PERSPECTIVES

The relevance of behavioural measures for functional-imaging studies of cognition

David Wilkinson and Peter Halligan

The psychological structure of cognition is often inferred from conjoint measures of behaviour (such as reaction time) and brain activation (such as cerebral blood flow). In many experiments these measures produce divergent results. One example is where a significant pattern of brain activation occurs without a corresponding change in overt behaviour. In such circumstances, can cognitive theory be informed from brain-activation data alone? Or, given the more established link between behaviour and cognition, is behavioural corroboration necessary?

The scientific study of psychology began towards the end of the nineteenth century, and set out to chart the mental processes that were assumed to drive human behaviour. Early researchers, such as W. Wundt and W. James, realized that, as the products of mental processing involve previous levels of 'unconscious' processing, it was necessary to develop objective indices to tap and distinguish the differential outputs of these hidden mental processes. These studies traditionally used the most accessible indicators of performance — response accuracy and reaction time. In the 1970s, neuropsychological evaluation of patients with brain damage provided a complementary approach from which to infer and develop cognitive theory. On the basis of the assumption that the separable elements of mental processing could be reflected in distinct patterns of neural activation, the developments in functional brain imaging in the 1990s extended this neuropsychological line of enquiry by measuring dynamic, task-specific responses in the healthy, human brain.

On the assumption that cognition is modular in nature[1], functional imaging claimed that different kinds of mental state could be 'mapped' onto spatially distinct brain areas. However, the relationship between neural activations and traditional behavioural indices of underlying mental states is not straightforward. Many experiments do not show a clear correspondence between changes at the brain and behavioural levels. One example is where a given experimental manipulation, such as a change in stimulus information or task instruction, alters the underlying pattern of brain activation but shows no significant effect on behavioural performance. This result is ambiguous because it provides inconsistent information as to whether the experimental manipulation has engaged the relevant cognitive processes. In such circumstances, a central question is the extent to which the profile of behavioural performance, as indexed by reaction time or response accuracy, should be used to constrain all interpretations of observed brain activity. If we consider that our current understanding of neural activation is still too vague to enable inferences to be drawn about the structure of cognition from activation data alone, then the presence of a corroborative behavioural effect is most important. By contrast, if we believe that brain imaging can offer independent, valid insights into the nature of cognition, then corroboration, although still important, becomes less relevant. Despite the importance of these issues for brain imaging studies, neither view has attained precedence within the neuroscience community. Yet the issue needs to be resolved if the contribution of imaging to cognitive theory is to be properly established.

Here, we ask whether cognitive differences can be inferred from patterns of brain activation alone or whether these must be validated by allied behavioural effects. We discuss how it is not possible to measure all cognitive processes with reaction time and accuracy measures, and argue that some 'hidden' processes might be detectable using functional imaging. By demanding that all patterns of brain activation be paired with a demonstrable behavioural correlate, the risk is therefore run of excluding new, potentially informative indices. We conclude that the primacy that is traditionally afforded to behavioural measures should be relaxed if the full nature of cognition is to be understood.

This issue has practical relevance for an imaging experiment that we recently conducted in which significant imaging, but not behavioural, results were obtained[2]. We examined whether the way in which participants located the midpoint of a stimulus was affected by the presence or absence of stimulus symmetry. The pattern of concurrent brain activation observed in the anterior cingulate gyrus was strong evidence for a facilitatory effect of symmetry. However, the presence/absence of symmetry did not affect the speed or accuracy of judgements (FIG. 1). This raised the question as to whether the imaging data could be solely used to claim that the symmetry manipulation had worked. Without an allied behavioural effect, were the neural effects related to symmetry processing, as opposed to either an extraneous variable or random variance?

There is little guidance from the wider imaging literature about the need for corroborative behavioural data. In some studies,
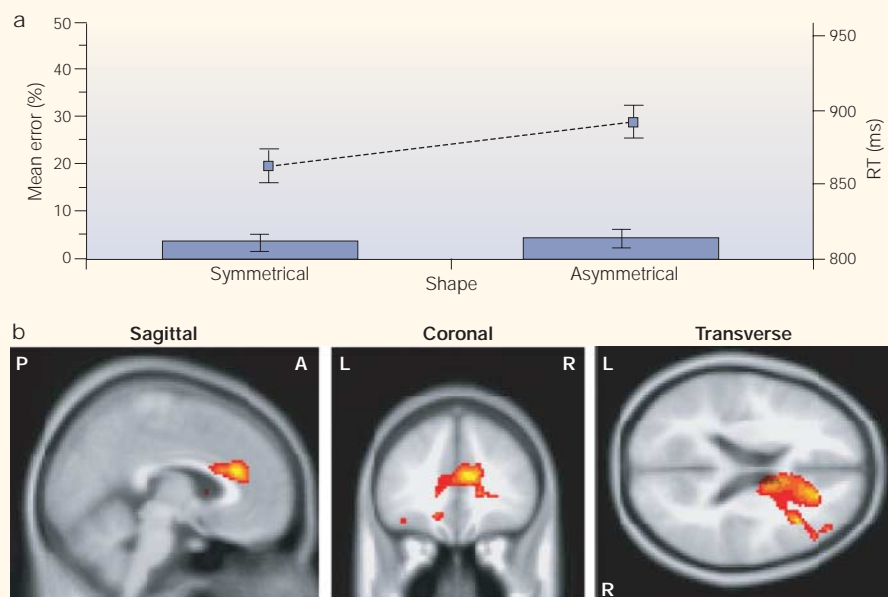
Figure 1 | **An example of statistically significant imaging data accompanied by non-significant behavioural data. a** | Mean percentage error and mean correct reaction time (with standard error bars) for bisecting symmetrical and asymmetrical shapes. Neither the response accuracy nor reaction time (RT) difference reached statistical significance. **b** | Significant activation of the cingulate gyrus associated with bisecting asymmetrical, relative to symmetrical, shapes. A, anterior; L, left; P, posterior; R, right. Reproduced, with permission, from REF. 2 © (2003) Elsevier Science.

interpretations of the imaging and behavioural data are closely interwoven — for example, when the specific aim is to map the neural basis of a well-established behavioural distinction, such as easy versus difficult visual search[3], successful versus unsuccessful memory retrieval[4] or certain kinds of visuomotor learning[5]. In other studies, however, the behavioural data are treated less respectfully, and inferences about cognition are drawn directly from the activation data, irrespective of overt performance. In some studies, behaviour is not even monitored, and it is simply assumed that subjects adhered to the task instructions from the presence of a 'predicted' pattern of activation[6,7]. In other studies, task adherence is confirmed by self-report, as opposed to a less introspective, objective means of response[8,9]. More relevant for the current discussion are imaging studies that have used an objective measure of behaviour (such as reaction time or response accuracy), and in which a given experimental manipulation has moderated responses at the imaging, but not the behavioural, level. This outcome is frequently reported, and encompasses a wide range of areas, including emotional processing[10], semantic categorization[11], lexical decision[12], visually guided movement[13], mental rotation[14], episodic memory[15], visual-object encoding[16], high-level conflict resolution[17] and lie detection[18] (but see REF. 19 for an

example of the reverse, in which significant effects are found in the behavioural, but not imaging, data).

Consider a recent study that examined the separability of rule-based (that is, propositional) and similarity-based (that is, perceptual) semantic categorization strategies[11]. Given that rule-based descriptions are typically harder to categorize than similarity-based descriptions, the authors sought to equate task difficulty by perceptually degrading stimuli in the 'similarity' condition. The imaging data indicated a high correlation between stimulus perceptibility and activation of the caudate nuclei. This was taken as evidence that, by increasing working memory load, perceptual degradation slowed overall cognitive processing. The problem is that perceptual degradation had no effect on the speed or accuracy of behavioural responses, which is clearly what would be expected if there had been general cognitive slowing. Imaging data were nevertheless deemed sufficient to support a perceptual-degradation hypothesis. The divergent results from this experiment raise two important questions: first, how do discrepancies arise between behavioural and imaging data and, second, does this undermine the degree to which distinct patterns of brain activation can be associated with specific cognitive processes?

Several articles have discussed some of the technological and theoretical problems that are inherent in functional brain imaging[20–23], and these do not need to be revisited. Our purpose here is to examine issues of a more practical nature that are concerned with the relationship between neurophysiological and behavioural responses that are evoked by task instructions and sensory manipulations. We first consider why behavioural corroboration is considered by some to constitute an essential feature of any meaningful imaging study. In this context, corroboration refers to the need for an appropriate or relevant behavioural effect (behavioural dependency) to be present for the meaningful interpretation of an effect found at the neural level.

In the second part of the article, we highlight how the failure to pair a significant pattern of brain activation with an allied behavioural effect might result from several reasons, such as random noise interference, poor design or inappropriate data analyses. More crucially, it might reflect an enduring aspect of the various structure/function mappings that characterize brain–behaviour relations. For example, the influence of some cognitive processes might be too subtle to be detected overtly, whereas others might mimic one another in the effects they exert on performance. In other cases, deliberately eliminating behavioural differences might be sought to refine neurophysiological interpretations, as in certain neuropsychological investigations. In all such instances, there is a danger in both over-determining the importance of behavioural data and in rejecting significant patterns of brain activation just because they are not mirrored behaviourally. Under these highly prescribed conditions, the demand for behavioural corroboration can be questioned. This is not to overlook or underplay the fact that behavioural indices are the most validated metric of cognition, or to suggest that the absence of behavioural corroboration is unproblematic. Rather, we want to point out that there are legitimate instances in which underlying cognitive differences are only captured by neural activation measures, and that this should not necessarily rule out sensible inferences from being made.

We conclude with a set of suggestions to consider when confronted with discordant behavioural and imaging effects. It should be pointed out that although brain imaging also embraces approaches derived from electrophysiology, such as single-unit recordings, our discussion will mainly concentrate on the application of functional imaging to cognitive science.

The case for corroboration

The need for behavioural corroboration is rooted in the belief that behavioural observation is the most informative index of human cognition. This idea is based on the assumption that the separable elements of cognition impose different time courses and degrees of error on performance, and can therefore be disclosed given the appropriate experimental manipulation[24]. It might also reflect simple historical standing, as until relatively recently there were remarkably few other ways with which to quantify cognitive activity. More generally, if the chief concern of cognition is to direct and control behaviour, then how informative can an experiment be in which there are no discernable behavioural outcomes? The precedence assigned to behavioural measures underlies a crucial position that might best be described as the 'behavioural-dependency criterion'. This assumes that a proven and well-tried path (behavioural observation) to cognitive theorizing has to be present to validate those findings of subsequent newcomers (such as neuroimaging). From this standpoint, behavioural measures offer the most powerful means of inference and therefore provide the yardstick by which imaging's own claim on cognition should be judged.

According to this view, a necessary step in any cognitive brain-imaging experiment is to ensure that it meets the behavioural criteria that confirm or constrain the range of cognitive processes that have been engaged. In doing so, the experimenter can be reasonably confident that the observed pattern of activation is of functional relevance. Proponents of this view might argue that the link between brain activation and a given cognitive state is currently too vague to be made without a relevant behavioural correlate. There is no need to confront the thorny issue of whether cognitive states can ultimately be reduced to specific brain states. Rather, it is suggested that the links between cell metabolism, the function of neuronal populations, cerebral blood flow and behaviourally relevant information are, at present, not clear enough to allow any meaningful exposition of cognitive theory on the basis of neural activations alone. Until a pattern of brain activation can be reliably shown to predict a given cognitive state, concurrent behavioural effects are therefore seen as an essential component of any functional-imaging study. As pointed out by Marshall and Gurd[25], "Current work on *in vivo* brain imaging is totally meaningless in the absence of functional theory drawn from data sources that are independent of neurophysiology; the number of degrees of freedom in the interpretation of rates of cerebral metabolism *per se* is little short of scandalous." One means of addressing this is to analyse activation data within the context of findings that do not rely on reductionist assumptions.

The argument for corroboration not only applies to cases in which no behavioural effect is reported, but also to instances in which a single or common behavioural effect is associated with multiple sites of brain activation. One means of deducing the precise role of each activated foci is to generalize in a *post hoc* manner from the findings of other studies that have reported similar activation (that is, the 'parity of precedence' method). This has been justified on the basis that other studies are of sufficient similarity to be of interpretative validity, and that the mere demonstration of a significant difference in the relevant gyrus or sulcus constitutes evidence in favour of the proposed hypothesis, regardless of the precise role of each activated area. This dependency on precedence should, however, not proceed without caution. Much has been written on the perils of conducting experiments that have so little explanatory power that the data can only be fully explained with reference to other studies[26]. If significant data can be interpreted with reference to other experiments, then under what conditions can they ever be given up as scientifically uninterpretable? Explanations must rely on the inherent logic found within the experiment, as determined by the initial research question, whereas acceptance of a hypothesis that rests on a host of anatomically distinct but functionally undifferentiated activations falls short of ensuring that each of the obtained effects are necessary or sufficient. As a recent case in point, functional-imaging studies of line bisection, a task commonly used to diagnose visual neglect, consistently reveal activation of the cerebellum[27–29], but damage to this region does not typically induce visual neglect.

The danger of over-determining uncorroborated patterns of brain activation is further highlighted, albeit indirectly, by the work of Marr[30]. He suggested that information processing can be described at three distinct levels. The top level specifies the aim or computational goal (for example, colour perception), the intermediate level specifies the means by which this aim is achieved, highlighting the processing steps or algorithms that are required to perform the necessary transformation (for example, discounting the illuminant), and the lowest level represents the physical implementation of these algorithms, which, put crudely, occurs in either brains or circuit boards. An important aspect of the framework is that the three levels are logically independent; there is no necessary or obligatory translation between them. Identifying the physical properties of a system does not readily disclose its informational content or purpose; even if we could fully simulate the design and workings of the nervous system, the overlying cognitive operations might still remain a mystery. This viewpoint borrows heavily from the artificial intelligence metaphor, which characterizes cognition as a set of formal operations that are regulated by rules and acted out across symbolic representations[1]. In this sense, the privileged level of description is at the algorithmic level, not the physical substrate on which these operations are performed, as it is the relationship between these symbols that are relevant and not the relationship between the physical states that implement them. This has led one commentator to suggest that the structure of cognition "can be given without regard to the material or hardware properties of the device on which these processes are executed", and that "in studying computation it is possible, and in certain respects essential, to factor apart the nature of the symbolic processes from the properties of the physical device in which it is realized"[31]. From this viewpoint, behavioural corroboration is important because a pattern of activation could conceivably stem from any number of cognitive algorithms. This 'behavioural-dependency criterion' should be distinguished from the more strident hyperfunctionalist view, which contests any role for imaging in the construction of cognitive theory (BOX 1).

The case against corroboration

Whereas few in the neuroimaging community would deny the value of corroborative behavioural effects, many would question the proposal that behaviour affords some unique or privileged insight into the psychological structure of cognition. Rather, behavioural performance is regarded as one index of underlying cognition, and brain activation as another, with neither form deserving precedence. Two assumptions are crucial to this viewpoint. The first is that imaging can potentially inform independently about the structure of cognition, and second, that behavioural measures might not be sensitive to all attributes of cognition, and so cannot serve as a gold standard for allied research methods.

The proposal that imaging can inform psychological theory is predicated on the idea that the separable elements of cognition can be paired with distinct neural markers. An initial aim for imaging science is therefore to identify

Box 1 | **The hyperfunctionalist view**

From this standpoint, the demand for allied behavioural effects is misplaced because it falsely assumes that imaging can, in principle, say something meaningful about cognitive architecture[65–67]. Central to this view is the assumption that cognitive science seeks to understand the nature of algorithms. In pursuit of this goal, it uses various techniques, including analysis of response time, accuracy and selective-interference techniques. As these techniques seek to address the same issue, corroboration among them is important. Neuroimaging seeks to understand the implementation of these algorithms through the application of techniques such as positron emission tomography and magnetic resonance imaging. Again, these measures seek to measure aspects of a common phenomenon (metabolic change), so corroboration is also important. However, as cognitive science and brain imaging are interested in fundamentally different questions, convergence is neither necessary nor possible (M. Coltheart, personal communication).

This does not mean that they are completely independent — if brain imaging addresses the neural implementation of algorithms, then it is reliant on those algorithms first being identified. Rather, no pattern of brain activation can ever be discrepant with a theory about algorithms because any given algorithm could be implemented in several ways. Suppose ten people are given a task to perform and that it is found that the task activates a different region in every person. Would this constitute evidence against the view that every person used the same algorithm? It is difficult to argue in the affirmative, simply because a given algorithm need not necessarily correspond with a given brain state (we are grateful to M. Coltheart for this example). As stated by Searle[68], "Even if my belief that Denver is the capital of Colorado is identical with a certain brain state, it seems too much to expect that everyone who believes that Denver is the capital of Colorado must have an identical neurophysiological configuration in his or her own brain." Proponents of the hard-line view would therefore argue that, as a neural dissociation cannot be taken as firm evidence for or against a psychological dissociation, functional imaging results can never confirm or disprove the operation of a given cognitive process. In support of this idea, its proponents would argue that imaging data has not led to a reformulation of existing cognitive theory. At best, they can guide ideas about algorithms, but this falls short of verifying that a specific algorithm has been implemented and, consequently, in its strongest form this perspective considers that there is little place for imaging science in the building of cognitive theory.

the neural correlates of known cognitive processes. This relies on cognitive psychology first specifying conditions in which such processes are engaged. However, once this has been achieved, imaging data might inform about some of the cognitive processes engaged in behaviours that are less well understood or in need of further explanation, as activation of a specific neural marker can be taken to imply that its associated cognitive algorithm has been implemented.

An illustrative example of this form of inference can be taken from the face-processing literature (for other examples, see REF. 32). A long-standing debate in cognitive psychology is whether faces are represented separately from other kinds of objects, such as cars and houses. Whereas much evidence has been derived from behavioural studies, the argument has focused on the role of fusiform gyrus, an area in the occipito-temporal cortex. Early studies seemed to show that damage to this region induces a specific deficit in recognizing faces, leaving object recognition comparatively intact[33,34]. This was taken as evidence for a separate face-processing module[35]. But later studies have cast doubt over the validity of some of these neuropsychological findings[36,37]. Coupled with the relative scarcity

of appropriate patients, the debate has become increasingly influenced by functional-imaging studies conducted on healthy subjects. These provide firm evidence that the fusiform gyrus is strongly activated by non-face stimuli. For example, ornithologists show strong fusiform activation when viewing specific species of birds[38], as do subjects who have acquired expertise in recognizing new non-face objects, such as greebles[39]. Fusiform activation is also apparent in untrained observers who must discriminate between visually similar objects[39]. These data indicate that the area might be more concerned with fine-grained discriminations and/or recognition expertise than with face processing *per se*. Although some researchers dispute the interpretation of these findings[40], the results have assumed importance in the argument against a distinct, face-processing module. In this way, imaging has provided useful constraints on theories of how cognition represents objects and faces.

More generally, not enough might yet be known about the structure of cognition to say how behavioural measures should constrain interpretations of brain activation. Consider the task of inferring the architecture of a computer from measurements of reaction time and response accuracy. Although this might

support a coarse description of the underlying organizational principles, it would seem unlikely that all component processes would be readily apparent, as different architectures can, in principle, produce similar output patterns[41]. The use of different lines of enquiry might be further justified by the fact that any number of cognitive models could potentially implement a specific goal or computation — a fact borne out by the varied forms of computer architectures that can simulate the same basic function (compare the model of Müller *et al.*[42] to that of Wolfe[43] for an example from the visual-search literature). On the one hand, the goal of cognitive science is to choose among these models. On the other, it seems undeniable that questions of how groups of neurons store information, respond to certain stimuli or modulate the operation of one another are in fact theories of representation and computation. Given the relative infancy of many cognitive theories, it seems sensible to make use of such biological data[44]. Observed discrepancies between behavioural and activation patterns could be taken not as evidence of a weak hypothesis, but as a point of positive departure with which to probe the validity of a pre-existing cognitive model[23]. Similarly, the association of a behavioural manipulation with multiple activation sites might indicate that the behavioural construct requires further fractionation. According to this view, functional imaging could potentially provide a source of empirical disconfirmation. As such, the absence of behavioural corroboration should not, in the first instance, be viewed as less meaningful than its presence.

A further difficulty with the behavioural-dependency criterion is that some aspects of cognition might not easily lend themselves to behavioural observation, and yet be readily apparent in the imaging data. This issue speaks to the variety of structure/function mappings that could characterize performance in a given task (FIG. 2). Consider a study comprising two experimental conditions, each of which is believed to tap different aspects of cognition. Responses are measured at both the behavioural and neural level. Comparisons between the two experimental conditions could produce any one of four outcomes: statistically significant effects in both the imaging and behavioural data; statistically significant effects only in the behavioural data; statistically significant effects only in the imaging data; or no statistically significant effects. The first outcome is perhaps the most desirable and the last outcome would, in most instances, constitute a failure. Inferences made on the basis of the

a One to one **b** One to many **c** Many to one **d** Many to many
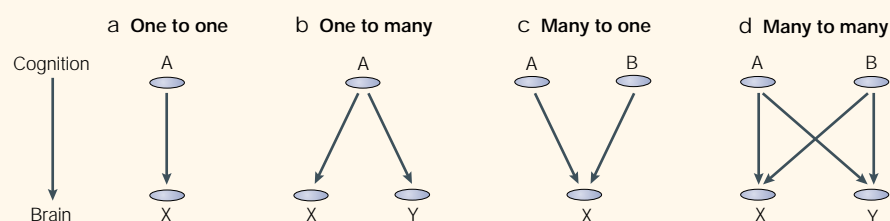
Cognition

Brain

Figure 2 | **Linking cognition to the brain: the effect of different structure/function mapping relations.** Logically speaking, cognition can be mapped to the brain in different ways. In some cases (**b**–**d**), components that are distinct at one stage will be integrated at another. This can prevent effects observed at the behavioural level from appearing at the neural level and vice versa.

second and third outcomes are more difficult to evaluate. On the one hand, they involve a null effect and so do not, by themselves, provide strong grounds for inference. On the other, they might faithfully characterize brain/behaviour mapping functions, and therefore be of potential insight.

In the case of the second outcome, the predicted dissociation is obtained in the behavioural data, but not in the imaging data. In theory, this pattern could arise if the same underlying cognitive strategy mediated responses in both experimental conditions, but at different degrees of behavioural efficiency. In other words, conditions 1 and 2 both recruit cognitive strategy A, which occupies anatomical location X, but condition 1 is performed at efficiency level n, whereas condition 2 is performed at efficiency level n + 1. This pattern would be consistent with a common cognitive-anatomical architecture that underlies several independent functions[41].

Of more relevance to the current discussion is the third outcome, in which the predicted effect is obtained in the imaging data, but not in the behavioural data. This pattern could arise if the two experimental conditions tapped different underlying cognitive strategies, both of which happened to impart similar time courses and patterns of error on task performance. In other words, condition 1 recruits strategy A, which occupies anatomical site X; condition 2 recruits strategy B, which occupies anatomical site Y; but both A and B operate at the same level of behavioural efficiency.

This form of organization seems to be present in certain parts of the visual system. Ablation studies in primates show that the perception of near objects can be differentially impaired from that of far objects after frontal-brain lesions[45]. This dissociation is also evident in brain-damaged patients, in which deficits have been observed for the bisection of stimuli in near space but not far[46], and in far space but not near[47]. Collectively, these data have been taken to endorse claims that the two forms of judgement recruit separate cognitive domains.

Support for the near/far distinction is, however, difficult to find in healthy volunteers. Five of six recent studies of line-bisection (the task most commonly used to elucidate near/far differences in neurological patients) failed to find any reliable behavioural difference[16,48–52]. More compellingly, recent imaging data have shown that the bisection of objects in near and far space recruits discrete brain networks in the ventral and dorsal visual streams, respectively[16,51]. However, this differential pattern of activity is not associated with distinct levels of reaction time or accuracy, which indicates that, although separate near/far representations might exist, both subsystems might normally operate at comparable levels of efficiency. This illustrates a problem for those who contend that patterns of brain activation can only be sensibly interpreted when accompanied by an allied behavioural effect.

The need for corroborative behavioural effects might also be questioned in cases in which the aim is to equate task difficulty across experimental conditions; for example, to finesse certain neurophysiological interpretations or to clarify the source of a behavioural dissociation. If a patient cannot perform the task, then we cannot know if the resulting patterns of activation caused the performance decrement, or if the performance decrement induced the abnormal pattern of activation[53,54]. This is not to deny the value of studying patients with performance decrement, as this can disclose the regions that are necessary to carry out a given task. However, this approach cannot easily ascertain the premorbid sufficiency of a damaged region in sustaining a particular function[55]. By contrast, by scanning patients on tasks that they can perform, inferences can be made about normative functional anatomy. Suppose that the network subserving a particular task is identified in normal subjects. If, owing to damage, patients can properly perform the task while activating only certain subsets of the network, then the regions that are not necessary for task completion can be identified. By implication,

a sufficient set of regions for the task comprises all those that are activated in normal people minus those that are not necessary to perform the task in patients (assuming that the intact parts of the network continue to operate as they did premorbidly)[53]. In such cases of underactivity, normal levels of performance in patients are therefore a prerequisite to inferring about the functional anatomy of healthy subjects. Overactivity, in which the task is performed normally but only with recruitment of additional areas, might indicate either the engagement of new strategies or the presence of degeneracy, in which a function is duplicated across different networks. Again, inferences can only be made about normal functional anatomy when the patients perform at a similar level to healthy subjects and, therefore, effects in the imaging data are not reproduced in the behavioural data.

Summary
The main point we wish to highlight is that reaction time/accuracy are imperfect measures of cognition. This implies that there will be legitimate circumstances in which divergence between behavioural and neural-activation measures can occur. Also, the relative importance of behavioural measures in formulating a unified cognitive theory cannot be known in advance. So, to treat these as the gold standard for related research methods might be overly conservative and prescriptive[56]. This point is underlined by recent evidence from the face-processing literature in which imaging data has led some cognitive psychologists to revise their models of how faces and objects are mentally represented. More practically, various structure/function mappings might give rise to divergence, as when separate cognitive processes are difficult to distinguish behaviourally, but produce distinct patterns of activation. There might also be specific circumstances in which performance differences are eliminated to hone interpretations of a given pattern of activation. Collectively, these arguments indicate that the absence of an allied behavioural effect (particularly in cases in which there is no relevant theory to predict the behavioural profile) does not render imaging data uninterpretable.

Implications for future studies
It is not our intention to argue that the absence of a behavioural effect is unimportant, or that it does not affect the strength of inference. Rather, we suggest that there are circumstances in which the absence of a corresponding behavioural effect should not be used to dismiss the potential contribution of a well-designed imaging study that addresses a

relevant question. Absence of evidence is not evidence of absence. The onus in such circumstances is to find an alternative source of corroboration (or falsification) as, although the absence of an effect might signal something important about an underlying mapping function, we note that little can be premised on a null effect.

If, after re-running an experiment, a concordant behavioural effect still fails to emerge, then three further steps could be taken to test if a given brain region is linked to a specific cognitive function: it should be activated in other tasks that are well known to tap it, but not in tasks that do not; incremental/ decremental increases in activation should occur when the process is incrementally/ decrementally taxed; and people with specific damage to the area of interest should show selective impairment for tests that incorporate the given process[57]. Recent advances in co-registration procedures might also, in some instances, enable temporal information to be simultaneously acquired using electro-encephalography, which might provide an alternative source of corroboration. Another way of teasing the two processes apart might be through the introduction of a new variable or task, which might interact with the processes of interest. In such a case, the design skills of the experimenter come to the fore.

The suggestion that neural effects might be more meaningfully interpreted when paired with behavioural data also assumes that the behavioural data have been appropriately analysed. However, accurate interpretation of reaction time and accuracy data is far from straightforward, and might rely on procedures that rarely seem to be considered in imaging experiments. For example, many imaging studies report processing dissociations that are based on the subtractive methodology proposed by Sternberg[58], in which discrete stages are inferred when two independent variables are shown to have non-interacting, additive effects on performance[59]. A problem is that additivity can also figure in other kinds of cognitive architecture, such as distributed parallel models. These models also assume discrete functional processes, but the temporal organization is quite different in that all processes can occur simultaneously, with an overt response occurring when a particular threshold is met[41]. Here, interaction effects might arise from two functionally independent processes[60]. Such concerns have led Sternberg himself to suggest that additive factor logic be used as a starting, rather than end point in data analyses[61]. A range of tests have been proposed to differentiate parallel from discrete stage models, involving analyses

of the higher-order moments of reaction-time distributions, such as variance, skewness, response priming, speed–accuracy trade-off analyses and sophisticated parallel guessing models[62]. On occasions in which these procedures are wrongly overlooked, the proposed number of cognitive elements comprising performance of any one task will be liable to miscalculation. In turn, this might affect the likelihood of finding a significant behavioural effect, and provide false constraints for the interpretation of any related imaging data.

Another shortfall of Sternberg's original model that still pervades some imaging studies is the tendency to overlook the response accuracy distributions in favour of reaction time. This approach is especially common in imaging studies in which error rates are low, and are therefore considered to convey little information about underlying processing differences. Small differences in error rate, however, can produce large differences in reaction time, especially if error rate is below 10% (REF 63). For this reason, joint consideration of both reaction time and error data is essential, especially if error rate interacts with experimental condition. One practice is to ensure that subjects perform at the same level of accuracy (as in some of the patient studies described above, and as advocated in a recent editorial[64]). This method, however, is potentially problematic in that the range of reaction times that can generate accurate performance is potentially unbounded, making any interpretation of reaction time difficult.

A further problem in equating a particular measure of performance across conditions is that this constitutes a new variable in itself which might or might not contribute to the underlying pattern of brain activation. In short, attempts to minimize error rates might contaminate the extent to which reaction time data reflect the operation of the underlying processes. Only with a greater understanding of the pitfalls inherent in reaction time and error measures can we be in a position to assess, in any imaging model, how useful these might be. Behavioural measures serve as a valid constraint only to the extent that they have received proper statistical treatment. Accordingly, lack of corroboration should not be confused with the failure to make proper sense of the behavioural data.

Conclusions

Behavioural observations have long formed the building blocks of cognitive theory. For this reason, it is perhaps inevitable that concerns are raised when an imaging study that is designed to explore the psychological basis of cognition seems to neglect the profile

of behavioural performance. The problem is that behavioural measures themselves are subject to a number of shortcomings and, as such, might be sensitive to only a few of the many cognitive attributes that mediate task performance. In this sense, it seems premature for behavioural indices to always take priority over neural ones. In placing too great an emphasis on behavioural indices, we also reduce the means by which newer and potentially more powerful methods can turn the spotlight back on the validity of these older, more trusted methods.

We therefore argue that the demand for behavioural corroboration in functional-imaging experiments should be tempered by the realization that meaningful corroboration can only begin when we know more about which parts of the cognitive process in question produce the observed behavioural responses, and which parts produce the observed neural activations. Imaging science, like all new sciences, must find converging support where and when it can (through, for example, electroencephalography, single-cell recording or magnetoencephalography) on the understanding that to overemphasize the relationship between brain activation and overt behaviour is to go beyond our current understanding of human cognition.

David Wilkinson is at the Geriatric Neuropsychology Laboratory, G.R.E.C.C. (182 JP), VA Boston Healthcare System, 150 South Huntington Avenue, Boston, Massachusetts 02130, USA.

Peter Halligan is at the School of Psychology, Cardiff University, Cardiff CF10 3YG, UK.

Correspondence to D.W.
e-mail: david_wilkinson@hms.harvard.edu

1. Fodor, J. A. *The Language of Thought* (Crowell, New York, 1975).
2. Wilkinson, D. & Halligan, P. W. Stimulus symmetry affects the bisection of figures but not lines: evidence from event-related fMRI. *NeuroImage* **20**, 1756–1764 (2003).
3. Nobre, A. C., Coull, J. T., Walsh, V. & Frith, C. D. Brain activations during visual search: contributions of search efficiency versus feature binding. *NeuroImage* **18**, 91–103 (2003).
4. Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L. & Schachter, D. L. Feeling-of-knowing in episodic memory: an event-related fMRI study. *NeuroImage* **18**, 827–836 (2003).
5. Toni, I., Rowe, J., Stephan, K. E. & Passingham, R. E. Changes of cortico-striatal effective connectivity during visuomotor learning. *Cereb. Cortex* **12**, 1040–1047 (2002).
6. Brooks, J. C., Nurmikko, T. J., Bimson, W. E., Singh, K. D. & Roberts, N. fMRI of thermal pain: effects of stimulus laterality and attention. *NeuroImage* **15**, 293–301 (2002).
7. Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y. & Shröger, E. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *NeuroImage* **15**, 167–174 (2002).
8. Inui, T. *et al.* Neural substrates for depth perception of the Necker Cube: a functional magnetic resonance imaging study in human subjects. *Neurosci. Lett.* **282**, 145–148 (2000).
9. Sterzer, P., Russ, M. O., Preibisch, C. & Kleinschmidt, A. Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *NeuroImage* **15**, 908–916 (2002).

10. Keightley, M. L. *et al.* An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli. *Neuropsychologia* **41**, 585–596 (2003).

11. Grossman, M. *et al.* The neural basis for categorization in semantic memory. *NeuroImage* **17**, 1549–1561 (2002).

12. Peng, D. *et al.* Neural basis of the non-attentional processing of briefly presented words. *Hum. Brain Mapp.* **18**, 215–221 (2003).

13. Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S. & Gazzaniga, M. S. Graspable objects grab attention when the potential for action is recognized. *Nature Neurosci.* **6**, 421–427 (2003).

14. Vingerhoets, G. *et al.* Regional brain activity during different paradigms of mental rotation in healthy volunteers: a positron emission tomography study. *NeuroImage* **13**, 381–391 (2001).

15. Suzuki, M. *et al.* Neural basis of temporal context memory: a functional MRI study. *NeuroImage* **17**, 1790–1796 (2002).

16. Weiss, P. H., Marshall, J.C., Zilles, K. & Fink, G. R. Are action and perception in near and far space additive or interactive factors? *NeuroImage* **18**, 837–846 (2003).

17. Paulus, M. P., Hozack, N., Frank, L. & Brown, G. G. Error rate and outcome predictability affect neural activation in prefrontal cortex and anterior cingulate during decision-making. *NeuroImage* **15**, 836–846 (2002).

18. Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L. & Yurgelun-Todd, D. A. Neural correlates of different types of deception: an fMRI investigation. *Cereb. Cortex* **13**, 830–836 (2003).

19. Grady, C. L., McIntosh, A. R., Beig, S. & Craik, F. I. An examination of the effects of stimulus type, encoding task, and functional connectivity on the role of right prefrontal cortex in recognition memory. *NeuroImage* **14**, 556–571 (2001).

20. Bub, D. N. Methodological issues confronting PET and fMRI studies of cognitive function. *Cogn. Neuropsychol.* **17**, 467–484 (2000).

21. Gusnard, D. A. & Raichle, M. E. Searching for a baseline: functional imaging and the resting brain. *Nature Rev. Neurosci.* **2**, 685–694 (2001).

22. Kosik, K. S. Beyond phrenology, at last. *Nature Rev. Neurosci.* **4**, 234–239 (2003).

23. Shulman, R. Interview. *J. Cogn. Neurosci.* **8**, 474–480 (1996).

24. Donders, F. C. On the speed of mental processes. *Acta Psychol.* **30**, 412–431 (1969).

25. Marshall, J. C. & Gurd, J. M. in *The Churchlands and their Critics* (ed. McCauley, R. N.) 176–191(Blackwell, Oxford, 1996).

26. Lakatos, I. & Musgrave, A. *Criticism and the Growth of Knowledge* (Cambridge Univ. Press, Cambridge, 1970).

27. Fink, G. R. *et al.* Line bisection judgements implicate right parietal cortex and cerebellum as assessed by fMRI. *Neurology* **54**, 1324–1331 (2000).

28. Fink, G. R., Marshall, J. C., Weiss, P. & Zilles, K. The neural basis of vertical and horizontal line bisection judgments: an fMRI study of normal volunteers. *NeuroImage* **14**, S59–S67 (2001).

29. Fink, G. R., Marshall, J. C., Weiss, P. H., Toni, I. & Zilles, K. Task instructions influence the cognitive strategies involved in line bisection judgments: evidence form modulated neural mechanisms revealed by fMRI. *Neuropsychologia* **40**, 119–130 (2002).

30. Marr, D. *Vision* (Freeman, San Francisco, 1982).

31. Pylyshyn, Z. Computation and cognition: issues in the foundation of cognitive science. *Behav. Brain Sci.* **3**, 111–134 (1980).

32. Downing, P., Liu, J. & Kanwisher, N. Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia* **39**, 1329–1342 (2001).

33. De Renzi, E. in *Aspects of Face Processing* (eds Ellis, H. D., Jeeves, M. A., Newcombe, F. & Young, A.) 243–252 (Martinus Nijhoff, Dordrecht, Netherlands, 1986).

34. Whiteley, A. M. & Warrington, E. K. Prosopagnosia: a clinical, psychological and anatomical study of three patients. *J. Neurol. Neurosurg. Psychiatry* **40**, 395–403 (1977).

35. Farah, M. J. Is face recognition 'special'? Evidence from neuropsychology. *Behav. Brain Res.* **76**, 181–189 (1996).

36. Gauthier, I., Behrmann, M. & Tarr, M. J. Can face recognition really be dissociated from object recognition? *J. Cogn. Neurosci.* **11**, 349–370 (1999).

37. Tovée, M. J. Is face processing special? *Neuron* **21**, 1239–1242 (1998).

38. Gauthier, I., Skudlarski, P., Gore, J. C. & Anderson, A. W. Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neurosci.* **3**, 191–197 (2000).

39. Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P. & Gore, J. C. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neurosci.* **2**, 568–573 (1999).

40. Kanwisher, N. Domain specificity in face perception. *Nature Neurosci.* **3**, 759–763 (2000).

41. Rumelhart, D. E. & McClelland, J. L. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol 1: Foundations* (MIT Press, Cambridge, Massachusetts, 1986).

42. Müller, H. J., Humphreys, G. W. & Donnelly, N. Search via recursive rejection (SERR): visual search for single and dual form conjunction targets. *J. Exp. Psychol. Hum. Percep. Perform.* **20**, 235–238 (1994).

43. Wolfe, J. M. Guided Search 2.0. A revised model of visual search. *Psychonom. Bull. Rev.* **1**, 202–238 (1994).

44. Churchland, P. S. & Sejnowski, T. in *Perspectives on Cognitive Neuroscience* (eds Lister, R. G. & Weingartner, H. J.) 3–23 (Oxford Univ. Press, Oxford, 1991).

45. Rizzolatti, G., Mattelli, M. & Pavesi, G. Deficits in attention and movement following the removal of the postarcuate (area 6) and prearcuate (area 8) cortex in macaque monkeys. *Brain* **106**, 655–673 (1983).

46. Halligan, P. W. & Marshall, J. C. Left neglect for near but not far space in man. *Nature* **350**, 498–500 (1991).

47. Cowey, A., Small, M. & Ellis, S. Visuospatial neglect can be worse in far than in near space. *Neuropsychologia* **32**, 1059–1066 (1994).

48. Barrett, A. M., Crucian G. P., Kim, M. & Heilman, K. M. Attentional grasp in far extrapersonal space after thalamic infarction. *Neuropsychologia* **38**, 778–784 (2000).

49. Cowey, A., Small, M. & Ellis, S. No abrupt change in visual hemineglect from near to far space. *Neuropsychologia* **37**, 1–6 (1999).

50. Tegnér, R. & Levander, M. The influence of stimulus properties on visual neglect. *J. Neurol. Neurosurg. Psychiatry* **54**, 882–887 (1991).

51. Weiss, P. H. *et al.* Neural consequences of acting in near versus far space: a physiological basis for clinical dissociations. *Brain* **123**, 2531–2541 (2000).

52. Wilkinson, D. & Halligan, P. W. The effects of stimulus size on bisection judgments in near and far space. *Visual Cogn.* **19**, 319–340 (2003).

53. Price, C. J. & Friston, K. J. Scanning patients with tasks they can perform. *Hum. Brain Mapp.* **8**, 102–108 (1999).

54. Price, C. J. & Friston, K. J. Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* **7**, 151–152 (2003).

55. Price, C. J., Mummery, C. J., Moore, R. S., Frackowiak, R. S. & Friston, K. J. Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients. *J. Cogn. Neurosci.* **11**, 371–382 (1999).

56. Churchland, P. S. *Neurophilosophy* (MIT Press, Cambridge, Massachusetts, 1986).

57. Smith, E. E. Research strategies for functional neuroimaging: a comment on the interview with R. G. Shulman. *J. Cogn. Neurosci.* **9**, 167–169 (1997).

58. Sternberg, S. The discovery of processing stages: extension of Donder's method. *Acta Psychologia* **30**, 276–315 (1969).

59. Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J. & Mazziotta, J. C. *Human Brain Function* (Academic Press, San Diego, 1997).

60. Meyer, D. E., Irwin, D. E., Osman, A. M. & Kounios, J. The dynamics of cognition and action: mental processes inferred from speed-accuracy decomposition. *Psychol. Rev.* **95**, 183–237 (1988).

61. Sternberg, S. in *Methods, Models and Conceptual Issues: An Invitation to Cognitive Science* (eds Scarborough, D. & Sternberg, S.) Vol. 4. 1–950 (MIT Press, Cambridge, Massachusetts, 1998).

62. Luce, R. D. *Response Times* (Oxford Univ. Press, Oxford, 1986).

63. Pachella, R. G. in *Human Information Processing: Tutorials in Performance and Cognition* (ed. Kantowitz, B.) 41–82 (Lawrence Erlbaum Associates, New York, 1974).

64. Editorial. Analyzing functional imaging studies. *Nature Neurosci.* **4**, 333 (2001).

65. Coltheart, M. in *Stevens' Handbook of Experimental Psychology* 3rd edn Vol. 4 (ed. Wixted, J.) 139–174 (John Wiley & Sons, New York, 2002).

66. Coltheart, M. Modularity and cognition *Trends Cogn. Sci.* **3**, 115–120 (1999).

67. Van Orden, G. C. & Paap, K. R. Functional neuroimages fail to discover pieces of mind in parts of the brain. *Philos. Sci.* **64**, S85–S94 (1997).

68. Searle, J. R. *The Rediscovery of the Mind* (MIT Press, Cambridge, Massachusetts, 1992).

🌐▶ **Online links**

FURTHER INFORMATION
Encyclopedia of Life Sciences: http://www.els.net/
Peter Halligan's homepage:
http://www.cf.ac.uk/psych/home/halliganpw/
Access to this interactive links box is free online.