# The probabilistic approach to human reasoning 

Mike Oaksford and Nick Chater


#### Abstract

A recent development in the cognitive science of reasoning has been the emergence of a probabilistic approach to the behaviour observed on ostensibly logical tasks. According to this approach the errors and biases documented on these tasks occur because people import their everyday uncertain reasoning strategies into the laboratory. Consequently participants' apparently irrational behaviour is the result of comparing it with an inappropriate logical standard. In this article, we contrast the probabilistic approach with other approaches to explaining rationality, and then show how it has been applied to three main areas of logical reasoning: conditional inference, Wason's selection task and syllogistic reasoning.


## Mike Oaksford*

School of Psychology, Cardiff University, PO Box 901, Cardiff, UK CF1 3YG
*e-mail:
oaksford@cardiff.ac.uk

## Nick Chater

Dept of Psychology,
University of Warwick,
Coventry, UK CV2 7AL.

In a standard reasoning task, performance is compared with the inferences peopleshould make according to logic, so a judgement can bemadeon the rational ity of people's reasoning. It has been found that peoplemakelargeand systematic (i.e. nonrandom) errors ${ }^{1,2}$, which suggests that humans might beirrational ${ }^{3,4}$. However, the probabilistic approach argues against this interpretation. Rather than view this behaviour as error proneit is argued that performance has been compared with the wrong normativestandard. When the comparison is with probability theory rather than logic, participants' reasoning is seen in a more positive light. This approach was first applied to Wason's selection task5, wheremany models have been proposed ${ }^{6-11}$ and investigated ${ }^{12-17}$. M ore recently the probabilistic approach has been extended totheother coreareas of thepsychology of reasoning: syllogisms ${ }^{18-20}$ and conditional inference ${ }^{21-24}$, wheretheneed for a probabilistic approach has been most apparent ${ }^{25-29}$.

Wefirst contrast the probabilistic approach with other approaches to human rationality. Wethen discuss theinadequacy of logic as an account of everyday reasoning, and show why a probabilistic approach looks more promising. Wethen review its application to explaining the coretasks in reasoning research (seeabove).

## Rationality and theories of reasoning

Both mental-logic approaches ${ }^{30}$ and mental-model approaches ${ }^{31}$ argue that systematic deviations from logic represent unavoidable performanceerrors. In both approaches limitations in working memory restrict people's reasoning abilities. In mental logic, the length of a formal derivation is the limiting factor. In mental models the number of models (quasi-pictorial mental representations of the meaning of a logical statement) that must beretained sets the limit. According to both approaches people
arerational in principlebut err in practice; that is, people's logical reasoning algorithms are sound but constrained by cognitivelimitations.

Theseapproaches arehard toreconcile with two facts. First, error rates can beas high as $96 \%$ (in Wason's selection task). Second, everyday rationality in guiding thought and action seems to behighly successful. How can this success be understood if peoples' reasoning system is proneto so much error?

Other theorists distinguish twotypes of rationality to resol vethis apparent conflict ${ }^{32}$. Everyday rationality does not depend on formal systems like logic and only formal rationality is constrained and error prone. But how then is thesuccess of everyday inference to beexplained? There would appear to be no obvious alternative, asidefrom arguing that everyday rationality is also based on formal principles of reasoning. But this seems to bring us full circle.

The probabilistic approach resolves this problem. Everyday reasoning is probabilistic and peoplemake errors in so-called logical tasks because they generalizethesestrategies tothelaboratory. This approach has been much influenced by Anderson's account of rational analysis ${ }^{32-36}$. Other authors have also questioned the appropriateness of the normative theories used to assess human reasoning ${ }^{37,38}$. Accordingtotheprobabilistic approach much experimental research in the'psychol ogy of deductive reasoning' does not engage people in deductive reasoning at all but rather engages strategies suitable for everyday reasoning. In thenext section weexplain why everyday reasoning is best characterized probabilistically rather than logically.

## Everyday inference: logic or probability?

Everyday inferences are uncertain. Suppose someone believes that birds fly, it is not clear how to capture the patterns of inference that this everyday generalization permits. For example, standard logic, which mental logic and mental models assume to be normative, will not do because it allows 'strengthening of the antecedent'. Formalizing this generalization as 'if something is a bird it flies', entails that if someone knows that Tweety is a bird, they can infer that Tweety flies. However, strengthening of the antecedent means that when given further information, like Twety is an Ostrich, they should still infer that Tweety flies. But, intuitively, this new piece of information should defeat the previous conclusion: everyday inferences are defeasible. There

## Box 1 Logic and everyday reasoning

Standard logic cannot account for conditionals in everyday inference. An alternative account of conditionals is given by the Lewis-Stalnaker possible-worlds semantics ${ }^{\text {a,b }}$. The intuitive idea for a counterfactual conditional such as 'if Tweety had been a bird, he would have been able to fly' is that in the world maximally similar to the actual world ( $\alpha$ ) but in which Tweety was a bird, he could fly. This semantics can be applied to the indicative conditional (where we simply don't know whether Tweety is a bird). When this is done it is clear that strengthening of the antecedent (see main text) cannot hold. For example, 'if it's a bird, then it flies' does not imply that 'ifit's a bird and it's an ostrich, then it flies'. The worlds in which the antecedents are evaluated will differ - the world most similar to $\alpha$ in which something is a bird is not the same as the world most similar to $\alpha$ in which something is an ostrich. In particular, in the first world, the thing will most probably fly (because most birds fly); but in the second world, the thing will not fly (because ostriches can't fly). However, for psychological purposes we need an account of the formal processes that can implement this semantics. The programme of attempting to mechanize reasoning about the way the world might be, has been taken up by the study of knowledge representation in artificial intelligence (AI). However, it is far from clear that formal attempts in AI can capture the Lewis-Stalnaker semanticsc.

Problems arise when the inferences that can be made from one antecedent intuitively conflict with the inferences that can be made from another. For example, knowing that 'Tweety is a sparrow' leads to the conclusion that 'Tweety flies', whereas knowing that 'Tweety is one second old' leads to the conclusion that 'Tweety cannot fly'. This leads to the problem of what we infer when we learn that 'Tweety is a one-second-old sparrow', that is, when the antecedent is strengthened. It is intuitively obvious that a one-second-old sparrow cannot fly. However, formally, it is not obvious how to capture this conclusion. We can formally regard these two pieces of information as two conditional rules: if something is a bird it can fly, and ifsomething is one second old it cannot fly. Formal proposals in AI appear to be unable to break the symmetry between these rules and specify which of these conflicting conclusions we should accept ${ }^{d}$. There have been various alternative proposals in AI that attempt to deal with this problem of strengthening the antecedent, or default reasoning ${ }^{d-g}$. However, these approaches ${ }^{\text {h-j }}$ all seem to fall foul of similar problems. Moreover, mental logics ${ }^{k}$ and mental models' also fail to address these issues because they formalize the conditional using standard logic.
However, as we have seen, standard logic is unable to capture the use of conditionals in everyday inference.

## References

a Lewis, D. (1973) Counterfactuals, Oxford University Press
b Stalnaker, R. (1968)A theory of conditionals. In Studies in Logical Theory (Rescher, N., ed.), pp. 98-112, Oxford University Press.
c Pylyshyn, Z.W. and Ford, K.M. (1995) TheRobot's Dilemma Revisited, Ablex Publishing
d Reiter, R. (1985) On reasoning by default. In Readings in KnowledgeRepresentation. (Brachman, R. and Levesque, H., eds), pp. 401-410, M organ K aufmann
e McCarthy, J.M. (1980) Circumscription: a form of non-monotonic reasoning. Artif. Intel. 13, 27-39
f McDermott, D. and Doyle, J . (1980) Non-monotonic logicl. Artif. Intell. 13, 41-72
g McDermott, D. (1982) Non-monotonic logic: II . Non-monotonic modal theories. J. Assoc. Comput. Mach. 29, 33-57
h McDermott, D. (1987)A critique of pure reason. Comput. Intell. 3, 151-160
i Shoam, Y. (1987) Reasoning About Change, MIT Press
j Shoam, Y. (1988) Efficient reasoning about rich temporal domains. J . Philos. Logic 17, 443-474
k Rips, L.J . (1994) ThePsychol ogy of Proof, MIT Press
I J ohnson-Laird, P.N. and Byrne, R.M.J . (1991) Deduction, Erlbaum
are a number of logical proposal sto get around this problem (seeBox 1) but noneappears to succeed ${ }^{39-44}$.

Theprobabilistic approach handles this case naturally by interpreting conditionals using conditional probability. Thus, thestatement birdsfly claims that the conditional probability of something flying, given that it is a bird, is high. Probability theory naturally captures defeasibility. If all that is known is that something is a bird, then the probability that it flies might be, say, $0.9[\mathrm{P}$ (flies| bird) $=0.9]$. However, the probability of it flying given that it is a bird and an ostrich is near 0 [ P (flies| bird, ostrich) $\approx 0$ ], and the probability of it flying given that it is a bird and a parrot may be, say $0.96[P(f l i e s \mid$ bird, parrot $)=0.96]$. All thesestatements are compatibleaccordingto probability theory. Sotheresult of strengthening the antecedent leads to acceptableresults. This approach tothemeaning of conditional statements was proposed in philosophy by Adams ${ }^{45,46}$, and has been used in artificial intelligence by Pearl ${ }^{47,48 \text {. }}$

## Everyday inference in the laboratory

Wenow review in detail how the probabilistic approach has been applied in the threemain areas of human reasoning research. We present them in the order of inferential complexity (least first). We concentrateon our own research becausethemodels wehaveproposed all explicitly aim toshow how biases in human reasoning can have a rational basis. Thus, they attempt to resol ve the paradox of why such a successful organism should appear soirrational in thelaboratory.

## Conditional inference

Conditional inferenceinvolves presenting participants with a conditional premise, if $p$ then $q$, and one of four categorical premises, p, not-p, q, or not-q. Logically, given $p$ participants should draw the conclusion q and given not-q they should draw the conclusion not-p. Thesearethelogically valid inferences of modus ponens (MP) and modus tollens (MT) respectively. Moreover, given not-p participants should not draw the conclusion not-q and given q they should not draw the conclusion $p$. These arethe logical fallacies of denying the antecedent (DA) and affirming the consequent (AC) respectively. So, logically, participants should endorseMP and MT equally and they should refusetoendorseDA or AC. However, in fact they endorse MP significantly more than MT and they endorseDA and AC at levels significantly above zero.

Following other researchers in this area ${ }^{25-29}$ we proposed a model of conditional reasoning based on conditional probability ${ }^{21}$. Thegreater the conditional probability of an inferencethemoreit should be endorsed. The meaning of a conditional statement was defined in a two by two contingencytable ${ }^{7,21}$ (Fig. 1).

The contingency tablerepresents a conditional rule, if $p$ then $q$, wherethere is a dependency between thep and q that may admit exceptions ( $\varepsilon$ ). a is the probability of the antecedent, $P(p)$. b is the probability of the

Fig. 1. The contingency table for a conditional rule, if $p$ then $q$, where there is a dependency between the $p$ and $q$ that may admit exceptions ( $\varepsilon$ ). $a=P(p), b=P(q)$, and $\varepsilon=P($ not $-q \mid p)$.

| $p$ | $q$ | not-q | a |
| :---: | :---: | :---: | :---: |
|  | $a(1-\varepsilon)$ | aع |  |
| not-p | $b-a(1-\varepsilon)$ | $1-b-a \varepsilon$ | $1-a$ |
|  | $b$ | $1-b$ |  |
|  |  | $\overline{\text { TRENDS }}$ in Cognitive Sciences |  |

consequent, $\mathrm{P}(\mathrm{q}) . \varepsilon$ isthe probability of exceptions; that is, the probability that q does not occur even though $p$ has, $P$ (not-q| $p$ ). Conditional probabilities for each inferencewerethen derived. For example, the conditional probability for $M P$, that is, $P(q \mid p)=1-\varepsilon$, depends only on the probability of exceptions. With few exceptionstheprobability of drawing theMP inference will behigh. However, the conditional probability of MT [that is, $P($ not-p| not-q) $=(1-b-a \varepsilon) /(1-b)]$ depends on the probability of the antecedent, $\mathrm{P}(\mathrm{p})$, and the probability of theconsequent, $P(q)$, as well as the probability of exceptions. As longas thereare exceptions ( $\varepsilon>0$ ) and the probability of the antecedent is greater than theprobability of the consequent not occurring $[P(p)>1-P(q)]$, then the probability of MT is less than MP $[P($ not-p| not-q) $<P(q \mid p)]$. F or example, if $P(p)=0.5, P(q)=0.8$ and $\varepsilon=0.1$, then $P(q \mid p)=0.9$ and $P($ not-p| not-q $)=0.75$.

This model accounts for the people's preference for MP over MT. The conditional probabilities associated with DA and AC also depend on these parameters, which means that they can benon-zero. Consequently themodel al so predicts that the fallacies should be endorsed to some degree. This model has also been applied to a variety of other effects in conditional inference ${ }^{21,22}$ (see Box 2).

In the conditional probability model, and in others ${ }^{25,27,28}$, the conditional premisealoneis regarded as uncertain. But what happens when the categorical premise is also uncertain? Peopletend to endorse conclusions that areonly as probableas the least probablepremise ${ }^{26,29}$. This is consistent with probability theory because the probability of the conclusion can beregarded as the product of the probabilities of thetwo premises ${ }^{26,29}$. For example, for DA, the probability that the conclusion not-q is endorsed is P (not-q| not-p).P(not-p). Of course, when the categorical premise, not-p, is certain [ $\mathrm{P}($ not- $p)=1$ ], then this account reduces to the model above. H owever, some recent research ${ }^{26}$ has shown behaviour less consistent with probability theory and more consistent with alternativeformalisms for dealing with uncertainty, such as Dempster-Shafer theory ${ }^{49}$.

## Wason's selection task

The probabilistic approach was originally applied to Wason's selection task, in which participants must select cards tofind out whether a rule[e.g. 'if there is an A on oneside of the card ( $p$ ) there is a 2 on the
other side (q)'] is true or false ${ }^{6,50,51}$. Participants see four cards, onewith A showing (p), onewith K (not-p), one with 2 (q) and one with 7 (not-q). They aretold to select only those they must turn in order tofind out whether theruleis trueor false.

In theinformation-gain model ${ }^{6}$ peopleare assumed to select evidence (i.e. turn cards) to determine whether $q$ depends on $p$, as in Fig. 1 (the dependence hypothesis, $\mathrm{H}_{\mathrm{D}}$ ), or whether p and $q$ are statistically independent (the independence hypothesis, $\mathrm{H}_{1}$ ). Participants are looking for evidence that provides the most discrimination between these two hypotheses. Initially, participants aremaximally uncertain about which is true; that is, theprior probabilities of $H_{D}$ and $H_{1}$ are each 0.5.

Theparticipants'goal is to select evidence (turn cards) that would beexpected to producethegreatest reduction in this uncertainty. This involves cal culating the posterior probabilities that the hypotheses, $\mathrm{H}_{\mathrm{D}}$ or $\mathrm{H}_{1}$, aretruegiven someevidence. These probabilities are calculated using Bayes' theorem, which requires information about prior probabilities $\left[P\left(H_{D}\right)=P\left(H_{1}\right)=0.5\right]$ and thelikelihoods of evidence given a hypothesis; for example, the probability of finding an $A$ when turning the 2 card, assuming $H_{D}\left[P\left(A \mid 2, H_{D}\right)\right]$. Theselikelihoods can be cal culated directly from the contingency tables for each hypothesis: for $\mathrm{H}_{\mathrm{D}}$, the contingency table in Fig. 1, and for $\mathrm{H}_{1}$, theindependence model, wherethe cell values are simply the products of the marginals. Theexpected reduction in uncertainty by turning any of thefour cards can then becalculated.

Assuming that themarginal probabilities $P(p)$ and $P(q)$ aresmall (the 'rarity assumption'), the $p$ and the q cards would beexpected to provide thegreatest reduction in uncertainty about which hypothesis was true. Consequently, the selection of cards that seemingly demonstrates human irrationality might reflect a rational data-selection strategy. I ndeed this strategy might be optimal in an environment where most properties arerare- for example, most things are not black, not ravens, and not apples - but this has been disputed ${ }^{9,50}$.

This model can account for most of thefindings on the selection task (see Box 3) and it has been defended ${ }^{2,51}$ against a variety of objections ${ }^{9,10,17,52-56}$. Therehas also been much research testing the empirical predictions of this and alternative probabilistic models ${ }^{10,11,13-16}$. Alternative models ${ }^{50}$ takeeither a 'disinterested' approach to human inquiry ${ }^{6,11}$ or a 'decision-theoretic' approach ${ }^{9,10}$. The disinterested approach was outlined above. Disinterested approaches, liketheinformation-gain model, makeno assumptions about the value people might or might not placeon particular types of evidence. By contrast, decision-theoretic approaches explicitly introducetheseutilities intothedecision to select a card. They make the same predictions for most of the data but diverge when peopledo not believea hypothesis. On thedecision-theoretic view,

## Box 2. Explaining conditional inference

The behaviour of the conditional-inference model is shown in Fig. I. This behaviour explains the two principal effects observed in conditional inference: negative conclusion bias, and suppression effects. Negative conclusion bias arises in Evans' negation paradigm ${ }^{\text {a-c }}$. Here, negations are used in the antecedents and consequents of the rules to createfour taskrules (whereA = Affirmative, $\mathrm{N}=$ Negative): if $p$ then $q(\mathrm{AA})$; if $p$ then not-q (AN ); ifnot-p then $q(N A)$; and if not-p then not- $q$ (NN). This manipulation means that half the conclusions of any inference will be affirmative and half of them will be negative. 'Negative conclusion bias' is observed when participants endorse more inferences with a negative conclusion than with an affirmative conclusion. Negated categories have a higher probability than their affirmative counterparts; for example, $P(x$ is a dog $)<P(x$

is not a dog) de. . Consequently if a conclusion is negated then it corresponds to a high-probability conclusion. The probability of the conclusion is on the $x$-axis in Fig. I. It is clear that as this probability increases so the probability with which the model predicts an inference should be endorsed also increases. So the model can explain negative conclusion bias. Recent experiments are broadly consistent with this account ${ }^{f-h}$.

Suppression effects occur when further information reduces the degree to which an inference is endorsed. For example, if someone is told that 'if the key is turned the car starts' and that 'the key is turned', they are likely to infer that 'the car starts', by modus ponens (MP). However, if they are also told that 'if the petrol tank is notempty the car starts', they are less likely to endorse
that conclusion because the car might not start if the petrol tank is empty. The petrol tank being empty provides an exception to the rule. These cases have been called 'additional antecedents', and it has been shown that they suppress the valid inferences modus ponens, MP, and modus tollens, MT (Ref. i). These effects are directly predicted by the probabilistic model, as can be seen from Fig. la. Additional antecedents correspond to high values of the exceptions parameter, $\varepsilon[P($ not- $q \mid p)]$. As Fig. la shows, as this parameter increases, the probability that the MP inference should be drawn decreases. This also happens for MT. More counter-intuitively, the model predicts that increases in this parameter should also decreaseendorsements of the conclusions 'denying the antecedent' (DA) and 'affirming the consequent' (AC) (see main text, and Ref. j), and this effect has been observed ${ }^{k}$.

## References

a Evans, J St-B.T. (1977) Linguistic factors in reasoning. Q.J . Exp. Psychol. 29, 297-306
b Evans, J St-B.T. etal. (1995) Bias in conditional inference: implications for mental models and mental logic. Q.J . Exp. Psychol. A 48, 644-670
c Evans, J St-B.T. and Handley, S.J . (1999) The role of negation in conditional inference. Q.J . Exp. Psychol. 52, 739-770
d Oaksford, M. and Stenning, K. (1992) Reasoning with conditionals containing negated constituents. J . Exp. Psychol. Learn. Mem. Cognit. 18, 835-854
e Oaksford, M. and Chater, N. (1994)A rational analysis of theselection task as optimal data selection. Psychol. Rev. 101, 608-631
f Oaksford, M. etal. (2000) Probabilities and polarity biases in conditional inference. J. Exp. Psychol. Learn. Mem. Cognit. 26, 883-899
g Schroyens, W. et al. (2000) Conditional reasoning with negations: implicit and explicit affirmation or denial and the role of contrast classes. Think. Reason. 6, 221-251
h Schroyens, W. etal. (2000) Heuristic and analytic processes in conditional reasoning with negatives. J . Exp. Psychol. Learn. Mem. Cognit. 26, 1713-1734
i Byrne, R.M.J . (1989) Suppressing valid inferences with conditionals. Cognition 31, 1-21
j Oaksford, M. and Chater, N. Probabilities and pragmatics in conditional inference: suppression and order effects. In Reasoning and Decision Making (Hardman, D. and Macchi, L., eds), J ohn Wiley \& Sons (in press)
k George, C. (1997) Reasoning from uncertain premises. Think. Reason. 3, 161-190

Fig. I. How the probability that a conclusion should be drawn ( $y$-axis) varies as a function of the probability of the premise, $P$ (premise), and conclusion ( $x$-axis) for DA (b), for AC (c) and for MT (d). The probability that an MP inference should be drawn (a) relies only on the exceptions parameter $\varepsilon[P(n o t-q \mid p)]$. Where no value appears this is because it violates the assumptions of the probability model ( $\varepsilon=0.25$ ). Values of $P$ (premise) are: blue squares, 0.1 ; violet squares, 0.3 ; green squares, 0.5 ; blue circles, 0.7 ; red squares, 0.9 .
participants should now select not-q cards, regardless of rarity. According to the disinterested approach, however, card sel ection is independent of believability.

Wehave argued that the current evidence favours the disinterested approach with respect tothe standard selection task ${ }^{50}$. However, there aretask

## Box 3. Explaining the selection task

The behaviour of the 'information-gain model' in shown in Fig. I. In modelling performance on the selection task, perhaps the most important pointto observe from these density plots is that when $P(p)$ and $P(q)$ are both small there is a region where the probability that the $q$ card should be selected is greater than the probability that the not-q card should be selected. When participants do the selection task, the most frequent response is to select the $p$ and the $q$ cards only. This behaviour is usually regarded as

irrational. However, according to the information-gain model, if the probabilities of the antecedent and consequent are quite small ('the rarity assumption') then this selection of cards is the rational selection: these two cards are more informative about which hypothesis is true. That the probabilities of the antecedent and consequent should be low is consistent with the observation that the categories of natural language divide the world up quite finely. So, for example, very few things are tables,


$\overline{\text { TRENDS }}$ in Cognitive Sciences

Fig. I. The probabilities with which a card should be selected as a function of the probabilities of the antecedent $[P(p), x$-axes ] and the consequent $[P(q), y$-axes] according to the revised information-gain model. The lighter the region the greater the probability that a card should be selected. These probabilities were calculated by transforming the information gains using a logistic selection-tendency functiona. The prior probabilities, $P\left(M_{1}\right)$ and $P\left(M_{D}\right)$ were set to 0.5 and the exceptions parameter $(\varepsilon)$ was set to 0.1 . Points in the lower-right triangular region in black violate the assumptions of the dependence model that $P(q)>P(p)(1-\varepsilon)$.
cars or gorillas. There is now good evidence that rarity is the default assumption when people aretesting ${ }^{\text {b,c }}$ or framing hypotheses ${ }^{d}$. The version of the model reported here meets a variety of theoretical objections, provides more rigorous fits to the data, and can account for more of the recent evidence on data selection.

In particular, the model provides detailed fits to the data obtained when negations are varied in the antecedents and consequents of the rules used in the selection task (see Box 2 ). For example, when this is done, participants select far more not- $q$ cards for the if $p$, then not- $q$ rule than the if $p$, then qrulee. Given the contrast set account of negation (see Box 2), Fig. I reveals why this should be the case. It is clear from the lower two panels in Fig. I that as the probability of the consequent, $P(q)$, increases, so the probability that the $q$ card should be chosen decreases but the probability that the not-q card should be chosen increases. It has also been shown that card selections vary in this way when $P(p)$ and $P(q)$ are varied rather than negations ${ }^{f}$.

## References

a Hattori, M. (1999) Theeffects of probabilistic information in Wason's selection task: an analysis of strategy based on theODS model. Proc. 16th Annu. MtgJ pn. Cognit. Sci. Soc. 16, 623-626
b Anderson, J.R. and Sheu, C.F. (1995) Causal inferences as perceptual judgments. Mem. Cognit. 23, 510-524
c McKenzie, C.R.M. and Mikkelsen, L.A. (2000) The psychol ogical side of Hempel's paradox of confirmation. Psychonomic Bull. Rev. 7, 360-366
d McKenzie, C.R.M. et al. Doconditional statements target rare events? Organiz. Behav. Hum. Decision Process. (in press)
e Evans, J .St-B.T. and Lynch, J .S. (1973) Matching bias in the selection task. Br.J. Psychol. 64, 391-397
f Oaksford, M. et al. (1999) Probabilistic effects in data selection. Think. Reason. 5, 193-243
manipulations ${ }^{57}$, particularly involving reasoning about practical action ${ }^{58}$, that require a decisiontheoretic approach. Indeed, in our original paper on theselection task ${ }^{6}$ the decision-theoretic approach was taken to thesetask versions, and we still believe that this is appropriate. Nonethel ess, whichever approach onetakes, despitesome contrary results ${ }^{17}$, it is now firmly establ ished that, '...no account of the selection task is sufficiently general if it cannot take account of the set size of $p$ and the set size of $q$ or the probability judgements which reflect these' ${ }^{\prime 15}$.

## Syllogistic reasoning

The probabilistic approach has also been extended to syllogistic reasoning, in which a conclusion is drawn from two premises (e.g. 'Someof theartists are bedkeepers'; 'All thebeekeepers arechemists'; therefore, 'Someof theartists arechemists'). In the probability heuristics model ${ }^{19}$ (PHM) the probabilistic interpretation of conditionals is extended to quantified claims: All, Some, None, and Some..not. In the contingency table of Fig. 1, if there areno exceptions, then the probability of the consequent

## Box 4. Probabilistic heuristics for syllogistic reasoning

There are three generation heuristics (G1-3) in the probability-heuristics model:
(G1) The min-heuristic: choose the quantifier of the conclusion to be the same as the quantifier in the least informative premise (the min-premise).

The most informative conclusions that can validly follow almost always follows this rule.

Some conclusions probabilistically-entail ('p-entail') other conclusions. For example, if All $X$ are $Y$, then it is probable that Some $X$ are $Y$ (this will follow as long as there are some $X$ s). Thus, the second heuristic is:
(G2) p-entailments: the next most preferred conclusion will be the $p$-entailment of the conclusion predicted by the minheuristic (the 'min-conclusion').

Heuristics G1 and G2 specify the quantifier of the conclusion. The third heuristic, the attachment-heuristic, specifies the order of end terms in the conclusion:
(G3) Attachment-heuristic: if justone of the possible conclusion subject noun phrases matches the subject noun phrase of just one premise, then the conclusion has that subject noun phrase.* (where subject noun is: $A\|X, A\| Z$, Some $X$, Some $Z$, etc.)

Crucially, Some $X$ are not $Z$ and Some $X$ are $Z$ have the same subject noun phrase: 'Some $X$ '. We illustrate these heuristics with an example (where '_' stands as a place holder for a subject of predicate term in the mental representation of the putative conclusion):

| Some $Z$ are $Y$ | (min-premise) |
| :--- | :--- |
| All $Y$ are $X$ | (max-premise) |
| Some_are_-_ | (by min) |
| Some $\bar{Z}$ are | (by attachment) |
| Some_arenot_- | (by p-entailment) |
| Some arenot $^{\text {a }}$ | (by attachment) |

By the min-heuristic, the conclusion is Some. The min-premise has an end term ( $Z$ ) as its subject. Therefore, by attachment, the conclusion will have $Z$ as its subject term, and the form Some $Z$ are $X$. The $p$-entailment of the min-conclusion is Some $Z$ are not $X$ which is predicted to be the next mostendorsed conclusion.

G1-G3 generate syllogistic conclusions. It is assumed that people are rarely able to test these conclusions for $p$-validity (or, for that matter, logical validity). However, it is argued that people use two test heuristics that provide a fast and frugal estimate of how likely the conclusion generated by G1-G3 is to be informative and $p$-valida:
(T1) The max-heuristic: be confident in the conclusion generated by G1-G3 in proportion to the informativeness of the most informative premise (the max-premise).
(T2) The Some...not-heuristic: avoid producing or accepting Some...not conclusions, because they are so uninformative relative to other conclusions.
*Wethank Geoff Goodwin (Brisbane University, Australia; pers. commun.) for pointing out an error in theoriginal formulation of G3 (Ref. a). Although the exampleis consistent with the definition we providehere, the original definition was too narrow.

## Reference

a Chater, N. and Oaksford, M. (1999) The probability heuristics model of syllogistic reasoning. Cognit. Psychol. 38, 191-258
given theantecedent, ( $\mathrm{P}[\mathrm{q} \mid \mathrm{p}]$ ), is 1 . Theconditional and the universal quantifier All have the same underlyinglogical form: $\forall x(P(x) \Rightarrow Q(x))$. Consequently universal daims, such as All Ps areQs, were interpreted as asserting that the probability of the predicate term $(Q)$ given the subject term $(P)$ is 1 ; that is, $P(Q \mid P)=1$. Probabilistic meanings for the other quantifiers are then easily defined: None, $P(Q \mid P)=0$; Some, $P(Q \mid P)>0$; Some...not, $P(Q \mid P)<1$.

Given these probabilistic interpretations it is possibleto provewhich condusions follow, probabilistically, for all 64 possi ble syllogisms (i.e. which syllogisms are 'p-valid'). M oreover, given theseinterpretations, and again making the rarity assumption (seeaboveon theselection task), the quantifiers can beordered interms of how informativethey are: All > Some> N one> Some...not. It turns out that a simpleset of heuristics defined with respect totheinformativeness of the premises can successfully predict the p-valid condusion if there is one (see Box 4).

Themost important is themin-heuristic, which states that the conclusion will havetheform of the least informative premise. Sofor example, a p-valid syllogism, such as All B areA; SomeB arenot C, yields the conclusion SomeA arenot C. This simple heuristic captures theform of the conclusion for most p-valid syllogisms (see Box 5). M oreover, if overgeneralized to the invalid syllogisms, the conclusions it suggests match the empirical data very well. Other heuristics determinethe confidence that people have in their conclusions and theorder of terms in the conclusion (see Box 4).

The most important feature of PHM is that it can generalize to syllogisms containing quantifiers, such as Most and Few, that have no logical interpretation. In terms of Fig. 1 thesequantifiers are used when thereare some (M ost) or many (F ew) exceptions. So themeaning of Most is: $1-\Delta<\mathrm{P}(\mathrm{Q} \mid \mathrm{P})<1$, and the meaning of F ew is: $0<P(Q \mid P)<\Delta$, where $\Delta$ is small. These interpretations lead to the fol lowing order of informativeness: All > M ost > F ew > Some > N one> Some...not. Consequently, PHM uniquely makes predictions for the 144 syllogisms that are produced when M ost and Few arecombined with the standard logical quantifiers. Weshowed previously that these heuristics pick out thep-valid condusions for these new syllogisms, and reported experiments that confirm the predictions of PHM when Most and Few areused in syllogistic arguments ${ }^{18}$.

There has already been some further work on syllogistic reasoning consistent with PHM (Refs 19,20). For example, themin-heuristic captures the novel distinction between strong and weak possible conclusions ${ }^{19}$. Takethe syllogism All Y areX; SomeZ areY. The conclusion SomeZ areX, follows necessarily from thesepremises; NoZ areX, is impossible, but Some $Z$ arenot $X$ and $A l l Z$ are $X$ are possible. Some possiblecondusions are endorsed as strongly as necessary conclusions (e.g. SomeZ are

## Box 5. Explaining syllogistic reasoning

Table I shows the results of a meta-analysis showing the weighted average of correct responses to the logically valid syllogisms ${ }^{\text {a-d }}$. The pairs, for example, (X,Z), in the syllogistic premises are ordered to show their status as the subject (first) or the predicate (second) term. The logically valid conclusion and the number of mental models required to reach that conclusion are also shown.

These data reveal how the heuristics of PHM can explain the data. Note that all the syllogisms above the red line are drawn more often than all those below it (with one exception). Moreover, all the syllogisms above the blue line are drawn more often than all of those below it. All the syllogisms above the red line conform perfectly to the min-heuristic (G1; see Box 4); those below, although not violating the min-heuristic (the conclusion is less informative than the minconclusion), do not conform to it. All those syllogisms that are below the blue line only have a very uninformative Some...not conclusion (T2) and interestingly this conclusion is drawn only as often as Nonewhich is the min-conclusion. Moreover, Some...notis a p-entailment of None (G2). Finally, all those syllogisms above the red line also have very informative max-premises (T1). Thus, it would seem that a set of very simple heuristics defined with respect to the informativeness of the premises can explain the differences in performance for the logically valid syllogisms.

## References

a Chater, N. and Oaksford, M. (1999) The probability heuristics model of syllogistic reasoning. Cognit. Psychol. 38, 191-258
b Dickstein, L.S. (1978) The effect of figureon syllogistic reasoning. Mem. Cognit. 6, 76-83

Table I. Meta-analysis of syllogisms

| Syllogism | Conclusion | No. of mental models | Mean ${ }^{\text {a }}$ |
| :---: | :---: | :---: | :---: |
| $A \\| /(\mathrm{Y}, \mathrm{X}), \mathrm{Al} /(\mathrm{Z}, \mathrm{Y})$ | All | 1 | 89.87 |
| All( $\mathrm{X}, \mathrm{Y}), \mathrm{Al}(\mathrm{Y}, \mathrm{Z})$ | All | 1 | 75.32 |
| All( $\mathrm{Y}, \mathrm{X}$ ), Some(Z, Y$)$ | Some | 1 | 86.71 |
| Some(Y, X$), A l /(\mathrm{Y}, \mathrm{Z})$ | Some | 1 | 87.97 |
| All( $\mathrm{Y}, \mathrm{X}$ ), Some $(\mathrm{Y}, \mathrm{Z})$ | Some | 1 | 88.61 |
| Some( $\mathrm{X}, \mathrm{Y}$ ), Al/(Y,Z) | Some | 1 | 86.71 |
| $N o(Y, X), A l l(Z, Y)$ | No | 1 | 92.41 |
| $A \\|(X, Y), N o(Z, Y)$ | No | 1 | 84.81 |
| $N o(X, Y), A l l(Z, Y)$ | No | 1 | 88.61 |
| All( $\mathrm{X}, \mathrm{Y}), \mathrm{No}(\mathrm{Y}, \mathrm{Z})$ | No | 1 | 91.14 |
| All (X,Y), Some...not(Z,Y) | Some...not | 2 | 67.09 |
| Some...not $(X, Y), A l l(Z, Y)$ | Some...not | 2 | 56.33 |
| All (Y, X$)$, Some...not $(\mathrm{Y}, \mathrm{Z})$ | Some...not | 2 | 66.46 |
| Some...not $(\mathrm{Y}, \mathrm{X}), \mathrm{Al/(Y,Z)}$ | Some...not | 2 | 68.99 |
| Some(Y, X$), \mathrm{No}(\mathrm{Z}, \mathrm{Y})$ | Some...not | 3 | 66.46 |
| No(Y, X$)$, Some(Z, Y$)$ | Some...not | 3 | 16.46 |
| Some(X, Y$), \mathrm{No}(\mathrm{Z}, \mathrm{Y})$ | Some...not | 3 | 51.90 |
| No(X,Y), Some(Z,Y) | Some...not | 3 | 30.38 |
| Some(Y, X$), \mathrm{No}(\mathrm{Y}, \mathrm{Z})$ | Some...not | 3 | 48.10 |
| No(Y,X), Some(Y,Z) | Some...not | 3 | 32.91 |
| Some(X, Y), No(Y,Z) | Some...not | 3 | 26.58 |
| No(X, Y$)$, Some(Y,Z) | Some...not | 3 | 44.30 |

aThe means in the final column are weighted by sample size.

C J ohnson-Laird, P.N. and Bara, B.G. (1984) Syllogistic inference. Cognition 16, 1-62
d J ohnson-Laird, P.N. and Steedman, M. (1978) Thepsychology of syllogisms. Cognit. Psychol. 10, 64-99
not X) and some are endorsed as weakly as impossible conclusions (e.g. All Z areX). Possiblestrong conclusions all conformtothemin-heuristic; that is, they either match themin-premiseor areless informativethan themin-premise. Possibleweak conclusions all violatethemin-heuristic (bar one); that is, they have conclusions that aremore informativethan themin-premise.

At first sight, mental-models theory also explains this result. Strong possible conclusions arethose licensed in theinitial model but not in subsequent models. Weak possibleconclusions are thoselicensed only in non-initial models. This means that the conclusions recommended by themin-heuristic correspond to thoselicensed by initial mental models. This is an interesting coincidence, as neither theory was constructed with the distinction between strong and weak possible conclusions in mind. However, it is unclear why mental-models theory licenses theseinitial models. The only suggestion that has been made is that it is an 'emergent
property' of the computer program that embodies the mental models theory. But no account of the principles underlying initial model construction that can explain why this property emerges has been forthcoming. In summary, PHM would appear to be gainingsomeempirical support.

## Everyday and logical inference revisited

Finally, we consider a potential problem for the probabilistic approach. Some people do reason logi cally some of the time ${ }^{37,38}$ and they tend to score more highly on IQ tests ${ }^{38}$. This observation has been interpreted as supporting a dual-process view ${ }^{10}$. 'System-1' processes are automatic, unconscious and based on implicitly acquired world knowledge. 'System-2' processes are controlled, analytical and based on explicitly acquired formal rules. The probabilistic approach provides a computationallevel theory of System-1 processes in which the probabilities invol ved are considered as summary statistics computed over world knowledge. On this

## Questions for future research

- The probabilistic approach has largely concentrated on the computational level of explanation. A challenge is to move to the algorithmic level and to propose theories of the actual mechanisms that implement these processes in the mind/brain. In this regard Bayesian networks ${ }^{47,48}$ look promising.
- We regard the probabilities in the models as deriving from world knowledge. But providing an account of the storage and retrieval of world knowledge on the scale required to explain human cognition and reasoning is not a tractable problem ${ }^{60}$. Major breakthroughs in the computer science of knowledge-based systems will probably be required before complete explanations of human reasoning will be possible.
- Other accounts of human reasoning, in particular mental models, have been extended beyond the coretasks discussed in this paper. The viability of the probabilistic approach depends on showing that it can be extended to these areas and continue to make novel predictions on the way.
- How human reasoning relates to information processing in the brain has not been seriously addressed by any theory (even though studies of the relationship between reasoning and the brain have begun ${ }^{61,62}$ ). The probabilistic approach might be helpful here because of the relationship between Bayesian networks and neural networks ${ }^{63}$. That is, probability theory could provide the right language in which to connect 'higher' level cognition to brain-style computation.
view, most reasoning invol ves only System-1 processes. However, some people, especially the more intelligent, might acquire explicit logical rules, either culturally or by explicit tuition. This is consistent with the probabilistic approach wherethe possibility has already been raised that some people might use System-2 processes to test conclusions generated by System-1 (Ref. 18).

Thecritical question is the bal ance of System-1 versus System-2 processes in human reasoning. Most recent theorizing has been about System-2 processes ${ }^{30,31}$. However, results from the selection task ${ }^{59}$ suggest that, at most, $10 \%$ of university students arecapable of engaging System-2 processes when reasoning. If, as this result suggests, most reasoning invokes only System-1 processes, then surely this is wherereasoning researchers should be looking. A probabilistic approach tothese processes explains people's performance in the laboratory as a rational attempt to make sense of thetasks they are set, by applying strategies adapted for coping with the uncertainty of theeveryday world. It is these strategies that create the appear ance of biased and irrational reasoning when compared with the standard provided by formal logic.

## References

1 Evans, J .St-B.T. et al. (1993) Human Reasoning, Erlbaum
2 Manktelow, K. (1999) Reasoning and Thinking, Psychology Press
3 Stein, E. (1996) Without Good Reason, Oxford University Press
4 Stich, S. (1985) Could man bean irrational animal? Synthese64, 115-135
5 Wason, P.C. (1968) Reasoning about a rule. Q.J. Exp. Psychol. 20, 273-281

6 Oaksford, M. and Chater, N. (1994)A rational analysis of the selection task as optimal data selection. Psychol. Rev. 101, 608-631
7 Oaksford, M. and Chater, N. (1998)A revised rational analysis of the selection task: exceptions and sequential sampling. In Rational Models of Cognition (Oaksford, M. and Chater, N., eds), pp. 372-398, Oxford University Press
8 Hattori, M. (1999) The effects of probabilistic information in Wason's selection task: an analysis of strategy based on theODS model. Proc. 16th Annu. MeetingJ pn. Cognit. Sci. Soc. 16, 623-626
9 Klauer, K.C. (1999) On thenormativejustification for information gain in Wason's selection task. Psychol. Rev. 106, 215-222
10 Evans, J .St-B.T. and Over, D.E. (1996) Rationality in theselection task: epistemic utility versus uncertainty reduction. Psychol. Rev. 103, 356-363
11 Nickerson, R.S. (1996) Hempel's paradox and Wason's selection task: logical and psychological puzzles of confirmation. Think. Reason. 2, 1-32
12 Oaksford, M., et al. (1999) Probabilistic effects in data selection. Think. Reason. 5, 193-243
13 Oaksford, M. et al. (1997) Optimal data selection in thereduced array selection task (RAST). J. Exp. Psychol. Learn. Mem. Cognit. 23, 441-458

14 Green. D.W. and Over, D.E. (1997) Causal inference, contingency tables and the selection task. Curr. Psychol. Cognit. 16, 459-487

15 Green, D.W. and Over, D.E. (2000) Decision theoretical effects in testing a causal conditional. Curr. Psychol. Cognit. 19, 51-68
16 Manktelow, K.I.etal. (1995) Probabilistic factors in deontic reasoning. Think. Reason. 1, 201-220
17 Oberauer, K. et al. (1999) Bayesian rationality for the Wason selection task? A test of optimal data selection theory. Think. Reason. 5, 115-144
18 Chater, N. and Oaksford, M. (1999) The probability heuristics model of syllogistic reasoning. Cognit. Psychol. 38, 191-258
19 Evans, J .St-B.T. et al. (1999) Reasoning about necessity and possibility: a test of the mental model theory of deduction. J . Exp. Psychol. Learn. Mem. Cognit. 25, 1495-1513
20 Newstead, S.E.et al. (1999). Falsifying mental models: testing the predictions of theories of syllogistic reasoning. Mem. Cognit. 27, 344-354
21 Oaksford, M. et al. (2000) Probabilities and polarity biases in conditional inference. J . Exp. Psychol. Learn. Mem. Cognit. 26, 883-899
22 Oaksford, M. and Chater, N. Probabilities and pragmatics in conditional inference: suppression and order effects. In Reasoning and Decision Making (Hardman, D. and Macchi, L., eds), J ohn Wiley \& Sons (in press)
23 Schroyens, W. et al. (2000) Conditional reasoning with negations: implicit and explicit affirmation or denial and the role of contrast classes. Think. Reason. 6, 221-251
24 Schroyens, W. et al. (2000) Heuristic and analytic processes in conditional reasoning with negatives. J. Exp. Psychol. Learn. Mem. Cognit. 26, 1713-1734
25 Chan, D. and Chua, F. (1994) Suppression of valid inferences: syntactic views, mental models, and relativesalience. Cognition 53, 217-238
26 George, C. (1997) Reasoning from uncertain premises. Think. Reason. 3, 161-190
27 Liu, I. etal. (1996)A probabilisticinterpretation of 'If-then'. Q.J.Exp. Psychol. A 49, 828-844

28 Stevenson, R.J . and Over, D.E. (1995) Deduction from uncertain premises. Q.J.Exp. Psychol. A 48, 613-643
29 George, C. (1999) Evaluation of the plausibility of a conclusion from several arguments with uncertain premises. Think. Reason. 5, 245-281
30 Rips, L.J . (1994) ThePsychol ogy of Proof, MIT Press
31 J ohnson-Laird, P.N. and Byrne, R.M.J . (1991) Deduction, Erlbaum
32 Evans, J .St-B. T. and Over, D. (1996) Rationality and Reasoning, Psychology Press
33 Chater, N. and Oaksford, M. (1999) Ten years of therational analysis of cognition. Trends Cognit. Sci. 3, 57-65
34 Anderson, J.R. (1990) TheAdaptiveCharacter of Thought, Erlbaum
35 Anderson, J.R. (1991) Is human cognition adaptive? Behav. Brain Sci. 14, 471-517
36 Oaksford, M. and Chater, N., eds (1998) Rational Models of Cognition, Oxford University Press
37 Braine, M.D.S. and O'Brien, D.P. (1998) Mental Logic, Erlbaum
38 Stanovich, K.E. and West, R.F. (2000) Individual differences in reasoning: implications for the rationality debate. Behav. Brain Sci. 23, 645-726
39 Oaksford, M. and Chater, N. (1998) Rationality in an Uncertain World, Psychology Press
40 Oaksford, M. and Chater, N. (1995) Theories of reasoning and the computational explanation of everyday inference. Think. Reason. 1, 121-152
41 Oaksford, M. and Chater, N. (1991)Against logicist cognitive science. Mind Lang. 6, 1-38
42 Oaksford, M. and Chater, N. (1992) Bounded rationality in taking risks and drawing inferences. Theory Psychol. 2, 225-230
43 Oaksford, M. and Chater, N. (1993) Reasoning theories and bounded rationality. In Rationality (Manktelow, K.I and Over, D.E., eds), pp. 31-60, Routledge

44 Chater, N. and Oaksford, M. (1993) Logicism, mental models and everyday reasoning. Mind Lang. 8, 72-89
45 Adams, E. (1966) Probability and the logic of conditionals. In Aspects of I nductiveLogic (Hintikka, J . and Suppes, P., eds), pp. 265-316, North Holland
46 Adams, E. (1975) TheL ogic of Conditionals: An Application of Probability Theory to Deductive Logic, Reidel
47 Pearl, J . (1988) Probabilistic Reasoning in Intelligent Systems, M organ K aufmann
48 Pearl, J . (2000) Causality, CambridgeUniversity Press
49 Shafer, G. (1976) A Mathematical Theory of Evidence, Princeton University Press
50 Chater, N. and Oaksford, M. (1999) Information gain vs. decision-theoretic approaches to data selection. ResponsetoKlauer. Psychol. Rev. 106, 223-227

51 Oaksford, M. and Chater, N. (1996) Rational explanation of the selection task. Psychol. Rev. 103, 381-391
52 Laming, D. (1996) On theanalysis of irrational data selection: a critique of Oaksford and Chater (1994). Psychol. Rev. 103, 364-373

53 Almor, A. and Sloman, S.A. (1996) Is deontic reasoning special? Psychol. Rev. 103, 374-380
54 Green, D.W. (2000) Review of Rationality in a Uncertain World: Essays on theCognitiveScience of Human Reasoning. Q.J . Exp. Psychol A 53, 281-283
55 Green, D.W. and Over, D.E. (1998) Reaching a decision. A reply to Oaksford. Think. Reason. 4, 231-248
56 Evans, J .St-B.T. (1999) Rational analysis of illogical reasoning (Review of Rationality in an Uncertain World: Essays on theCognitiveScience of Human Reasoning). Contemp. Psychol. 44, 461-463

57 Sperber, D. et al. (1995) Relevancetheory explains theselection task. Cognition 57, 31-95
58 Manktelow, K.I. and Over, D.E. (1993) Rationality: Psychological and Philosophical Perspectives, Routledge
59 Stanovich, K.E. and West, R.F. (1998) Cognitive ability and variation in selection task performance. Think. Reason. 4, 193-230
60 Pylyshyn, Z.W. and Ford, K.M. (1995) TheRobot's Dilemma Revisited, Ablex Publishing
61 Wharton, C.M. and Grafman, J . (1998) Deductive reasoning and thebrain. TrendsCognit. Sci. 2, 54-59
62 Houde, O. et al. (2000) Shifting from the perceptual brain tothelogical brain: theneural impact of cognitiveinhibition training. J . Cogn. Neurosci. 12, 721-728
63 McClelland, J.L. (1998) Connectionist models and Bayesian inference. In Rational Models of Cognition (Oaksford, M. and Chater, N., eds), pp. 21-53, Oxford University Press

# Shedding light on brain function: the event-related optical signal 

Gabriele Gratton and Monica Fabiani


#### Abstract

One of the basic goals of cognitive psychology is the analysis of the covert processes that occur between stimulus and response. In the past 20-30 years, the tools available to cognitive psychologists have been augmented by a number of imaging techniques for studying the 'brain in action' in a non-invasive manner. These techniques have their strength in either temporal or spatial information, but not both. We review here recent advances of a new approach, the event-related optical signal (EROS). This method allows measurements of the time course of neural activity in specific cortical structures, thus combining good spatial and temporal specificity. As an example, we show how EROS can be used to distinguish between serial and parallel models of information processing.


## Gabriele Gratton*

 Monica Fabiani Beckman Institute, University of Illinois, 405 N. Matthews Ave, Urbana, IL 61801, USA.*e-mail:
GrattonG@uinc.edu

A new brain imagingtechnique, theevent-related optical signal (EROS), combines spatial resolution of better than a centimetrewith temporal resolution of theorder of milliseconds, which makes it ideal to study thetime course of neural activity in localized cortical areas. Theimagingsignal reflects changes in the optical scattering properties of brain tissuethat are concurrent with neuronal activity. Becausethesame instrument that measures EROS can al so detect hemodynamic phenomena that occur subsequent to neuronal activity, thistechniqueis uniquely suited tostudy therel ationship between neuronal and vascular effects (neurovascular coupling). Presently, limitations of thetechniqueincludereduced penetration (sothat only structures within $3-5 \mathrm{~cm}$ from the surface of thehead can bestudied) and relatively low signal-to-noiseratio, which in reality requires that data is averaged across subjects.

Two major classes of functional brain imaging methods aremost used at present: hemodynamic techniques (such as PET and fMRI), which are particularly useful for visualizing whereneural activity occurs; and electrophysiol ogi cal techniques (such as event-related brain potentials-ERPs, and magnetoencephalography-MEG), which are particularly useful for determining when activity occurs ${ }^{1}$. Both classes of techniquehave provided important data for understanding the brain processes underlying cognitivefunction ${ }^{2-4}$. However, each of them special izes in a particular domain, but provides limited information in the other domain. For this reason, several investigators have proposed that they becombined to providespatio-temporal maps of brain activity ${ }^{5,6}$. This approach requires the assumptions that the different signals are co-localized and respond similarly to experimental manipulation, assumptions that need to be val idated in each case.

Light scattering changes associated with neural activity EROS is based on a different approach from previous methods, and in particular on thefact that lightscattering properties of neural tissuechange when thetissueis active. This phenomenon was first demonstrated more than 50 years ago $^{7}$, and has been used to study the behavior of largenumbers of individual neurons in parallel ${ }^{8}$. Rector et al. have shown that the samemethod can beused tostudy

