Borgia, G. Dynamic mate searching tactic allows female satin bowerbirds (*Ptilonorhynchus violaceus*) to reduce searching. *Proc. R. Soc. Lond. B* **267,** 251–256 (2000).

13. Uy, J. A. C., Patricelli, G. L. & Borgia, G. Loss of preferred mate forces female satin bowerbirds (*Ptilonorhynchus violaceus*) to increase mate searching. *Proc. R. Soc. Lond. B* **268,** 633–638 (2001).

14. Widemo, F. & Sæther, S. A. Beauty is in the eye of the beholder: causes and consequences of variation in mating preferences. *Trends Ecol. Evol.* **14,** 26–31 (1999).

15. Jennions, M. D. & Petrie, M. Variation in mate choice and mating preferences: a review of causes and consequences. *Biol. Rev.* **72,** 283–327 (1997).

16. Gibson, R. M. & Langen, T. A. How do animals choose their mates. *Trends Ecol. Evol.* **11,** 468–470 (1996).

17. Höglund, J. & Alatalo, R. *Leks* (Princeton Univ. Press, Princeton, 1995).

18. Borgia, G. & Gore, M. A. Sexual competition by feather stealing in the satin bowerbird (*Ptilonorhynchus violaceus*). *Anim. Behav.* **34,** 727–738 (1986).

....

Variation in behaviour promotes cooperation in the Prisoner's Dilemma game

John M. McNamara[1]**, Zoltan Barta**[1] **& Alasdair I. Houston**[2]

[1]*Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK*
[2]*School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*

....

The Prisoner's Dilemma game[1–4] **is widely used to investigate how cooperation between unrelated individuals can evolve by natural selection. In this game, each player can either 'cooperate' (invest in a common good) or 'defect' (exploit the other's investment). If the opponent cooperates, you get R if you cooperate and T if you defect. If the opponent defects, you get S if you cooperate and P if you defect. Here $T > R > 0$ and $P > S$, so that 'defect' is the best response to any action by the opponent. Thus in a single play of the game, each player should defect. In our game, a fixed maximum number of rounds of the Prisoner's Dilemma game is played against the same opponent. A standard argument based on working backwards from the last round**[1,5] **shows that**

defection on all rounds is the only stable outcome. In contrast, we show that if extrinsic factors maintain variation in behaviour, high levels of co-operation are stable. Our results highlight the importance of extrinsic variability in determining the outcome of evolutionary games.

In our version of the Prisoner's Dilemma game, the fixed maximum number of possible rounds, N, is known to both players. On each round before the last one, if either player defects then the game ends; only if both cooperate do they proceed to the next round. (This approach reflects the ability of mobile organisms to terminate an unfavourable interaction by leaving[6,7].) After N rounds, the interaction ends whatever decisions are made. The total payoff is the sum of the payoff from each round.

We assume that $T > (P + R)$. The standard arguments based on working backwards then lead to the conclusion that defection on all rounds is the only evolutionarily stable outcome in our game (see Supplementary Information). At evolutionary stability all population members behave in this way, so that there is no variation in the population, and unlike some iterated Prisoner's Dilemma games, variation cannot be maintained by frequency dependence. We believe, however, that in real populations there are always factors other than frequency dependence—such as mutation, immigration, recombination and epistasis—that maintain genetic variation. We show that maintaining variation fundamentally changes the nature of the game. The intuition behind this is as follows. In a population at the defect evolutionarily stable strategy, a player should defect on the first round. But if variability is maintained, and hence there is a chance that an opponent will cooperate, then there is the potential for a substantial gain, and it may be worth cooperating initially in the hope that the opponent is cooperative (compare ref. 1). Whether for this or for other reasons, humans do not defect as much as expected in the Prisoner's Dilemma and related games[5,8].

In our model, a strategy specifies the number of rounds n to cooperate before defecting. (The game may not last for $n + 1$ rounds because the opponent may terminate the game by defecting beforehand.) We consider the evolution of an infinite population with discrete generations. The number of offspring left by an individual in the next generation is the individual's total payoff plus a positive constant, K, which represents the contributions to fitness that come from outside the game. If a parent adopts strategy n then each offspring is also n with probability $1 - 2\varepsilon$. Genetic variability is maintained by mutation; with probability ε offspring are $n - 1$ and with probability ε they are $n + 1$. We use the standard
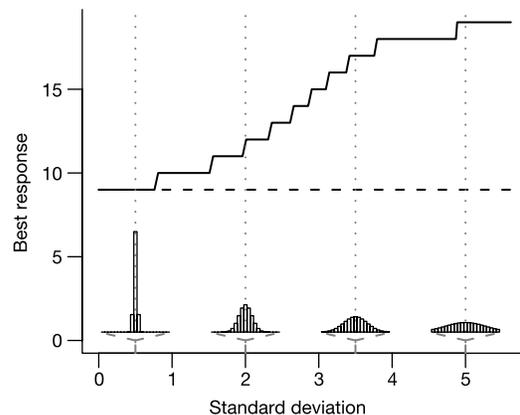
Figure 1 The best response (continuous line) as a function of the variation in the degree of cooperation in the population. In all cases, the mean number of rounds of cooperation before defection is $E(n) = 10$. The distribution of n about this mean is pseudo-normal with the standard deviation indicated. Four actual distributions are illustrated. The dashed line gives the best response when each round is chosen against a new opponent, randomly selected from the population. In these calculations, $N = 20$.

payoffs for the Prisoner's Dilemma game as our baseline, that is, $T = 5$, $R = 3$, $P = 1$, $S = 0$.

If you know that the other player will cooperate for n rounds and defect on the $(n + 1)$th round, then the best thing to do is to pre-empt this defection by cooperating for $n - 1$ rounds and defecting on the nth round (see Supplementary Information). It follows that if all members of the population adopt the strategy n, then the best response (that is, the best value of n) for a mutant is $n - 1$. The situation changes if the population consists of a mixture of individuals each with different strategies. If the population mean is $E(n)$, then the best response can be greater than $E(n)$ when there is variation in the population (see Supplementary Information). This is because some population members cooperate for more than $E(n)$ rounds, and the potential benefit of interacting with such an opponent outweighs the cost incurred if the opponent is the first to defect. Figure 1 shows the best response for a mutant in a population for which $E(n) = 10$. As the variation about the mean increases, so does the best response. This effect is not restricted to the game that we consider; it can occur in any game in which the benefit of interacting with a cooperative opponent is significantly greater than the cost of being exploited by an uncooperative opponent. Box 1 gives an analytic example based on a continuous trait.

Figures 2 and 3 illustrate the evolutionarily stable outcomes when a population evolves from an initial population that consists entirely of animals that defect on the first round. (Computations show that the same stable outcome is reached from any initial population.) Fig. 2 shows that the outcome depends on the

mutation rate ε. If ε is below a critical value, then the population is essentially non-cooperative, with a small amount of cooperation maintained by mutation. In contrast, if ε is above a critical value, then the population is essentially cooperative, but with some variation. In this case the unique best response to the population is more cooperative than the mean level of cooperation. Cooperation does not, however, increase further because of biased mutation at the upper limit of possible cooperation. As Fig. 3 shows, the sharp transition from a non-cooperative solution to a cooperative solution as ε increases is a general feature of our model.

Figure 3a shows the effect of the parameter K. The bigger the background fitness contribution K, the less important is the payoff from the game in determining an individual's lifetime fitness, and hence the more genetic variation there is likely to be in a population for a given value of ε. Thus when K is large, the transition to cooperation occurs at a lower value of ε. This means that we predict that cooperation becomes less likely to evolve as the game payoff constitutes a greater proportion of the lifetime reproductive success of the organism. Figure 3b shows the effect of the maximum possible number of rounds, N. The transition to cooperation is almost independent of N, provided that $N > 2$. Sensitivity analysis reveals that our conclusions are robust to changes in the values of the payoffs R, P, S and T.

We have illustrated our argument using mutation as a source of genetic variation, but similar effects are found if phenotypic variation is produced by errors in decision-making or developmental noise (see Fig. 2b). In natural populations, variation is likely to be maintained by multiple factors.

Box 1
Mutation rate determines the direction of evolution

We consider a large (infinite) population in which individuals differ in their cooperativeness—a continuous trait. Population members meet at random and engage in a contest. The payoff to an individual with cooperativeness x in a contest with an opponent with cooperativeness y is:

$$w(x,y) = \exp\left\{ \alpha y - \frac{1}{2}(y - 1 - x)^2 \right\}$$

This is maximized when the individual is less cooperative than the opponent by one unit. The other key feature is that simultaneously increasing the cooperativeness of the focal individual and the opponent, while maintaining a constant difference between the two, increases the focal individual's payoff. The rate of increase increases with the parameter α.

The fitness of an individual with cooperativeness x, $W(x)$, is the average of $w(x, y)$ across the cooperativeness y of population members. Assuming that cooperativeness is normally distributed with mean μ and variance σ^2 in the population, it can be shown that:

$$W(x) \propto \exp\{ - [(x + 1) - (\mu + \alpha\sigma^2)]^2 / 2(\sigma^2 + 1)\}$$

This is maximized when $x = x^*$, where $x^* = \mu - 1 + \alpha\sigma^2$.

Thus if all population members have cooperativeness $y = \mu$ (that is, $\sigma^2 = 0$), fitness is maximized when $x^* = \mu - 1$. However, if the population variance σ^2 exceeds $1/\alpha$, x^* exceeds the population mean μ.

Evolution is modelled by discrete-generation replicator dynamics, with error. An individual of cooperativeness x leaves $W(x)$ offspring in the next generation. The distribution of offspring cooperativeness is normal, with mean x and variance β^2. Then, given that cooperativeness is normally distributed in generation t, it is also normally distributed in generation $t + 1$, with successive means and variances related by $\mu_{t+1} = \mu_t + v_t(\alpha\sigma_t^2 - 1)$ and $\sigma_{t+1}^2 = \beta^2 + v_t(\sigma_t^2 + 1)$, where $v_t = \sigma_t^2/(2\sigma_t^2 + 1)$. Over time, the variance tends to the limiting value $\sigma_\infty^2 = \beta^2 + \beta\sqrt{\beta^2 + 1}$. At this limit:

$$\mu_{t+1} > \mu_t \Leftrightarrow \beta^2 > \frac{1}{\alpha(\alpha + 2)}$$

Figure 2 The proportion of each type of individual (as specified by n) in the population at evolutionary stability. **a**, Results from the model presented in the text. Two cases, corresponding to two values of the mutation parameter ε, are shown. **b**, Results when there is also a distribution of phenotypes (with standard deviation D) for a given genotype. Results are presented for four combinations of ε and D. For each value of D, the corresponding value of ε is the minimum at which cooperation occurs. $N = 10$ and $K = 20$ throughout.

Maintaining a small number of defecting individuals in a population playing the iterated Prisoner's Dilemma can influence the evolutionary outcome[9,10]. In the examples of refs 9 and 10, a population in which individuals cooperate is not evolutionarily stable in a strict sense because several strategies do equally well and their proportions can change by drift. Introducing small numbers of defectors into the population removes this neutrality, preventing the drift towards strategies that could be exploited by defectors. In contrast, our model does not involve neutrality. When a population is at the equilibrium where all individuals defect immediately, any other strategy is strictly worse. Introducing variation changes the fitness landscape so that strategies that initially cooperate are now on the other side of a valley and do better than the strategy of always defecting (Fig. 1). At the stable equilibrium with cooperation (Figs 2 and 3) there is still no neutrality. Instead, there is a unique best response to the population that is more cooperative than the mean level of cooperation.

Our version of the Prisoner's Dilemma game has a similar structure to the centipede game[5,8], in that non-cooperation ends the game and cooperating with an opponent that cooperates leads to a high payoff. As in our game, non-cooperation on the first round is the only stable outcome. It is well documented, however, that humans do not adopt this strategy[8,11]. Models that attempt to account for this behaviour include a certain fraction of completely cooperative individuals in the population[8,11]. Unlike our model, this fraction is imposed by the modeller rather than emerging from the evolutionary dynamics. It is clear that applying our approach to this game would lead to the evolution of high levels of cooperation even if the mutation rates are very low.

Sources of variation in models of cooperation include stochastic strategies, errors in decision making, mutations and differences in quality[9,12–23]. All these sources of variation may favour cooperation.

When variability is maintained by mutation, as in our model, not all strategies present at evolutionary stability maximize fitness. This means that the outcome of evolution cannot be found by the standard approach in which the endpoint is a best response to itself.

The strategies that are used to play the Prisoner's Dilemma game are often characterized in terms of personality traits (for example, 'Tit-for-tat' is said[24] to be nice, can be provoked but forgives). In our game, players differ in a single aspect of personality: their degree of cooperation. Because a series of rounds is played against the same individual, a player can learn about the opponent's persistence in cooperating. If the opponent has cooperated for n rounds, the probability that it will cooperate on round $n + 1$ is greater than the probability that a randomly chosen population member will cooperate on round $n + 1$. Repeated play against the same opponent is crucial for the evolution of cooperation in our model. If the opponent on each round is chosen anew, then cooperation will not evolve (see Fig. 1).

A possible explanation for the diversity of personality types within social groups[7,25] is that each type has equal fitness and variation is maintained by frequency dependence (compare alternative male mating strategies[26,27], producers and scroungers[28]). Another possibility is that the success of the group is crucial[28–30], and this success depends on the range of types. In contrast, our approach assumes that variability is maintained by various genetic and environmental factors. As a result, a population will always have a range of personalities. Our argument is that an adequate model of the evolution of not just cooperation, but social behaviour in general, should take this into account. Our results show that the outcome may be very different from the outcome when variability is not maintained by extrinsic factors. □
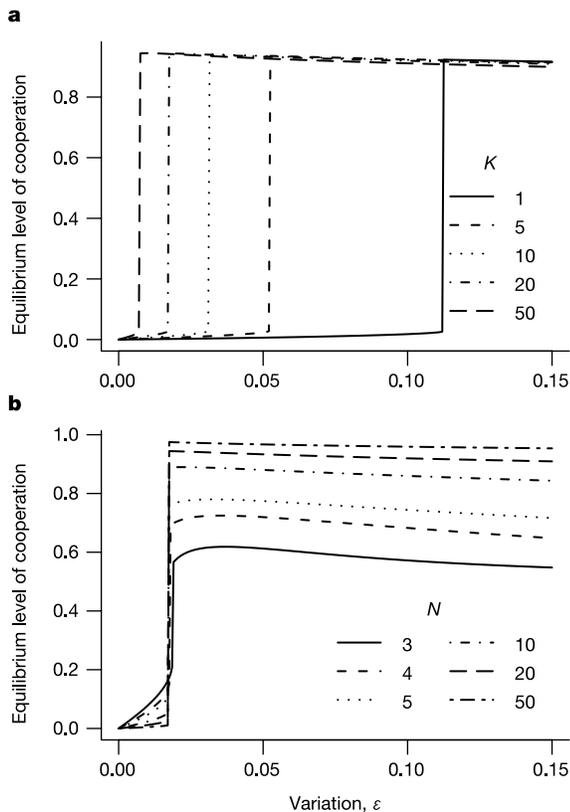
1. Luce, R. D. & Raiffa, H. *Games and Decisions* (Wiley, New York, 1957).
2. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211,** 1390–1396 (1981).
3. Axelrod, R. & Dion, D. The further evolution of cooperation. *Science* **242,** 1385–1390 (1988).
4. Brembs, B. Chaos, cheating and cooperation: Potential solutions to the Prisoner's Dilemma. *Oikos* **76,** 14–24 (1996).
5. Colman, A. M. Cooperation, psychological game theory, and the limitations of rationality in social interaction. *Behav. Brain Sci.* **26,** 139–198 (2003).
6. Enquist, M. & Leimar, O. The evolution of cooperation in mobile organisms. *Anim. Behav.* **45,** 747–757 (1993).
7. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425,** 785–791 (2003).
8. Camerer, C. *Behavioral Game Theory* (Princeton Univ. Press, Princeton, 2003).
9. Crowley, P. H. & Sargent, R. C. Whence tit-for-tat? *Evol. Ecol.* **10,** 499–516 (1996).
10. Sherratt, T. N. & Roberts, G. The importance of phenotypic defectors in stabilizing reciprocal altruism. *Behav. Ecol.* **12,** 313–317 (2001).
11. McKelvey, R. D. & Palfrey, T. R. An experimental study of the centipede game. *Econometrica* **60,** 803–836 (1992).
12. Boyd, R. Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game. *J. Theor. Biol.* **136,** 47–56 (1989).
13. Fudenberg, D. & Maskin, E. Evolution and cooperation in noisy repeated games. *Am. Econ. Rev.* **80,** 274–279 (1990).
14. Nowak, M. & Sigmund, K. The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Appl. Math.* **20,** 247–265 (1990).
15. Bendor, J., Kramer, R. M. & Stout, S. When in doubt. Cooperation in a noisy prisoner's dilemma. *J. Conflict Resolut.* **35,** 691–719 (1991).
16. Nowak, M. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355,** 250–252 (1992).
17. Nowak, M. & Sigmund, K. Chaos and the evolution of cooperation. *Proc. Natl Acad. Sci. USA* **90,** 5091–5094 (1993).
18. Nowak, M. & Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* **364,** 56–58 (1993).
19. Nowak, M. A., Sigmund, K. & Elsedy, E. Automata, repeated games and noise. *J. Math. Biol.* **33,** 703–722 (1995).
20. Wu, J. Z. & Axelrod, R. How to cope with noise in the iterated prisoner's dilemma. *J. Conflict Resolut.* **39,** 183–189 (1995).
21. Wahl, L. M. & Nowak, M. A. The continuous prisoner's dilemma: II. Linear reactive strategies with noise. *J. Theor. Biol.* **200,** 323–338 (1999).
22. Lorberbaum, J. P., Bohning, D. E., Shastri, A. & Sine, L. E. Are there really no evolutionarily stable strategies in the iterated prisoner's dilemma? *J. Theor. Biol.* **214,** 155–169 (2002).
23. Fishman, M. A., Lotem, A. & Stone, L. Heterogeneity stabilizes reciprocal altruism interactions. *J. Theor. Biol.* **209,** 87–95 (2001).
24. Axelrod, R. More effective choice in the Prisoner's Dilemma. *J. Conflict Resolut.* **24,** 379–403 (1980).
25. Heinsohn, R. & Packer, C. Complex cooperative strategies in group-territorial African lions. *Science* **269,** 1260–1262 (1995).
26. Brockmann, H. J. The evolution of alternative strategies and tactics. *Adv. Stud. Behav.* **30,** 1–51 (2001).

Figure 3 Levels of cooperation in the population, measured as $E(n)/N$, at evolutionary stability, shown as a function of the mutation parameter ε. **a**, Various values of K for the case $N = 20$. **b**, Various values of N for the case $K = 20$.

27. Henson, S. A. & Warner, R. R. Male and female alternative reproductive behaviors in fishes: a new approach using intersexual dynamics. *Annu. Rev. Ecol. Syst.* **28**, 571–592 (1997).

28. Giraldeau, L. A. & Caraco, T. *Social Foraging Theory* (Princeton Univ. Press, Princeton, 2000).

29. Sober, E. & Wilson, D. S. *Unto Others* (Harvard Univ. Press, Cambridge, Massachusetts, 1998).

30. Boyd, R. & Richerson, P. J. Group selection among alternative evolutionarily stable strategies. *J. Theor. Biol.* **145**, 331–342 (1990).

....

Neural activity predicts individual differences in visual working memory capacity

Edward K. Vogel & Maro G. Machizawa

Department of Psychology, University of Oregon, Eugene, Oregon 97403-1227, USA

....

Contrary to our rich phenomenological visual experience, our visual short-term memory system can maintain representations of only three to four objects at any given moment[1,2]. For over a century, the capacity of visual memory has been shown to vary substantially across individuals, ranging from 1.5 to about 5 objects[3–7]. Although numerous studies have recently begun to characterize the neural substrates of visual memory processes[8–12], a neurophysiological index of storage capacity limitations has not yet been established. Here, we provide electrophysiological evidence for lateralized activity in humans that reflects the encoding and maintenance of items in visual memory. The amplitude of this activity is strongly modulated by the number of objects being held in the memory at the time, but approaches a limit asymptotically for arrays that meet or exceed storage capacity. Indeed, the precise limit is determined by each individual's memory capacity, such that the activity from low-capacity individuals reaches this plateau much sooner than that from high-capacity individuals. Consequently, this measure provides a strong neurophysiological predictor of an individual's capacity, allowing the demonstration of a direct relationship between neural activity and memory capacity.

To measure the neural correlates of visual memory capacity, we recorded event-related potentials (ERPs) from normal young adults while they performed a visual memory task. On each trial they were presented with a brief bilateral array of coloured squares and were asked to remember the items in only one hemifield, which was indicated with an arrow (Fig. 1a). Memory was tested one second later with the presentation of a test array that was either identical to the memory array or differed by one colour. Subjects pressed one of two buttons to indicate whether the two arrays were identical or different. We have used variations of this paradigm previously and have found that observers are accurate for array sizes of up to three

to four items, and that performance is not significantly influenced by perceptual or verbal processes[1,3].

In the first experiment, we recorded ERPs to the onset of a four-item memory array so that we could observe the sustained electrophysiological response during the memory retention interval. A few previous ERP studies have observed a sustained response during working memory tasks for foveally presented stimuli, but did not examine lateralized effects[13,14]. In contrast, we took advantage of the primarily contralateral organization of the visual system by presenting lateralized stimuli in each hemifield so that we could measure the spatially specific hemispheric responses to memory arrays that were either contralateral or ipsilateral with respect to electrode position[15,16]. Approximately 200 ms after the onset of the memory array, we found a large negative-going voltage over the hemisphere that was contralateral to the memorized hemifield, and this response persisted throughout the duration of the memory retention interval (Fig. 1b). This response was focused primarily over the posterior parietal and lateral occipital electrode sites and strongly resembled delay activity recorded from individual neurons in monkey visual cortex[12,17].

Numerous processes contribute to visual memory performance, and we sought to determine which aspects of processing are reflected by the contralateral delay activity. Although this effect seems to reflect the maintenance of object representations from the memory array, it is necessary to rule out the possibility that it reflects executive processes[18] involved in performing the task, or even more general processes such as increased effort or arousal[19–21]. In the second experiment, we tested this by varying the number of items in the memory array to establish whether the amplitude is sensitive to the number of representations that are being held in visual memory. Memory arrays in this experiment varied from one to four items in each hemifield (average capacity in this task is normally around three items[3,7]). To compare directly the magnitude of activity across array sizes, we constructed 'difference waves' in which the ipsilateral activity was subtracted from the contralateral activity for each array size, which removes the contribution of any nonspecific, bilateral ERP activity.

As shown in Fig. 2a, the amplitude was highly sensitive to the number of items in the memory array. Indeed, increasing an array
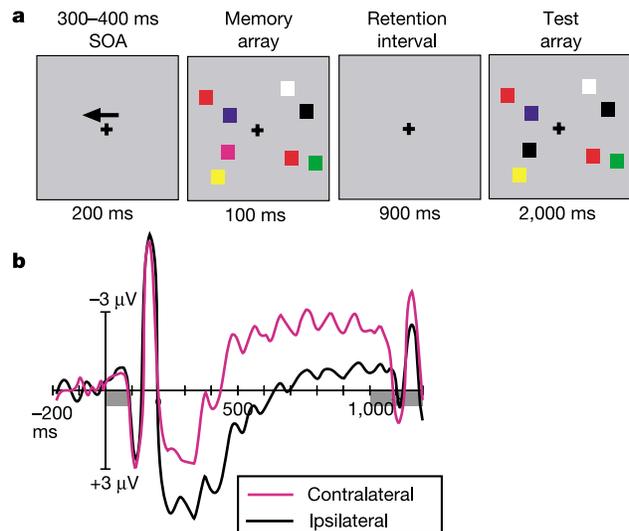
Figure 1 Stimuli and results from experiment one. **a**, Example of a visual memory trial for the left hemifield. SOA, stimulus onset asynchrony. **b**, Grand averaged ERP waveforms time-locked to the memory array averaged across the lateral occipital and posterior parietal electrode sites in experiment one. The two grey rectangles reflect the time periods for the memory and test arrays, respectively. Note that, by convention, negative voltage is plotted upwards.