Double dissociation, modularity, and distributed organization

John A. Bullinaria[a] and Nick Chater[b]
[a]*Department of Psychology, Birkbeck College, London WC1E 7HX, United Kingdom;* [b]*Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom.* **johnbull@ed.ac.uk; nick@psy.ox.ed.uk**

Abstract: Müller argues that double dissociations do not imply underlying modularity of the cognitive system, citing neural networks as examples of fully distributed systems that can give rise to double dissociations. We challenge this claim, noting that such double dissociations typically do not "scale-up," and that even some single dissociations can be difficult to account for in a distributed system.

Müller provides an extremely valuable and wide-ranging analysis of the biological evidence concerning the basis of human language, but we believe that he overstates the difficulties of inferring modular organization from neuropsychological evidence (sects. 5.1 and 6.2). In traditional cognitive neuropsychology, double dissociation (DD) is assumed to imply modularity of function. This inference and its numerous exceptions have been discussed at length by Shallice (1988; see also *BBS* multiple book review Précis of *From neuropsychology to mental structure,* 14(3) 1991). Contrary to this, Müller suggests that fully distributed systems can also give rise to double dissociations, citing a number of connectionist simulations (including a study by one of the present authors, Ganis & Chater 1991).

Recent work suggests that producing DDs in distributed systems is not that straightforward. Consider the classic DD in reading: surface dyslexics can read nonwords much better than irregular words, whereas phonological dyslexics show the opposite profile. This has been taken to imply that there are distinct lexical and sublexical routes in reading (e.g., Coltheart et al. 1993). Early single-route connectionist models of reading (e.g., Seidenberg & McClelland 1989) that can pronounce both regular and irregular words and nonwords cast doubt on the need for a dual route model, and on the DD inference more generally (sects. 5.1 and 6.2). However, a more careful analysis of how these connectionist models operate and respond to damage reveals a natural regularity effect that leads to realistic surface dyslexia symptoms, but no possibility of exhibiting symptoms anything like phonological dyslexia (Bullinaria 1994). A similar pattern arises for the corresponding developmental effects. Thus there are no reliable DDs in these models.

A more detailed study (Bullinaria & Chater 1995) of quasiregular (rules plus exceptions) mappings in connectionist models showed how the DDs that could be obtained in small-scale networks evaporated into single (exceptions lost) dissociations when they were scaled up to more realistic "fully distributed" networks in which no single "neuron" had a significant effect on the network's output. For a wide range of types of damage, the networks showed selective impairment of exceptions, while regulars were relatively preserved, but never the reverse pattern. These effects were seen to be largely a simple consequence of associations being learned more quickly and accurately the more regular they were. In this sense, "regularity" is a measure of rate of occurrence between contradictions, so high-frequency associations must be considered more regular than low frequency associations. A further complication arises because rich error-correcting representations will naturally be more robust than minimal sparse representations. We therefore have the possibility of complementary dissociations in a single globally damaged network resulting from differential rates of performance degradation due to separate "regularity" and "representation richness" effects. Such trend dissociations are well known, however, and can be ruled out by careful definition of DD requiring "crossover" (see Shallice 1988, sect. 10.5). Thus, there are no reliable DDs in these models.

This does not mean that we can only obtain DD by the specific loss of one of two "modules" that function *in parallel,* as in traditional dual-route models of reading. It is quite possible for the DD to occur as a result of damage to one of two modules operating *in series* (Shallice 1988, sects. 10.5 and 11.2). For example, Plaut (1995) shows how damage at two different locations of a single-route recurrent network model can result in a DD between the reading of concrete and abstract words. The DD does not imply that the concrete and abstract words are processed separately, but rather that they are differentially susceptible to damage at different levels of processing. Whether one refers to the two levels (i.e., network locations) as different modules is simply a matter of terminology – as long as there is no confusion, it does not matter. Unfortunately, there often is confusion and, as a result, doubt is cast upon the DD inference in situations where such doubt is not justified. As far as we are aware, no recent advances in connectionist modelling require alterations to the general exposition of Shallice (1988).

As noted elsewhere (e.g., Bullinaria & Chater 1995; Plaut 1995), however, connectionism allows us to study neuropsychological phenomena in finer detail than does the traditional "box and arrow" account. Moreover, it promises to help us model compensatory factors (such as relearning) in the performance of neuropsychological patients, which, as Müller correctly points out, may cause considerable difficulties in interpreting neuropsychological data (e.g., see Plaut 1994).