Theoretical Relevance of Neuropsychological Data for Connectionist Modelling

Mauricio Iza
University of Málaga-Avda
Cervantes, 2 - 29071 Málaga, Spain, tel. 952 13 10 00
iza@uma.es

Alexandra Konstenius
University of Helsinki, Finland

user2@yourdomain.com

Abstract:
The symbolic information-processing paradigm in cognitive psychology has met a growing challenge from neural network models over the past two decades. While neuropsychological evidence has been of great utility to theories concerned with information processing, the real question is, whether the less rigid connectionist models provide valid, or enough, information concerning complex cognitive structures. In this work, we will discuss the theoretical implications that neuropsychological data posits for modelling cognitive systems.
Keywords: Connectionist representation, Neuropsychological deficits, computational neuroscience

1. Introduction

During the past 20 years neural network models has proved to be a growing challenge to the previous theoretical approach, the computer-inspired, symbolic information-processing paradigm. The connectionist approach is in some part grounded on neuropsychological empirical evidence. Neuropsychology has provided detailed information of the patterns of breakdown of cognitive systems resulting from brain damage, and demanded that theories of normal performance should be able to answer to such findings [8,43]. Neural network models are constructed from a transparently perturbable "computational substrate", which permits the performance to break down gradually. Neural networks can therefore meet the demand of modelling both normal performance as well as system failure. The same terrain has, in general, proved to be rather rocky for the traditional symbolic approach, since such models are not obviously perturbable in a graded mood [5,6,18,28,36].

There is an additional demand, placed on the cognitive theory, namely, that to explain the encoding and use of language by the physical organism, i.e., to make the correlation between function and anatomy. This results in a break down between the discipline boundaries [42,31]. In this new interdisciplinary framework, computation can play a crucial role [1,7,11]. Previously, computer models of normal language processes have addressed the questions of modularity and cognitive subprocessing at a descriptive level [15-17,32]. Recently, these models have been developed to meet the questions of both cognitive constraints and neurological implementation, particularly with respect to language disorders. As such, any given computational description of a language processing task must now account for both cognitive constraints (e.g., memory space and time limits) and neuroanatomical ones (e.g., impairment naming animals but not tools on patchy bitemporal infarctions from herpes simplex encephalitis).

In this work, we will discuss neuropsychological data and its relevance for cognitive theory in more detail. For example, connectionist models of two subcomponents of the spelling process, which explain, rather than describe, the relevant neuropsychological evidence [19] operate serially, and thus fall within a domain that has been the stronghold of symbolic modeling. The neuropsychological support that emerges for such models is therefore of particular interest: whether the promise shown by connectionist research in understanding disordered cognition can be extended to the problem of serial processes.

2. Theoretical relevance of neuropsychological data

Before 1970, neuropsychology was a subject that has essentially of interest only to clinicians. At that time, cognitive neuropsychology began to develop in a sharply different way from the domain practice. The differences concerned the overall aim of the research, its theoretical paradigm and its empirical practice. The theoretical paradigm mainly employed in cognitive neuropsychology was the information-processing approach, which had developed in human experimental psychology in the preceding decade. Transmission routes which, ideally, carried out no processing as such connected subsystems with verbally specified functions or producing verbally specifiable representations.

Complementing this new theoretical approach was a reversion to the empirical approach. The empirical findings considered optimal or obligatory were no longer the average performance of a large group of broadly similar patients, but the results of detailed case studies of individual patients. This differed from its previous precursor in being based on quantitative experimental paradigms from human experimental psychology. The approach was assumed to provide privileged information on the structure of the cognitive system through the analysis of highly selective deficits in patients who exhibit striking characteristics, in particular so-called "strong" or "classical" dissociation [12,35].

The grain of the empirical analysis fitted well with that of information processing. In theory, there were many different ways in which dissociation, or better a double dissociation, might be explained. In practice, the assumption that they arose from selective damage to a subset only of subsystems within an information processing system worked well. Apparently, the reason that the approach worked was because the cognitive system, or at least part of it, is in some broad sense modular. In this case, the (classical) dissociation was held to provide a privileged view into the modular structure.

Now, the period where modularity as a basic explanatory assumption is being challenged, at least for some aspects of behavior, by connectionist models that are not obviously amenable to modular decomposition, at least at some level of explanation. A priori, the range of types of behavior that occur when given connectionist model is lesioned would be great. In that case, observation of a patient whose behavior was within that range would not provide strong support for the model. More, the relation between the nature of the model and its performance when lesioned could be far from transparent, so that neuropsychological dissociation would lose their heuristic quality. These are additional problems to those frequently alleged against connectionist modeling, such as that the number of variables that need to be set "hand-coded" makes a model's confrontation with empirical data more like curve fitting that theory testing.

However, this suggests that two levels of support might be provided by neuropsychological evidence for a connectionist model. The first would occur when the range of distinct possible behaviors that can theoretically result from a "lesion" is large. Then, the existence of a patient whose performance fits with one of the theoretically possible "lesioned" states would not be of interest unless there are other classes of model which do not predict the existence of the behavior. The second is where variations in lesion type and size produce few quantitative differences in the expected behavior of the model. Here observation of the behavior provides strong support for the model if lesioning other classes of model does not have the same restricted set of expected behaviors.

The implicit assumption of the above argument is that the set of impaired behaviors that a connectionist model may produce when lesioned is larger than its set of unimpaired behaviors. And the observation of impaired performance after a lesion is less useful for model testing than the observation of normal behavior. This assumption may well turn out to be incorrect. Theoretically, a connectionist net, at least a non-recurrent multi-layer net, may be considered as a device for mimicking a particular input-output transformation. So, the observation that the behavior of a non-recurrent multi-layer net provides a good fit to observed human behavior is little more than the concrete realization of the perspective just presented. The situation with respect to impaired behavior is quite different. There is no reason to assume that an ideal statistical estimator of an output domain given an input domain could, when damaged, produce a particular pattern of performance. Thus the observation of the pattern is the direct support for the internal structure of the model.

An empirical support for this argument is the spelling-to-sound processing in reading. Here, models of both connectionist and symbolic varieties give a very good account of the existing data. Thus, the evidence on normal nonword reading that proved very problematic for the first generation of connectionist models in this domain is easily accounted by current models. However, the neuropsychological evidence has proved much more contradictory. Now, it is argued that the normal reader is not capable of reading all pronounceable nonwords through the operation of orthographic-to-phonological processes alone [34]. Even partial semantic and pragmatic mediation is also claimed to be required. This fact suggests that neuropsychological inference may be more useful for distinguishing between different models. Indeed, its utility may even rival that which it had for the earlier generation of classical information-processing models.

Many examples exist of interactive activation or connectionist models that give a reasonable attempt when "lesioned" to syndromes found in the neuropsychological literature [10,25,33]. In other domains where little is known about the space of possible behaviors that the lesioned connectionist model can produce, except that one part of the space corresponds to behaviors observed in patients, there exist symbolic models alternatives. In this way, forms of anomie and deep dysphasia may be modeled in terms of damage to connectionist models that map from the semantic to the phonological representation of words. In this domain, symbolic model alternative exists. However, a domain exists where the sets of possible behavior following "lesions" are both few and well described empirically and no plausible symbolic model alternative exists. This is the domain of orthographic-to-semantic representations, where lesions to a range of connectionist models of an attractor type give rise to a set of behaviors that correspond to the syndrome "deep dyslexia". Thus, a reasonable case can be made for the utility of neuropsychological data in domains where processing can be viewed as the mapping from a single input to a single output pattern.

3. Functional level in computational neuroscience

It can be argued that brain-imaging studies of cognition make no sense until we have identified undecomposable modules –called primitive processors. What makes a processor primitive? One answer is that for primitive processors, the question "How does the processor work?" is not a question for cognitive science to answer. The line between the upper levels of processors and the level of primitive processors is the same as the line between cognitive science and one of the "realization" sciences such as electronic or physiology. The functional analysis of human intelligence will bottom out in primitive processors in the brain.

The term 'functional analysis' refer to abstract theories as 'global accounts', and contend that the only sciences that are currently providing global functional accounts of brain cells are sciences such as psychology. Precisely, claims about the localization of language in the brain are conceptually dependent upon, and make no sense in the absence of, an abstract theory of –for example, an information-processing model of- how language-processing is carried out.

The psycholinguistic analysis of reading disorders has priority over neuroanatomical approaches because, before one can begin to study the difference between one reading disorder and another in terms of lesion sites, one must define the disorders; this can only be done in psycholinguistic terms.

There is another way in which the abstract level has priority over the physical-instantiation level: it can be very hard to understand what a system is actually doing if one's only information about it is a description at the physical-instantiation level. A description at the abstract-theory level will be far more enlightening. In turn, if one wishes to study the physical-instantiation level, a prior description at the abstract level is needed; otherwise one will not know what to look for at the physical-instantiation level (as can be argued in connection with neuroimaging studies of language).

Arguments of this sort are reasons for rejecting the Biological Neuroscience Thesis, which is: (i) the science of the mind is (part of) the science of the cells that constitute the brain. Instead, we can advocate a different thesis, which is: (ii) the science of the cells that constitute the brain is part of the science of the mind. The other part of the science of the mind is, of course, at the abstract-theory (functional) level.

At the present, several models of memory have incorporated evidence from neuropsychological, neuroanatomical and computational modelling studies [20,21]. Each of these models tries to account for the relationship between the wider neocortex and the medial temporal lobe structures (that is, the hippocampus). On the one hand, we could argue that the brain can be thought as a series of hierarchically organized functions zones. At the bottom of the hierarchy are regions that undertake the most basic level of sensory motor processing and code for feature fragments. Next come convergence zones that receive inputs from the sensory motor regions and code for 'coherent entities' (objects) and finally there are zones that code for the co-occurrence of entities that constitutes an event. This latter zone is considered to include the hippocampus and related medial temporal lobe structures. Memory of an event involves the reactivation of the hippocampal-neocortical network initially established during initial encoding of the event. On the other hand, other model suggests that through repeated recollection (and posterior reactivation) over time, event memories become established at the neocortical level, while in the previous model, the hippocampus is seen as always being involved in the recollection of events. However, all the models agree that semantic or factual knowledge is acquired by the gradual formation of associations at the lower levels of the hierarchy, which can take place without the involvement of the higher levels. In that way, people with a damage or dysfunctional medial temporal lobe system can still learn new semantic or factual information, though much more slowly than individuals with an intact hippocampus.

Within this context, other further distinction in memory is that between anterograde and retrograde memory. The first one refers to the acquisition of new information following the onset of injury or disease. The second one refers to memory for facts that were acquired or events that took place before the onset of injury or disease.

The classical memory disorder is the amnesic syndrome that involves: (a) intact working memory, (b) intact general intellectual and semantic memory functioning, (c) anterograde amnesia, (d) a variable degree of retrograde amnesia, and (e) intact implicit memory. This pattern of impairment tends to follow from selective bilateral medial temporal lobe damage. Cases of pure amnesia are relatively uncommon, as most people with brain injury or disease have most extensive damage. For instance, people with traumatic brain injury tend to have less complete destruction of medial temporal lobe structures and, in that way, may have slightly milder memory impairment, but are also likely to have damage to frontal lobe structures and so have associated attentional and executive impairments, which will also impact on memory performance.

Although anterograde memory deficits sometimes occur very little in the way of retrograde impairment, the question of whether retrograde memory impairments can arise in the absence of anterograde memory deficits seem more controversial. On the one hand, a pattern of severe retrograde amnesia, accompanied by a normal ability to acquire new information, is associated with psychogenic forms of amnesia. On the other hand, there have been described some cases of focal retrograde amnesia arising from organic impairment [4].

Selective disorders of working memory are rare. Patients with selective phonological loop deficits are unable to hold in mind more than two digits, words or letters, compared with the normal seven (+/-2). Although they show normal language processing, these patients show normal learning of some information, but there is evidence that long-term memory is impaired when it taps the same information that is impaired in short-term memory [26].

Selective deficits in semantic memory can occur following damage to lateral temporal lobes, following head injury, infection or as a result of a focal dementia process referred to as semantic dementia. Within semantic memory, there is evidence that semantic memories for different categories break down in a dissociable manner. The most commonly reported dissociation is between living and non-living items [22].

4. Discussion

Most people requiring psychological intervention after neurological illness or injury will have some reduction in cognitive skills, depending on the nature, location and severity of damage. After generalized damage (for instance, head injury and multiple sclerosis) the most common cognitive difficulties are with attention, concentration, memory, speed of information processing and executive functioning (that is, planning, problem-solving, self-awareness and self-regulation). After more focal damage (for instance, stroke or tumour) specific deficits may also be seen in motor skills, visual perception, spatial judgment, language functions, etc. The subjective experience of cognitive impairment can be disturbing and perceptual and spatial deficits can be bewildering and frightening. Loss of the ability to communicate effectively (due to language impairment or severe dysarthria) may be immensely frustrating, as well as socially isolating. Marked impairment of memory disrupts daily living and compromises a sense of continuity, as well as perception of progress in rehabilitation. Of particular importance is the disruption to executive function, not only due to its direct effects on independence and self-control but also in the limitations imposed on insight, understanding, use of compensatory strategies and long-term adjustment [9].

The exponential rise in the sophistication of computational neuroscience has encouraged the consideration of their potential value in cognitive rehabilitation. Computational models offer a number of clear advantages. They can deliver highly structured, standardized (that is, replicable) training packages involving attractive graphics/sound and instant feedback on performance. Such packages can be delivered adaptively, at times that are most convenient for patients, and may be easily transferred from the hospital to the patient’s home. Of course, there also are potential disadvantages including whether patients feel at ease with the compensatory technology, whether a ‘one size fits all’ approach is merited for patients with potentially quite different difficulties and, in particular, whether benefits generalize from the computer to everyday life.

Where verbal memory is differentially impaired, the use of drawings to express emotions, the presentation of information in a visual form or encouraging the patient to visualize parts of the session may be useful. Moreover, utilizing ‘negative automatic imagery’ and ‘adaptive imaginary responses’ may be more effective than working with thought-based material. Imagery may have the added advantage of being more emotionally potent due to its ability to ‘simplify meaning into sensory experience’ in a more direct way. On the contrary, relaxation training may be useful in minimizing high general arousal levels, which may impede memory functioning.

In summary, compensatory memory strategies provide support at retrieval and reduce the ‘effort’ or cognitive ‘load’ on explicit memory and attention. The development of a ‘prosthetic environment’ with signs and cues to assist the person and the current growth of assistive devices or ‘smart house technology’ on design and architecture on dementia are examples of broad-based approaches designed to reduce cognitive load and maximize function in people with dementia. The efficacy of these techniques has yet to be established taking into account the above limitations of connectionist modelling. In so far as they minimize the potential for divided attention during an activity, thereby reducing load on central executive and working memory systems, and limit the distractions of environmental ‘noise’, they are likely to generate some benefits.

References

[1] Ajjanagadde, V. & Shastri, L. (1991) Rules and variables in neural nets, Neural Computation 3, 121-134.
[2] Best, W. & Nickels, L. (2000). From theory to therapy in aphasia: Where are we now and where to next? Neuropsychological Rehabilitation 10, 231-247.
[3] Burguess, P.W. & Alderman, N. (2004). Executive dysfunction. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 185-209. Chichester: Wiley.

[4] Cicerone, K.D. (2005). Rehabilitation of executive function impairments. In W.M. High, A.M. Sander, M.A. Struchen & K.A. Hart (eds), Rehabilitation for traumatic brain injury, pp. 71-87. New York: Oxford University Press.
[5] Clark, A. (1989). Microcognition. Cambridge, MA: MIT Press.
[6] Clark, A. (1993). Associative engines. Cambridge, MA: MIT Press.
[7] Clark, A. (1997) Being there. Cambridge, MA: MIT Press.
[8] Cummins, R. (1989). Meaning and mental representation. Cambridge, MA: MIT Press.
[9] Dell, G.S., Lwaler, E.N., Harris, H.D. & Gordon, J.K. (2004). Models of errors of omission in aphasic naming. In M. Behrmann & K. Patterson (eds), Words and things, pp. 125-146. New York: Psychology Press.
[10] Dell, G.S., Shwartz, M.F., Martin, N., Saffran, E. & Gagnon, D.A. (1997).Lexical access in aphasic and non-aphasic speakers.Psychological Review 104, 811-838.
[11] Dennett, D.C. (1996). Kinds of minds. NY: Basic Books.
[12] Ellis, A.W. & Young, A.W. (1992). Human cognitive neuropsychology. London: LEA.
[13] Evans, J.J. (2004). Disorders of memory. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 143-164. Chichester: Wiley.
[14] Evans, J., Wilson, B.A., Needham, P. & Brentnall, S. (2003). Who makes god use of memory aids? Results of a survey of people with acquired brain injury. Journal of the International Neuropsychological Society 9, 925-935.
[15] Fodor, J.A. (1975). The language of thought. NY: Crowell.
[16] Fodor, J.A. (1983). The modularity of mind. Cambridge, MA: MIT Press.
[17] Fodor, J.A. & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. Cognition 28, 3-71.
[18] Franklin, S. (1995). Artificial minds. Cambridge, MA: MIT Press.
[19] Hanley, J.R., Dell, G.S., Kay, J. & Baron, R. (2004). Evidence for the involvement of nonlexical route in the repetition of familiar words. In M. Behrmann & K. Patterson (eds), Words and things, pp. 147-158. New York: Psychology Press.
[20] Hart, J., Berndt, R.S. & Caramazza, A. (1985). Category specific naming deficit following cerebral infarction. Nature 316, 439-440.
[21] Hart, T., Hawkey, K., & Whyte, J. (2002). Use of a portable voice organizer to remember therapy goals in traumatic brain injury rehabilitation: a within-subjects trial. Journal of Head Trauma Rehabilitation 17(6), 556-570.
[22] Laiacona, M. & Caramazza, A. (2004). The noun/verb dissociation in language production: varieties of causes. In M. Behrmann & K. Patterson (eds), Words and things, pp. 103-124. New York: Psychology Press.
[23] LoPresti, E.F., Mihailidis, A. & Kirsch, N. (2004). Assistive technology for cognitive rehabilitation: State of the art. Neuropsychological Rehabilitation 14(1/2), 5-39.
[24] Manly, T. & Mattingley, J.B. (2004). Visuospatial and attentional disorders. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 229-252. Chichester: Wiley.
[25] Martin, N., Dell, G.S., Saffran, E. & Schwartz, M. (1994). Origins of paraphasias in deep dysphasia: testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. Brain and Language 47, 609-660.
[26] Mayes, A.R. (2000). Selective memory disorders. In E. Tulving & F.I.M. Craik (eds) The Oxford Handbook of Memory Disorders, pp. 427-440. Oxford University Press: Oxford.
[27] McCarthy, R.A. & Warrington, E.K. (1988). Evidence for modality-specific meaning systems in the brain. Nature 334, 428-430.
[28] McClelland, J.L. & Rumelhart, D.E. (eds.) (1986). Parallel distributed processing. Cambridge, Massachusets: MIT Press.

[29] McKenna, P. (2004). Disorders of language and communication. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 165-184. Chichester: Wiley.
[30] Moniz-Cook, E. & Rusted, J. (2004). Neurorehabilitation strategies for people with neurodegenerative conditions. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 405-420. Chichester: Wiley.
[31] Moorhouse, M. & Barry, P. (2004). Bioinformatics, biocomputing and Perl. Chichester, Sussex: Wiley.
[32] Pylyshyn, Z. (1984). Computation and cognition. Cambridge, MA: MIT Press.
[33] Saffran, E., Dell, G.S. & Schwartz, M.F. (2000). Computational modelling of language disorders. In M. Gazzaniga (ed), The new cognitive neuroscience, pp. 933-948. Cambridge, MA: MIT Press.
[34] Schwartz, M.F., Wilshire, C.E., Gagnon, D.A. & Polansky, M. (2004). Origins of nonword phonological errors in aphasic picture naming. In M. Behrmann & K. Patterson (eds), Words and things, pp. 159-186. New York: Psychology Press.
[35] Shalice, T., McLeod, P. & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? Quarterly Journal of Experimental Psychology, 37A, 507-532.
[36] Smolensky, P. (1988). On the proper treatment of connectionism. Behavioral and Brain Sciences, 11, 1-23.
[37] Sohlberg, M.M. (2005). External aids for management of memory impairment. In W.M. High, A.M. Sander, M.A. Struchen & K.A. Hart (eds), Rehabilitation for traumatic brain injury, pp. 47-70. New York: Oxford University Press.
[38] Sohlberg, M.M. & Mateer, C.A. (2001). Cognitive Rehabilitation: A neuropsychological perspective. New York: Guildford Press.
[39] Sohlberg, M.M., Ehlhardt, E., Fickas, S. & Sutcliff, A. (2003). A pilot study exploring electronic mail in users with acquired cognitive-linguistics impairments. Brain Injury 17(7), 609-629.
[40] Stanovich, K.E. (2004). The robot's rebellion. Chicago: University of Chicago Press.
[41] van der Broek, M.D., Downes, J., Johnson, Z., Dayus, B. & Hilton, N. (2000). Evaluation of an electronic memory aid in the neuropsychological rehabilitation of prospective memory deficits. Brain Injury 14(5), 455-462.
[42] van Gelder, T. (1990). Compositionality: a connectionist variation on a classical theme. Cognitive Science, 14, 355-384.
[43] Von Eckardt, B. (1993). What is cognitive science? Cambridge, MA: MIT Press.
[44] Wilson, B.A. (2004). Theoretical approaches to cognitive rehabilitation. In L.H. Goldstein & J.E. McNeil (eds), Clinical neuropsychology: A practical guide to assessment and management for clinicians, pp. 345-365. Chichester: Wiley.
[45] Wilson, B.A., Emslie, H., Quirk, K. & Evans, J.J. (2001). Reducing everyday memory and planning problems by means of a paging system: A randomised control crossover study. Journal of Neurosurgery, Neurology and Psychiatry 70, 477-482.
[46] Wright, P., Rogers, N., Hall, C., Wilson, B.A., Evans, J.J., Emslie, H. et al. (2001). Comparison of pocket-computer memory aids for people with brain injury. Brain Injury 15, 787-800.
[47] Ylivsaker, M., Hanks, R. & Jonson-Greene, D. (2003). Rehabilitation of children and adults with cognitive-communication disorders after brain injury. ASHA Supplement 23, 59-72.

Acknowledgements

This work has been partially supported by the project <<Modularidad, Razonamiento y Cognición Social. Un análisis critico de la Psicología evolucionista (FFI2009-08999)>>.