Australian Journal of Psychology

Australian Journal of Psychology 2012; **64**: 29–36
doi:10.1111/j.1742-9536.2011.00041.x

APS Australian Psychological Society

What could cognitive neuroscience tell us about recognition memory?

Michael L. Kalish[1] **and John C. Dunn**[2]

[1]*Institute of Cognitive Science, University of Louisiana at Lafayette, Lafayette, Louisiana, USA and* [2]*Department of Psychology, University of Adelaide, Adelaide, SA, Australia*

Abstract

This paper is concerned with how the debate between single- and dual-process theories of recognition memory might be resolved. We argue that this is only possible if the theories concerned are competing to offer an explanation for the same phenomenon. We distinguish two kinds of explanations of recognition memory—roughly, one that explains what a person does to recognise an item, and another that explains what the brain does in order to enable a person to recognise an item. Our first point is that single- and dual-process theories typically, and perhaps counter-intuitively, do not offer competing explanations. Our second point is that this suggests two clear roles for neuroscience to play in the debate. Adjudicating between constitutive explanations would, we argue, require new experimental designs. Adjudicating between causal explanations requires prior determination of what function the brain is performing (an agreed psychological theory) before neuroscience could tell us how the brain is producing that performance.

Key words: biological correlates, cognitive processes, memory, theoretical and methodological issues

'And suddenly the memory revealed itself.' *Proust, Swann's Way*

The study of recognition memory is currently dominated by debate between two theoretical perspectives, generally referred to as *single-process* and *dual-process* models (for summaries from different sides of the debate, see Wixted, 2007; Yonelinas & Parks, 2007). Both models seek to account for the phenomena that occur in a single type of experiment; that in which participants study a list of items (words, picture, etc.) and are then tested on their ability to judge that those items (either as types or tokens), and not other items, were on the studied list. This debate, which shows little sign of immediate resolution, raises important questions concerning the nature of the claims regarding recognition memory, and the kinds of evidence that can address those claims. In particular, evidence from the neuroscience domain has been interpreted as having the potential to adjudicate between the two sides of the debate (Rugg & Yonelinas, 2003; Vilberg & Rugg, 2008). Our aim is not to review this evidence nor to try to settle the debate, one way or the other, but rather to conduct an analysis of the nature of the theoretical claims that have been, and could be made, about such constructs as *recognition*, *recollection*, and *familiarity*, and how this, in turn, affects the ways in which single- and dual-process models can be interpreted. In so doing, we answer the question of what cognitive neuroscience could tell us about recognition memory.

Our starting point is a set of arguments presented by Bennett and Hacker (2003) and further elaborated by Trigg and Kalish (2011). Both sets of authors draw attention to a fundamental distinction between *capacities* that humans and other agents may possess and the *causally enabling conditions* for their exercise. Capacities may be analysed as being constituted by other capacities. As an illustration, Trigg and Kalish offer the example of a person, Paul, who exercises a capacity to *open a door*. Analysis of this capacity shows it to be constituted by a set of other capacities such as the capacity to walk towards the door, to grasp the doorknob, to turn the knob, and so on. If Paul exercises the capacity to open the door then he also exercises the relevant constitutive capacities. As Cummins (2000) says, 'A cook's capacity to bake a cake analyzes into other capacities of the "whole cook" '.

In turn, the exercise of a capacity or set of constitutive capacities depends upon the presence of a set of causally enabling conditions. The fact that the door is unlocked is an example of one such condition. Conditions such as this inhere to the environment and offer the opportunity for the exercise of the capacity (Hyman, 1994). Other causally enabling conditions inhere to Paul himself, and provide a set

Correspondence: Michael L. Kalish, Institute of Cognitive Science, University of Louisiana at Lafayette, Lafayette, LA 70504-3772, USA. Email: kalish@louisiana.edu

Received 30 March 2011. Accepted for publication 2 September 2011.

of mechanisms that enable the capacity to be exercised (Kenny, 1976).[1] One such condition is that Paul has a hand to grasp the doorknob (if this is the method he uses) and a functional nervous system to activate the requisite muscle contractions to enable Paul to grasp, turn, and pull the doorknob appropriately. A functional nervous system also includes the relevant parts of Paul's brain that produce the nervous impulses that activate the required muscle contractions. The point is that the function of the brain is properly thought of as one of the causally enabling conditions (on the same footing as an unlocked door) and not as something that can exercise capacities constitutive of *opening the door*. It is Paul who opens the door—it is his brain whose function enables him to do it.

Trigg and Kalish (2011) point out that the distinction between capacities and their causally enabling conditions applies equally well to cognitive capacities such as thinking, reasoning, believing, and remembering. To remember something is thus to exercise a capacity. And, as with opening a door, in order for Paul to remember an event, the relevant causally enabling conditions must be in play. First, he must have an opportunity to remember the event in question—for example, he must have encountered it previously. Second, the appropriate mechanism(s) for the exercise of the capacity must be working effectively—for example, he should not be impaired by too much alcohol.

Some capacities are abilities and, in the sections to follow, we will refer to recognition memory as an *ability*. An ability is a capacity that admits to a performance criterion—it makes sense to say that an ability is performed well or poorly (Hacker, 2007). Recognition memory, at least as defined behaviourally, is an ability in this sense because it is routinely evaluated against a performance criterion (e.g., per cent correct). Other cognitive capacities, such as the capacity to hope, or to believe, are not abilities in this sense.

TWO KINDS OF EXPLANATION

Given the distinction between cognitive capacities and their causally enabling conditions, it is possible to distinguish two kinds of explanatory theory. The first kind, which we call a *constitutive* theory (what Cummins (2000) calls 'functional analysis'), explains a capacity by analysing it into other, perhaps more basic, capacities. An account of Paul's capacity to open a door in terms of his capacity to reach, grasp, and turn is an example of a constitutive theory. A second kind of theory, which we call a *causal* theory (what Cummins calls 'structural analysis'), explains a capacity by developing an account of the mechanisms that enable it. An account of the muscle contractions that enable Paul to reach, grasp, and turn is an example of at least part of a causal theory. These two kinds of theory are not in any necessary opposition. Indeed, it will often be the case that a constitutive analysis of a capacity may precede the development of an appropriate causal theory. However, they do answer different questions. A constitutive theory tells us *what* is accomplished, while a causal theory tells us *how* it is accomplished.

Clearly, constitutive theories are about capacities (including abilities) while causal theories are about causally enabling conditions (including mechanisms). As Trigg and Kalish (2011) point out, confusions necessarily arise when theories use explanations that do not conform to these alternative kinds of analysis. This may happen if a person's capacities are ascribed to one of their enabling mechanisms—for example, where a brain is said to remember or judge. It may also occur if the actions of a mechanism are ascribed to the person—for example, where a person is said to search memory or to retrieve a memory trace.

The point of this paper is to ascertain whether all theories of recognition memory can be meaningfully construed as being of just one (constitutive or causal) type, or if they are of different types. Performing this analysis on a cognitive theory is non-trivial, and, indeed, contentious. To foreshadow, we will argue that single- and dual-process theories are most sensibly construed as explanations of different kinds, and that this analysis can advance the debate between the single- and dual-process theories.

CONSTITUTIVE THEORIES OF RECOGNITION MEMORY

Recognition memory is a capacity that persons, such as Paul, are able to exercise given the necessary causally enabling conditions. Just what is this capacity? Mandler (1980) has proposed that recognition memory is just 'the judgment of previous occurrence'. This, however, offers at least three different interpretations referring to three very different capacities that we propose to call *recognition1*, *recognition2*, and *recognition3*. Recognition1 refers to the judgement that elements of the current environment have been experienced at least once before. In a police line-up, for example, the relevant question would be 'have you seen any of these people before?' Mandler considered this question in his (now) famous *butcher-on-the-bus* thought experiment. Here, he described an experience that is familiar to most people. A person is encountered in a particular situation (on the bus) and there is an immediate sense of recognition (recognition1)—the near certainty that this person has been encountered before. Mandler also described the experience of frustration that follows if the circumstances of those previous encounters cannot be immediately retrieved.

In the laboratory, participants are induced to exercise their capacity to recognise1 when asked to distinguish between familiar and unfamiliar stimuli such as words versus non-words, or photographs of famous versus non-famous people. However, in the study of recognition memory, such investigations are relatively infrequent and are often not viewed as central to the theoretical analysis of recognition memory as such. Instead, most attention is focused on a different judgement, *recognition2*: The judgement that elements of the current environment have been experienced at least once before *in a designated context*. In a police line-up, the relevant question would be, 'is one of these people the person who attacked you?' This is operationalised in laboratory studies by presenting some common items (usually words) that can be easily recognised1 in a study list and requiring participants to discriminate between these items and other unstudied (but also well recognised1) items at test. Participants are therefore required to judge which of these recognised1 items can also be recognised2 as having appeared in the context of the study list.

Returning to the butcher-on-the-bus scenario, Mandler (1980) has also suggested that a form of recognition also occurs when the individual on the bus, who has been recognised1 but whose identity is unknown, is then identified as the butcher—in other words, some of the relevant contexts in which this person had been previously encountered have been retrieved. We call this *recognition3*. In a police line-up the relevant questions might be 'Do you know who this person is?' or 'Can you recall this person's name?'

To summarise the results of our conceptual analysis, recognition admits to at least three interpretations, each related to a distinct ability. First, recognition1 is the (correct) judgement that an object or event (e.g., person, picture, word) has been previously encountered at least once. This ability is tested in the laboratory when people are asked to discriminate words from non-words or pictures of known faces from unknown faces. Second, recognition2 is the (correct) judgement that an event had been previously encountered in a designated context. This ability is tested when people are presented with a list of items and asked to discriminate studied items—those that appeared on the list—from other items that did not. Third, recognition3, or recollection, is the (correct) judgement that material associated with a previous encounter of an object or event has been retrieved. This ability may be tested in the laboratory using associative and source memory tasks.

Dual-process accounts

Dual-process accounts of recognition memory propose that such judgements are based on two 'processes', usually called *recollection* and *familiarity* (Parks & Yonelinas, 2007; Yonelinas, 1994, 2002). Proponents of dual-process theories have often motivated the distinction between recollection and familiarity by reference to the distinction between recognition3 and recognition1, respectively (Wixted & Mickes, 2010; Yovel & Paller, 2004). On this view, recollection refers to the capacity to remember information associated with the item, such as the fact that the person on the bus is the butcher. Familiarity refers to the capacity to merely recognise1 that the item in question had been encountered previously. Consistent with this, Yonelinas (1994) states that, 'studying an item temporarily increases the item's familiarity, such that old items will on average be more familiar than new items. Thus, an assessment of familiarity provides a good basis for recognition judgments. However, subjects may not be limited to assessments of familiarity. If some aspect of the study event can be recollected (e.g., "I remember seeing that word . . . It was the first one in the list"), this could also serve as a basis for recognition judgments' (p. 1341); see also Yonelinas (2001, 2002). Dual-process theory asserts that recognition2 ('memory judgements') is reducible to the combination of recollection and familiarity; that recognition2 may be done via recognition3, or via recognition1, or via a combination of both. In so doing, dual-process theory effectively claims that there in fact is no unique capacity for recognition2.

The fact that dual-process theory considers both recollection and familiarity to be abilities is further supported by the widespread use of remember/know instructions to disentangle them from recognition2 judgements (Gardiner, 1988; Vilberg & Rugg, 2008; Yonelinas & Jacoby, 1995). In this variant of a standard recognition2 task, participants are asked to classify each recognition2 judgement as either a 'remember' response, if they are able to recall any associated material from the study list, or a 'know' response otherwise. For this to be possible, the relevant recognition3 and recognition1 judgements must be the result of a person exercising their relevant abilities. Because it cannot be said of someone's brain that it recollected the last time it saw its butcher, or that a brain assessed the face of the butcher to be familiar, these capacities cannot be those of person's components, but of the person himself.

To summarise, our reading of dual-process theory is that it is a constitutive theory of recognition memory. We will call this reading DPT-A, for being a dual-process theory about abilities. It states that one ability, recognition2, is constituted of two other abilities, recognition1 (or familiarity) and recognition3 (or recollection). Research aimed at this claim would seek to discover if, indeed, people are using these abilities—and this research could take the prosaic form of asking them, because these abilities are not exotic constructs performed in the mind, but ordinary human activities being performed at will. If asking is not feasible, then research could determine what goes on (in the world, or in the head) while people recognise1 or recognise3, and then look to see if this is also going on when they recognise2.

Single-process accounts

The alternative to dual-process theory in accounting for recognition2 judgements is the single-process model. Although often described in terms of signal detection theory, single-process models of recognition2 memory receive their theoretical justification directly from computational models of memory (discussed below), which propose that recognition2 depends on some *global matching* process that measures how similar the test item is to the contents of memory. The match between a studied item and the contents of memory will, on average, be greater than the match of a non-studied item with the same contents of memory. The two sets of matches will therefore generate two distributions of *memory strength* or *evidence* on which a decision can be made by setting a response criterion as described by signal detection theory (Dunn, 2004). We suggest that there is an ability-level reading of this theory.

One tempting, but we think mistaken, reading is to take the usage of 'memory', 'similarity', 'evidence', and 'decision' as employing their ordinary meanings, and say that it is the person who does the computations during recognition2. Such computations would then be the capacities that constitute recognition2. But this would require that persons must literally compute the similarity of a test item to items in a memory, compute evidence strength, and set and employ a criterion on that strength. This reading is quite easy to fall into, but it is clearly impossible for people (many of whom cannot do arithmetic, for example) to be doing these things when making a recognition2 judgement. That is, the flaw in this reading is that it relies on people having capacities that they do not have (Trigg & Kalish, 2011).

A second reading is less problematic. It holds simply that recognition2 is a distinct ability, and that people accomplish recognition2 without relying on either recognition1 or recognition3. In other words, recognition2 is constitutive of itself (just as Paul's reaching for the doorknob may be similarly self-constituted). This does not deny that either recognition1 or recognition3 judgements may also be made in the context of a recognition2 judgement, but it does deny that they are constitutive of recognition2. We call this reading SPT-A, because it is about abilities and states that recognition2 is an ability constituted of itself.

CAUSAL THEORIES OF RECOGNITION MEMORY

As noted earlier, single-process models of recognition memory owe their theoretical justification to global matching models of memory such as TODAM (Murdock, 1982), MINERVA2 (Hintzman, 1984), SAM (Gillund & Shiffrin, 1984), BCDMEM (Dennis & Humphreys, 2001), and REM (Shiffrin & Steyvers, 1997). Such models provide a computational account of recognition2 which can be interpreted as essentially a causal theory of this ability. Such a theory seeks to explain recognition2 by first postulating that there is a mechanism that provides this ability, and by then by describing this mechanism's functional form.

The single-process causal theory of recognition2 memory, we argue, describes the set of non-conscious computations required for the global match which, in turn, describes the goodness of the reason, or the strength of the evidence, that a person has for claiming that they recognise2 the test item. The person concerned does not direct this process, and need have no inkling as to how it works. To say that 'when a person is recognizing an item, a global match is computed by pattern matching in the memory system' is much like saying 'when a person is turning a doorknob, abduction of the wrist is produced by flexion of the *extensor carpi ulnaris*' in that both are (coarse) statements of the structure and dynamic function of the mechanism required for exercise of the relevant capacities. We call this reading SPT-C, because it is about causal preconditions.

We can now restate SPT-C: The probability that a given attempt at recognition2 will be of the kind that gives one reason to judge that recognition2 has occurred can be predicted by a single computation. On this view, 'strength of evidence' should be read as 'probability that the attempt will lead to a positive judgement'.

Analogous to SPT-C, we can construct a reading of dual-process theory as being about the nature of the causally relevant conditions; call this DPT-C. According to this reading, recollection and familiarity are technical terms, indicating two different computations that are the causal preconditions for recognition2. If one also accepts DPT-A, then these two computations could be those required for recognition1 and recognition3, respectively. An example of a theory that appears to combine DPT-C and DPT-A is the source activation confusion model proposed by Reder and associates (Reder et al., 2000).

However, there is also another reading of DPT-C that is consistent with SPT-A. On this view, the computations that are causal of recognition2 consist of the calculation of quantities that may be functionally identified as 'recollection' and 'familiarity'. This denotes new meanings for these terms as two components in a model of recognition2 which may or may not have anything to do with the abilities to recollect (recognise3) or judge familiarity (recognise1); they would simply be the names given to components of a computational description of the causally enabling conditions for people to exercise their capacity to recognise2. A model of the remember/know procedure recently proposed by Wixted and Mickes (2010) is open to this interpretation. The term 'dual-process model' would then simply refer to a particular choice of model architecture based on these components, in contrast to distinctions based on the presence or absence of other components such as 'global match', or 'distributed memory vector', or 'likelihood ratio'.

To summarise, both dual-process and single-process models can be understood as offering either a constitutive or a causal explanation of recognition memory. However, because of the ways in which these models have been developed, the principal empirical claim of dual-process models is constitutive (i.e., DPT-A) as there has been relatively little development of causal (i.e., computational) models of recollection and familiarity. In contrast, the principal empirical claim of single-process models is causal (i.e., different claims concerning the nature of SPT-C) with the development of these accounts being predicated on the validity of SPT-A. This difference has meant that, to some extent, the debate between single-process and dual-process models has pitted theories of different things against each other and that this has led to confusion about the kinds of evidence that are or could be relevant to 'testing' them. This confusion, we suggest, has been most apparent in the neuroscience domain.

THE EVIDENCE FROM NEUROSCIENCE

In the previous sections, we have distinguished two kinds of constitutive theory of recognition memory and two kinds of causal theory. The constitutive dual-process theory (DPT-A) explains the exercise of recognition2 by theorising that to do it, people must exercise two other capacities—judgements of familiarity (recognition1) and recollection (recognition3). The constitutive single-process theory (SPT-A) explains the exercise of recognition2 by theorising that to do it, people must exercise a single capacity. The causal dual-process theory (DPT-C) explains the exercise of recognition2 by theorising about the mechanism that enables it, in particular by postulating that this mechanism generates two quantities, defined within the theory as 'recollection' and 'familiarity'. The causal single-process theory (SPT-C) also explains the exercise of recognition2 by theorising about the mechanism that enables it but, in this case, by postulating that this mechanism generates a single quantity, sometimes called 'strength of evidence'. In this section, we argue that the kind of explanation that is offered has important implications for the role of evidence from neuroscience in testing psychological theories.

Let us start by clarifying what we believe cognitive neuroscience can and cannot do. When we know which capacity an agent is exercising, neuroscience can potentially tell us what the neural mechanism of that capacity is. For example, it may tell us how the functions of the visual system, consisting of the optic nerve, the lateral geniculate nucleus, the optic radiation, the striatum, etc., enable seeing (see van Eck, De Jong, & Schouten, 2006, on the role of neuroscience in vision). Neuroscience cannot, when asked to explain what a capacity is, provide that analysis. If we did not know what seeing was, nothing we discovered in the brain could help us to find out. If the argument between SPT-A and DPT-A is one of what recognition2 memory *is*, then neuroscience evidence cannot be of assistance. No science, including neuroscience, can tell us what it means to remember, any more than it can tell us what it means to see, or to be human. On the other hand, if the argument is one of whether or not people are covertly exercising constitutive abilities while making recognition2 judgements, then neuroscience may have something to offer.

How could neuroscience adjudicate between the constitutive theories, DPT-A and SPT-A? As alluded to earlier, this would seem to be possible if there was agreement as to the circumstances under which persons may be said to be exercising their capacities to recognise1 or to recognise3, in which case there may be brain correlates of each that neuroimaging could detect. We earlier suggested that, for recognition1, this might be presentation of words versus nonwords, or famous versus novel faces. For recognition3, this might be source memory judgements (such as asking when the last time a given word was encountered) or a cued recall task. The neural activity accompanying the exercise of these capacities, DPT-A would predict, must be the same as the activity that occurs when people recognise2. This evidence would be analogous to the observation that Paul's opening the door consists of the same muscle contractions as those involved in Paul's reaching for the doorknob and Paul's turning the same.

Could neuroscience adjudicate between the causal theories, DPT-C and SPT-C? This is a more difficult question although it might also be the more interesting. Computational accounts propose functional analyses of the mechanisms underlying recognition2. The difference between DPT-C and SPT-C turns on alternative functional accounts. To the extent that the functional components of these different accounts map relatively simply onto different brain structures (Henson, 2005), then it might be possible to determine which causal model is most plausible. However, as single- and dual-process models are detailed and complex and because these descriptions of what the brain does are couched in vague computational formalisms (with respect to neural function), we believe it may be impossible for neuroscience alone to unambiguously discriminate between them.

Our analysis thus sheds light on current memory research in the neuroscience domain. We suggest that much of this research assume DPT-A to be true and, on this basis, that therefore DPT-C must be true, and finally that there is unambiguous mapping of DPT-C constructs to neural constructs. Yet this research strategy cannot, in principle, yield evidence that might adjudicate between DPT-A and SPT-A (because they are not causal theories) and cannot, in practice, adjudicate between DPT-C and SPT-C (because they are not neutrally specified). We demonstrate this in the context of a particular study that illustrates the problem we have identified.

A large proportion of cognitive neuroscience studies of recognition memory have employed a methodology based on the remember/know paradigm or a variant of it (Vilberg & Rugg, 2008). Using present terminology, this constitutes an attempt to analyse a recognition2 judgement into its component recognition1 and recognition3 judgements according to DPT-A along with a set of bridging assumptions that link this theory to the remember/know paradigm itself.[2] A recent study by Daselaar, Fleck, and Cabeza (2006) serves to illustrate this approach.[3] Participants in this study were given a standard recognition memory test and asked to use a 6-point scale to rate their confidence that a test item had been presented in the study list. This is a very common way of measuring recognition memory and can be described by mathematical models based on either DPT-A or SPT-A (e.g., Heathcote, Raymond, & Dunn, 2006; Yonelinas, 1994). Daselaar et al. analysed this scale in a similar manner to that applied to remember/know responses. They argued that if participants are able to recollect an item from the study phase (recognition3) then they will be very confident that the item was studied. Therefore, all responses based on recollection will be in the highest confidence category (category 6 on the scale). If participants cannot recollect the item then they must rely on its familiarity. The feeling of familiarity is said to be graded, and its strength will lead participants to decide to use any one of the six response categories, including the highest. It follows from this version of DPT-A that responses in categories 1–5 reflect different degrees of familiarity while responses in category 6 reflect both recollection and high levels of familiarity. Daselaar et al. further argued that brain regions involved in the computation of familiarity would show an increasing (linear) pattern of activity across categories 1–6 (what they called a 'familiarity-related' pattern) while brain regions involved in the computation of recollection would show a relatively flat response across categories 1–5 coupled with a significant increase in activity for category 6 (a 'recollection-related' pattern). Similar methods have been developed and applied by other researchers (e.g., Woodruff, Hayama, & Rugg, 2006; Yonelinas, Otten, Shaw, & Rugg, 2005).

Daselaar et al. (2006) found one set of brain regions that showed a recollection-related pattern and another set of brain regions that showed a familiarity-related pattern (they also observed a third 'novelty-related' pattern in which activity decreased linearly from categories 1 to 6). They concluded that their 'findings demonstrate the existence of different brain regions that are differentially involved in recollection, familiarity, and novelty processes. This finding supports the recollection/familiarity distinction and suggests that these processes are independent' (p. 1910).

The Daselaar et al. (2006) study demonstrates precisely the conceptual difficulties we have identified. Their conclusion is either a claim about DPT-A, in relation to SPT-A, or it is a claim about DPT-C in relation to SPT-C. If it is a claim about DPT-A then it is a claim about how recognition2 is *constituted*. But nothing they have found speaks to the constitution of recognition2. All that they have found is that when recognition2 memory responses are partitioned in a particular way (by 'confidence'), different brain regions appear to be differentially involved. This is relevant to the validity of DPT-A only if we knew, from other evidence, that these regions are selectively related to recognition1 and recognition3. However, it is not known a priori if the relevant regions reflect familiarity (recognition1) or recollection (recognition3). The only evidence that is provided is that they are correlated with a partition of recognition2 responses based on the validity of a particular instantiation of DPT-A. The circularity of this reasoning is apparent.

Can the conclusions reached by Daselaar et al. (2006) be construed as a claim about DPT-C? If so, then it is a claim both about how recognition2 is *causally enabled* and about the brain regions involved in this. Critically, DPT-C uses the terms recollection and familiarity as technical terms, divorced from their ordinary meanings of things people do. But, as noted earlier, the kinds of functional models that we currently possess (of these terms as things brains do) have not been worked out to a level of detail in relation to their implementation (how brains do these things) that would allow us to say, for example, that one theory (DPT-C) predicts that brain regions A, B, and C must be involved while another theory (SPT-C) predicts that brain regions X, Y, and Z must be involved. Without these bridging assumptions, it is difficult to conclude that the results found by Daselaar et al. selectively support one or the other theory. As Wais (2008) has noted, Daselaar et al.'s data are equally consistent with SPT-C; their result can be taken to have identified regions that are simply differentially sensitive to items with high global matches, or 'strong memories' in the technical sense. It thus appears that these sorts of imaging studies have no bearing on the debate between DPT and SPT regardless of whether they are -A or -C type theories.

CONCLUSION

In the behavioural domain, we observe the behaviour of people. In the neuroscience domain, we observe the behaviour of brains. Cognitive capacities, such as recognition memory, can be explained in two different ways: constitutively, by analysing them in terms of other cognitive capacities that people may exercise, and causally, in terms of the existence of functional mechanisms (computations) or associated brain behaviours that constitute some of the relevant causally enabling conditions. Dual-process and single-process models of recognition may therefore be understood as offering either constitutive or causal explanations and, we have argued, in the current state of development of the field, dual-process theories have been primarily concerned with a constitutive explanation while single-process theories have been primarily concerned with causal explanations. We have further argued that neuroscience investigations often fail to connect with this debate at either level. If the debate is between different constitutive accounts (DPT-A vs SPT-A), neuroscience investigations need to be predicated on a well-accepted understanding of the conditions under which people exercise their abilities to recognise1, recognise2, and recognise3. Analyses based on the remember/know paradigm or its variants are inadequate because they depend upon the prior acceptance of DPT-A, the validity of which is the subject of investigation. Similarly, in order to connect with the debate between different causal accounts (DPT-C vs SPT-C), neuroscience investigations would need to focus on the mutual constraints of brain structure and function predicated on a well-accepted understanding of the relationship between different computational algorithms (e.g., of global match or of recollection) and their implementation in associated brain regions.

If our analysis of cognitive models as being either constitutive (type 'A' theories) or causal (type 'C' theories) is correct in general, then we can extend our conclusion about the role of neuroscience in testing psychological theories. To the extent that the theories in question are constitutive, there is no role for neuroscience to play. To the extent that the theories in question are causal, but are specified only functionally, there is again no role for neuroscience to play. To the extent, however, that there are psychological theories that make claims about the way neural processes form mechanisms that provide functional processes that in turn allow people to possess and exercise their cognitive capacities, then, and only then, can neuroscience test psychological theories.

ACKNOWLEDGEMENTS

The support of the Australian Research Council (Grant DP: 0877510 awarded to the authors and Ben Newell) is gratefully acknowledged. We also thank Jonathon Trigg and Ben Newell for many useful discussions on and about the issues outlined in the paper.

NOTES

1. By 'mechanism', we mean what Kenny (1976) identifies as the *vehicle* for the possession of a capacity. For a capacity to exist, some system (a natural object, an animal, a person) must possess it. A vehicle is whatever it is about a system such that it can possess it. Kenny's classic examples include alcohol, which is the vehicle for whisky's capacity to intoxicate; shape, which is the vehicle for key's capacity to fit a lock; an efficient carburettor, which is the vehicle for a car's capacity to go faster than some other car. The concept of a vehicle is an extension of Aristotle's point that for every potentiality (in this case, a capacity) there is an actuality (a vehicle). The familiar structural analysis of a system (Cummins, 2000) is an explanation by vehicle. We use the term 'mechanism' because of its familiarity.

2. Specifically, if it is assumed that 'remember' responses are based on recollection (recognition3) and 'know' responses are based on familiarity (recognition1) in the absence of recollection.

3. The risk of choosing just one study is to loose generality; the risk of choosing more than one is to loose clarity. We hope that the general nature of our conceptual analysis will serve to contextualise this one chosen example.

REFERENCES

Bennett, M., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Blackwell.

Cummins, R. (2000). How does it work vs. what are the laws: Two conceptions of psychological explanation. In F., Keil & R., Wilson (Eds.), *Explanation and Cognition*. (pp. 117–144). Cambridge, MA: MIT Press.

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *Journal of Neurophysiology*, *96*(4), 1902–1911.

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*(2), 452–478.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*(2), 524–542.

van Eck, D., De Jong, H. L., & Schouten, M. K. D. (2006). Evaluating new wave reductionism: The case of vision. *The British Journal for the Philosophy of Science*, *57*(1), 167–196.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309–313.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1–67.

Hacker, P. M. S. (2007). *Human nature: The categorical framework*. Oxford: Blackwell.

Heathcote, A., Raymond, F., & Dunn, J. C. (2006). Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language*, *55*(4), 495–514.

Henson, R. (2005). What can functional neuroimaging tell the experimental psychologist? *The Quarterly Journal of Experimental Psychology*, *58A*(2), 193–233.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*, 96–101.

Hyman, J. (1994). Vison and power. *The Journal of Philosophy*, *91*(5), 236–252.

Kenny, A. (1976). *Will, freedom and power*. New York: Barnes & Noble.

Mandler, G. A. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609–626.

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). *Psychological Review*, *114*(1), 188–202.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect of word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(2), 294–320.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*(7), 313–319.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145–166.

Trigg, J., & Kalish, M. (2011). Explaining how the mind works: On the relation between cognitive science and philosophy. *Topics in Cognitive Science*, *3*(2), 399–424.

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia*, *46*(7), 1787–1799.

Wais, P. E. (2008). fMRI signals associated with memory strength in the medial temporal lobes: A meta-analysis. *Neuropsychologia*, *46*, 3185–3196.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of Remember/Know judgments. *Psychological Review*, *117*(4), 1025–1054.

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*(1), 125–135.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(6), 1341–1354.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology. General*, *130*(3), 361–379.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, *34*(5), 622–643.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, *133*(5), 800–832.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience*, *25*(11), 3002–3008.

Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *Neuroimage*, *21*(2), 789–800.