

DISCUSSION

INFERENCE AND EXPLANATION IN COGNITIVE NEUROPSYCHOLOGY

Max Coltheart¹ and Martin Davies²

(¹Macquarie Centre for Cognitive Science, Macquarie University; ²Philosophy Program,
Research School of Social Sciences, Australian National University)

The question posed by Dunn and Kirsner (D&K) is an instance of a more general one: What can we infer from data? One answer, if we are talking about logically valid deductive inference, is that we cannot infer theories from data. A theory is supposed to explain the data and so cannot be a mere summary of the data to be explained. The truth of an explanatory theory goes beyond the data and so is never logically guaranteed by the data. This is not just a point about cognitive neuropsychology, or even about psychology in general. It is a familiar point about all science.

Suppose that T1 is a theory about the structure of some cognitive system, that D is a set of data including a double dissociation, and that a cognitive neuropsychologist, C, offers D as evidence in support of T1. D&K want to encourage debate about the use of double-dissociation arguments in cognitive neuropsychology. But it is very unclear just what claim they would take C to be advancing. In their article, at least four possible claims seem to be countenanced:

- (a) D *invites* the inference that T1 is true
- (b) D *supports* the inference that T1 is true
- (c) D *implies* that T1 is true
- (d) If D is observed, T1 *must be* true

These are different claims, but D&K make no attempt to distinguish between them. So it is difficult to know just what they would like to have a debate about.

Unless cognitive neuropsychology proceeds in a way that is different from all the rest of science, C cannot be claiming that D logically guarantees the truth of T1. So claims (c) and (d) are not at issue. Nor is C claiming that T1 can be extrapolated from D. The inference of theory from data is not like the inference from several swans being black to the next swan being black or to all swans being black. The inference from data to theory is neither deductive nor inductive, but abductive. It is *inference to the best explanation* (Lipton, 1991).

Our cognitive neuropsychologist, C, is claiming that T1 provides a good explanation of the data in D – the double dissociation, in particular, is just what one would expect if T1 were true – and that no other available candidate theory provides an explanation that is as good. This may be because there is no competing theory that is even consistent with D. But T1 might provide the best explanation of D even though there are competing theories that are consistent with D. It might be, for example, that one competing theory, T2,

offers an explanation of D that is implausible on independent grounds while another competing theory, T3, though strictly speaking consistent with D, would make the occurrence of that data extraordinarily improbable (recall Bayes's theorem).

It surely cannot be that D&K are denying that one theory might provide a better explanation than competing theories for a set of data including a double dissociation. So we remain unclear what *general* feature of cognitive neuropsychological practice it is that they are concerned about. We also consider that some of D&K's specific claims are incorrect; space permits us to mention only three of these

Patients need to be pure cases and tasks need to be process pure for inferences to be made from double dissociations. Both of these claims are wrong. Take the double dissociation between exception word reading and nonword reading that is argued to provide support for a dual-route model of reading (Coltheart et al., 2001). Selective impairment of exception word reading is surface dyslexia: a pure case of surface dyslexia would be one in which all tasks other than exception word reading were performed normally. No such patient has ever been reported. Selective impairment of nonword reading is phonological dyslexia: a pure case of phonological dyslexia would be one in which all tasks other than nonword reading were performed normally. No such patient has ever been reported. If cognitive neuropsychology required that all cases be pure cases, the discipline could not exist. There are no pure cases because the ways in which the brain can be damaged simply do not allow the result of brain damage to be impairment of only one cognitive function. D&K say nothing about why they think pure cases are required, nor do they seem to be aware that pure cases are never found. Nor does process purity hold. A dual-route model of reading does not assert that the tasks of reading exception words aloud and reading nonwords aloud use completely different modules. The letter identification module is used by both reading routes; so is the speech production (phoneme level) module.

Since any two tasks, different enough to be called different, cannot recruit exactly the same mental functions in the same way, it is inevitable that they will eventually yield a dissociation. This is also wrong. Consider the tasks of spelling nouns to dictation and spelling verbs to dictation. Are they different enough to be called different? What about the tasks of reading aloud words of Romance origin versus reading aloud words of Germanic origin – are they different enough? Or reading words containing the letter A versus words not containing the letter A? Which of these task pairs are different enough to yield a dissociation? This is not something that can be determined *a priori*. Whether a dissociation can occur is determined by the nature of the relevant cognitive processing system; it cannot be predicted simply by considering the nature of the tasks. Reading nonwords aloud and reading exception words aloud can be defined as two different tasks, but some have argued that these two tasks recruit exactly the same mental information-processing system (Friedman and Kohn, 1990; Seidenberg and McClelland, 1989).

It would be fallacious to conclude that since double dissociations exist, modules must also exist. The claim that D&K make here is correct; but what is

incorrect is the suggestion that it needs to be made. As we have already noted, cognitive neuropsychologists never argue that a particular theory about cognitive architecture *must* be true given the occurrence of a particular double dissociation. Studies in which connectionist networks are lesioned sometimes show that it is possible for a double dissociation to arise from damage – even random damage – to a single network. Such a demonstration shows that a theory that does not postulate distinct components implicated in the two tasks is consistent with the occurrence of the double dissociation. But it certainly does not show that a theory of that kind provides the best explanation of the double dissociation. So it does not show that any cognitive neuropsychologist's abductive inference is fallacious.

In our view, D&K have not presented an accurate picture of the way in which double dissociations are in practice put to use in cognitive neuropsychology. So we provide an example that illustrates the typical use. Plaut et al. (1996) constructed a computational model of reading aloud in which there was a single procedure for converting print to speech which read both exception words and nonwords accurately. They appreciated the fact that the double dissociation between impaired exception word reading (surface dyslexia) and impaired nonword reading (phonological dyslexia) was *prima facie* evidence against this single-route model. So they attempted to discover ways in which the trained connectionist model could be lesioned to generate selective impairment of exception word reading and selective impairment of nonword reading. But their attempts to lesion the model so as to generate impaired exception word reading, with regular word and nonword reading intact, were unsuccessful. So they «acknowledged that the dramatic pattern of pure surface dyslexia (i.e. normal reading aloud of regular words and nonwords coupled with a severe, frequency-sensitive deficit on exception words) seems unlikely to arise from damage to the kind of single, direct [orthography-to-phonology] pathway developed thus far. 'Lesions' to the network sufficiently severe to reproduce the appropriate degree of impairment on exception words also disrupt the model's performance on regular and nonwords ... lesions to such networks do not reproduce the pattern of pure surface alexia. We therefore acknowledge that another 'pathway' or, as we have characterised it, another source of input to phonology appears to be necessary to model surface alexia» (Patterson et al., 1996, pp. 181 and 188-189). Here the dissociation data are used to reject one model of reading aloud and support another in which an additional component is postulated; and the inference seems to us methodologically impeccable. We wonder what D&K might find problematic in this piece of cognitive-neuropsychological science.

REFERENCES

- COLTHEART M, RASTLE K, PERRY C, LANGDON R and ZIEGLER J. DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108: 204-256, 2001.
- FRIEDMAN RB and KOHN SE. Impaired activation of the phonological lexicon: Effects upon oral reading. *Brain and Language*, 38: 278-297, 1990.
- LIPTON P. *Inference to the Best Explanation*. London: Routledge, 1991.
- PATTERSON K, PLAUT DC, McCLELLAND JL, SEIDENBERG MS, BEHRMANN M and HODGES JR. Connections and disconnections: A connectionist account of surface dyslexia. In JA Reggia, E

- Ruppin and RS Berndt (Eds), *Neural Modelling of Brain and Cognitive Disorders*. New York: World Scientific, 1996.
- PLAUT DC, SEIDENBERG MS, MCCLELLAND JL and PATTERSON K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103: 56-115, 1996.
- SEIDENBERG MS and MCCLELLAND JL. A distributed, developmental model of word recognition and naming. *Psychological Review*, 96: 523-568, 1989.

Max Coltheart, Macquarie Centre for Cognitive Science, Macquarie University, Sydney, NSW, 2109, Australia.
e-mail: max@maccs.mq.edu.au