Cell PRESS

A unitary signal-detection model of implicit and explicit memory

Christopher J. Berry[1], David R. Shanks[1] and Richard N.A. Henson[2]

[1] Division of Psychology and Language Sciences, University College London, Gower Street, London WC1E 6BT, UK
[2] MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, CB2 7EF, UK

Do dissociations imply independent systems? In the memory field, the view that there are independent implicit and explicit memory systems has been predominantly supported by dissociation evidence. Here, we argue that many of these dissociations do not necessarily imply distinct memory systems. We review recent work with a single-system computational model that extends signal-detection theory (SDT) to implicit memory. SDT has had a major influence on research in a variety of domains. The current work shows that it can be broadened even further in its range of application. Indeed, the single-system model that we present does surprisingly well in accounting for some key dissociations that have been taken as evidence for independent implicit and explicit memory systems.

Implicit and explicit memory

A popular view of memory is that there are functionally and neuroanatomically distinct explicit and implicit memory systems in the brain [1,2]. Explicit (declarative) memory is thought to be accessible to awareness, whereas the contents of implicit (non-declarative) memory are unconscious. The majority of evidence in favour of this 'multiple-systems' view has attempted to show that performance on particular tasks, thought to be driven by either implicit or explicit memory, can be dissociated from one another. The two most commonly compared tasks measure long-term repetition priming (henceforth priming, taken to index implicit memory) and recognition (taken to index explicit memory). Priming refers to a change in behavioural response to a stimulus after re-exposure. This change often takes the form of an improvement in performance, such as shorter identification times, or increased identification rates of stimuli presented in degraded form. Recognition refers to the capacity to judge whether an item has been previously encountered in a particular context.

Given the large amount of research that has been conducted over the past 20 years (for a review of the topic see Ref. [3]), it is surprising that there are few computational models that have been applied to both implicit and explicit memory. Computational models offer many benefits: they promote theoretical transparency and can be used to generate quantitative predictions that can be tested. Furthermore, they can often indicate alternative interpretations of dissociations [4–6]. Here, we propose that many dissociations between priming and recognition, which, on the surface, seem indicative of multiple-memory systems, are in fact not inconsistent with a single-system account. To illustrate, we review recent work that we have conducted with a simple (and easy-to-implement) single-system model. The model extends signal-detection theory (SDT) of recognition (Box 1) to priming (and also to fluency, another traditional implicit memory phenomenon) [7,8] (Figure 1). A single memory-strength signal drives priming and recognition in the model but, crucially, this signal is subjected to independent sources of non-memorial noise for each task. Even if ultimately shown to be inadequate, we believe that our exploration of a single-system model demonstrates the value of using formal models in implicit and explicit memory research.

Simulating dissociations

Functional dissociations

Consider the often reported dissociation that a manipulation affects recognition but has little or no effect on priming. In normal adults, examples of variables that produce this result are depth of processing manipulations (e.g. making a semantic versus non-semantic judgment about a word at encoding) [9,10] or attentional manipulations (e.g. encoding words with or without a concurrent distractor task) [11]. A common interpretation of this dissociation is that explicit memory is selectively influenced by the manipulation, whereas implicit memory is not. Often, in cases in which a (similar but smaller) effect is also detected on priming, the effect is explained by saying that the priming measure is 'contaminated' with explicit memory. Both of these interpretations postulate more than one memory system (or source of memory) to explain the dissociation.

The single-system model can explain this type of dissociation by postulating only one source of memory [8]. In the model, each item in the test phase is associated with a single memory strength value (f) that is sampled from a normal distribution, the mean of which is assumed to be greater for old (studied) than new (unstudied) items. The value of f of an item is used to generate its recognition judgment and its priming measure. Crucially, this value of f is combined with one randomly sampled noise value for each task. These sources of noise are independent of the memory signal and can, therefore, be conceptualized as non-memorial influences on task performance.

The functional dissociation described earlier can be produced by the model if it is assumed that the variance of the noise associated with the priming task (perceptual

Corresponding author: Berry, C.J. (christopher.berry@ucl.ac.uk).

Box 1. Theories of recognition memory

SDT [26] has played an important part in theories of recognition (see Ref. [27] for a review). In standard signal-detection models of recognition, old (studied) and new (non-studied) items are represented as overlapping Gaussian distributions on a single 'strength of evidence' continuum (Figure I). Because of the influence of the study phase, the mean strength of the old item distribution is assumed to be greater than that of new items. Typically, a participant is assumed to decide whether an item is old or new by assessing its strength relative to a decision criterion located at some point along this continuum. If the strength of the item exceeds the criterion then it will be judged old, otherwise it will be judged new.

A controversial issue in recognition memory research is whether there are qualitatively distinct familiarity and recollection processes in recognition. Familiarity is thought to be context-free, whereas recollection involves retrieval of specific details of the study episode. In an influential dual-process model, familiarity is represented as a continuous strength variable (as in standard SDT) and recollection is represented as an independent high-threshold component [28]. Methods have been proposed to enable one to obtain separate estimates of familiarity and recollection (e.g. analysis of receiver operating characteristics and remember and know judgments). However, the dual-process model has been challenged by 'single-process' SDT models, in which recognition is based upon a single dimension of memory strength and the variance of old and new item strengths are unequal (e.g. Refs [27,29]; although see Ref. [30]). The unidimensional SDT model has been successful in accounting for a wide range of data previously taken to support the dual-process model.

Given the success of SDT in accounting for recognition, it seems logical to extend SDT to account for implicit memory phenomena such as repetition priming. According to some dual-process theories of recognition, repetition priming is one basis of familiarity [9,18]. Understanding the extent to which implicit memory can be accounted for by SDT will therefore have implications for theories of recognition.

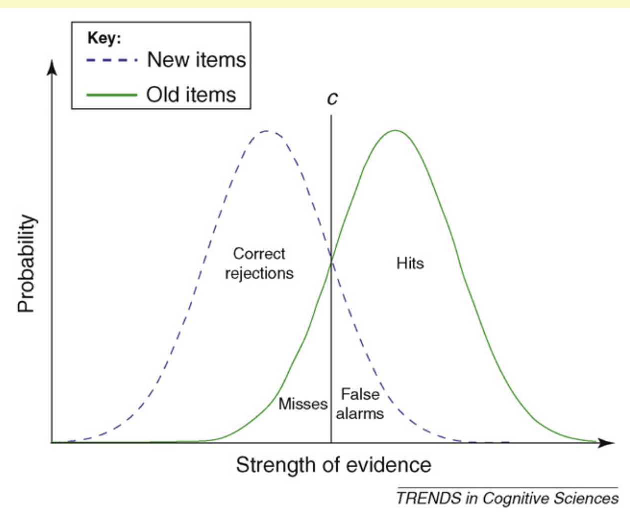

Figure I. Standard SDT of old and new recognition judgments. *c* represents the decision criterion. Items to the right of *c* will be judged old, and those to the left of *c* will be judged new. An old item is classified as a hit if judged old, or a miss if judged new; a new item is classified as a false alarm if judged old, or a correct rejection if judged new.

identification) is usually greater than that of recognition. This assumption is supported by the typically lower inter-trial reliability of priming tasks, compared with recognition [12] (Figure 2a), and is consistent with the view that the influence of non-memorial factors is greater in priming tasks than in recognition [13]. From this assumption it follows that manipulations which increase overall memory strength (i.e. increase the difference between mean f values for old and new items) will be less likely to affect priming than recognition. In sum, this dissociation can arise in the model because the memory signal in priming is overshadowed by a higher degree of noise.

The assumption that the variance of the noise associated with priming is typically greater than that associated with recognition leads the model to make another interesting prediction: when tasks are comparable and performance is measured on the same response metric, it predicts that the sensitivity (e.g. d') of priming tasks cannot exceed that of recognition. Evidence against this prediction would therefore constitute evidence against the model; it would also support the notion that the content of memory supporting priming is unconscious (Box 2).

Other kinds of functional dissociations have been taken to support the multiple-systems view. For example, changes in modality between study and test can produce larger reductions for priming than for recognition, whereas generating versus reading an item at study can improve recognition, but impair priming (see Ref. [3]). Particularly compelling is the demonstration of such crossed double dissociations in parallel with associations between priming and recognition tasks (e.g. Ref. [14]) – so-called 'reversed associations' [15].

The model has not yet been applied directly to these dissociations; however, they could be explicable by a simple extension of the model to two or more distributions of memory strengths. One distribution could reflect an amodal, conceptual memory signal, whereas others could reflect modality-specific memory signals. When combining the signals from these different sources to make a decision, recognition tests might typically place more emphasis on the conceptual signal, whereas priming tests might place more emphasis on modality-specific signals. This could explain the earlier described dissociations between recognition and priming which are produced by changes in modality of presentation between study and test, or by reading versus generating items at study. Importantly, however, none of these memory signals would correspond to an explicit or implicit memory system *per se*; under other conditions (e.g. different task instructions), recognition decisions could be more heavily influenced by a modality-specific memory signal, and priming decisions could be more heavily influenced by an amodal memory signal. This proposal resembles the 'transfer appropriate conceptual–perceptual processing' account that has also questioned the implicit–explicit memory systems account (e.g. Ref. [16]).

In sum, as others have pointed out (e.g. Refs. [12,15]), many functional dissociations do not necessarily imply distinct implicit and explicit memory systems.

Amnesia

Damage to the medial temporal lobe (MTL)/hippocampal regions results in amnesia and impairments in recognition, but leaves priming relatively unaffected (compared with controls). This striking dissociation is often considered to be the most compelling evidence for the multiple-systems view. From a multiple-systems perspective, it indicates that the MTL is the site of an explicit memory system that drives recognition but not priming (see also Box 3).