BEHAVIORAL AND BRAIN SCIENCES (2002) **25,** 727–788
Printed in the United States of America

Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling

Michael Thomas
Neurocognitive Development Unit, Institute of Child Health and School of Psychology, Birkbeck College, University of London, London WC1E 7HX, United Kingdom
m.thomas@psychology.bbk.ac.uk

Annette Karmiloff-Smith
Neurocognitive Development Unit, Institute of Child Health, London WC1N 1EH, United Kingdom
a.karmiloff-smith@ich.ucl.ac.uk
http://www.ich.ucl.ac.uk/units/ncdu/NDU_homepage.htm

Abstract: It is often assumed that similar domain-specific behavioural impairments found in cases of adult brain damage and developmental disorders correspond to similar underlying causes, and can serve as convergent evidence for the modular structure of the normal adult cognitive system. We argue that this correspondence is contingent on an unsupported assumption that atypical development can produce selective deficits while the rest of the system develops normally (Residual Normality), and that this assumption tends to bias data collection in the field. Based on a review of connectionist models of acquired and developmental disorders in the domains of reading and past tense, as well as on new simulations, we explore the computational viability of Residual Normality and the potential role of development in producing behavioural deficits. Simulations demonstrate that damage to a developmental model can produce very different effects depending on whether it occurs prior to or following the training process. Because developmental disorders typically involve damage prior to learning, we conclude that the developmental process is a key component of the explanation of endstate impairments in such disorders. Further simulations demonstrate that in simple connectionist learning systems, the assumption of Residual Normality is undermined by processes of compensation or alteration elsewhere in the system. We outline the precise computational conditions required for Residual Normality to hold in development, and suggest that in many cases it is an unlikely hypothesis. We conclude that in developmental disorders, inferences from behavioural deficits to underlying structure crucially depend on developmental conditions, and that the process of ontogenetic development cannot be ignored in constructing models of developmental disorders.

Keywords: Acquired and developmental disorders; connectionist models; modularity; past tense; reading

1. Introduction

Behavioural impairments found in developmental disorders and in cases of acquired brain damage provide a source of information about the structure of the cognitive system. Historically, the logic of deriving implications about cognitive structure from behavioural impairments was formulated in the domain of acquired disorders in adults (see, e.g., Shallice 1988). It was argued that under some circumstances, highly selective patterns of impairment after damage could de-monstrate the relative independence of different cognitive processes, predicated on an a priori assumption of modular structure within the cognitive system. Ultimately, this was thought to lead to the identification of the components of cognition.

Recently, behavioural impairments found in developmental disorders have often been interpreted within the same cognitive neuropsychology framework (see, e.g., Baron-Cohen 1998; Leslie 1992; Temple 1997). In this case, there is an inference that selective behavioural impairments reveal discrete components of the cognitive system that have not developed properly, for example, the purported defective "theory of mind" processor in autism (Leslie 1992), or the defective phonological processor in dyslexia (Frith 1995). However, the extension of the cognitive neuropsychology framework to interpret developmental disorders has proved controversial. Indeed, some researchers (Bishop 1997a; Karmiloff-Smith 1997; 1998) have argued that the process of development itself violates key assumptions of the static cognitive neuropsychology model and thus invalidates the direct inference from impairment to cognitive structure.

Our aim in this article is to evaluate this debate from the perspective of connectionist modelling of cognitive processes. This is a useful perspective because such models

© 2002 Cambridge University Press 0140-525X/02 $12.50
https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

have been employed to capture both acquired deficits (when models of adult performance are damaged) and developmental deficits (when initial computational constraints are altered in models of typical development). Connectionist models therefore provide a concrete computational basis on which to anchor a debate on the potential causes of each type of deficit.

From the stance of the behavioural outcome, the impairments found in cases of acquired and developmental disorders can look very similar. For example, in relation to types of dysgraphia, dyslexia, and dyscalculia, Temple (1997, p. 324) comments that "were one to give the data from adult and child cases to a cognitive neuropsychologist and ask the question, which are the adults and which the children, there are no apparent criteria by which to distinguish them."

One might take this similarity in behaviour as an indication that the two types of disorder are linked at a deeper level, namely, that they share a similar underlying cause. For example, where one can appeal to a static information-processing model of the adult system, one might characterise an impairment in the adult case as corresponding to selective damage to one (or more) processing components, and an impairment in the developmental case as a failure of one (or more) components to be properly acquired. Temple (1997) offers just such a characterisation for cases of developmental prosopagnosia (pp. 139 and 141), as well as two subtypes of developmental dyslexia (pp. 192 and 206), developmental disorders of spelling (pp. 238 and 244), and developmental dyscalculia (pp. 285–86).

In this target article, we argue that such a causal link between acquired and developmental disorders can only occur if, for a given domain, a very particular kind of developmental account holds true. In most cases where researchers have linked acquired and developmental disorders, the required developmental account has not been argued for, but merely assumed.

Our aim in the computational part of this article (sect. 5) is to characterise the conditions that must hold in a developing cognitive system for acquired and developmental disorders to be linked at a causal level. We demonstrate by simulation that, in the absence of a precise developmental account of a cognitive system, behavioural data alone may be insufficient to infer underlying functional structure from a pattern of impairments. As a result, we argue that researchers working with developmental disorders must compare their data against developmental rather than static models, even though those static models may be appropriate for explaining patterns of acquired deficits in normal adults.

First, however, we introduce two concrete examples of domains in which explicit links have been drawn between acquired and developmental impairments. These are the domains of dyslexia and English past tense formation. For current purposes, these areas are important not only because they illustrate how explicit the claims have been about the relation between acquired and developmental impairments, but also because both areas have been the focus of substantial computational modelling work exploring the possible underlying causes of those impairments.

Michael Thomas completed his D. Phil at Universitiy of Oxford on bilingual language processing and connectionist modelling. Previously a Senior Research Fellow at the Neurocognitive Development Unit, Institute of Child Health, London, he has recently accepted a lectureship in Cognitive Neuropsychology at Birkbeck College, University of London. His research focuses on typical and atypical language development, combining empirical and computational approaches to mechanisms of learning.

Annette Karmiloff-Smith is Professor of Neurocognitive Development at the Institute of Child Health, University College London, where she heads the Neurocognitive Development Unit. Originally a student and then collaborator of Piaget at Geneva University, she is author of 8 books, over 200 articles/chapters on infant/child cognitive development and genetic disorders, and in 1995 won the British Psychological Society's book award for excellence in the literature of psychology. Her writings have been translated into Japanese, Chinese, Hungarian, Hebrew, Arabic, Italian, Spanish, French, German, Danish, Norwegian, Dutch and Greek. In the 1990s she was elected Fellow of the British Academy and of the Academy of Medical Sciences.

2. Comparisons of acquired and developmental deficits in two domains

2.1. Dyslexia

When adults experience difficulty in reading following brain damage, their patterns of behavioural impairments can be described according to several subtypes. Two subtypes are of particular relevance. In acquired phonological dyslexia, patients demonstrate particular difficulty in reading nonwords. In acquired surface dyslexia, they show a deficit in reading *exception* words, where the pronunciation cannot be predicted from the usual letter-to-sound correspondence. For these exception words, patients tend to display errors of regularisation, for example, *gauge* pronounced as "gorge," *trough* as "truff," *come* as "kome," and *quay* as "kway" (Shallice et al. 1983).

Cognitive neuropsychologists have interpreted these two patterns as reflecting specific damage to independent subcomponents of the skilled reading system. The traditional information-processing model of the skilled reading system proposes that three processing routes link print to sound (see, e.g., Patterson & Shewell 1987; Temple 1997). One route decomposes written words into their component graphemes and constructs a pronunciation via a system of grapheme-to-phoneme correspondences. This is called the *nonlexical* or *phonological* route. A second *direct* or *lexical* route recognises the whole written word form and uses this representation to recover the whole-word pronunciation. A third, *semantic*, route uses the written word form to recover the word's meaning, and this semantic representation is then used to recover the word's pronunciation.

Assuming this model, acquired phonological dyslexia can be interpreted as a normal adult reading system that has experienced damage to the grapheme-to-phoneme processing route; existing words can be read via the whole-word recognition routes, but reading of nonwords is impaired. Acquired surface dyslexia, however, can be interpreted as an adult system that has experienced damage to both whole-word recognition routes. Words can only be read via decomposition into component graphemes and the appli-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

cation of grapheme-phoneme correspondences, resulting in the regularisation of exception words.

The patterns of errors defining these two subtypes of dyslexia have been reported in children, both in single case studies and in group studies (see, e.g., Castles & Coltheart 1993; Manis et al. 1996). Researchers have employed similar explanations in the developmental case – once more appealing to the structure of the adult reading system, but now replacing the notion of "specific damage" with the notion of "a specific failure to develop." Therefore, in developmental phonological dyslexia, children may be "having difficulty with [acquiring] one or more components of the nonlexical route" (Coltheart et al. 2001) or they may exhibit an overall system with "relatively normal development of semantic, lexical, and direct reading systems but with impairment in the acquisition of the phonological reading route" (Temple 1997, p. 206). In developmental surface dyslexia, children may be having difficulty acquiring "one or more components of the lexical route" (Coltheart et al. 2001), or they may exhibit a reading system in which the "direct and semantic reading routes have failed to become established properly" (Temple 1997, p. 192).

2.2. Past tense formation

Our second example comes from the domain of inflectional morphology, and in particular the formation of the English past tense. Once again, a model has been proposed for the functional structure of the adult system in which separate subcomponents tackle different aspects of the task (Pinker 1991; 1999). One component is claimed to be responsible for forming the majority of past tenses that conform to a rule ("add -ed") and for generating past tenses for novel verbs (wug-wugged). A second component memorises individual past tense forms, particularly those that are exceptions to the rule (e.g., go-went, sleep-slept, hit-hit, etc.).

In cases of acquired aphasia and in neurodegenerative diseases, adults can exhibit dissociations between performance on regular and exception past tense formation. Patients with nonfluent aphasia can be worse at producing and reading regular past tense forms than exception forms, whereas patients with fluent aphasia can be worse at producing and reading exception forms than regular forms (e.g., Tyler et al. 2002a; 2002b; Ullman et al. 1997; in press; though, see Bird et al. 2002). Similarly, patients with Parkinson's disease can make more errors producing regular and novel -ed forms than exception forms, whereas patients with Alzheimer's disease can make more errors producing exception past tense forms than regular past tense forms (Ullman et al. 1997; Ullman, in press). Assuming the dual-mechanism model of the adult system, these patterns of acquired deficit are taken to reflect selective damage to either the rule-processing component or the exception memorisation component.

Once again, parallel impairments have been reported in the developmental domain, in this case in two developmental disorders with a genetic origin, Specific Language Impairment (SLI) and Williams syndrome (WS). Ullman and Gopnik (1999) and van der Lely and Ullman (2001) reported that children with SLI perform poorly on past tense formation tasks and show a much smaller advantage of regular past tense formation over exception past tense formation – interpreted as a relative impairment in regular past formation. However, Clahsen and Almazan (1998) reported that children with WS exhibit a specific difficulty with generating exception past tense forms.[1] Pinker (1999) offered an interpretation of these respective findings in terms of the adult model: SLI represents a case where the mutation of certain genes interferes with the development of the ability to inflect new and uncommon regular verbs. WS represents a case where the rule-based computational mechanism is intact but the memory mechanism for storing exception verbs is specifically impaired. Together these disorders are argued to represent a "genetic double dissociation . . . the first group of children rarely generalise the regular pattern; the second group of children generalise it freely" (Pinker 1999, p. 262).

Both the examples of dyslexia and past tense formation illustrate the way in which developmental impairments are often interpreted by appealing to the structure of adult models. They show, too, how the central double dissociation logic of adult cognitive neuropsychology has been extended to developmental cases. Dissociable behavioural impairments are taken as evidence of independent underlying mechanisms, by virtue of the claimed independent failure of those mechanisms to develop properly. We turn now to consider why the validity of this extension is questionable, before examining specific computational implementations of deficits in these target domains.

3. Is the cognitive neuropsychology framework appropriate for the interpretation of developmental disorders?

3.1. Development in a "static" framework

When acquired damage causes selective cognitive deficits in normal adults, these deficits occur against a background of hitherto normal function. (This is also the case for acquired deficits in children, at least at the time of insult.) Such cognitive systems are hence discussed in terms of the cognitive mechanisms or processes that have become *impaired* compared to those that have remained *intact.*

When selective behavioural deficits are identified in developmental disorders, they are frequently characterised in the same way, in terms of developmental impairments against a background of normal development. We will refer to the second half of this characterisation as the *assumption of Residual Normality.* This is the assumption that, in the face of a selective developmental deficit, the rest of the system can nevertheless develop normally and independently of the deficit. It is this developmental assumption that allows researchers to relate patterns of deficits in developmental disorders to static models of the normal cognitive system. Because patterns of deficits are usually identified in older children, adolescents, or adults with the developmental disorder, static models of the normal adult system are often deemed an appropriate point of reference. In principle, however, deficits identified in the younger child could be compared against a static model of the normal system for the appropriate stage in development, were such a model to exist. In either case, the essential point here is that the assumption of Residual Normality permits developmental deficits to be compared against functional models that *themselves have no developmental component.*

The assumption of Residual Normality has been widely deployed in the study of atypical development, including in the case of disorders such as autism, WS, SLI, dyslexia, dyscalculia, Gilles de la Tourette syndrome, and develop-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

mental prosopagnosia. The following quotes illustrate three explicit renditions of the claim for Residual Normality in developmental disorders:

> I suggest that the study of mental retardation would profit from the application of the framework of cognitive neuropsychology (e.g., McCarthy & Warrington 1990; Shallice 1988). In cognitive neuropsychology, one key question running through the investigator's mind is "Is this process or mechanism intact or impaired in this person?" . . . In fact researchers in mental retardation have been searching for *intact* versus impaired cognitive processes for quite some time without discussing this in terms of modularity. (Baron-Cohen 1998, p. 335, emphasis added; and Footnote 1)

> The analysis of the developmental dyslexias offered by the dual-route model is that, just as each of the two routes can be selectively affected by brain damage with the other remaining intact, it is possible for a child to have difficulty acquiring one of the routes, *with the other being acquired at a normal rate.* (Coltheart et al. 1993, p. 591, emphasis added)

> Within modular theories, the linguistic performance of subjects with [developmental] language impairments may reflect *the architecture of the normal system* but with selective components of this system under- or over-developed. (Clahsen & Temple 2002, emphasis added)

It is interesting that Residual Normality (henceforth RN) is less frequently deployed as a developmental hypothesis in pediatric neuropsychology. For children with acquired brain damage, the clinically driven focus is usually on recovery. Researchers tend to eschew static models and explore the effect of cerebral insult on the potentially plastic process of development. Because structural damage is seen in the context of a dynamic and interactive developmental process, there is recognition of the possible influences of compensation within the cognitive system and of disruption to the acquisition of further cognitive skills, as well as family and social factors (Anderson et al. 2001b; see Thomas 2003, for discussion). If the undamaged part of the cognitive system compensates or alters across development in response to the part that has suffered a selective deficit, the undamaged part may not follow the normal path of development, in which case RN would not hold.[2]

However, researchers in developmental disorders of a genetic origin routinely deploy RN in their explanations, as we have seen in the cases of SLI, WS, and dyslexia. This is probably because of the fact that such disorders are used (in part) for theoretical purposes within the cognitive neuropsychology framework, as a source of evidence about the structure of the normal cognitive system (and because of the genetic origins, about the potential innateness of that structure). RN is an assumption (often implicit) about how development takes place. But, is it likely to be correct? In section 3.2, we consider two opposite claims. First, we examine the claim that *no answer to the preceding question is necessary:* From the perspective of cognitive neuropsychology, development can be ignored in the study of behavioural deficits in developmental disorders. Second, we consider the claim that, not only must development be incorporated, but that when it is, the assumptions required to use the cognitive neuropsychology framework are fatally undermined. We then propose a resolution of these two opposing positions.

3.2. Development and modularity

Jackson and Coltheart (2001) have recently defended the use of the cognitive neuropsychology framework for studying developmental disorders. They have argued that the process of development is not relevant to identifying intact and impaired processes in a cognitive system, so long as modularity can be assumed for that system. In their view, the framework is equally suitable in both acquired and developmental cases for establishing what they call the *proximal cause* of the behavioural impairment. By this they mean "what is wrong with the cognitive system right now," irrespective of whether the original cause was brain damage, atypical development caused by a genetic abnormality, or even poor schooling. These latter causes are what Jackson and Coltheart term *distal.* They maintain that the distal causes of an impairment are potentially *independent* of the common proximal cause, allowing one to consider acquired and developmental deficits within the same framework. Although Jackson and Coltheart agree that a full explanation will involve both proximal and distal causes, they defend the cognitive neuropsychology framework as the appropriate way to reveal the proximal cause of any behavioural impairment, independent of distal causes.

Jackson and Coltheart discuss the case of phonological dyslexia, which, as we saw earlier, is defined by a difficulty in reading novel words. They argue that, in relation to the traditional cognitive model of reading, both acquired and developmental phonological dyslexia can be assigned the same proximal cause, namely, a problem with the processing route that maps graphemes to their respective phonemes (the "GPC" route). What differs in the acquired and developmental cases is the distal cause, respectively, brain damage and some developmental (perhaps genetic) anomaly. In short, these authors argue strongly that synchronic similarities in behavioural deficits between acquired and developmental cases *can be linked by a common underlying cause at a cognitive level of description,* and that this cognitive cause can be established by the methods of cognitive neuropsychology.

The extension of the cognitive neuropsychology framework to developmental disorders has, however, been criticised on three main grounds (Bishop 1997a; Karmiloff-Smith 1997; 1998). The first criticism is that the framework unnecessarily warps the types of data that are collected in the developmental case by focusing on the search for specific deficits, only superficially examining areas of presumed intactness. The second criticism is that the framework is unable to comment on one of the key contributory causes of the patterns of behavioural impairments found in developmental disorders, namely, the process of development itself. Where different developmental hypotheses exist for a given impairment, the cognitive neuropsychology framework cannot distinguish between them. The third criticism is that the assumption of a universal modular structure in the cognitive system on which the framework relies may not hold in the developmental case. We briefly look at each claim in turn.

First, Bishop (1997a) has argued that developmental and acquired disorders require empirical approaches with different emphases. Although researchers in adult cognitive neuropsychology look for *single cases* showing *dissociations* between cognitive abilities (as existence proofs of their dissociability), developmental disorders are likely to show patterns of *associated impairments* as a consequence of cascading effects of early deficits on subsequent development. Particular developmental disorders will be best identified by seeking consistent patterns of associated impairments in

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

group studies. Karmiloff-Smith (1997; 1998) argues that, particularly for disorders of a genetic origin, behavioural impairments at the end of development are likely to be the outcome of an extended atypical developmental trajectory, determined in part by the initial structural anomalies in the cognitive system and in part by the interactions of that system with its environment. She argues that, methodologically, researchers should not search exclusively for selective deficits at the end of the developmental process but should also seek differences in infancy, where the origins of the atypical trajectory may be revealed.

Second, both Bishop and Karmiloff-Smith claim that the cognitive neuropsychology framework is impoverished in the developmental domain by its exclusion of the developmental process as an explanation of patterns of behavioural impairments in the adult state. For Bishop (1997a), these processes include top-down as well as bottom-up interactions between cognitive subsystems during development, compensatory processes, and timing differences that may lead to changes in the patterns of impairment over time. As an example of the latter, she points to the hypothesis that in children with SLI, early problems in auditory discrimination that occur at a crucial stage in language development cause a lasting legacy of language impairment, even if the auditory problems subsequently resolve themselves and are undetectable. Karmiloff-Smith (1997; 1998) argues that the causes of behavioural impairments in developmental disorders are likely to be found in the low-level computational properties of the neonate brain, such as atypical neuronal firing levels or local connectivity. Such low-level differences can only lead to behavioural impairments via the developmental process, a process that may exaggerate some initial computational differences but attenuate others, depending on the nature of the domain. In the former case, the developmental process itself must be considered a key cause of the subsequent impairments.

In response to this criticism, it is worth noting Jackson and Coltheart's (2001) position that the cognitive neuropsychology framework is not designed to comment on distal causes of deficits such as development, merely on the proximal causes, that is, the functional deficits shown in the current state. As such, the cognitive neuropsychology framework simply does not have the power to address the questions of concern to Bishop and Karmiloff-Smith.

The third criticism of the extension of the cognitive neuropsychology framework to developmental disorders is potentially the most serious. The only necessary a priori assumption required to employ that framework is that in many domains the cognitive system is modular, so that selective deficits in behaviour may be traced to independent functional components. Bishop (1997a) has suggested that two of the defining properties of modules – innateness and imperviousness to top-down feedback – are clearly challenged by processes of development. Some cognitive abilities acquired by children are not innate, and top-down processing is used a great deal by children when they are performing cognitive tasks. There are problems with this criticism that concern the precise definition of what constitutes a module. Fodor (1983) identified several possible characteristics of modules (that they be domain specific, innately specified, informationally encapsulated, fast, hardwired, autonomous, and not assembled). However, none of these was stipulated as a *necessary* property, rather, those *likely* to be associated with modular processing (Coltheart 1999). There has been significant disagreement concerning the key properties of a module, if indeed modularity is to remain a single explanatory concept (Thomas & Karmiloff-Smith 1999). For example, for Fodor (2000), the most important property is encapsulation; for Coltheart, it is domain specificity. Until there is agreement on what constitutes a functional module, it will prove difficult to demonstrate whether development violates the necessary conditions and so clearly undermines the use of the cognitive neuropsychology framework to explain developmental disorders.

However, a more grave criticism lies in wait. Even if we accept a (loose) notion of modularity, it may be that in some types of developmental disorders, individuals do not share the same set of functional modules as in the normal cognitive system. Karmiloff-Smith (1998) argues that neurobiological evidence of the development of neocortex in infants strongly suggests that genes do not code directly for high-level cognitive modules, but that processing structure is emergent and experience-dependent, the *outcome* of a developmental process. The implication is twofold. First, if modular structure is the product of development (Karmiloff-Smith 1992), in cases of atypical development, the resultant modular structure may not be the same as in the normal adult case. Second, even if early damage is limited to a specific cognitive component, if the modular structure of the cognitive system is sensitive to experience, compensation may occur elsewhere in the system, altering the function of the initially intact components.

We can illustrate this idea with reference to Jackson and Coltheart's own example, phonological dyslexia. Recall that these authors attribute the impairment in nonword reading in the developmental and acquired cases of this disorder to a common proximal cause, an impairment to the GPC route in the traditional model of adult performance. However, in the pure case of the disorder, this common proximal cause is actually shorthand for "the GPC route is impaired *and the lexical routes are functioning normally.*" In cases of adult brain damage, this seems possible. But in the developmental case, Bishop and Karmiloff-Smith's position is that problems with the GPC route may also lead to differences in the way in which the lexical routes themselves develop. According to this view, the developmentally disordered system could comprise a GPC route and two lexical routes,[3] *all* of which are functioning atypically. Together, however, these routes would then manifest a behavioural impairment in nonword reading.

The extent to which modular structure can vary in cases of atypical development is currently an open question (see Tager-Flusberg 2000 for discussion). Indeed, the degree of plasticity across different cognitive systems and their underlying neural substrates is an area of active investigation (Thomas 2003). For example, it remains to be seen whether limits to plasticity are different in cases of acquired damage in early childhood than in individuals with genetic developmental disorders. Nevertheless, if a common modular structure cannot be assumed, it is evident that developmental disorders cannot be straightforwardly related to static models of the normal cognitive system. It is therefore of key importance to understand how modular structure emerges and to what extent this process can be disrupted.

3.3. A reconciliation

Jackson and Coltheart's (2001) claim that it is possible to study independently the endstate of a developmental dis-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

order (proximal cause) and the developmental process by which it was reached (distal cause) must be considered carefully. Although it may be possible to *study* them independently, they are not in fact independent but are mutually constraining. It would be unwise to characterise the endstate of a cognitive system in a form that could not be reached by a feasible developmental process (Piaget 1971). This point does not just apply to the study of developmental disorders. In the same way, theories within normal cognitive psychology and normal developmental psychology must be mutually constraining, despite existing as separate fields of inquiry.

Is Residual Normality a feasible type of developmental account? Is it realistic to expect developmental patterns of specific deficits to stand against a background of normal modular function? Karmiloff-Smith and colleagues have argued that a priori, the effects of genetic abnormalities are likely to be widespread throughout the brain and unlikely to be isolated to single high-level cognitive modules (Karmiloff-Smith 1998; Karmiloff-Smith et al. 2002). When marked behavioural deficits arise in a single domain, it is likely that the cognitive processes underlying apparently intact performance in other domains are also atypical in subtle ways – which may go undetected without the sensitive testing of abilities outside of the main behavioural impairment. Such investigations are prompted only by a realistic developmental hypothesis. In support, Karmiloff-Smith cites examples such as Williams syndrome, where ostensibly intact face recognition was subsequently shown to be achieved by atypical cognitive processes (see below), and developmental dyslexia, where motor deficits have been found in children previously thought only to have a highly selective problem in reading (e.g., Bishop 1990; Fawcett et al. 1996; Hill 1998).

Despite a priori leanings, one might view the issue as one to be determined merely on empirical grounds. Are there or are there not selective cognitive deficits in developmental disorders? Unlike modularity, RN is not an assumption that is a priori required to employ the cognitive neuropsychology framework. Instead, it is a hypothesis invoked to explain a particular set of empirical data. In this sense, Jackson and Coltheart are right that the characterisation of the current deficit in a disorder could be blind to the causes of that disorder. Should RN, therefore, be seen as a neutral hypothesis, simply "calling the data"? The answer is no. The reason is that theory and data collection are clearly not independent. Disorders are typically first investigated by application of a range of standardised tests to establish which areas show behavioural deficits and which show behaviour in the normal range. If one has a predilection to believe that RN is true, the risk is that scores in the normal range will be accepted as final evidence of normal underlying processes, and data collection cut short prematurely. If, however, one is more suspicious of RN, as developmentalists usually are given the interactive nature of the developmental process, then there is a motivation to perform more fine-grained analyses to establish whether apparently normal behaviour is actually being achieved by atypical underlying processes. If so, then deficits are not specific.

An example illustrates the point. Despite deficits in visuospatial processing, face recognition in Williams syndrome was initially reported as a "spared" ability, on the basis that scores on standardised tests fell within the normal range (Bellugi et al. 1994; Udwin & Yule 1991). This prompted claims that the development of systems underlying spatial reasoning is disrupted in WS, but the systems underlying face perception develop normally (Pinker 1999). If one were happy to invoke RN, one would stop at this point and perhaps use WS in combination with developmental prosopagnosia as a double dissociation implying the independence of face processing structures from general visuospatial processing.

However, suspicion of RN in genetic developmental disorders actually led to further investigation of this apparently intact ability. Closer examination of the items within the standardised tests on which individuals with WS performed well, and those on which they performed poorly, suggested that their recognition of faces proceeded atypically. Specifically, individuals with WS were better at recognising faces that could be identified by single features than those that required computation of configurations of features; control participants showed no such distinction (Karmiloff-Smith 1997).

Subsequent research with specially designed face stimuli and geometrical patterns supported the hypothesis that face processing follows an abnormal developmental course in WS (Deruelle et al. 1999; Humphreys et al. 2002). Electrophysiological brain imaging studies also indicate anomalous underlying processing in WS, including reduced sensitivity to inverted faces compared to normal faces, and an absence of the progressive developmental pattern of right hemisphere localisation found in typically developing controls (Grice et al. 2001; Mills et al. 2000). In short, when examined in detail, a superficially intact ability turned out to be associated with quite atypical cognitive and brain processes.

In reconciling the two opposing positions, a subtler picture emerges. The static cognitive neuropsychology framework is not *in principle* inappropriate for the study of developmental disorders. For example, there is nothing intrinsic to the nondevelopmental approach espoused by Jackson, Coltheart, and others that would prevent it from verifying whether RN is true in a particular child, domain, or disorder. Empirical data will eventually reveal if there are selective cognitive deficits in a given case. If there are, it will be necessary to construct and test a developmental account in which RN holds. However, a tendency simply to *assume* RN – conditioned by research in adult cognitive neuropsychology where a background of normal function can often indeed be assumed – leads to inadequate data collection. These data are then insufficient to establish beyond reasonable doubt that normal processes underlie behavioural scores that fall within the normal range. This bias has impeded progress in the study of developmental disorders, and particularly in building links to developmental cognitive neuroscience, developmental neurobiology, and, ultimately, the genetic anomalies that underlie many disorders.

This state of affairs has arisen because insufficient attention has been paid to the process of development itself in the study of developmental disorders. In particular, because RN is typically implicitly assumed, there has been no elucidation of the necessary conditions under which it would actually hold.

4. Under what developmental conditions would we expect to see similarities between developmental and acquired disorders?

Acquired deficits and developmental deficits can be related to the same model of the normal cognitive system if a com-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

ponent that can be damaged in the endstate can also fail to develop in isolation of the entire developing cognitive system. This will occur under the following conditions:

• The process of development does not contribute to the pattern of behavioural deficits, so damage produces an identical effect at any stage.

• There are pre-existing modules that develop independently; so damage to one does not affect development of the others.

• If it is the case that modules emerge through a developmental process (modularisation), this process can be disrupted in such a way that some modules emerge normally and some develop atypically.

By contrast, similarities between developmental and acquired disorders will *not* occur under the following conditions:

• The process of development always contributes to the precise pattern of deficits in developmental disorders (either attenuating or exaggerating the effects of early damage) and produces patterns not found in acquired damage.

• The system has sufficient plasticity that compensation occurs; undamaged systems take on the function of damaged systems, perhaps suboptimally, perhaps also at some expense to their normal function.

• Module emergence is disrupted to produce a different functional structure that shows a behavioural pattern not found in acquired damage.

Similarities may occur between developmental and acquired disorders that *cannot* be related to the same model of the normal cognitive system under the following conditions:

• Module emergence is disrupted to produce a different functional structure that nevertheless shows a similar *behavioural* pattern to that found in acquired damage.

• The patterns of behavioural breakdown are specified not by the structure of the cognitive system but by features of the problem domain (e.g., in all suboptimal systems, performance degrades on the hard parts of the problem before it degrades on the easy parts).

What conditions *actually* hold in cognitive development? To address this issue, we need to know the answer to several further questions. First, how does the process of development interact with damage to a cognitive system to produce endstate behavioural impairments? Second, does the process of development always play a central role in producing the impairments – and if so, does development tend to produce a different pattern of endstate impairments to acquired damage in the endstate? Third, what is the origin of the specialised functional components stipulated in adult models? If components are not innately specified, how do they emerge through a process of development, as in Karmiloff-Smith's (1992) theoretical notion of emergent modularisation? Finally, how can such a process of emergent specialisation be affected or unaffected by disruption to the computational conditions existing in the early cognitive system?

These are difficult questions, and the field of developmental cognitive neuroscience is some way from having answers to all of them. In what follows, we explore potential answers to these questions by examining computational models of cognitive development. To retain a focus, we restrict our investigation to computational models applied to our two target domains, reading and past tense formation, and to research based on a single, influential class of computational learning systems, that is, connectionist networks.

5. Computer modelling

The modelling section comprises three parts. In the first, we compare the methods that researchers have used to extend connectionist models of normal development and adult function to cases of developmental and acquired impairments in reading and past tense formation. In the second, we introduce new simulations to gauge the contribution of the developmental process to producing patterns of endstate impairments, within the framework offered by those models. Specifically, we investigate the extent to which the process of development itself is a causal factor of the specific pattern of impairments shown in a developmental disorder. In the third, we use new simulations to examine the assumption of Residual Normality. In a system with emergent specialisation of function (i.e., one exhibiting modularisation), how viable is RN? Specifically, when one component of the system is prevented from developing normally, does the rest of the system nevertheless develop independently and normally? If not, what are the conditions under which learning systems *would* show RN, so that developmental impairments could be interpreted in terms of selective deficits to an adult model? From a behavioural perspective, how does assuming RN affect the inferences we can make from dissociations in behaviour to underlying structures?

6. Connectionist models of acquired and developmental deficits

Connectionist networks have been widely used in recent years to construct models of cognitive processing in adults (see, e.g., McLeod et al. 1998; Rumelhart et al. 1986). Because one of the main strengths of these networks is their ability to learn input-output functions, they have increasingly been used to model the development of cognitive processes (Elman et al. 1996; Thomas & Karmiloff-Smith 2002).[4]

When there exists a working model of a normal adult system, the validity of the model can be further tested by investigating its ability to capture patterns of acquired deficits when the model is damaged in various ways. Connectionist models have been used to capture deficits in a number of acquired disorders, including dyslexia, aphasia, alexia, prosopagnosia, epilepsy, phantom limbs, stroke, frontal lobe damage, unipolar depression, Parkinson's disease, Alzheimer's disease, and schizophrenia (see, e.g., Reggia et al. 1996b; Stein & Ludik 1998). Where knowledge is encoded in the connectionist network through a training process, acquired deficits are modelled by damaging the network *after that training process is complete.*

Where connectionist networks have been used to model phenomena within cognitive development, this has permitted the investigation of developmental disorders when development is made to follow an atypical trajectory (Oliver et al. 2000; Thomas & Karmiloff-Smith 2002). Although work in this area is relatively new, models have already been put forward attempting to capture behavioural deficits in developmental dyslexia, SLI, WS, and autism (see Thomas & Karmiloff-Smith 2002 for a review). In developmental models, knowledge is encoded in network systems via a training process, whereby the model aims to simulate both the developmental trajectory and the endstate abilities of the system. In contrast to models of acquired deficits,

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

changes in models of atypical development are made to the network or to the way it learns *prior to the training process.* These changes in the computational constraints of the learning system lead to atypical trajectories of development and an endstate performance that may exhibit behavioural impairments.

The contrast between connectionist models of acquired and developmental disorders is a fairly clear one. In the acquired case, damage of some sort is applied to the model at the end of a training process. In the developmental case, it is applied prior to the training process.[5] Immediately one might ask, do modellers use the same kind of damage in each case, and does this damage cause the same behavioural impairment? If the answer to both is yes, one might conclude that the developmental process plays a limited causal role in generating the pattern of behavioural impairments.[6] However, if the answer is no, and different impairments result from the same damage in the two cases, then the implication would be that the developmental process is an important component in determining the pattern of impairments in developmental disorders. In sections 6.1 and 6.2, we consider this question in relation to connectionist models of reading and past tense formation.

6.1. Connectionist models of reading

Connectionist models of reading assume that the computational problem in this domain is to learn to map between representational codes of the written form of a word, the spoken form of a word, and the word's meaning (Plaut et al. 1996; Seidenberg & McClelland 1989). Typically, this involves three connectionist networks, one to map from orthography to phonology, one to map from orthography to semantics, and one to map from semantics to phonology (although in many models only the first of these networks is implemented; see Harm & Seidenberg 2001 for an exception). Each network has a three-layered structure, comprising an input layer, an output layer, and a layer of hidden units in between. Some models employ recurrent connections that allow cycling activation patterns so that the model will settle into a stable output state. Sometimes a layer of "clean-up" units is connected to the output layer to aid this settling process (see, e.g., Harm & Seidenberg 1999; 2001).

Within these models, acquired dyslexia is produced by different kinds of damage to the trained model. Acquired surface dyslexia, a deficit in reading exception words, has been modelled by damaging the network that maps orthography to phonology, via the removal of hidden units or the severing of connections (Patterson 1990; Patterson et al. 1989). Such damage produces a greater impairment on reading exception words than regular words. However, failures of this approach to fit more extreme patterns of surface dyslexia subsequently led to the claim that exception word reading might be achieved via an indirect semantic route, particularly in the case of low-frequency words. Acquired surface dyslexia might represent damage to this indirect route, so naming must proceed via the orthography-to-phonology route alone, a route that has not learned to name low-frequency exception words and, as a result, regularises them (Patterson et al. 1996; Plaut et al. 1996; see Coltheart et al. 2001 for discussion).

Previously, acquired phonological dyslexia was not explicitly simulated within these learning models because, in theoretical terms, it corresponded to selective damage to the entire orthography-to-phonology network; because most models only implemented the orthography-to-phonology network itself, such a lesion was outside their scope. Theoretically, lesioning the direct orthography-phonology route would mean that reading must be accomplished primarily or exclusively via the semantic route, so that novel words without a stored meaning would be severely impaired. However, recently Harm and Seidenberg (2001) have implemented the full connectionist reading model, including pathways between phonology, orthography, and semantics. The authors report a manipulation intended to simulate acquired phonological dyslexia, whereby noise is added to processing within the phonological component of the model (i.e., the phonological output units and associated clean-up units). Harm and Seidenberg demonstrate how this impairs the nonword reading of the model much more severely than its reading of words in the training set (Harm, personal communication, June 2001). It also accounts for several effects found in the nonword reading of acquired phonological dyslexics that were previously taken as support for the traditional model of reading.

Doubts have been raised as to the full developmental validity of several of these connectionist reading models. Nevertheless, we consider them here for their insight into systems that acquire representations appropriate for a cognitive domain through a learning process. Patterns of developmental dyslexia have been simulated by applying the relevant damage prior to this learning process. Thus surface dyslexia has been simulated in a number of models by removing units from the hidden layer in the orthography-to-phonology network prior to training (Harm & Seidenberg 1999; Plaut et al. 1996; Seidenberg & McClelland 1989). This damage produces a greater impairment in exception word reading than regular word reading or nonword reading, particularly for low-frequency exception words. Because exception word performance generally lags behind regular word performance during training, some authors have also simulated poorer performance in reading exception words simply by giving the network less training or less efficient training (Bullinaria 1997; Harm & Seidenberg 1999).

Developmental phonological dyslexia has been simulated in two main ways. The first approach reflects a prior claim that developmental phonological dyslexia may correspond to phonological representations (and perhaps orthographic representations as well) that have insufficient componentiality (Manis et al. 1996; Plaut et al. 1996). Harm and Seidenberg (1999) implemented this proposal by restricting the computational properties of the phonological component of their model (the phonological output layer, its recurrent connections, and its clean-up units). Their manipulations included the removal of the clean-up units and severing half the recurrent connections between the phonological units, or restricting the size of the weights in the recurrent connections, or making computations within the phonological component more noisy. All of these resulted in poorer nonword reading, and some impacted on exception word reading as well. Brown (1997) also demonstrated that when both orthographic representations and phonological representations are deliberately constructed with reduced componentiality, reduced nonword reading is found at the end of training.

The second approach seeks to constrain the nature of the computational function that can be learned between orthog-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

raphy and phonology. For example, Zorzi et al. (1998) have argued that the reading system is better conceived of as including a network in which orthographic representations are directly connected to phonological representations. Lack of these direct connections in the initial architecture, and the presence only of a route mediated by hidden units, prevented their network from learning a simple function relating orthography and phonology, and so generalisation was reduced. Brown (1997) used another constraint on the computational function by employing several three-layer networks with progressively reduced numbers of hidden units and comparing them when their performance on regular and exception words was matched. (By necessity, this meant that the networks with fewer hidden units had experienced more training.) Networks with fewer hidden units were unable to learn a robust function linking orthography and phonology and so showed poorer nonword reading.

In summary, both surface and phonological dyslexia permit a direct comparison between simulations of the acquired and developmental forms. In the case of surface dyslexia, initial approaches used the *same* method, the removal of hidden units and/or connections, to simulate the same impairment, namely, a deficit in exception word reading. Recently, acquired accounts have appealed to the entire lesion of unimplemented routes. In the case of phonological dyslexia, several methods have been used to simulate the developmental impairment, including either altering phonological and/or orthographic representations, or constraining the computational function that the network can use to link orthographic and phonological codes. One method used by Harm and Seidenberg (1999), the addition of noise to processing within the phonological component of the model during training, was also used by Harm and Seidenberg (2001) to simulate acquired phonological dyslexia, where such noise was added to a normally trained model. Indeed, Harm and Seidenberg (2001) specifically comment that the "form of impairment is identical" in the two cases (p. 80). In short, on the basis of connectionist models of reading, one might conclude that the same form of damage before and after training creates the same behavioural deficit – as if the developmental process itself contributed nothing to the nature of that behavioural deficit.

6.2. Connectionist models of past tense formation

Connectionist models of past tense formation assume that the computational problem in this domain is to learn to map between representational codes of the phonological form of the stem of a verb and a phonological form of that verb's past tense (Plunkett & Marchman 1993; 1996; Rumelhart & McClelland 1986), sometimes in the presence of the verb's semantic representation (Hoeffner 1992; Joanisse & Seidenberg 1999), and sometimes in the presence of more restricted semantic information (MacWhinney & Leinbach 1991; Plunkett & Juola 1999).

Joanisse and Seidenberg (1999) sought to simulate two kinds of acquired deficits in their model of past tense formation, either an impairment in producing regular past tense forms (found in cases of nonfluent aphasia and Parkinson's disease), or an impairment in producing exception past tense forms (found in cases of fluent aphasia and Alzheimer's disease). Impairment to the formation of regular past tenses was achieved by randomly severing connections between the phonological output layer and that layer's bank of clean-up units in the trained network. However, the fit to patient data was not ideal here, because the model showed a much larger decrement in extending the past tense rule to novel items (e.g., *wug-wugged*) than on the formation of existing regular past tenses (e.g., *talk-talked*). Patients, however, can show similar decrements to both (e.g., Ullman et al. 1997).[7] It is possible that formation of existing regular past tenses was driven too much by word-specific information because the model was trained only on a single inflectional paradigm. In a larger model in which individual words can be inflected in several different ways, regularities may be pushed further into the phonological part of the network, such that regular verbs would also be amenable to selective damage. Impairment to the formation of exception past tenses, however, was achieved by randomly severing connections between the semantic representations and their clean-up units in the trained model, while adding noise to the semantic activation level. This gave a good fit to patient data (though, see Tyler et al. 2002a).

Developmental problems with regular past tense formation have been reported in SLI, although recently Ullman and collaborators have argued that the deficit is relative, in that the normal advantage for regular verbs over exception verbs is much reduced and most past tense forms are uninflected (Ullman & Gopnik 1999; van der Lely & Ullman 2001). Hoeffner and McClelland (1993) sought to simulate the developmental regular verb deficit by altering the phonological representations of their model prior to training. The phonological representations were changed in line with a hypothesis that individuals with SLI have difficulty processing fast-changing auditory signals, which particularly impairs perception of phonemes such as /t/ and /d/ (e.g., Tallal & Stark 1981; though, see Bishop et al. 1999b; Joanisse & Seidenberg 1998). Both these phonemes are involved in marking the regular past tense form in English. In the model, word final stops and fricatives were given weaker representations in the normal case to reflect their lower salience. In the impaired model, the overall strength of the phonological representations was weakened, exaggerating the disadvantage of word final stops and fricatives. When the model was trained with these altered representations, the result was poorer performance on past tense formation, such that regular past tenses showed a greater impairment than exceptions, and the predominant error pattern was a failure to inflect the verb stem. Moreover, just as in SLI, the model showed an impairment on morphemic phonemes (e.g., the final /d/ in cared) but not phonologically identical phonemes that were nonmorphemic (e.g., the final /d/ in card). The model was able to produce a differential impairment for regular verbs in its trained state, but did not successfully simulate the very low and equal performance on both regular and exception verbs (see Ullman & Gopnik 1999 for further discussion of the model).

Joanisse (2000) attempted to simulate the pattern of SLI data by applying processing noise to the phonological representations of his past tense model throughout the training process. Here, the model's level of correct performance on regular, exception, and novel verbs was closer to that shown in recent empirical data (van der Lely & Ullman 2001), with low scores on all types. However, the model did not reproduce the predominant error pattern of uninflected stems found in SLI, suggesting that a future model needs to incorporate aspects of both the Joanisse and the Hoeffner & McClelland models.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

In Williams syndrome, it was initially reported that there was a selective deficit in forming the past tense of exception verbs (Clahsen & Almazan 1998). However, a larger study suggested that this apparent deficit was actually a consequence of language delay, because performance on exception verbs lags behind that on regular verbs in normal development, and language is typically delayed in WS (Thomas et al. 2001). When language delay was controlled for in this latter study, the greater deficit on exception verbs in WS disappeared. The Thomas et al. study did, however, reveal reduced generalisation of the past tense rule to novel forms in WS, a pattern which persisted even when language delay was controlled for. Using a past tense network that mapped from verb stem to past tense form in the presence of semantic information, Thomas and Karmiloff-Smith (in press) explored the manipulations to the normal model that could reproduce this pattern of developmental data. Various claims have been made that there are subtle deficits to the language system in Williams syndrome. These include the proposals that language development may be "hyper-phonological," relying to a greater extent on phonological than lexical-semantic information (Grant et al. 1997; Vicari et al. 1996a; 1996b; Volterra et al. 2001), that the phonological representations themselves may be atypical and perhaps rely on sensitive auditory processing (Karmiloff-Smith et al. 1997; Majerus et al. 2001; Neville et al. 1994), that lexical-semantic representations may be atypical (Clahsen & Almazan 1998; Rossen et al. 1996; Temple et al. 2002), or that lexical-semantics may be poorly integrated with phonology (Frawley 2002; Karmiloff-Smith et al. 1998).

Thomas and Karmiloff-Smith found that a manipulation of the phonological representations that reduced their similarity and redundancy was sufficient to reproduce the delay for regular and exception past tense forms as well as the reduction in generalisation. However, the pattern could also be reproduced when noise was added to the information coming from the lexical-semantic system. By contrast, slowed learning failed to produce a reduction in generalisation, suggesting that delay alone was insufficient to explain the data. Although elimination or weakening of the lexical-semantic contribution produced a selective delay (but no final impairment) for exception verbs, it also failed to show the reduction in generalisation. In short, manipulations to phonology or to the integration of phonology and lexical-semantics could simulate the WS data, but a manipulation to lexical-semantics alone could not.

What if the WS data had shown a selective deficit in exceptions as initially reported in the syndrome – could the model have shown this pattern? Performance on exception verbs could be preferentially delayed under at least two conditions: by attenuating lexical-semantic input, or by restricting the computational complexity of the representations the system could learn (e.g., by employing a two-layer network or by reducing the numbers of hidden units by a certain calibrated amount). However, in both cases, the delay was not associated with an endstate impairment. The only way to achieve such a final deficit in exception verbs was to combine manipulations (e.g., attenuating lexical-semantic information and slowing down learning/terminating training at a point where regulars had reached ceiling but exceptions had not, or attenuating lexical-semantic input while restricting computational complexity).

In summary, for inflectional morphology, we have direct comparisons of attempts to simulate acquired and developmental deficits to both regular and exception verbs. Impairments to regular verbs were simulated by damage to phonology either prior to or following training. It is worth noting that in the Hoeffner and McClelland model, a specific regular deficit in the developmental case was achieved by effectively targeting information that encoded the regular rule. However, a deficit to regular performance in the acquired case was achieved with more general damage (putting aside, for a moment, the fact that acquired damage impaired generalisation of the rule rather than its application to existing verbs). For exception verbs, an acquired impairment was simulated by damaging the input from semantics. Similar damage in a developmental model delayed the learning of exception past tense forms but, importantly, failed to produce an impairment at the end of training. Broadly, then, phonological damage targeted regular inflection/generalisation, whereas semantic damage targeted exception inflections.

6.3. Summary

What can we conclude from the detailed comparison of models of acquired and developmental deficits in these two domains? The results are somewhat contradictory. For both surface and phonological dyslexia, acquired and developmental approaches employed the same kind of damage to produce the same impairment – the intervention of the developmental process did not appear to contribute to the pattern of impairments. For impairments to regular past tense formation, however, more specific damage was required prior to training than at the end of training to generate a specific impairment in regular past tense formation – as if the developmental process risked changing the nature of the impairment. And, indeed, for the impairment of exception past tense forms, damage to semantic input only produced a delay in acquiring these forms, whereas damage at the end of training produced a marked behavioural deficit. In other words, in this case the developmental process overcame initial damage to produce a successful outcome via an altered developmental trajectory.

Our ability to gauge the contribution of the developmental process to the final impairments is compromised by the fact that in each of the preceding cases, the comparisons have involved separate models whose implementations have differed in detail. No models have afforded a direct comparison of the outcome of the same damage carried out prior to versus following the training process. For this reason, in section 7 we describe a simulation designed specifically to make such a direct comparison, something never hitherto undertaken in the literature.

7. Simulation one: Comparing startstate and endstate damage

7.1. Introduction

The design of the following simulations is relatively straightforward. We take a given problem domain and model architecture and train the model on the domain. This establishes its "normal performance." We then run the model in two conditions. We either damage the model prior to its training process, or damage it following its training process. Any difference in the pattern of impairments in the two cases

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

must arise from the contribution of the training process, that is, from development.

Three forms of damage are reported: (1) removal of a proportion of connections from the network; (2) addition of noise to the activation levels in the network; and (3) alteration of the discriminability of the processing units, that is, the ability of a unit to produce large changes in its activation state in response to small changes in the input it receives. All manipulations have been widely used in modelling both acquired and developmental deficits. For example, lesioning of network structure has been used to model dyslexia (e.g., Hinton & Shallice 1991; Patterson et al. 1996; Plaut et al. 1996; Plaut & Shallice, 1993), alexia (Mayall & Humphreys 1996), phantom limbs (Spitzer 1996), stroke (Reggia et al. 1996a), Alzheimer's disease (Ruppin et al. 1996), prosopagnosia (Farah et al. 1993), schizophrenia (Hoffman 1996), and autism (Cohen 1998). Addition of noise to processing has been used to model dyslexia (Harm & Seidenberg 1999; 2001), SLI (Joanisse 2000), language in Williams syndrome (Thomas & Karmiloff-Smith, in press), and Alzheimer's disease (Joanisse & Seidenberg 1999). Alteration of unit discriminability has been used to model schizophrenia (Cohen & Servan-Schreiber 1992), executive dysfunction (Levine 1996) and the effects of ageing (Li & Lindenberger 1999).

For comparison with the preceding discussion, we test the contribution of the training process in a domain analogous to past tense formation.

7.2. Simulation details

Architecture: A three-layered feedforward network was used, with the architecture as shown in Figure 1a.

Training set: The training set was taken from Plunkett and Marchman (1993) and comprises an artificial language set constructed to reflect the most important structural features of English past tense formation. There were 500 monosyllabic verbs, constructed using consonant-vowel templates and the phoneme set of English. Phonemes were represented over six articulatory features, and separate banks of units were used to represent the initial, middle, and final phonemes of each monosyllable. The output layer incorporated an additional two features to represent the affix for regular verbs. This corresponds to a network with 18 input units and 20 output units. However, the current simulations involved removing connection weights, and Bullinaria and Chater (1995) have argued that when network models are lesioned, resulting patterns of impairments can be artefactual if very small networks are used. In an attempt to avoid this, the representational scheme was duplicated five times, with the addition of a small amount of noise (whereby the binary features in each duplication had a 20% chance of flipping their state). This preserved the nature of the computational problem faced by the network, but increased the network's size to 90 input units and 100 output units. Fifty hidden units were used in the hidden layer.

There were four types of verbs in the training set: (1) regular verbs that formed their past tense by adding one of the three allomorphs of the +ed rule, conditioned by the final phoneme of the verb stem (e.g., *tame-tamed, wrap-wrapped, chat-chatted*); (2) exception verbs whose past tense form was identical to the verb stem (e.g., *hit-hit*); (3) exception verbs that formed their past tenses by changing an internal vowel (e.g., *hide-hid*); (4) exception verbs

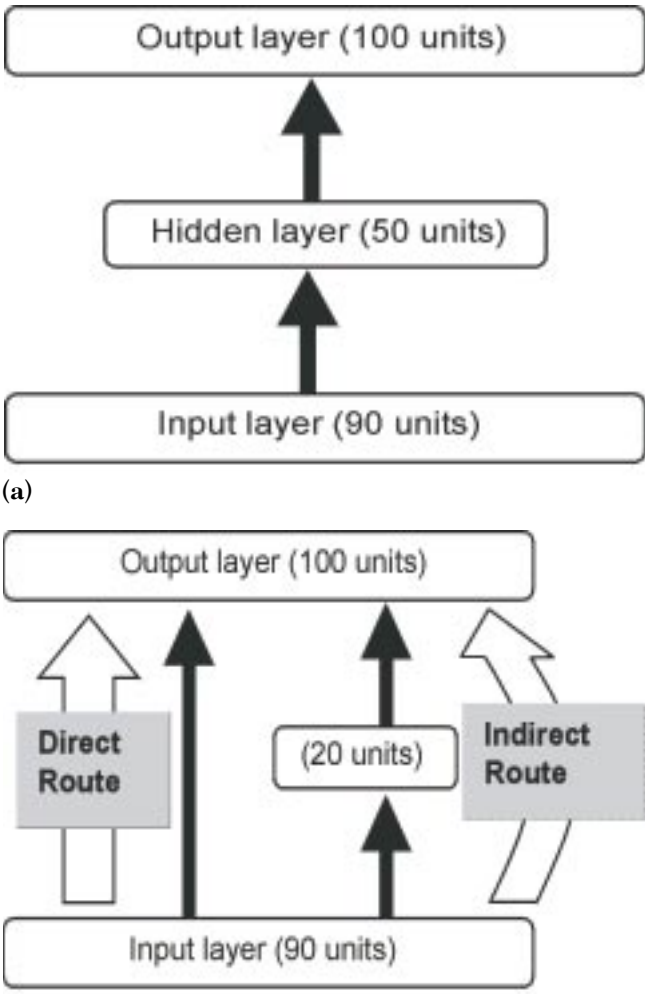


Figure 1. Architectures of the models used in Simulations One and Two. (a) Three-layer pattern associator. (b) Dual-route pattern associator.

whose past tense form bore no relation to its verb stem (e.g., *go-went*). The token frequency of this last type of exception verb had to be higher for the network to learn them successfully (see Plunkett & Marchman 1991), as is the case in real languages. As a result, this verb type experienced three times as much training as the other types. There were 410 regular verbs, and 20, 68, and 2, respectively, of each exception verb type.

A separate set of novel verbs was constructed to evaluate the generalisation performance of the network. These verbs could differ depending on their similarity to items in the training set. For simplicity, 410 novel verbs were used, each of which shared two phonemes with one of the regular verbs in the training set. Generalisation was evaluated depending on the proportion of these novel verbs, which were assigned the correct allomorph of the regular past tense rule.

Learning algorithm: The network was trained with the backpropagation learning algorithm, using cross-entropy between the output and target as the error signal (Hinton 1989). The learning rate was 0.01, and momentum was 0. The entire corpus was presented on each epoch, and pattern update was used. Networks were trained for 5,000 epochs.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Performance measure: A nearest-neighbour method was used to evaluate network performance, using a Euclidean distance metric. For each position in the output, the phoneme that the set of activation values most resembled was taken as the intended output for that position. If the resulting output string was the target output, it was marked as correct. Scores were therefore percentage correct for each verb type.

7.3. Implementation of damage

Lesioning: Weights were probabilistically set to zero throughout the network. The probability level determined the severity of the lesion. A probability of 0.3 would on average lesion 30% of the connection weights. Because of differences in sensitivity, probability levels of .01, .025, .05, .1, .2, and .3 were used for lesions applied at the end of training (with no retraining after damage), whereas probability levels of .5, .6, .7, .8, .9, and .95 were used for lesions applied at the beginning of training.

Noise: Noise was added to the activation levels of the units in the hidden layer, with a Gaussian distribution with mean zero and a standard deviation that determined the severity of the damage. Standard deviations of .025, .05, .0625, .075, .0875, .1, .2, .3, .4, .5, .6, and .7 were used. The baseline condition included no noise. Units had a maximum activation level of 1 and a minimum of 0, and noise could not take the activation state of a unit outside of these limits.

Discriminability: The activation of each processing unit in the hidden and output layers was determined by the following equation:

$$Output = \frac{1}{1 + e^{-(netinput/Temperature)}}$$

where *net input* is the summed activation arriving at the unit including its bias, and where the *Temperature* parameter controls the steepness of this sigmoid function (see e.g., Hinton & Sejnowski 1986). High temperatures correspond to low discriminability, whereas low temperatures correspond to high discriminability. Values of 4 and .25 were used.

Replications. For the baseline model and for cases of damage prior to training, results were averaged over six runs of each network using different random seeds. Initial weights were randomised within the range ±0.5 and pattern presentation during training was random without replacement. For cases of damage at the end of training, results were averaged over damage to each of the six baseline networks. For the addition of noise and the probabilistic lesioning of connection weights at the end of training, results were averaged over 10 repetitions of the damage for each of the six baseline networks. Graphs include standard error bars across the network replications as an indication of variability.

7.4. Results

The following graphs show performance on regular verbs, performance on the vowel-change exception verbs, and performance on generalisation of the regular rule. Results for the other two exception types were similar, and so are omitted. Baseline performance on the regular, exception, and rule pattern types was 100%, 100%, and 77%, respectively. Figure 2a shows the effect of lesioning weights before training ("startstate" damage) and after training ("endstate" damage). For each pattern type, Figure 2 shows the relative performance of the startstate and endstate conditions for increasing levels of damage.

The results here indicate a similar pattern of impairment for both startstate and endstate damage on regular and exception verbs. Figure 2b demonstrates that in both cases, exception patterns suffer a greater impairment than regular patterns, echoing the surface dyslexia simulations. For novel items, startstate lesioning initially improves generalisation of the rule, whereas endstate lesioning is only deleterious. At higher levels of startstate lesioning, however, generalisation declines here, also. There are two major points to note from this simulation. First, very much greater damage is required in the startstate than in the endstate to produce an equivalent amount of behavioural impairment. Therefore a lesion of 2.5% of network connections in the endstate reduces performance on regular patterns to approximately 90%, whereas a lesion of 80% of the connections in the startstate is required to produce an equivalent deficit. Despite the fact that the same damage produces a similar behavioural impairment here, the training process creates a huge difference in *sensitivity* to damage between the startstate and endstate conditions. This is because the trained network is losing connections that have already stored specific knowledge, whereas the untrained network is reduced in its potential to learn and uses the remaining potential to acquire the domain as best it can.

The second finding is that the relationship among regular, exception, and rule performance in the startstate and endstate differs. For example, for a given level of performance on regular patterns in the damaged networks, the startstate network will show lower exception performance and higher generalisation performance. Despite broadly equivalent behavioural impairments, in detail the patterns of deficit are different in the acquired and developmental case.

Figure 3 illustrates the effect of adding noise to activation levels within the network, either once training is complete or throughout training. The results once more show a differential pattern of sensitivity, but now in the reverse direction to the lesioning condition. The network is much more sensitive to noise occurring in the startstate than it is to noise occurring in the endstate. In the endstate, the network has established its knowledge and, as a result of the nonlinear processing units, is able to tolerate noise in processing. As a result, performance has not yet reached floor when noise is added with a standard deviation (SD) of .7. However, when damage occurs in the startstate, the network is never provided with a reliable rendition of the knowledge it must learn. When noise is added with an SD of as little as .2, no learning is possible at all. Although the acquired and developmental phonological dyslexia models of Harm and Seidenberg (1999; 2001) are not directly comparable to each other, it is interesting to note that the acquired impairment was simulated by the addition of noise an order of magnitude greater than that used to simulate the developmental impairment, in line with the current findings. And a similar indirect comparison of Joanisse's (2000) model of SLI and Joanisse and Seidenberg's (1999) model of aphasia in the past tense domain indicates a comparable requirement for greater noise in the endstate than the startstate to produce an equivalent level of deficit.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

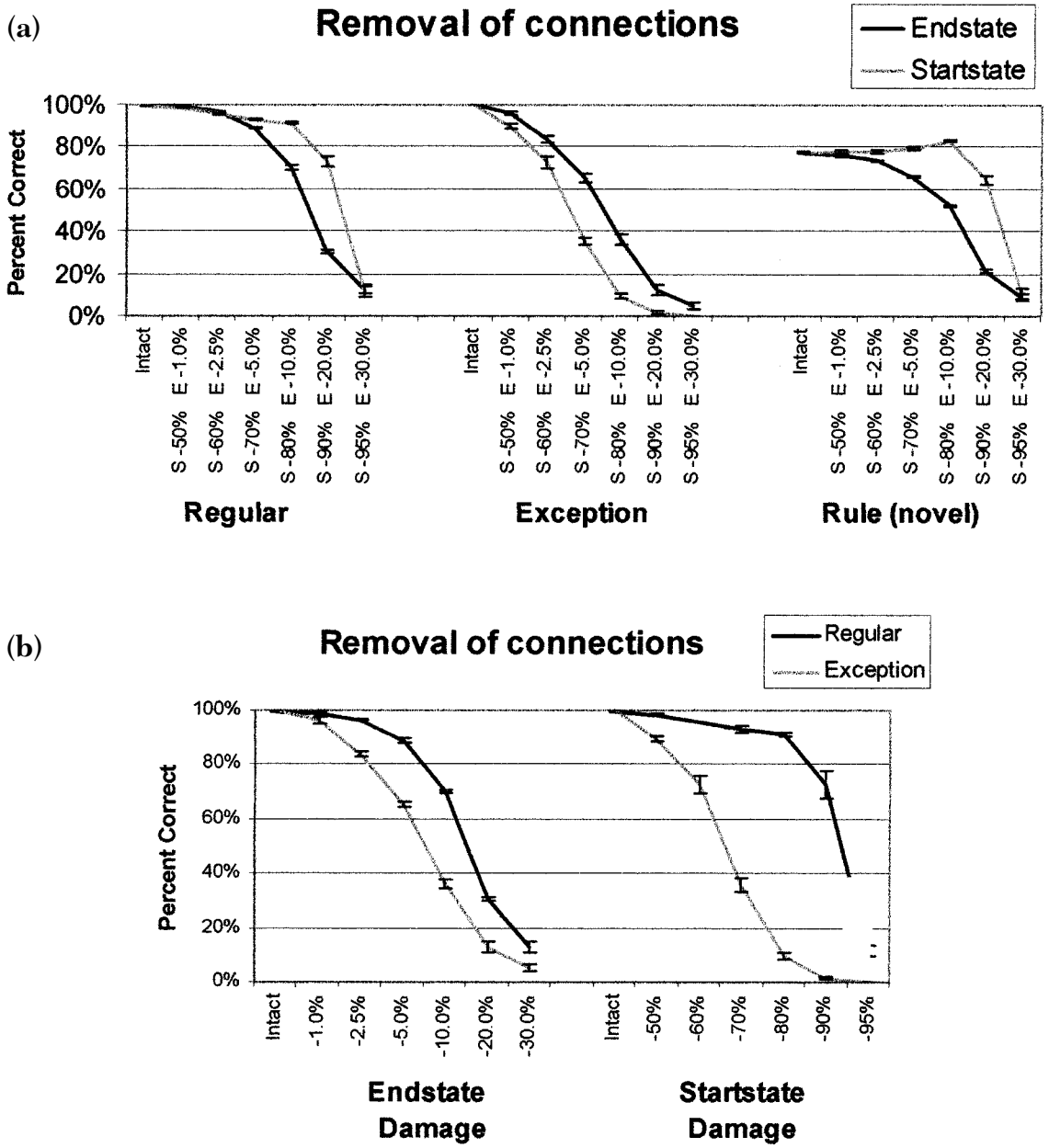


Figure 2. (a) The effect of removing connections from the network either prior to (startstate) or following (endstate) the training process for regular, exception, and novel patterns. The *x*-axis plots increasing levels of damage, with much greater startstate damage (*S*) required to produced an equivalent impairment than endstate (E). (b) Direct comparison of impairments on regular and exception patterns following startstate and endstate damage.

Noise added to the endstate led to a roughly uniform decrement across regular, exception, and novel patterns (e.g., from SD levels of .3 onwards). However, noise added to the startstate led to greater impairments to exception patterns than regular patterns. As an indication of this effect, when regular pattern performance was roughly comparable in the two conditions (startstate 90% with SD = .075, endstate 87% with SD = .25), exception patterns had fallen to 69% in the startstate against 79% in the endstate. The similarity of the mapping between regular patterns as well as their majority in the training set allows them to better overcome the addition of noise in training than the unique and minority exception patterns. In contrast, the ability of the network to deal with noise in the trained state depends on the nonlinear functions within the hidden units – units that, broadly speaking, are shared by all patterns. Based on the relation of these different components of performance, once more one must conclude that the detailed pattern of impairments in the acquired and developmental cases was different.

In sum, the addition of noise could produce effects that were *uniform* and *global* in effect (across regulars, exceptions, and generalisation), as in the endstate, or that were *differential,* as in the startstate; but, most clearly, effects were much stronger when the damage occurred in the *startstate* than when it occurred in the endstate.

Figure 4 shows the effect of changing the discriminability of the processing units within the network. Reduced dis-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

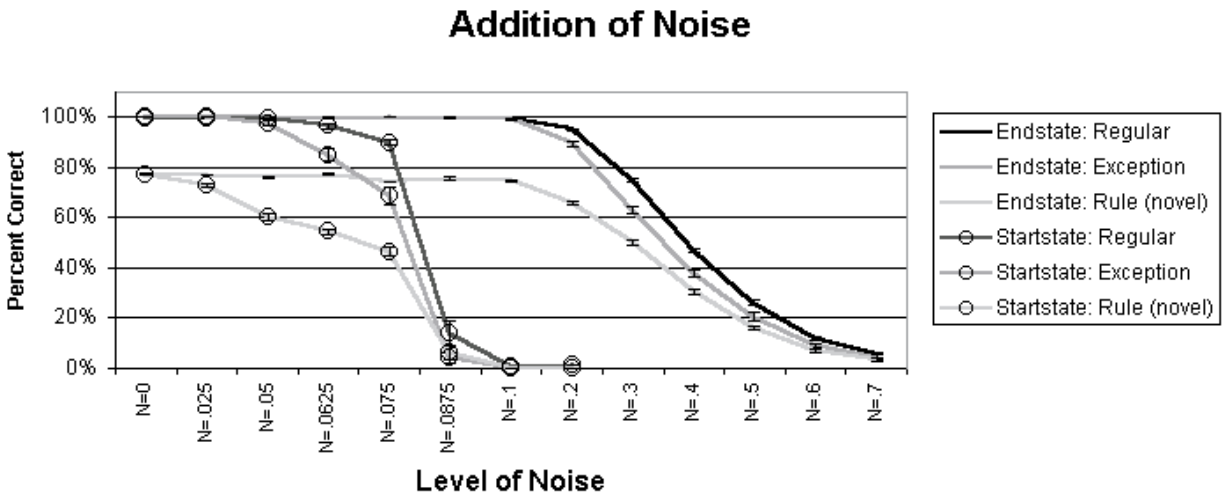


Figure 3. The effect of adding noise to activation levels within the network either prior to (startstate) or following (endstate) the training processing for regular, exception, and novel patterns.

criminability has little effect when applied either to startstate or endstate. The network can evidently compensate for it during training or tolerate this disruption at the end of training. An increase in discriminability also produces little effect when applied to the startstate – once more, the network can evidently compensate during training. However, if an increase in discriminability is applied to the endstate, the result is a marked and selective deficit in performance on exception patterns, dropping in performance from 100% to 30%. Meanwhile, regular patterns only suffer a minor dip in performance, and generalisation increases slightly. Evidently, exception patterns rely more on a certain level of discriminability in the processing units than do regular patterns. On the whole, this type of damage produces an impairment that is *selective* in the behaviour it impairs and only occurs when the damage is applied to the *endstate.*

7.5. Discussion

This simulation addressed two issues. First, does the process of development contribute to the pattern of deficits? Second, does the process of development produce patterns of deficits that are the same as those produced in acquired damage?

With regard to the first issue, direct comparison of the effects of identical damage at startstate and endstate demonstrated a complex relation between these two conditions, indicating a significant role for the process of training. Three different forms of damage produced three different possible relations. First, damage could produce a similar pattern of impairment for the startstate and endstate conditions, but the two conditions could vary in their sensitivity to the damage (removal of connections, addition of noise). Second,

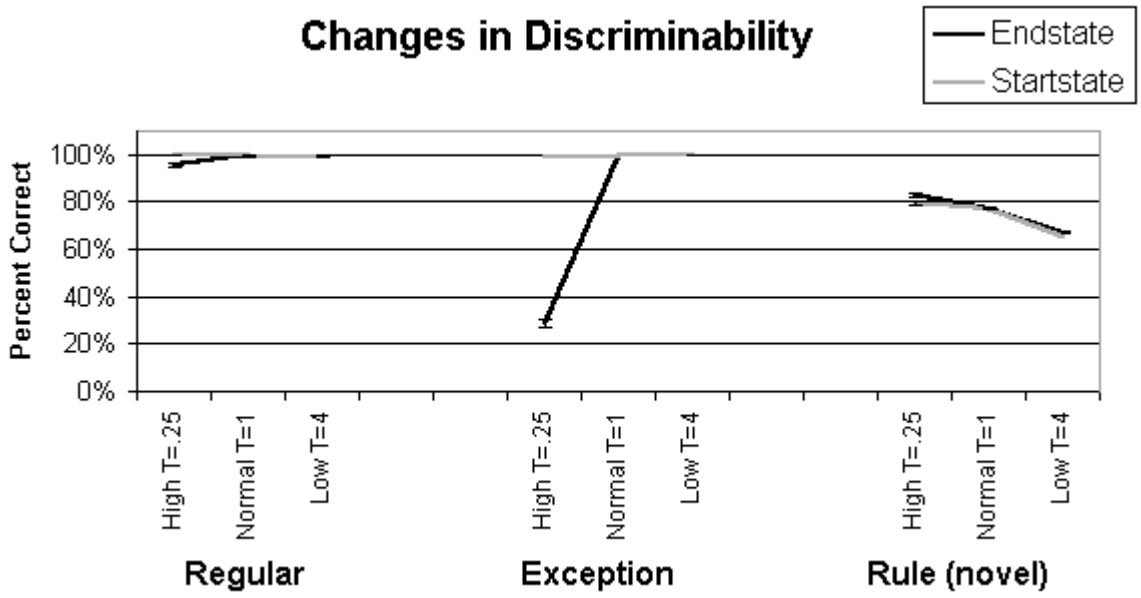


Figure 4. The effect of altering unit discriminability within the network either prior to (startstate) or following (endstate) the training processing.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

damage could produce impairments predominantly in the startstate (addition of noise) or predominantly in the endstate (increase in discriminability, removal of connections). Third, damage could produce impairments that were global (addition of noise, removal of connections) or selective (increase in discriminability). This complex relationship exists because the training process can sometimes play a crucial, compensatory role after damage occurs to the startstate of a learning system. However, it can play this role only to the extent that resources permit, and only to the extent that the representation of the domain remains reliable. To the extent that this is a valid model of cognitive development, the suggestion here is that the process of development will indeed contribute to patterns of deficits in a single system, but that the exact contribution will depend on the type of damage and the structure of the problem domain.[8]

With regard to the second issue, the results suggested that, at most, broad similarities were evident in the deficits caused by startstate and endstate damage, for example in the cases of lesioning and adding noise. However, in both cases, detailed examination revealed that the patterns of deficits were different – the behavioural impairments across related measures (regular, exception, and rule) did not line up. Again, to the extent that this is a valid model of cognitive development, the results do not support the idea that developmental and acquired deficits will produce precisely the same patterns in a single system.

It is instructive to see why this was the case. Take the example of lesioning the model. In the case of endstate deficits, the decline of regular and exception patterns was more closely tied because both pattern types shared a representational space that was being damaged. In the case of startstate deficits, the potential representational space was reduced, but the training process allowed the regular patterns to dominate the space that was available. The result was a system in which exception patterns were eventually squeezed out. These two impaired systems did not share a common final deficit because there is a distinction between a process of deleting parts of a representational space that is already occupied and the outcome of a process of occupying a representational space in which the initial size has been reduced.

However, the results by no means rule out the possibility that learning systems can be damaged in *different* ways prior to and following training, such that they exhibit identical endstate behavioural impairments. Generally, one must be very cautious about assuming identical causes in the case of identical outcomes. It is certainly the case that in connectionist models of developmental disorders, different forms of startstate damage can produce similar endstate behavioural impairments, as we saw in the case of phonological dyslexia and past tense formation in Williams syndrome. However, we have not yet unearthed any convincing examples in our own work or in the literature that startstate and endstate damage can separately produce identical endstate deficits in these networks. Time will tell on this point.

8. Simulation Two: Testing the assumption of Residual Normality in a simple connectionist learning system

8.1. Introduction

Many claims for RN relate to static adult models containing multiple independent, functionally specialised components. These components are supposed to fail separately under both acquired and development damage. Claims about developmental damage, however, are quite inappropriately applied to such models, because they are not models of development (nor do they pretend to be). In this simulation, we address the issue of RN in models with specialised components that are the product of a learning process.

How do specialised processing components arise in the cognitive system? Most connectionist models of cognitive processes have focused on single domains – in effect, they have been models of components within a *modular* system (see discussion in Karmiloff-Smith 1992). Less work has examined how specialised components may actually emerge from an initially undifferentiated computational substrate. Evidence from the study of brain processes suggests that the neocortices of newborns are less structurally differentiated compared to those of adults, and that cognitive processes are less localised in this early substrate (e.g., Johnson 1999). However, the key question regarding specialised structures has been whether their emergence during development reflects the unfolding of a maturational blueprint, whether their emergence depends entirely on experience, or whether it reflects a gradual process of modularisation that lies somewhere between these two extremes (Elman et al. 1996; Karmiloff-Smith 1992).

In a recent review, Jacobs (1999) argues that the evidence points to the experience-sensitive theory of specialisation. He discusses three computational approaches that have sought to model the experience-dependent emergence of structure. In the first approach, called *mixture of experts,* the initial computational system is assumed to be computationally heterogeneous. There are components that, while not dedicated to processing any particular content, have different computational properties. These components compete to perform the computations corresponding to a new cognitive domain. The component whose computational properties best fit the demands of the domain, known as a *structure-function correspondence,* will win the competition and come to specialise in processing that domain in the future (see Jacobs 1997; Jacobs et al. 1991). In the second approach, called *neural selectionism* or *parcellation,* the initial computational system has a surplus of connections. However, during learning, many of these connections are weeded out, whereas others are stabilised depending on usage. In addition, a locality constraint favours the stabilisation of connections between nearby processing units. The result is that nearby units communicate with each other and come to perform the same functions, whereas those far apart do not communicate and come to specialise in different functions (Jacobs & Jordan 1992; Johnson & Karmiloff-Smith 1992; see Plaut 2002, for a recent application to a cognitive model of naming and gesturing). In the final approach, called the *wave of plasticity,* the initial computational system experiences differential responsiveness to learning, both spatially and temporally. Conceived of as a sheet of computational units, plasticity is reduced over time with one side of the sheet losing its plasticity earlier than the other. The result is that the later maturing units can use the functions computed by earlier maturing units as input, and derive more complex and abstract computational functions from them – in essence, the later maturing units specialise on the more abstract or high-level aspects of the problem domain (Shrager & Johnson 1996).

Let us assume that the outcome of normal development

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

is a set of specialised components in the endstate, which can be revealed by adult neuropsychology. For our purposes, the relevant question is, if the developmental process itself is pushed off course in a developmental disorder, could this also alter the nature of the specialised structures that are the outcome of development? Little computational work has explored this question (see Oliver et al. 2000 for some preliminary work). It seems possible that specialisation could be disrupted in any of the above computational approaches. An alteration in the initial set of computational primitives or in the competition process could disrupt specialisation in the mixture-of-experts approach. An alteration in the method or timing of pruning long connections could disrupt specialisation in the neural selectionism approach. An alteration in the timing of plasticity changes could disrupt specialisation in the wave of plasticity approach.

Decisive evidence has yet to be put forward demonstrating radical differences in specialisation in developmental disorders, although some hints have been made in this direction. For example, Karmiloff-Smith (1998) has speculated that the cognitive processes of individuals with Down syndrome may be characterised by insufficient specialisation, perhaps because of a failure to prune long connections during development. However, conceptually, it is not yet clear to what extent one could compromise the emergence of specialised cognitive structures in a disordered state and still produce a viable cognitive system.

To extend the static adult damage model to developmental disorders is to make a more precise claim, however, that damage may be highly selective and thwart the development of a single specialised module. The question now becomes, if specialisation is not predetermined, under what conditions will the rest of the system develop normally despite this early selective damage? Fortunately, some existing models of reading and past tense formation allow us to explore this question. In these models, *structure-function correspondences* have been used to generate emergent specialisation in connectionist learning systems with multiple processing routes. Such networks include two processing routes in an initially content-free network.[9] The routes have different computational properties, and these properties line up respectively with the computational requirements of learning regular and exception patterns (see sect. 8.2). The result in both a reading model (Zorzi et al. 1998) and a past tense model (Plunkett et al. 2001) was partial specialisation of the two routes to processing regular and exception patterns (see Westermann 1998 for a related constructivist approach). We sought to evaluate the assumption of RN using this dual-route model, and in particular to answer the following question: Does disruption to one route prior to training alter the function that the initially intact route takes on at the end of training?

8.2. Simulation details

Architecture. The architecture of the dual route network is shown in Figure 1b. The feedforward network included an input layer and an output layer connected via two processing routes. The *Direct* processing route comprised a set of connections linking the input and output layer. The *Indirect* processing route connected these two layers via an intermediate layer of 20 hidden units.

Structure-function correspondences. When exception patterns must be learned in the face of a majority of regular patterns, additional computational resources are necessary. Specifically, although a two-layer network can learn the mappings for a set of purely regular patterns, hidden units are necessary to mark out the inconsistency of the exception patterns, typically involving the use of the three-layer architecture. These effects are not all or nothing. A two-layer network can tolerate a small proportion of exception patterns in the training set; exception patterns themselves can be more or less inconsistent with the regular patterns and therefore more or less demanding of hidden units to mark out their inconsistency. (In computational terms, the role of hidden units in overcoming the inconsistency between regular and exception patterns is a question of *linear inseparability* – see Elman et al. 1996, Ch. 2 for an introductory discussion.) Furthermore, the disadvantageous effect of inconsistency can be mitigated by increasing the frequency of exception patterns in the training set. Broadly, then, in a network combining a two-layer architecture and a three-layer architecture in separate routes, the two-layer route will be best fitted to learn the regular patterns, and the three-layer route will be required to learn the exceptions, more so with the greater the inconsistency of the exception patterns with the regular patterns. Structure-function correspondences can drive specialisation in error correction networks with multiple routes, because there is competition between each route to reduce the disparity between output activations and the training target. If one route succeeds in reducing the disparity, no error signal is left to change the weight strengths in the other route(s).

Our training set includes three types of exception pattern. Those based on the No Change past tense paradigm (*hit-hit*) are the least inconsistent with regular verbs, because as with regulars, the verb stem is reproduced, whereas the affix is omitted. We term these exceptions EP1. Exception patterns based on the Vowel Change paradigm (*hide-hid*) are more inconsistent, because in addition to the omission of an affix, the central vowel of the verb stem must be transformed. These exceptions we term EP2. Finally, exception patterns based on the Arbitrary past tense paradigm (*go-went*) are the most inconsistent with the regular patterns, because the verb stem must be entirely transformed and the affix omitted. We might expect these patterns to be most dependent on the hidden units of the Indirect processing route. However, in the past tense domain, it is argued that arbitrary past tenses can only be retained in English if they are of very high token frequency. In the current training set, arbitrary mappings were presented three times as often as other forms. We term this most inconsistent exception type EP3f, to reflect the fact that high frequency may modulate patterns of specialisation.

Training set. The training set was identical to that in Simulation One, except that instead of two arbitrary patterns in the training set, there were 10 such patterns. This permitted a more sensitive evaluation of performance on this pattern type.

Learning algorithm. The learning algorithm was the same as that in Simulation One.

Residual Normality condition: In addition to the normal training scheme, for comparison the model was also trained under a Residual Normality condition. This condition assumed "guided specialisation" (see sect. 9). Here, the Direct route was trained on Regulars alone, and the Indirect route trained on Irregulars alone. Guided specialisation in a mul-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

ticomponent model requires an external control system to coordinate the subsequent function of the trained components (see, e.g., the "Blocking" device in Pinker's [1991] dual-mechanism past tense model). For simplicity, the control system was assumed in the RN condition. Routes were trained and tested independently.

Performance measure. Performance was measured using a nearest-neighbour calculation based on output activations, and scores marked as percentage correct. Specialisation of a particular pattern type to a particular route was evaluated by selectively lesioning the network at the relevant point in training. The Direct route and the Indirect route were separately given a probabilistic lesion of their weights with $p = .5$ (50% of all weights in that route). If damage to the Direct route caused more impairment on a pattern type than damage to the Indirect route, it was assumed that the function for this pattern type was specialised to the Direct route, and vice versa. Note that removal of 50% of the connections may not have equivalent effects on the Direct and Indirect routes, because the latter has two layers of weights. However, we were concerned here with differential effects between pattern types rather than routes. We did check for interactions, that is, the possibility that pattern types might show differential sensitivity to damage in each route whereby, for example, a pattern type may appear to be specialised to one route at 50% damage but the other route at 10% damage. This possibility was explored by carrying out endstate lesions with probabilities of .025, .05, .1, .2, .25, .5, and .75. Although, overall, the Indirect route showed greater sensitivity to damage than the Direct route, there was very little modulation of *relative* specialisation levels of the pattern types across damage levels. However, the *absolute* level of specialisation was affected by whether the level of damage was so great or small that it produced floor or ceiling effects in regular or exception performance. A level of 50% lesioning was used to assess specialisation because this was in the midrange of sensitivity for both regular and exception patterns.

Implementation of pretraining damage. This simulation sought to explore the implication of specific damage to either of the two routes prior to training. This was achieved by removing different proportions of the weights in each route. Probability levels of .6, .75, .9 and 1 (removal of entire route) were applied. In addition, this level of damage was performed on both routes simultaneously, as a control. After initial damage, training proceeded as normal, and levels of specialisation were then assessed in the endstate.

Replications. Results were averaged across six networks with different initial random seeds for each level of startstate damage. Probabilistic lesions were carried out 10 times and the results averaged. Graphs include standard error bars as an indication of variability across the six network replications.

8.3. Results

Figure 5c represents an index of specialisation, where positive values represent specialisation to the Direct route, and negative values represent specialisation to the Indirect route. This index corresponds to the differential impairment caused by lesioning to a single route. The first point to note is that specialisation is only partial. Using this measurement technique, in Figure 5c damage to the Direct route can cause a maximum decrement in performance only 53% greater than damage to the Indirect route. Damage to the Indirect route can cause a maximum decrement 29% greater than damage to Direct route. Secondly, most patterns show a shift towards using the Indirect route later in training.

To understand this latter point, it is important to realise that the two routes of the network do not just differ in their computational properties, but also in their plasticity. By virtue of the learning algorithm, weights from the input layer to the hidden units change more slowly than the weights directly connecting input and output layers. In effect, this network comprises one relatively more "stupid" but fast-changing route, and one relatively more "clever" but slow-changing route. Early on in training, successful performance is largely a result of the Direct route, and this performance is best on regular patterns, generalisation of the rule, and the EP1 exceptions – those that are least inconsistent with the regular patterns. Subsequently, the slower-changing Indirect route increasingly contributes to performance, especially for the exception patterns. However, by the end of training, both regular patterns and generalisation of the rule rely more on the Direct route, whereas all exception patterns rely more on the Indirect route. As expected, the more inconsistent EP2 patterns turn out to rely more on the Indirect route than the EP1 patterns. However, the higher frequency of the EP3f patterns means that their greater inconsistency does not lead to Indirect route specialisation any more than that shown by EP2; higher frequency allows these patterns to recruit more processing from the Direct route in the face of the dominance of regular patterns.

In summary, this network shows emergent specialisation of function of different pattern types to separate processing structures. Although this specialisation is not complete, our concern here was to establish a baseline level to explore the effect of initial, route-specific damage to this process. Figure 6 shows (from left to right) the result of startstate lesions to the whole network, to the Indirect route in isolation, and to the Direct route in isolation. Figure 6a illustrates the effect of these startstate lesions on endstate performance, while Figure 6b shows their effect on the profile of specialisation in the endstate network.

Figure 6a demonstrates that when the entire Indirect route is removed, endstate performance on regular patterns and generalisation is only mildly impaired. However, performance on the exception patterns reveals a marked decrement, particularly for EP1 and EP2, which are not protected by increased frequency. When the entire Direct route is removed, regular patterns are impaired to a greater extent. Exception patterns are also impaired, but less than when the Indirect route is lost. Removal of the Direct route, however, produces a marked deficit in rule generalisation. In isolation, then, each of the two routes will attempt to acquire both regular and exception patterns, but each does so less efficiently than in the dual-route system. The routes in isolation both produce decrements in exception performance, but relatively speaking, the Indirect route is less able (but not unable) to learn the regular rule, and the Direct route is less able (but again, not unable) to learn the exception patterns.

These results mark the maximum compensation that is available to the network. Figure 6b demonstrates the specialisation when there is residual processing capacity in the damaged route.[10] Here the data are unambiguous. When there is initial damage to the Indirect route, specialisation of

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

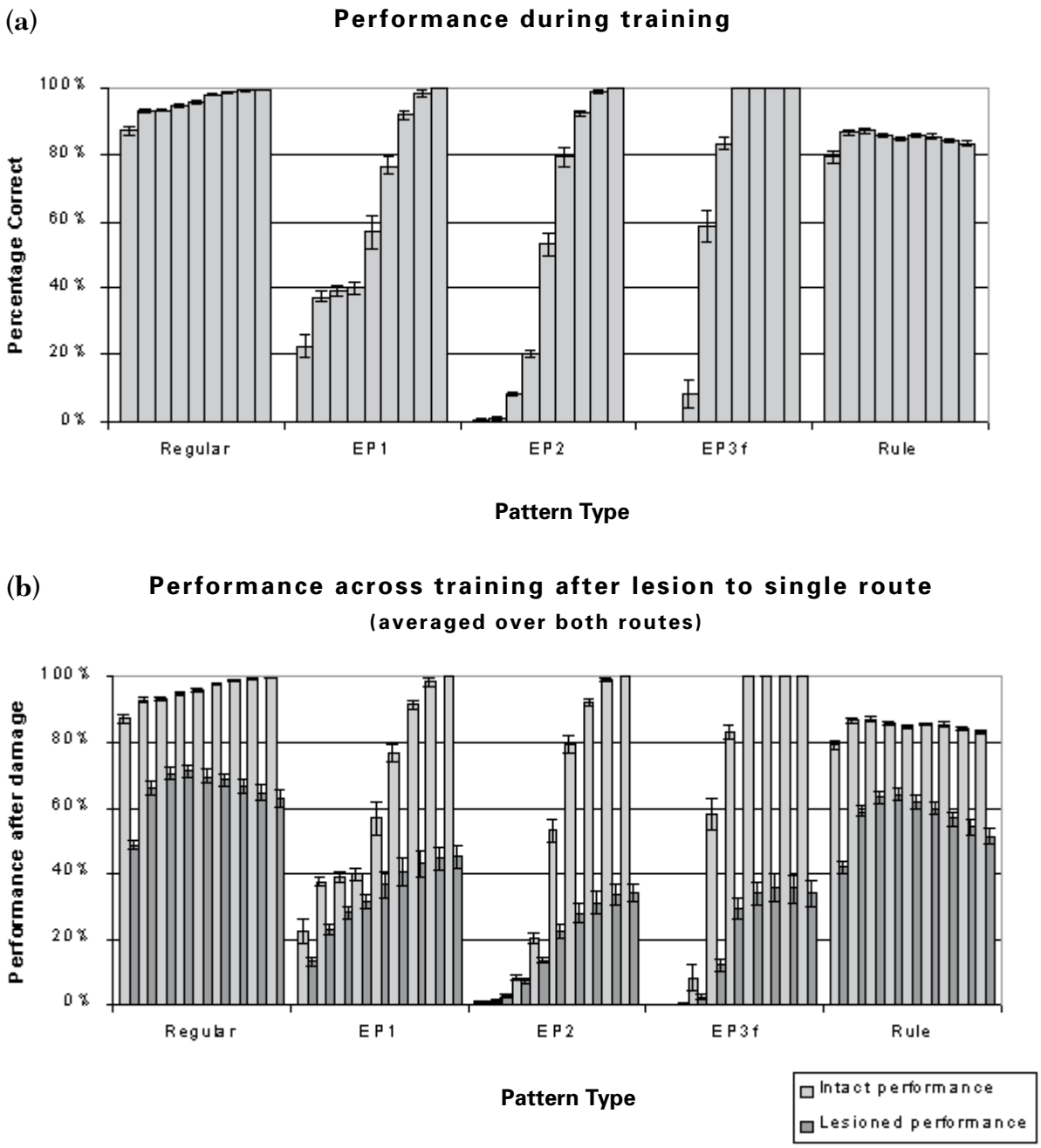


Figure 5. (a) Performance on each pattern type during training in the dual-route network, at 10, 25, 50, 100, 250, 500, 1,000, 2,000, and 5,000 epochs of training. See Simulation Details (Sec. 8.2) for description of each pattern type. (b) Performance after a 50% lesion to each route, averaged over the two routes, carried out at each point in training. The 50% lesion was used to measure specialisation of function to each route. (c) Specialisation of function for each pattern type during training, indexed by the differential impairment caused by damaging each route in isolation. Positive values indicate specialisation to the Direct route; negative values, to the Indirect route. Error bars show standard errors across network replications.

function increasingly moves over to the intact Direct route. When there is initial damage to the Direct route, specialisation of function increasingly moves over to the intact Indirect route. A crucial lesson is demonstrated by this simulation: The assumption of Residual Normality does not hold in this learning system. Damage one route and the other route will *not* develop normally. It will compensate, and take on part of the function of the damaged route, at the cost of poorer performance across all pattern types. It cannot be taken for granted that every learning system will show RN.

This simulation is interesting in two other respects. In the condition where both routes experienced initial damage, the overall outcome was reduced specialisation. A system with uniformly reduced computational resources does not have the luxury of allocating functions to different components. The effect of resources on specialisation is a point to which we will return shortly. Second, the 50% lesion used to measure specialisation produced greater deficits in systems that had experienced startstate damage to either or both routes than those that had not. Importantly, even if

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

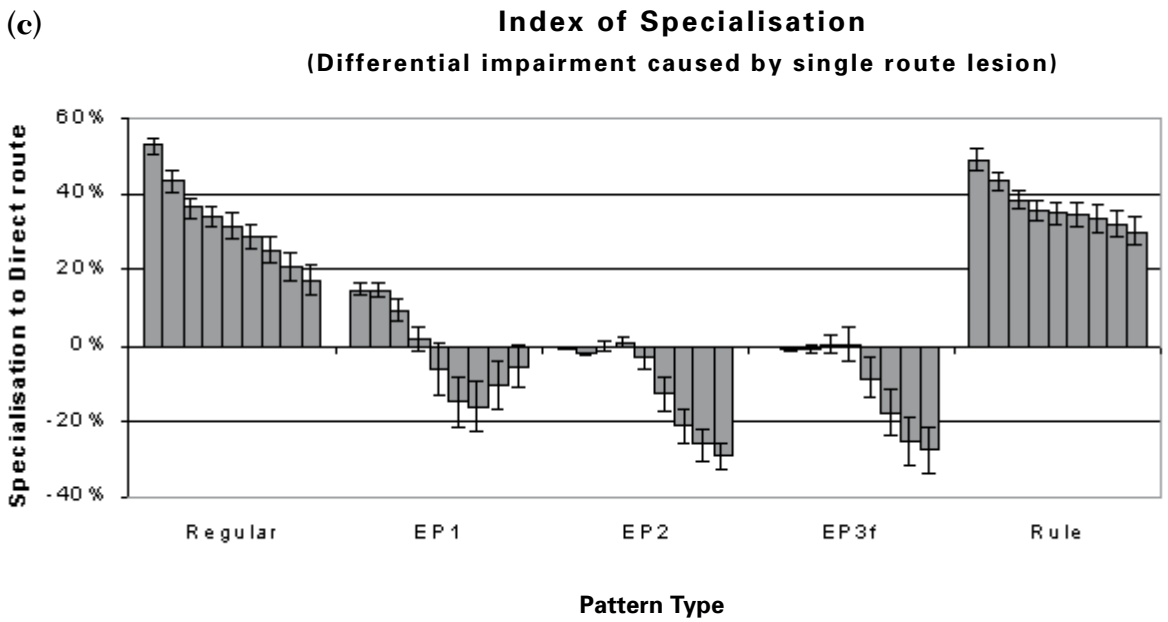


Figure 5. (*continued*)

systems that experience early damage achieve reasonable endstate performance, they remain more vulnerable to subsequent disruption.[11]

8.4. Discussion

Given the compensatory characteristics of the training process evident in Simulation One, it is perhaps unsurprising to find similar compensation here. When the two routes of a dual-route network show endstate specialisation, startstate damage to either route results in compensation by the intact route and poorer performance overall. If RN had held in this model, damaging a route in the startstate would have led to the (endstate) loss of the same function for which that route was responsible in the endstate of the normal unimpaired model, with normal function in the initially undamaged route. This is precisely the claim that is made (for example) in the case of phonological dyslexia, where a developmental problem in the GPC route of the reading system is assumed to lead to the same pattern of behavioural impairments as damage to the GPC route in the adult state (Coltheart et al. 2001), with the lexical routes functioning normally in both cases.

The feedforward network presented in Simulation Two is similar to connectionist systems that have been used to successfully capture a wide range of developmental phenomena. These systems plainly do not demonstrate RN. Nevertheless, as we saw in the Introduction, RN is frequently postulated (albeit implicitly) in many explanations of developmental disorders. Perhaps, then, despite their success, current connectionist models are not the right sort of learning system to explain how the structure of the adult cognitive system comes about. Can other sorts of learning system show the emergence of specialised components while exhibiting RN after initial damage?

To date, artificial neural networks are the computational systems that have been most widely applied to the study of cognitive development. However, it is certainly possible that other approaches will come to the fore in the future, such as decision-tree learning, Bayesian methods, production systems, reinforcement learning, instance-based learning, genetic algorithms, or indeed other types of artificial neural networks. In the meantime, from the perspective of developmental disorders, it is vital to stipulate what additional constraints any such learning systems would need to incorporate to achieve RN. We will discuss five (somewhat overlapping) notions: (1) stronger structure-function correspondences, (2) stronger competition, (3) early commitment, (4) guided specialisation, and (5) restrictions on computational resources. Note that all of these notions are based on the assumption that endstate cognitive structure is experience-sensitive. RN can of course be *stipulated*, as per accounts that propose that modular structure in the cognitive system is prespecified and that if components develop, they do so independently. (Those who stipulate innate modularity would then need to justify this claim with evidence from developmental cognitive neuroscience – evidence that we currently believe to be wanting.)

9. Ways to achieve Residual Normality in systems with emergent modularisation

9.1. Stronger structure-function correspondences

Each of the routes in the dual-route network was able to show a fair degree of compensation for the functions of the other, suggesting that the correspondences between the functions of the two routes and the regular/exception structure of the learning problem were partially overlapping. One way to assure RN would be to have much stronger structure-function correspondences, whereby the computational properties of each route were entirely inappropriate for learning the patterns on which the other route specialised.

We might illustrate this idea by stepping outside our two-example domains for a moment and considering a connectionist model of the development of object-oriented behav-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

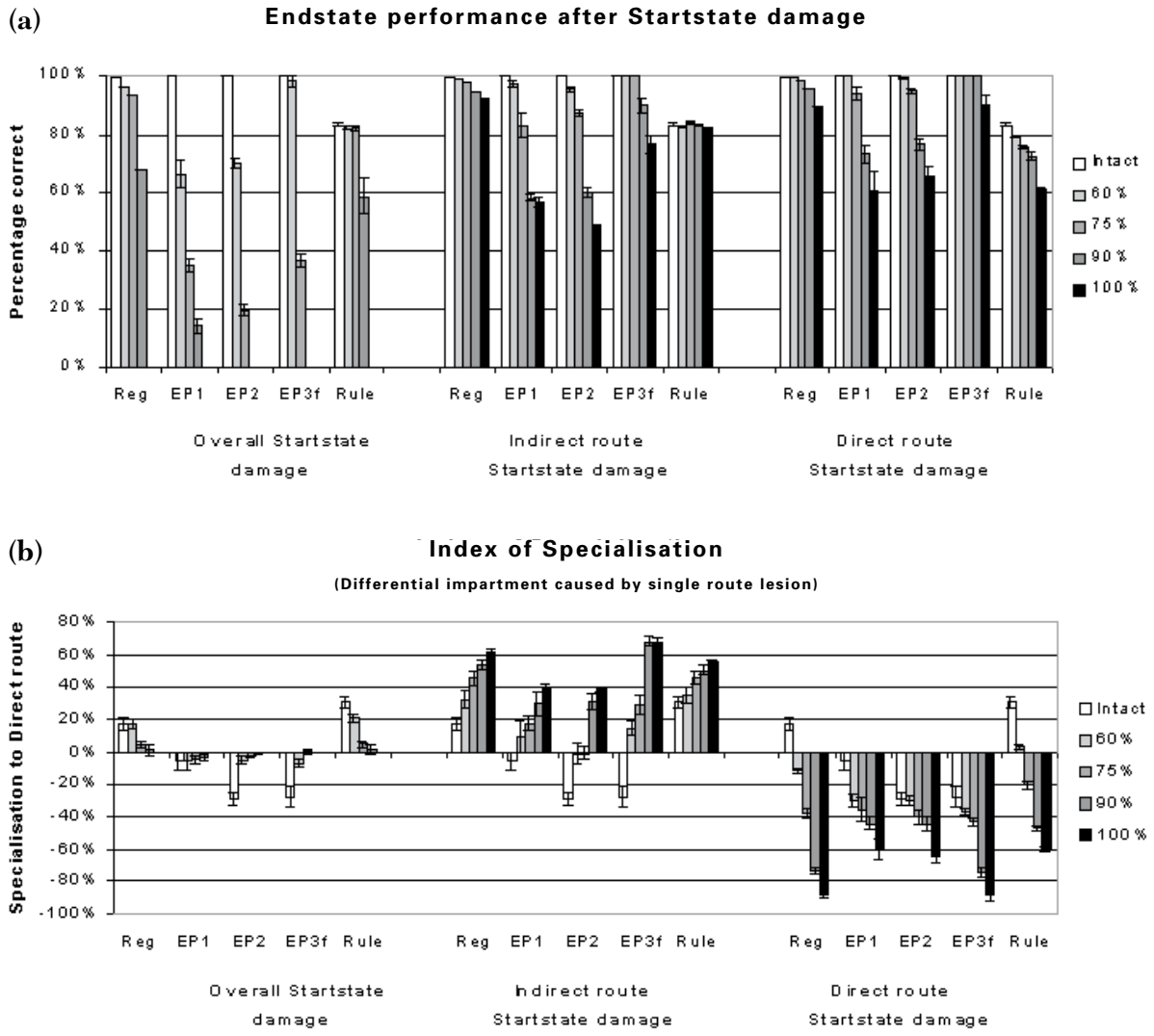


Figure 6. (a) Performance at the end of training following initial damage to both routes, to the Indirect route only, or to the Direct route only. (b) Patterns of endstate specialisation for each pattern type after initial damage. Error bars show standard errors across network replications.

iour proposed by Mareschal et al. (1999). This model used an input retina to represent the trajectories of various sorts of moving objects. The system had to learn to reach for some of the objects it saw but not for others. To achieve this task, it was given two routes, one that processed spatiotemporal information about the position of each object (the "where" channel), the other that processed featural information about the identity of each object, such as its colour and shape (the "what" channel). A final layer of the system combined the two processing routes to achieve reaching behaviour. The relevance of this model is that the computational properties of each processing route were very different. The "what" channel utilised competitive learning to achieve translation-invariant feature recognition across the entire retina. The "where" channel employed recurrent circuits to encode time-varying information about position. If either of these channels were to be damaged prior to the training process, the remaining route simply would not have the appropriate computational primitives to compensate for the function of the damaged route: Translation-invariant featural information contains no clues to location, and spatial trajectory information contains no clues about object identity.[12] The result would be a system with RN, where the remaining initially intact route would develop the only skill it had the capacity to learn and nothing else.[13]

9.2. Stronger competition

In the mixture-of-experts approach to specialisation, separate components compete to represent a new domain. The component best able to represent the domain is given "sole rights" to it, and the other components are inhibited. In the dual-route network, however, both routes worked in harness to learn the appropriate mappings – neither was prevented from adjusting its weights to improve performance on any pattern. The result was to encourage cooperation between the routes. Much stronger competitive processes might permit a component to claim sole rights only to patterns that best suited its computational properties, and might inhibit it from acquiring patterns outside of that set. If the same component were to win the competition in both normal and atypical development (a big "if"), then the component would exhibit RN and no compensation.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

9.3. Early commitment

If plasticity reduces rapidly over time, then early commitment may contribute to RN. In this scenario, early commitment of separate components to functions must occur before the developmental disruption takes place. This can be either because the damage has not yet occurred, or because the underlying damage does not make itself apparent because it is not relevant for the cognitive processes appropriate for the current stage of development (see discussion in Thomas 2003). Early and irreversible (yet experience-dependent) specialisation could contribute to a modular system where RN holds (see, e.g., Miller & Erwin 2001), provided an account exists of the required delay in the emergence of the developmental disruption.

9.4. Guided specialisation

In the dual-route network, prior analysis of the computational properties of two- and three-layer networks suggested that the Direct route would be better suited for learning regular patterns and the Indirect route would be better suited to learning the exception patterns. In principle, we could have determined to label each pattern as Regular or Exception in advance, and then only allowed the Direct route to alter its weights in response to Regular patterns and the Indirect route to alter its weights in response to Exception patterns. Unsurprisingly, the result of this form of guided training would be independent specialisation. Clearly if one route were damaged prior to training without any change to the advanced labelling system, this route would fail to learn the patterns assigned to it. The other route would be unaffected and would hence show RN.

In such a case, we are of course left with the burning question of where the advanced labelling information comes from. Some would argue it is innate. Alternatively, the labelling information could be the product of an earlier phase of learning, in which an analysis of the target domain identified the presence of regular and exception patterns *before* the dual-route system was engaged. It should be evident here that if one chooses to appeal to guided specialisation to support RN, then claims about the availability of a priori knowledge need serious substantiation.

9.5. Restrictions on computational resources

We saw in the results of Simulation Two that specialisation was eliminated under severe resource limitation. Resource limitations may also have implications for limits on compensation and, indirectly, for RN. This idea can be illustrated with an example from the past tense domain and claims made for the (computationally unimplemented) traditional model. The traditional model comprises two mechanisms, one employing rule-based representations and nominally responsible for learning regular inflections, and the other employing an associative memory and nominally responsible for learning exception inflections. If these two mechanisms are to specialise appropriately on their respective inflections (in the absence of external guidance; see Pinker 1999 for discussion of a possible mechanism for guided specialisation in this theoretical model), it is important that neither mechanism be too powerful. Given sufficient "rule" resources, for example, all past tenses could be learned in terms of a large set of rules (see, e.g., Ling & Marinov 1993; Taatgen & Anderson 2002). In contrast, given sufficient "associative memory" resources, all past tenses could be learned as specific instances. Generalisation to novel exemplars could be achieved in either case by an analogy-based strategy, such as similarity-to-the-nearest-known exemplar. Given two overly powerful mechanisms, the result would be duplicated processing systems, each able to perform the whole task. To achieve specialisation, therefore, one needs to restrict the resources of each mechanism. For example, the English past tense has one rule and about 150 exceptions. If the relevant mechanisms were restricted to these limits, then they would show little compensation in the event of damage to the other mechanism.

The case of SLI illustrates the point. Children with SLI show low levels of inflection on both regular and exception verbs, and poor generalisation of the regular rule to novel strings. Appealing to the traditional two-mechanism model of past tense formation, Ullman and Gopnik (1999), Pinker (1999), and van der Lely and Ullman (2001) have all argued that there is a startstate deficit to the mechanism intended to learn the regular rule. All that remains is the exception mechanism, which learns the past tenses of some exception and some regular verbs, and (presumably by analogy) can struggle to offer a few correct generalisations to novel verbs. The implicit claim here is that there is a *particular limit in computational resources* in the exception mechanism that prevents it from learning more than a handful of regular past tenses by way of compensation. Specifically, the exception mechanism must be able to learn the couple of hundred exceptions to explain its performance in normal development. But, to explain the lack of compensation in SLI, it must be unable to learn a couple of thousand high-frequency regular past tenses that would be sufficient to get by in everyday language use and therefore mask the regular-mechanism deficit. The explanation of the endstate impairment crucially relies on this precise memory stipulation which constrains compensation. (It is interesting that no empirical support is offered for such a memory limitation in the preceding accounts of SLI.)

In short, resource limitations may be a necessary component of RN but not a sufficient one.

10. Residual Normality and the inference from behaviour to structure

In section 4, we identified two developmental conditions where behavioural similarities between acquired and developmental disorders should not lead to the inference that they reflect selective impairments to the same components of a static adult model. One condition was when features of the problem domain determine the pattern of breakdown rather than features of the processing structure. In the simulations, this situation is reflected in the greater vulnerability of exception patterns, irrespective of damage type. However, closer inspection revealed different patterns of deficit to regulars, exceptions, and generalisation for different damage types, implying an effect of residual processing structures on performance.

The second condition was when similar patterns of behaviour in atypical development are generated by a different underlying structure of specialised components. In section 9, we have argued that the emergent pattern of functional specialisation depends on computational constraints operating during cognitive development. However,

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Jackson and Coltheart (2001) maintain that developmental disorders in their endstate are potentially no different from acquired disorders and can be used with reference to the normal adult model quite independently of the nature of the developmental process that produced them. One reading of their position is that, at a given moment in time, inferences can be made from behavioural impairments directly to underlying structure, irrespective of the developmental processes that produced the system.

Computational modelling allows us to explore this claim, because we can simultaneously generate patterns of behavioural deficits while knowing the underlying cause and the background developmental account in each case. Figure 7 illustrates a behavioural impairment generated from the dual-route model following selective damage to one of its routes. Two versions of the model were damaged, the one we have already studied in which RN does not hold, and a second version in which RN *does* hold (in this case, RN was achieved by guided specialisation of regulars and exceptions to the two routes. See simulation details in section 8.2). Figure 7 shows that both versions of the model can generate similar behavioural impairments after startstate and after endstate damage, (although in the absence of RN, regular performance is not quite at ceiling). However, as summarised in Table 1, the inferences that one can make from intact behaviour to intact underlying process, or from impaired behaviour to impaired underlying process, crucially depend on the developmental constraints of the system. With regard to *developmental deficits,* where RN holds, intact behaviour implies intact underlying process and a dissociation of independent structures. Where RN does not hold, intact behaviour implies atypical underlying process, in a system that has experienced compensation during training. One has to know the developmental story with regard to *acquired deficits,* however, the developmental constraints are not key in inferring underlying structure, because specialised processing components have formed prior to damage.

This simulation illustrates that the second condition identified in section 4 is a quite viable one. A pattern of intact and impaired behaviour can be produced by atypical underlying processes in the developmental case. It supports the idea that in developmental disorders, underlying structure cannot be inferred from behavioural deficits without explicit reference to the developmental constraints under which processing structures have emerged. In contrast, the logic of acquired deficits in the endstate is agnostic as to whether specialisation was prespecified or emergent.

11. Conclusion

In developmental cognitive neuropsychology, the hypothesis of Residual Normality is widely deployed, explicitly or implicitly, in explanations of developmental deficits, but with little supporting evidence. The motivation is clear – it allows developmental disorders to be interpreted according to static adult models that themselves have no developmental component. We have argued that this tendency to focus on specific deficits and take RN for granted has had serious consequences for the type of data collected in characterising developmental disorders. Researchers have been too prone to assume that superficially normal behaviour (e.g., in standardised tests) corresponds to normal underlying processes, irrespective of the developmental plausibility of RN in each cognitive domain. We believe this tendency has also impeded the discipline in building links to developmental cognitive neuroscience and developmental neurobiology, because the use of RN with static models stipulates rather than explores the developmental origins of deficits.

There may be no universal causal link between developmental and acquired disorders. Rather, RN may hold for some cognitive domains and not for others. Compensation and plasticity mitigate against it on a fine-scale decomposition of cognitive domains, but RN may obtain across some broader functional distinctions, such as ventral versus dorsal

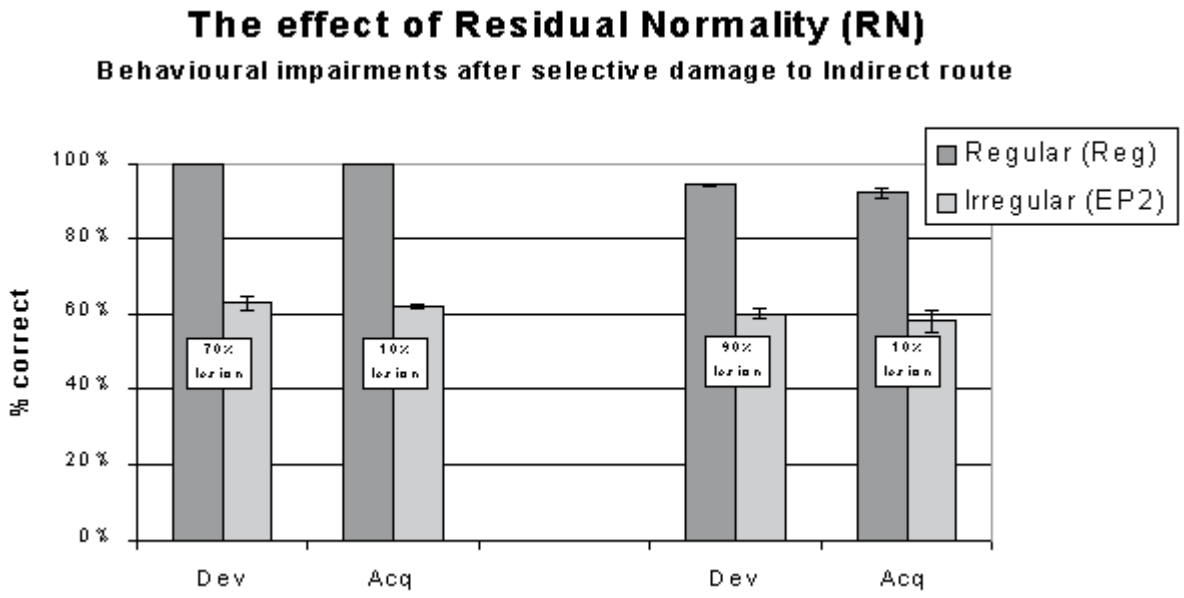


Figure 7. Selective behavioural impairment caused by startstate (Developmental) and endstate (Acquired) lesion to the Indirect route of the Dual-Route network, in the cases when RN operates and when it does not (see Simulation Details, sect. 8.2). From left to right, the lesions to the Indirect routes were 70%, 10%, 90%, and 10% of connections, respectively.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Table 1. *Inferences from behaviour to structure. In two different models, selective damage to the Indirect route of the dual route system causes a similar selective behavioural impairment to Irregular patterns (EP2), with largely intact performance on Regular patterns (Reg), in both developmental and acquired cases. In one model, Residual Normality (RN) holds; in the other it does not. The correct inference from behaviour to underlying process depends on the developmental assumption (of whether RN holds)*

	What is producing normal (intact) *behaviour?* [Reg]		Does the dissociation imply independent underlying processes?	
	RN operates	RN does not operate	RN operates	RN does not operate
Acquired deficit	Normal process in Direct route	Mix of Normal process in Direct route and residual process in damaged Indirect route	Yes	Yes *
Developmental deficit	Normal process in Direct route	Atypical processes in both Direct and residual Indirect route after development	Yes	No
	What is producing impaired *behaviour?* [EP2]			
	RN operates	RN does not operate		
Acquired deficit	Impaired process in residual damaged Indirect route	Predominantly Impaired process in residual damaged Indirect route		
Developmental deficit	Impaired process in residual damaged Indirect route	Atypical processes in both Direct and residual Indirect route after development		

*Partial independence (see sect. 8.3)

systems, anterior versus posterior systems, or cortical versus subcortical systems. The plausibility of RN across various systems and domains is a topic in need of urgent investigation.

We conclude that RN cannot be taken for granted, but must be argued for on the basis of developmental evidence. It is essential that developmental *disorders* be compared against developmental *models,* and that the process of ontogenetic development itself be taken seriously.

ACKNOWLEDGMENTS

Many thanks to Max Coltheart for helpful discussions on the arguments within this article. We thank Denis Mareschal and Birkbeck College, and Kim Plunkett and the McDonnell-Pew Centre for Cognitive Neuroscience for the use of time on their computers. Thanks also to Michael Rutter, Denis Mareschal, Marc Joanisse, Lorraine Tyler, and Mike Harm for helpful comments. This work was funded by MRC Project Grant No. G9809880 and MRC Programme Grant No. G9715642 to Annette Karmiloff-Smith.

NOTES

1. With a much larger sample size, Thomas et al. (2001) subsequently found that this deficit appeared to be an artefact of delayed language development in Williams syndrome. Their results suggested that the main difference between the Williams syndrome group and controls was a reluctance to inflect novel forms. However, for current purposes, Clahsen and Almazan's (1998) claim serves to illustrate the explanatory framework within which those authors were operating.

2. Where we use the term *compensation* in this article, we do not mean that alterations in response to damage elsewhere in the system must necessarily produce a beneficial outcome. Undamaged (i.e., previously normally developing) components may have their development disrupted through attempts to compensate for the impairment, such that they function more poorly on the task that they normally acquire. For example, in the simulations presented in section 8.3, a component that normally specialises in learning regular patterns achieves this function less efficiently when, as a consequence of damage elsewhere in the system, it also has to learn exception patterns. Similarly, a component that normally specialises in exception patterns achieves this function less efficiently when, as a consequence of damage elsewhere in the system, it also has to learn regular patterns.

3. Or indeed more routes, less routes, or other routes.

4. Because of space restrictions, our discussion focuses on fixed architecture models. We suspect similar arguments will hold for constructivist models that change their architecture as a function of learning. However, the general interaction of early damage and constructivist processes of development requires further investigation (see Thomas & Karmiloff-Smith, in press, for discussion of constructivist models and developmental disorders in the specific case of past tense formation).

5. In principle, models of atypical development could employ damage that occurs *during* training. For example, in typical development, a computational constraint might change as a function of developmental time. Such constraints might include the rate at which connections are pruned, or the rate at which connection strengths are altered in response to error signals. Although constructivist neural networks add hidden units as part of the learning algorithm, thus far, limited modelling work has explored how typical development may *rely* on continuously changing computational constraints, let alone how atypical development may arise from dysfunctions in such changes.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

6. The developmental process is equated here, as in most connectionist models of development, to the training process – see Thomas and Karmiloff-Smith (2003) for discussion of this assumption.

7. A debate continues as to (1) whether patients can indeed show decrements in past tense performance whereby regular inflections for existing words are much poorer than exception inflections (see, e.g., Bird et al. 2002; Ullman et al. 1997); and (2), whether deficits in regular inflection for existing words are necessarily associated with phonological deficits, as predicted by the Joanisse and Seidenberg (1999) model (see e.g., Bird et al. 2002; Tyler et al. 2002b).

8. The results reported here were fairly robust to changes in hidden unit numbers and input-output coding schemes. The coding scheme determines the complexity of the mapping problem, and the architecture in some sense determines the computational capacity of the system. Networks closer to their capacity limit (for example, when a more complex encoding of the mappings is used or there are fewer hidden units) are more vulnerable to damage, but breakdown patterns appeared to be robust.

9. Although the network is content-free in terms of an absence of domain-specific knowledge, it nevertheless has knowledge of a sort, in terms of the computational mechanisms that constrain the content that it can learn (see Elman et al. 1996).

10. In Figure 6b, a value of specialisation is reported when one of the routes has experienced a 100% startstate lesion. On the face of it, this may seem confusing. This value is in fact derived by calculating the difference between the decrement in performance caused by a 50% lesion to the intact route and that caused by a 50% lesion to the (100%) damaged route. However, because no further damage is possible beyond 100%, the second term in this comparison is zero. The specialisation value therefore just indexes the decrement in performance caused by a 50% lesion to the single intact route.

11. With regard to the robustness of the findings for the dual-route network, decreasing the hidden units decreased functional specialisation, whereas increasing hidden units had little effect. Recalling that the coding scheme determined the complexity of the mapping problem and the architecture the computational capacity of the system, increasing the mapping complexity reduced specialisation for a fixed capacity. The network showed no critical period for specialisation when damage occurred at different points in training, provided training time was equated (Thomas, in preparation). Changes in learning algorithm, learning rate, discriminability, and noise were not examined in these simulations.

12. As implemented by Mareschal et al., this isn't strictly true. If the quality of featural information in this model were to vary to any extent across the retina, this variation could in principle be used to track position, so that the *what* channel would implicitly code information about location. Moreover, spatiotemporal codes must retain information about shape to predict the movement of all the parts of an object. If objects differ in shape, the *where* channel in the model would implicitly encode information about identity (Mareschal, personal communication, May 2001). However, it seems unlikely that these residual sources of information would be sufficient to generate robust object-reaching behaviour in an impaired model.

13. For illustration, as applied to Coltheart et al.'s (2001) non-developmental model of reading, structure-function correspondences would ensure RN under the following conditions: (1) The GPC route is constrained to learn context-sensitive rules that encode *just enough* context to enable this mechanism to relate graphemes to phonemes, but *not so much* context that it is able to learn whole exception words. For example, a priori, the system must be able to learn variously that *t* → /t/, that *h* → /h/, that *t* followed by *h* in *th* → /T/, that *i* → /i/, that *i* followed by *e* in *_i_e* → /I/, that *i* followed by *g* followed by *h* in *_igh* → /I/, but *not* that *i* preceded by *p* and followed by *n* and *t* in *pint*→/pInt/. The difficulty of calibrating initial computational constraints so precisely is illustrated by the fact that when Colheart et al. (1993) attempted to deliberately handcraft a rule-learning algorithm to acquire solely the GPC regularities (that is, the regular pronunciations alone), they could not prevent the GPC route from learning to pronounce 5 of the 47 exception words in the training set. (2) The lexical route is constrained to learn representations of written word forms that preserve zero similarity between the compositional structure of different input forms. Any remaining similarity would permit analogy-based generalisation of novel pronunciations, based on pronunciations of orthographically related lexical forms. For example, if *nog* produced partial activation of the lexical representations for *nod* and *dog*, the pronunciations /nod/ and /dog/ could be combined to generate a pronunciation for the novel word. This computational constraint entails that there is a strict winner-takes-all competition between lexical representations prior to the activation of phonemic output, where "no winner" implies no remaining activation of lexical representations bearing any similarity to the input. The psychological plausibility of these two computational conditions within the developing reading system remains to be seen.

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

The beauty of models for developmental disorders

J. Briscoe

Cardiff University, Tower Building, Park Place, Cardiff, CF10 3YG, United Kingdom. **briscoejm@cardiff.ac.uk**

Abstract: Cognitive frameworks provide important means for uniting concepts of specificity, cognition, and dynamic change in development. Two points are challenged by evidence from special populations: (1) that boundary constraints such as Residual Normality and a cognitive "endstate" compromise the use of cognitive models; and (2) the developmental process itself automatically rejects either Residual Normality or residual deviance from typical development.

Thomas & Karmiloff-Smith (T&K-S) question the validity of adopting cognitive frameworks derived from adult neuropyschology for motivating study of developmental disorders. They highlight the danger of invoking static cognitive models as an explanatory framework because they cannot be "straightforwardly related" (target article, sect. 3.2, para. 10). Here, it is argued that cognitive frameworks do enrich our understanding of overlap between "specificity" and "cognition." For example, children of premature birth may carry a developmental insult resulting in brain abnormalities and cognitive impairment (Briscoe et al. 2001b; Isaacs et al. 2000). By regression of everyday memory scores against language abilities, Figure 1 demonstrates the selective memory impairment of several children of early school age who were born prematurely.

According to recent modification of working memory (Baddeley 2000), one could invoke a working hypothesis of "episodic buffer" deficits to explain such selective memory impairment without requiring assumptions of "Residual Normality." In this case, a cognitive model provides a good benchmark for further empirical testing. T&K-S also argue that cognitive frameworks in developmental disorders are limited by assumptions of cognitive stability such as a stable "endstate." Here it is argued that the in-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

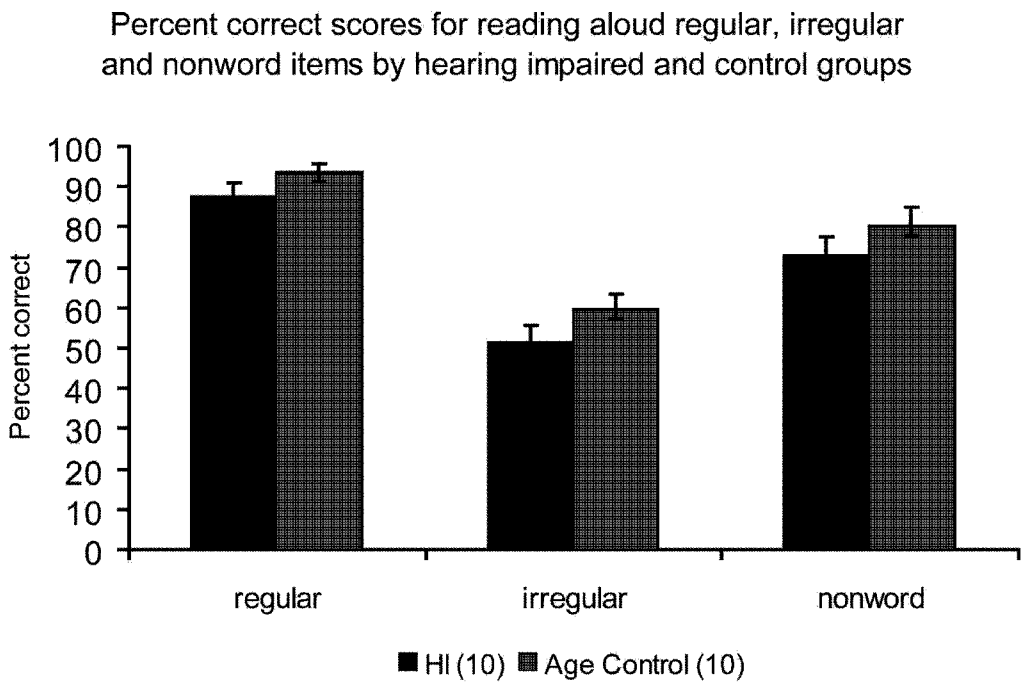


Figure 1 (Briscoe).

corporation of dynamic change in cognitive models has implications beyond developmental disorders. For example, how do cognitive models interpret performance in semantic dementia of a progressive type or dyslexia in older adults? This challenges the notion of a secure "endstate": It is not clear when the parameters of an "endstate" should be invoked for a given point during development of a modular system or when they should be revoked because of change in the aging adult brain. Understanding dynamic change will enrich our understanding of some facets of cognition beyond isolated examination of frozen "endstate" behaviour or "overdeveloped" learning systems in adults. For example, the developmental amnesic syndrome has enriched theories of cognitive memory by establishing core dissociations between semantic learning and episodic recall but not recognition, as previously suggested by adult models (Vargha-Khadem et al. 1997).

T&K-S argue that cognitive models are impoverished by the exclusion of developmental process (sect. 3.2, para. 5). However, it is not clear that including a developmental process necessarily provides a compelling addition to the cognitive model. One example is the role of auditory perceptual skills for reading acquisition. One route suggests that remediation of deficits in auditory perceptual processing underpins development of reading skills in children (Merzenich et al. 1996). Alternatively, developmental data from children with hearing impairment find that auditory perceptual impairment does not directly impair reading skill (Briscoe et al. 2001). Table 1 shows word and nonword reading aptitude for children with hearing impairment (HI), as compared to children with Specific Language Impairment (SLI) and age-matched controls (CON) (taken from Briscoe et al. 2001). These scores suggested reading aptitude despite widespread deficits at the phonological level and auditory perceptual deficits in the HI group. In this sense, increasing the specification of cognitive models of reading to include developmental change in auditory perception, for example, does not necessarily enrich our interpretation of reading difficulties and strengths in childhood.

An advantage of using cognitive models is precisely the nature of the boundary conditions that support explanations within the model. For example, the Castles and Coltheart (1993) model of reading can facilitate a hypothesis-testing approach that could benefit interpretation of the pattern of deficits. Figure 2 (Halliday & Briscoe 2002) demonstrates that children with hearing impairment do not produce a specific pattern of impairment across the word lists of Castles and Coltheart (1993). Therefore, reading strategies in this group cannot be sufficiently explained in terms of semantic, lexical, or phonological routes to reading according to this model. One does not need to revoke the use of cognitive models in developmental disorders or invoke a computational account of singular changes in a cognitive system to demonstrate the necessity for further investigation of routes to reading competence: The result is data-driven by directed empirical investigation in these children.

A second advantage of using a cognitive framework is the potential for inclusion of a performance-related account rather than making a priori assumptions about competence within a skill or domain. By the logic of T&K-S, the developmental process that would support reading competence in children with hearing impairment is construed to be part of the source of "deviant behaviour" that it *must* at some level be, because of their hypothesised interaction of a beginning state with environmental factors (sect. 3.2, para. 4). Because Table 1 shows reading aptitude in children with hearing impairment, it is argued here that the developmental process does not set boundary conditions for inferring a residual *lack of competence,* nor an assumed competence in any domain.

Table 1 (Briscoe). *Mean (SD) age and standard scores on standardized tests of reading for hearing-impaired, age-matched control, and specific language-impaired groups*

	HI N=19	SLI N=14	CON N=20
Age (years)	8.66 (1.36)	8.96 (1.15)	8.49 (1.55)
Word reading	106.56 (14.28)	76.00 (9.23)	110.74 (16.56)
Reading comprehension	100.22 (13.17)	74.14 (10.16)	109.95 (10.51)
Nonword reading	98.11 (14.61)	77.57 (13.12)	107.95 (14.13)

Source: Taken from Briscoe et al. 2001

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

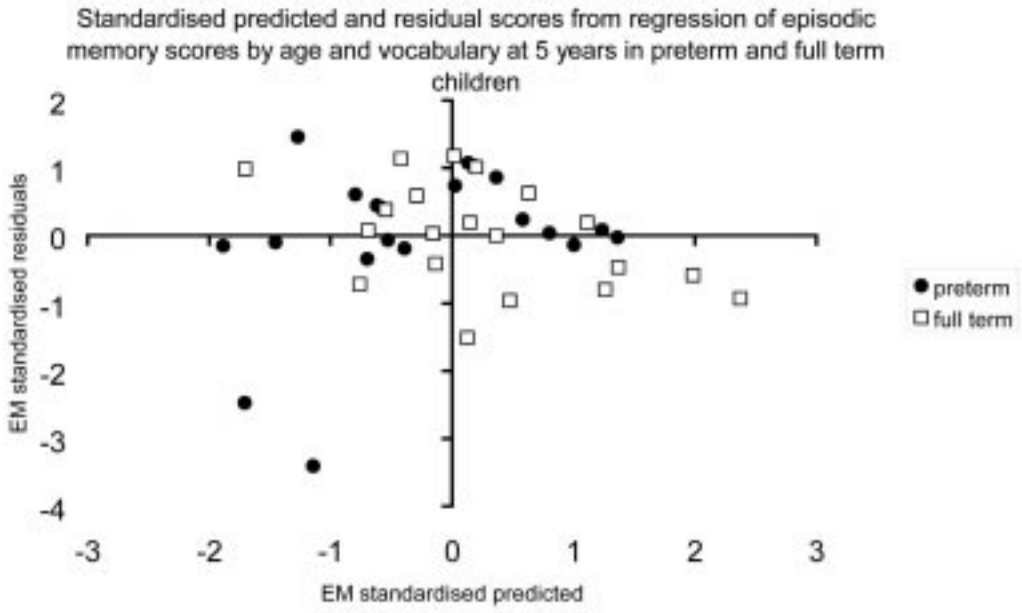


Figure 2 (Briscoe).

T&K-S are critical of assumptions of "Residual Normality," that is, that "in the face of a selective developmental deficit . . . the rest of the system can develop normally" (sect. 3.1) and that Residual Normality is an assumption regarding "how development takes place." A closer look at children with SLI shows that they are not selected directly from an underlying genotype. Instead, selectivity of language impairment occurs despite appropriate marking of developmental milestones, that is, "in the context of normal development" (Bishop 1997b, p. 21). Residual Normality is, under these circumstances, a rational application of a species-specific developmental pattern, not a misguided assumption from cognitive models. The problem of tying a cognitive framework to developmental disorders is really the problem of providing explanatory power at multiple levels of analysis, not the use of cognitive models per se. Ultimately, the truth behind the process of development lies with exceptional children whose abilities defy explanation by simplistic models and challenge cognitive psychology to provide answers.

Raising the bar for connectionist modeling of cognitive developmental disorders

Morten H. Christiansen,[a] Christopher M. Conway,[a] and Michelle R. Ellefson[b]

[a]*Department of Psychology, Cornell University, Ithaca, NY 14853;* [b]*Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom.* **mhc27@cornell.edu cmc82@cornell.edu M.Ellefson@warwick.ac.uk http://cnl.psych.cornell.edu/people/chris.html http://www.psych.cornell.edu/people/Faculty/mhc27.html http://www.warwick.ac.uk/fac/sci/Psychology/staff/research.html#ME**

Abstract: Cognitive developmental disorders cannot be properly understood without due attention to the developmental process, and we commend the authors' simulations in this regard. We note the contribution of these simulations to the nascent field of connectionist modeling of developmental disorders and outline a set of criteria for assessing individual models in the hope of furthering future modeling efforts.

Thomas & Karmiloff-Smith (T&K-S) make an important theoretical contribution to our understanding of cognitive developmental disorders. We find their arguments regarding the problems associated with the assumption of Residual Normality very compelling. In particular, we agree that the developmental process must be taken into account when considering the possible causes of cognitive developmental disorders.

Connectionist modeling plays a crucial role in the arguments put forward by T&K-S. As highlighted in the target article, connectionists have recently begun to model various developmental disorders, including dyslexia (Brown 1997; Brown & Loosemore 1995; Harm & Seidenberg 1999), autism (Cohen 1998; O'Laughlin & Thagard 2000), selective language impairment (Hoeffner & McClelland 1993), Williams syndrome (WS) (Thomas & Karmiloff-Smith, in press), mental retardation (Bray et al. 1997), and schizophrenia (Hoffman & McGlashan 1997). However, in a recent review of connectionist modeling of cognitive developmental disorders, we found that most models suffer from a variety of shortcomings (Conway et al., in preparation). In this regard, we believe that the simulation approach taken by T&K-S provides a positive step forward. Using their simulations as a starting point, we discuss how connectionist modeling of cognitive developmental disorders may be improved further.

When modeling cognitive developmental disorders, one of the critical questions concerns how a particular deficit should be implemented. That is, how should the disordered startstate differ from the "normal" case? T&K-S explored three different ways of manipulating the startstate of their networks: removal of weights (lesioning), addition of noise, and changes in unit discriminability. They found that different types of startstate damage can result in very similar patterns of endstate performance. This is an important result from the viewpoint of connectionist modeling of cognitive developmental disorders. Many connectionists typically report only a single type of startstate damage (though they may have investigated others). But if a particular startstate manipulation is to have theoretical significance in terms of explaining a developmental deficit, it is crucial to establish that *this* manipulation, and no other, is the right causal factor. We therefore think that it is imperative to follow the T&K-S example and explore several types of startstate damage.

Of course, there is more to the modeling of developmental disorders than manipulating the startstate. It is also important that the model captures relevant aspects of behavior given reasonable input and a plausible learning task. Although existence-proof

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

models are crucial to establish the feasibility of a particular psychological modeling approach, the long-term success of any such approach requires close and substantial coverage of empirical data (Christiansen & Chater 2001). To further the modeling of cognitive developmental disorders, we have proposed four criteria to assess the strengths and weaknesses of individual models: deficit implementation, data contact, task veridicality, and input representativeness (Conway et al., in preparation). We see the T&K-S simulations as being on a par with existence-proof models and will instead discuss our criteria in the context of the related model of past-tense formation in WS patients (Thomas & Karmiloff-Smith, in press) described in T&K-S.[1]

The first criterion, *deficit implementation,* refers to how well the manipulations used to simulate a particular deficit are motivated by theoretical and/or empirical research. Thomas and Karmiloff-Smith (in press) implemented several different types of deficits corresponding to different theoretical views of the underlying cause. They found that both manipulations to phonology and the integration of phonology and lexical-semantics allowed for the simulation of the appropriate pattern of WS past-tense formation, suggesting a good, theoretically informed deficit implementation.

Our second criterion, *data contact,* assesses how well a model provides a fit with the relevant psychological data. We further distinguish between primary data contact, which refers to contact with data from specific experiments, and secondary contact, referring to contact with general trends of a population. An example of the former is a model that directly simulates dependent measures from psychological experiments, such as reaction time; an example of the latter is a model that recreates the general trend that children with mental retardation perform worse on a memory task than normal children. With startstate damage to either phonology or the integration of phonology and lexical-semantics, Thomas and Karmiloff-Smith (in press) found close contact with primary human data.

The third criterion, *task veridicality,* is aimed at the extent to which the learning task given to the model is realistic relative to what human participants may face. We further differentiate training-task veridicality – which is the degree to which the network training task maps onto what the target population faces – from test-task veridicality – the degree to which the network test task maps onto what participants do in the actual human experiment(s). In the case of the Thomas and Karmiloff-Smith (in press) study, test-task veridicality is good because the networks are faced with a close approximation of the human test task. Although the stem-to-past-tense training task is typical in connectionist modeling of past-tense morphology, it is arguably not the primary way in which children acquire their inflectional skills,[2] leading to a relatively poor training-task veridicality.

Our final criterion, *input representativeness,* refers to the degree to which the information provided to a model matches the input available to a child. Although we realize that it is often necessary for practical reasons to simplify the input, models should be made as realistic as possible. The phonological input to the Thomas and Karmiloff-Smith (in press) models is based on an artificial language of 500 triphonemic verb stems taken from Plunkett and Marchman (1991). The lexical-semantic input came in different types of nonreferential formats (both localist and distributed). Overall, input representativeness is low, but both the phonological and lexical-semantic representations are not far from the state of the art in current connectionist past-tense modeling.

In conclusion, we see the T&K-S simulations as providing a valuable contribution to the field of connectionist modeling of developmental cognitive disorders. On a theoretical level, their simulations illuminate the problematic aspects of the Residual Normality assumption. On a connectionist modeling level, the simulations have pointed to the importance of exploring several deficit implementations. Although this emerging field of connectionist modeling is very much in its infancy, the simulations presented by T&K-S here and in Thomas and Karmiloff-Smith (in press) highlight the promise that it holds. We hope that the criteria we have outlined here may additionally help raise the bar for connectionist modeling of developmental cognitive disorders.

NOTES

1. Briefly, in this model the phonological pattern of a verb stem was mapped onto a phonological form of its past tense through a set of hidden units – in some versions of the model, lexical-semantic information was additionally provided as input.

2. The version of the model in which both phonological and lexical-semantic information is mapped onto a past-tense form may be less prone to this criticism.

The residual normality assumption and models of cognition in schizophrenia

Ruth Condray and Stuart R. Steinhauer

Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213.

condrayr@msx.upmc.edu **sthauer@pitt.edu**

Abstract: Thomas & Karmiloff-Smith's (T&K-S's) argument that the Residual Normality assumption is not valid for developmental disorders has implications for models of cognition in schizophrenia, a disorder that may involve a neurodevelopmental pathogenesis. A limiting factor for such theories is the lack of understanding about the nature of the cognitive system (modular components versus global processes). Moreover, it is unclear how the proposal that modularization emerges from developmental processes would change that fundamental question.

In their target article, Thomas & Karmiloff-Smith (T&K-S) make several important arguments. The primary contribution is their critical evaluation of the assumption that atypical development can produce selective cognitive deficits while the rest of the system develops normally (the Residual Normality assumption). Additionally valuable is their illustration that behavioral outcomes may be influenced jointly by the types of task and central nervous system (CNS) damage, which in turn may be further qualified by the developmental period at the time of CNS damage (before vs. after training).

We will focus our comments on the relevance of T&K-S's challenge for current theories of cognition in schizophrenia. There are two important points of contact between these two areas of research. The first is the modular versus global distinction made regarding cognition in both sets of discussions; the second concerns the assumption of a developmental etiology. Schizophrenia is a severe psychiatric disorder that affects approximately one percent of the general population, includes cognitive dysfunction as an important clinical characteristic, and may involve a neurodevelopmental pathogenesis. Adult schizophrenia patients exhibit performance decrements on a wide range of cognitive tasks. However, their reduced performance differs significantly from that of nonclinical controls on only some, not all, tasks.

A current trend is to view this overall pattern in a polarized fashion, which commonly takes the following form: Do patients exhibit reduced performance on *X, Y, Z* tasks because of compromises to specific cognitive components, or is their suboptimal performance merely due to an overall reduction in cognitive ability that is manifested differently as a function of task difficulty? Most writers agree that an accurate partitioning of the relative contributions of global and specialized dysfunction will advance our understanding about cognition in schizophrenia. Writers differ, however, about what constitutes acceptable theoretical and methodological approaches, with much of the disagreement concerning how task difficulty is viewed; as a process-oriented variable or as a nuisance variable requiring psychometric solutions (see Chapman & Chapman 2001; Knight & Silverstein 2001; Strauss 2001). We suggest that T&K-S's proposed framework, in which task is an independent factor that may interact with etiology and developmental phase, represents a more interesting formulation, from a cognitive

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

perspective, and may supply a more informative approach for examining the nature of cognitive dysfunction in schizophrenia.

T&K-S correctly emphasize that a major difficulty for the Residual Normality assumption is the current lack of understanding about which aspects of the cognitive system are global and which are modular (for a recent discussion of this general issue, see Fodor 2000). Data from our studies (Condray et al., in press) illustrate the importance of such uncertainty for models of cognition in endstate schizophrenia. Accuracy on a grammatical parsing task (who did what to whom?) was reduced for schizophrenia patients, compared to nonclinical controls. More important, grammatical parsing accuracy was correlated with measures of general intelligence and semantic knowledge for healthy controls, but these factors were *not* correlated for schizophrenia patients. Why the coupling of grammatical parsing accuracy with general intelligence and semantic knowledge in the healthy nonpatient participants? Why the disconnection in the patient group? It is possible that the disconnection in patients' responding merely underscores the relative independence (modularity) of at least some portion of the grammar parsing system (e.g., Chomskian movement). And although it could be argued that our nonpatient control data (association of intelligence, semantics, and grammar) are problematic for such an independence account of parsing, it is also possible that controls' performance merely reflects the smooth interconnectivity of systems that work optimally. Minimally, grammatical parsing appeared to be connected with controls' global cognitive processes; patients' grammatical parsing appeared more local (compartmentalized). Hence, patients were able to perform the task, albeit less accurately, but they may have accomplished that performance in a different way. The source of this grammatical parsing problem is unknown; patients' compromised function could be due to problems with that grammar (sub)system, other cognitive systems, the connective linkages (interfaces) between cognitive systems, or the combined effects of some or all of the above. Finally, fundamental to the position of T&K-S is the possibility that damage to cognitive systems during childhood may produce different developmental trajectories, which could affect how such systems are organized and function during adulthood. It is unclear, however, whether T&K-S's theory of emergent modularization (specialized components that emerge through developmental processes) would necessarily change the central question concerning which cognitive functions are modular and which are global. Hence, while it is possible that atypical development could transform a module into something else (or vice versa), such a possibility would not alter interest in compartmentalized versus global functions and their interconnectivity. If our data are representative for schizophrenia, then an important question concerns how modular and global functions may be interconnected (e.g., some type of binding mechanism) in this patient population.

The pathogenesis of adult schizophrenia may originate *in utero* or during early childhood and may affect brain development (Gooding & Iacono 1995; Marenco & Weinberger 2000). Some version of the scenario described by T&K-S may therefore apply to this disorder. The precise etiology is yet to be determined, however, with strong candidates including abnormalities of selected neural circuitry and neurochemistry function (Cohen & Servan-Schreiber 1992; Weinberger 1996); cytoarchitecture (Barbeau et al. 1995; Weinberger 1996); and cell membrane structure and function (Horrobin 1998). Hence, models of cognitive dysfunction in schizophrenia will require reconciliation with the mechanism(s) of pathogenesis eventually revealed. An etiology that involves a general mechanism (e.g., abnormality of cell membrane structure) might influence the cognitive system in a pervasive manner so that disruption occurs at multiple levels (e.g., specific, global, and binding functions). In contrast, an etiology involving more restricted circuitry (e.g., prefrontal cortex) might disrupt more specialized functions.

We believe the idea that types of cognitive task and CNS compromise should be considered within the context of developmental phase is important. We wish to emphasize, however, that questions about modular function and interconnectivity remain of interest despite the hypothesized alterations in developmental trajectory.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Mental Health (MH 50631) and with resources and the use of facilities at the VA Pittsburgh Healthcare System, Highland Drive Division.

Two closely related simulations provide weak limits on Residual Normality

Richard P. Cooper

School of Psychology, Birkbeck College, London, WC1E 7HX, United Kingdom. **R.Cooper@bbk.ac.uk**
http://www.psyc.bbk.ac.uk/people/academic/cooper_r/

Abstract: Thomas & Karmiloff-Smith (T&K-S) correctly identify Residual Normality (RN) as a critical assumption of some theorising about mental structure within developmental psychology. However, their simulations provide only weak support for the conditions under which RN may occur because they explore closely related architectures that share a learning algorithm. It is suggested that more work is required to establish the limits of RN.

Within cognitive neuropsychology, the assumption of noncompensation is acknowledged to be critical in making inferences about the structure of mental processes from their breakdown. Loosely, this assumption states that following brain injury unaffected processes function normally – their functioning does not change to compensate for the malfunctioning of damaged processes. Thomas & Karmiloff-Smith (T&K-S) are correct in identifying Residual Normality (RN) as playing an analogous role within attempts to deduce the structure of mental processes from developmental disorders.

T&K-S provide two connectionist simulation studies that they suggest demonstrate why RN cannot be assumed. The first compares the effects of startstate and endstate damage within a three-layer feed-forward network using a standard connectionist learning algorithm (back-propagation of error). When lesioned through removal of units early in its training history, the network is able to compensate and, notwithstanding subtle but important differences in responses to different classes of test item, still achieves good overall performance on the test data. Further simulations demonstrate the limits of this compensation. The second set of simulations using a dual-route architecture provides a situation in which RN does not hold. The lack of RN in these simulations derives from the compensatory nature of the architecture: The learning process allows one route to partially compensate for the failure of the other route (provided failure occurs early in learning).

Although the degree of compensation shown in the simulation studies is of theoretical interest, it provides only weak support for the position adopted by T&K-S. The scope of the simulations is highly restricted: Both sets of simulations employ learning by back-propagation of error within feed-forward networks. This combination of architecture and algorithm is known to be capable of compensating at least in part for differences in startstate. That such compensation can undermine RN should not be surprising. More convincing support for the implausibility of RN would come from consideration of alternative architectures and alternative learning algorithms, but whereas T&K-S do discuss alternative learning systems (sect. 8.4), that discussion is limited to little more than an acknowledgment of their existence.

The importance of considering alternative learning algorithms is critical, given the apparent role of back-propagation of error in partially compensating for differences in startstate (simulation 1) and for differences in computational properties of different routes

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

(simulation 2). It is also critical given that back-propagation is generally acknowledged to be biologically and psychologically implausible (e.g., Crick 1989). Although the implausibility of back-propagation may be irrelevant when the object of study is the endstate of a system, it cannot be so easily dismissed where simulations aim to explore developmental phenomena and particularly where they explore developmental dysfunction. In such simulations, detailed properties of the learning algorithm are likely to be critical. In defence of the use of back-propagation, T&K-S may rightfully argue that whatever the biological learning mechanism may be, it is likely to exhibit compensation because such compensation is likely to be evolutionarily advantageous (if not evolutionarily necessary). However, biological learning mechanisms may show more or less compensation than back-propagation, and either outcome could alter the conclusions one might draw. Significantly more compensation would undermine the presupposition that developmental deficits could be explained by differences in startstate, whereas significantly less compensation would argue in favour of RN.

The importance of considering alternative architectures parallels the importance of considering alternative learning algorithms. Are similar levels of compensation available to other forms of connectionist network (e.g., networks with recurrent connections) or to production system architectures? Even within the domain of past-tense acquisition there are numerous competing models. Does startstate and endstate damage affect the behaviour of alternative models in similar ways? Notwithstanding suggestions made by T&K-S on ways to achieve RN, many more simulations would appear to be necessary to provide a definitive answer to these questions.

The above issues raise questions about the degree to which the simulation studies may be taken as support for the implausibility of RN. Further questions are raised by two additional assumptions apparent in the simulation work. First, learning (via back-propagation) is equated with development. This view is not universally held within the developmental literature. Indeed, T&K-S make passing reference to constructivist connectionist approaches in which development involves more than simply adjusting network weights. Second, developmental disorders are assumed to arise from damage to the startstate of the system rather than, for example, damage to its learning mechanism. Both of these assumptions limit the inferences that may be drawn from the simulation studies. In short, does RN remain implausible if development is more than just weight tuning or if developmental disorders result from damage to the learning mechanism rather than damage to the system's startstate?

T&K-S attempt to address issues relating to the limits of their specific simulations in section 9, where they speculate on some ways in which RN may be achieved. Each of these effectively constrains the degree of compensation available via learning to the damaged system, effectively by building in the seeds of modularity before the developmental disorder has any effect. T&K-S emphasise that some of these ways of achieving RN raise more questions than they answer (particularly guided specialisation), but others arise from apparently plausible mechanisms that would seem to deserve further investigation. Again, further simulations are necessary to demonstrate the viability of, for example, strong competition or early commitment.

The criticisms made here are really criticisms of the methodology of computational modelling. Although the modelling methodology adopted by T&K-S has some rare strengths – it is, for example, hypothesis-driven experimentation rather than data-driven parameter fitting – it also suffers from a common modelling difficulty - the reported simulation studies are likely to have taken months of computing time, yet they represent only one small region of a large, multidimensional parameter space of models, and generalisation from that region to the entire parameter space is not straightforward. In short, T&K-S are correct to highlight the issue of RN, but much work remains to identify fully the conditions under which RN may or may not hold.

Development is also experienced by a personal self who is shaped by culture

Michel Ferrari

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, M5S 1V6, Canada. **mferrari@oise.utoronto.ca**

Abstract: I agree with Thomas & Karmiloff-Smith (T&K-S) in their critique of Residual Normality. However, first-person data must be integrated into their account of neurobiological development of disabilities. Furthermore, psychological development itself is not only about an individual's brain and how it interacts with the world; rather, development depends crucially on the sociocultural context in which (normal and abnormal) children develop.

Michael Thomas and Annette Karmiloff-Smith (henceforth T&K-S) provide an important critique of a widespread idea in the study of developmental disorders – Residual Normality. Their evidence shows fine-grained differences in observed deficits when a simulated cognitive system is damaged, depending on whether that damage occurs before training (startstate; – analogous to infancy) or after training (endstate; – analogous to adult brain damage). In cases of startstate deficits or abnormalities, the implicit assumption of "residual normality" is undermined by compensation or alteration by the rest of the system in an attempt to cope with this damage. They conclude that "ontogenetic development cannot be ignored in constructing models of developmental disorders" (target article, Abstract).

This apparently simple claim strikes at the heart of how best to study psychology as a natural science, a question that has haunted psychology from its beginnings. As George Trumbull Ladd wrote in his review of William James's *Principles of Psychology,* "QUESTIONS concerning the nature, problem, and legitimate methods of Psychology, and concerning the relations it sustains to other forms of science and to metaphysics, are apparently far from being settled to the entire satisfaction of all parties to the controversy" (Ladd 1892). And the controversy continues, as we see by the present article. The very idea of a normal child and normal child development is a relatively recent view for science (Hacking 2002). Indeed, the scientific study of development really only took off in the nineteenth century, as shown by Crombie (1994).

Because how to study developmental disorders is still being debated, let me state my view up front. I agree entirely with the authors that we need to take development seriously in order to truly understand neurobiological disorders. In fact, I will go further and say that two additional dimensions are left out of T&K-S's account.

The first addition needed is a principled consideration of how first- and third-person data can be integrated into their account of neurobiological development of disabilities. Some may agree to this point, in principle, but feel that it cannot be studied experimentally within a neuroscientific framework. However, I would like to hear the authors' views on the research program launched by the late Francesco Varela and continued by Antoine Lutz and Evan Thompson, among others, which they called "neurophenomenology." According to Lutz: "The neurophenomenological program encourages researchers to pay attention not only to neuronal or physiological data but also to the data produced by accounts of subjective experience" (Lutz 2002, p. 134). For Lutz and those working in this research tradition, connectionist neurobiological models (even developmental models) of how the mind works can provide information about how the brain/mind works in and of itself, but not about how it can work *for* itself. They believe we also need to consider first-person experiences (e.g., of perception, pain, and possibilities for action) as they are experienced by a developing subject. For Lutz and many others, there is an explanatory gap between first- and third-person accounts that Varela was one of the few directly to address within a neuroscientific framework; he proposed a pragmatic integration of lived experience and neurobiological measurements of brain activity.

The upshot of this program is that while it is important to con-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

sider "local phase synchrony" in the neural activity of particular regions of the brain, there is also evidence of "large-scale synchronization" among neural assemblies in different regions of the brain (the analogy used is of several computers linked through a program like Napster over the World Wide Web). This program, I think, adds ammunition to the claims made by T&K-S against a simple modular view of brain development, but also challenges them to find ways to integrate first-person experiences and aims into their account of atypical development.

There may also be critical "turning points" that orient how biological development will unfold in cases of genetic abnormalities, just as there are turning points in personal autobiography (Bruner 2001). This is especially plausible when one considers that autobiography in general is initially coconstructed with parents (Nelson 1993) and so is not divorced from the cultural context in which such constructions take place. This coconstruction is just as true in cases of abnormal development, as we have shown in a detailed biographical study of an autistic child (Vuletic & Ferrari, in press). Our study shows that how a particular autistic child develops is deeply influenced by how his parents conceive of his life and his possible development and not by neurobiological considerations alone. Of course, parents must adapt to the reality of how their child is acting and experiencing the world; but that adaptation can range from making every attempt to assure a self-sufficient and happy life for their child, as these parents and grandparents do, to despair at making any positive impact on their developing child, as many doctors schooled in the modular view of fixed potential sadly recommend.

This leads to the second point of this commentary. Psychological development itself cannot be understood as a uniquely individual thing involving only an individual's brain and how that brain interacts with the world. Development depends crucially on the sociocultural context in which (normal and abnormal) children develop. In cases of infant attachment, for example, an infant's inner working model of his or her relationship to an adult caregiver is constructed from the history of his or her interactions with that caregiver, as both are integrated into a sociohistorical/cultural context (Bretherton 1996). According to Hacking (1998) and Morris (1998), cultural context is one of the factors that helps give a specific form to an illness or disability as it interacts with biological abnormality through development. In other words, while it is true that neurobiology is an important aspect of human development, it is equally true that how particular children are treated, and the general sociocultural conditions under which they develop (which extend to basic nutrition and medical care as well as to parental expectations about children's physical and social obligations as part of their community), will all have a huge impact on how children with genetic abnormalities develop (Keating & Miller 1999). Hacking (1998) shows how an entire type of mental illness disappeared when social conditions changed – even if trace elements of that illness remain and have been reclassified.

Both of these points are meant only to add to what is a very strong and important critique of the widespread notion of Residual Normality. I think that it is time to broaden our understanding of psychology beyond the limited neurobiological framework that many psychologists have already found too constricting for a natural science of psychology over a century ago.

Modularity in developmental disorders: Evidence from Specific Language Impairment and peripheral dyslexias

Naama Friedmann[a] and Aviah Gvion[b]

[a]*School of Education, Tel Aviv University, Tel Aviv, 69978, Israel;* [b]*School of Education, Tel Aviv University, and Reuth Medical Center, Tel Aviv, 69978 Israel.* **naamafr@post.tau.ac.il** **aviahg@post.tau.ac.il**
http://www.tau.ac.il/~naamafr

Abstract: Evidence from various subtypes of Specific Language Impairment and developmental peripheral dyslexias is presented to support the idea that even developmental disorders can be modular. However, in developmental letter position dyslexia and neglect dyslexia we show that additional errors can occur because of insufficient orthographic-lexical knowledge.

One of the intriguing questions regarding disorders of language and reading is whether these disorders are modular. Thomas & Karmiloff-Smith (T&K-S) suggest that although acquired disorders might be modular, developmental disorders are not. Of course, there might be multifaceted cases of developmental aphasias and dyslexias, and many types of impairment can coexist in the same individual. Cases like these also exist in acquired aphasia and dyslexia. However, we believe that the main question is: Can modular impairments be found in cases prior to training? We believe there can. The structure of our argument will be as follows: first, we will demonstrate some evidence for modular impairments in developmental aphasia and dyslexia, focusing on their modular nature and the similarities between acquired and developmental forms of the same disorder, and then we will show, taking the reading domain as an example, that such a modular and selective impairment might cause a deficit in other reading abilities, stemming from the lack of exposure to written text and possibly from lack of well-formed input to the orthographic lexicon during reading acquisition.

Starting with developmental aphasia or Specific Language Impairment (SLI), it is possible to distinguish various types of impairments, differing, just as in acquired aphasia, in the linguistic module which was impaired. Friedmann and Novogrodsky (2003) reported 10 children with syntactic SLI (G-SLI) whose syntactic deficit manifests in chance-level performance on the comprehension of object relative sentences and referential object questions, and the inability to produce relative clauses and sentences that include verb movement; these children are unimpaired in naming and their phonological abilities are normal for their age, as tested by pseudoword repetition, nonword judgment and phonological awareness tasks.

A group of children with a severe naming disorder (stemming either from lexical-semantic or lexical-phonological deficit) was also identified, who show intact syntactic abilities, as manifest by their normal comprehension of relative clauses, normal production of relative clauses (except for paraphasias, which preserve the syntactic structure of the sentence), and normal production of sentences that include verb movement (Novogrodsky & Friedmann 2002). This dissociation indicates that it is possible to see linguistic modules such as syntactic movement, phonology, and lexical access selectively impaired even in developmental aphasias.

The same selective and modular nature is also manifest in dyslexia. Acquired dyslexias can appear in very selective forms, stemming from a circumscribed deficit in one of the processes or lexicons that participate in the reading process or in the connection between them (Ellis 1993; Ellis & Young 1996; Marshall & Newcombe 1973; Patterson 1981). This is the way modularity reflects in reading disorders. What about developmental dyslexias? Indeed, children and adults with dyslexia of early onset – before reading acquisition – can show a myriad of reading deficits and sometimes additional disorders such as attentional and phonological deficits. But does this mean that there is no modularity in developmental reading deficits? We believe it does not. A group of

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

studies has now established cases that show a very selective reading deficit parallel in nature to acquired dyslexias (Stuart & Howard 1995; Temple 1984; 1997; Temple & Marshall 1983).

Peripheral dyslexias, which have been less widely explored with respect to developmental origin, also show modular impairments in developmental forms. Letter position dyslexia (LPD), a selective deficit in letter position encoding with unimpaired letter identification, was initially reported in acquired cases (Friedmann & Gvion 2001). This dyslexia has now been identified in 12 Hebrew-speaking children and adolescents whose reading patterns show striking similarities to acquired LPD (Friedmann & Gvion 2002; Friedmann & Rahamim 2002; Precel & Friedmann 2002).

Both acquired and developmental LPD manifest modular impairment, as only one submodule of the visual analysis system, letter position encoding, is impaired, whereas another submodule, letter identification, is unimpaired. Both in acquired and in developmental LPD, individuals show the dissociation between making predominantly letter migration errors within words (reading "bread" for "beard") and very few or no letter substitutions in a wide variety of tasks. In reading aloud, individuals with acquired LPD had 21% letter-order errors and only 1% letter substitutions; individuals with developmental LPD had 15% letter-order errors and 1% letter substitutions. In a same-different task, individuals in both groups could detect differences in letter identity between words but failed to detect letter-order differences: The individuals with acquired LPD made 48% errors in letter order but only 7% errors in letter identification; the 12 individuals with developmental LPD had 37% versus 4% errors.

Moreover, exactly the same pattern of migration errors occurs in acquired and developmental LPD. In both cases the migration errors occur almost exclusively in medial-letter positions. In word reading and definition tasks, individuals with acquired LPD made 16% medial-letter position errors, compared to 0.01% exterior-letter position errors. Similarly, the individuals with developmental LPD had 15.1% medial errors and 0.8% exterior errors.

However, many of the individuals with developmental dyslexias show errors in addition to the errors that characterize their reading deficits. We argue that these stem from their incomplete orthographic-lexical knowledge because of incomplete and flawed input to the lexicon and lack of sufficient exposure to written words and texts that is only a side effect of their reading impairment. For example, the individuals with developmental LPD also showed, in addition to letter migrations, 5.7% errors in reading aloud that resulted from insufficient lexical-orthographic knowledge: They made errors on heterographic homophones (or heterophones with homophonic letters), which can be read correctly only after being lexicalized, as well as regularization errors and errors of the vocalic pattern of unvoweled words (in Hebrew, vowels are underrepresented, and so lexical knowledge is required for reading). The same was true for the Hebrew-speaking neglect-dyslexic child in Friedmann and Nachman-Katz (in press), who made neglect errors in 50% of the words. He too had homophone and vocalic-pattern errors that are not characteristic of the reading of individuals with acquired neglect dyslexia. These error types were not found in the reading of the individuals with acquired LPD in Friedmann and Gvion (2001) or in acquired neglect dyslexia, probably because the onset of the dyslexia in the acquired cases followed reading acquisition and the loading of the orthographic input lexicon, whereas the developmental or early-onset cases did not have the chance to fill up their orthographic lexicon.

To conclude, selective impairments that indicate the modular nature of developmental disorders are reported both in SLI and in peripheral dyslexias. Additional errors can result from the lack of lexical knowledge rather than lack of Residual Normality.

ACKNOWLEDGMENTS

We thank Rama Novogrodsky, Ivana Nachman-Katz, Einav Rahamim, and Karen Precel for interesting discussions and data regarding the modularity of developmental disorders. The study was supported by the Joint German-Israeli Research Program grant GR01791 (Friedmann).

Nativism, neuroconstructivism, and developmental disorder

Philip Gerrans

Department of Philosophy, University of Adelaide, Adelaide, SA 5005, Australia. **philip.gerrans@adelaide.edu.au**
http://www.arts.adelaide.edu.au/Philosophy/pgerrans/

Abstract: Either genetically specified modular cognitive architecture for syntactic processing does not exist (neuroconstructivism), or there is a module but its development is so abnormal in Williams syndrome (WS) that no conclusion can be drawn about its normal architecture (moderate nativism). Radical nativism, which holds that WS is a case of intact syntax, is untenable. Specific Language Impairment and WS create a dilemma that radical nativism cannot accommodate.

Nativists about syntax argue that syntax (1) is innate; (2) depends on rule-based processing; (3) is modularised; *and* (4) Williams syndrome (WS) and Specific Language Impairment (SLI) are a double dissociation (DD) which identifies a syntax module damaged in SLI and "spared" in WS. Thus, Pinker explains the presence of inflectional morphology in WS as follows:

> Their grammar is running smoothly but their word-fetcher doesn't have the usual bias to fetch frequent and appropriate words quickly. Irregular verbs survive on that basis, so occasionally an irregular form doesn't survive quickly enough and the rule is ready and waiting to step in. (Pinker 1999, p. 262)

The word-fetcher is a memory system whose functioning is statistical, not rule-based (hence the salience of irregulars), whereas the grammar module, presumed intact in WS, is a rule-based system. Nativists model the interaction of these two systems in a dual-route neural network which stores and retrieves stems and regular and irregular suffixes. If the word-fetcher cannot find an irregular suffix, the default regularisation mechanism automatically produces the regular construction.

Karmiloff-Smith and neuroconstructivist collaborators have undermined point 4 by producing a model that captures the linguistic performance of SLI and WS subjects in a single-route network whose functioning is statistical. The network approximates the linguistic performance of SLI if the noise/signal ratio in phonological input is increased and WS if the pattern is reversed and the signal/noise ratio is increased.

These findings have different consequences for different forms of nativism. They strongly undermine a reading of point 4 which depends on Residual Normality (RN). Heather van der Lely, for example, has argued that in SLI subjects have "*normal cognitive and auditory abilities* alongside impaired grammatical abilities" (van der Lely 1999, p. 286).

Thomas & Karmiloff-Smith (T&K-S) point out, however, that performing *within the normal range on standard tests* is not sufficient to justify the RN hypothesis of spared or intact function. In the case of phonological processing, fine-grained testing can disclose subtle deficits with major developmental consequences. For example, in the extraction of a signal from a variable acoustic stream, temporal interval as well as amplitude can affect performance. Hence, a standard hearing test may not detect subtle deficits in the ability to overcome the masking effects of the surrounding stream. One explanation of dyslexia is that it is primarily a result of such basic processing deficits rather than a problem with higher-level linguistic processing (Nagarajan et al. 1999; Wright et al. 1997a; 1997b). Normal can thus mean "test within the normal range," or (the RN hypothesis) "processed by the same mechanism(s) in the same way as normal subjects." It is this sense of normality as RN which is required by van der Lely.

A nativist cannot accept T&K-S's data and abandon RN without abandoning the hypothesis that WS and SLI are a DD which identifies a syntax module. A nativist might initially argue that SLI is a consequence of impairment to the syntax module, but then they face a dilemma regarding WS. Is the module intact in WS? If it is, why do language and syntax have an unusual developmental tra-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

jectory in WS? (Donnai & Karmiloff-Smith 2000; Laing et al. 2002). Possibly, increased sensitivity to phonology overrides the default mechanism for regularisation. So a WS subject's syntax module develops abnormally because the syntax module is crowded out. However, if this is the case, RN must be abandoned. Points 1 to 3 can be maintained, but point 4 cannot.

The nativist who abandons point 4 might then argue that abnormal syntax in WS is the result of absence of the syntax module. But in that case, WS should resemble SLI, when in fact they present initially as a DD. So the best nativist hypothesis is to abandon point 4 and retain the idea that what we see in WS is the anomalous development rather than absence of a syntax module. But distinguishing between these two hypotheses on the grounds of performance alone seems very difficult.

On either story – and the former is the most plausible – it seems that a nativist must eventually agree that apparent linguistic fluency in WS is essentially due to nonsyntactic factors, and hence that whatever the fate of nativism about syntax, the hypothesis of RN for SLI or WS is unsustainable.

The radical neuroconstructivist conclusion is, of course, that nativist modular hypotheses about high-level cognitive processing such as syntax or theory of mind are mistaken. The appearance of such modularity is an artefact of interactive development of general cognition and low-level processing modules (in this case, phonology; in the theory-of-mind case, perhaps functions such as recognition of emotional and facial expression and intentional movement) (Gerrans 2002; 2003; Gerrans & McGeer, 2003).

The nice thing about T&K-S's article is that it shows that the resolution of the issue requires a theory that models the neural implementation of computational properties considered essential to processing the domain in question. A simple inference from performance to cognitive architecture of high-level, abstract cognition is not licensed by the data.

Distinguishing proximal from distal causes is useful and compatible with accounts of compensatory processing in developmental disorders of cognition

Nancy Ewald Jackson[a] and Max Coltheart[b]

[a]*Department of Educational Psychology, University of Iowa, Iowa City, IA 52242;* [b]*Macquarie Center for Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia.* **nancy-jackson@uiowa.edu max@maccs.mq.edu.au http://www.rosella.bhs.mq.edu.au/~max/**

Abstract: Models of the architecture of mature cognitive systems can inform the study of normal and disordered cognitive development, if one distinguishes between proximal and distal causes of performance. The assumption of residual normality need not be made in order to apply adult models to performance early in development, because these models can be modified to reflect the results of compensatory processing.

Thomas & Karmiloff-Smith (T&K-S) agreed with us (Jackson & Coltheart 2001; hereafter J&C) that one can distinguish proximal causes of abnormal performance (explanations in terms of the current architecture of a cognitive system) from distal causes (a category in which we include everything else, including developmental history). J&C argued that this proximal-distal distinction permits models of the architecture of mature cognitive systems to inform the study of normal and disordered cognitive development. We did not assert that adult models should be *assumed* adequate to describe children's behavior or used to restrict the search for differences between developmental and acquired impairments. However, we remain convinced that models of mature performance are useful starting points for describing both typical and atypical performance during the development of skills such as reading. For example, if T&K-S had not been familiar with dual-route models of skilled performance, would they have attempted to falsify a hypothesis of similarity between early- and late-damaged systems?

We have *not* assumed that those components in an abnormally developing system that are not directly implicated in abnormal performance will show Residual Normality (RN). On the contrary, we explicitly disavowed the assumption of RN in developing systems. Indeed, rather than arguing for RN, we proposed that: "[i]n a developing system, single deficits might be more likely to have broad implications than is the case when a previously intact system has been damaged" (J&C, p. 152), and went on to speculate about ways in which impaired development of either a lexical or nonlexical route might cause problems throughout a child's developing reading system:

> If each kind of deficit impedes development, we should not be surprised to find that children who are atypical at exception-word reading often are atypical at pseudoword reading, and vice versa, even though these two tasks depend on different parts of the reading system. If what we have proposed about how learning based on each route's operations contributes to the development of the other route is true, we would expect deficits in both types of reading to be the most common, but not the invariable pattern. Indeed this is what has been found. (J&C, p. 193)

What, then, of T&K-S's claim that inferences from observed behavior to underlying cognitive structure must be conditional on "developmental constraints under which processing structures have emerged" (sect. 10, para. 4)? Given the likelihood of compensatory processing in early-damaged systems, how can models of skilled cognition be useful for understanding developmental disorders?

T&K-S assumed that static models cannot deal with phenomena such as compensatory processing. They argued that if RN does not hold, then behaviors that appear similarly intact in acquired and developmental disorders could reflect operation of an intact processing module in the acquired case but *qualitatively different* compensatory processing in the developmental case. But what constitutes a qualitative difference? Is compensatory processing at a particular point in development really beyond the scope of nondevelopmental models such as the Dual Route Cascaded Model of Word Recognition and Reading Aloud, or DRC (Coltheart et al. 2001)? We counter that it is not and that *quantitative* changes in how the parameters of such a model are set can alter its function in a way that might be called strategic and compensatory.

For example, Rastle and Coltheart (1999; see also J&C) showed that the DRC model can simulate strategic effects that might occur if a reader expected to read only nonwords. This is not a developmental scenario; but such tuning of system parameters is analogous to what might happen if a child's prior experience in using either the lexical or the nonlexical route of a modular system like DRC had been repeatedly unsuccessful. J&C also sketched how the DRC might be modified to simulate the effects of the restricted orthographic lexicon and incomplete knowledge of grapheme-phoneme correspondence rules that characterize normal beginning readers (J&C, Fig. 5.1). The extent to which a modified version of a static model such as the DRC might account for observed reading performance at different points in development is an empirical question that has already been answered in the affirmative in some instances (Coltheart et al. 2001, Fig. 15).

Static models such as DRC say nothing about processes by which reading systems change over time as children develop reading ability. Nonetheless, such models can offer successive snapshots of what the system might look like, for normal and impaired readers, as reading develops. However, given that static models do not seek to explain processes of development, why do we still prefer such models to apparently more comprehensive and undoubtedly more dynamic connectionist models? To answer this question, we return to the distinction between proximal and distal cause.

Distal causes that have been proposed for developmental reading failure are diverse, ranging from genetic anomalies to restric-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

tions in the child's oral language environment to inadequate instruction in print. How various genetic and environmental causes of reading failure fit together is the focus of considerable current research (e.g., Castles et al. 1999). Unlike T&K-S, we do not find it helpful to try to map these distal effects onto manipulations of connectionist models such as lesioning network structure or adding noise to processing. Doing so requires leaps of faith we prefer not to take. What, for instance, is the appropriate analogue, for a connectionist learning algorithm, of a home environment restricted in oral language input? Of the sources of reading failure for a garden-variety poor reader who enters school with deficits in both oral vocabulary knowledge and phonological recoding ability? Of the compensatory mechanisms used by a child whose ability to use grapheme-phoneme correspondences to read nonwords is impaired, perhaps because of lack of direct phonics instruction, but who reads real words adequately (Thompson & Johnston 2000)? Metaphorical links between manipulations in connectionist models and various developmental impairments in reading can be made (e.g., Harm & Seidenberg 1999), but where is their empirical justification? Pending that justification, we prefer to keep a fence between distal questions about causes of developmental change and proximal questions about what the architecture of a child's reading system is like after those causes have exerted their effects.

Encapsulating architecture and encapsulating processes

Patrick Juola

Department of Mathematics and Computer Science, Duquesne University, Pittsburgh, PA 15282. **juola@mathcs.duq.edu**
http://www.mathcs.duq.edu/~juola

Abstract: Thomas & Karmiloff-Smith (T&K-S) raise the excellent and, in retrospect, obvious point that in a dynamic learning environment where feedback is possible, we should expect networks to adapt to damage by altering details of their behavior. We should therefore not expect that developmental disorders should result in "normal" modules. The implications of this point go much further, since interprocess dependency in the brain does not rely only on learned neural connections. This argues strongly against behavioral and process-related definitions, as opposed to structural and architecture-related definitions, of mental modularity.

One mark of a good scientific paper is that it raises new issues. By this standard, Thomas & Karmiloff-Smith (T&K-S) have written an excellent paper. The issue of "Residual Normality" should be crucial to any analysis where neuroplasticity or development is an issue. In a dynamic and complex system, one expects damage, alterations, or perturbations to one facet of the system to be reflected in changed behavior of both the system as a whole and of other facets – this is part of the fundamentals of complexity (Nichols & Prigogine 1989). Within broad limits, neural tissue can adapt itself to changes in its environment, including changes in the surrounding tissue. And although researchers can comfortably expect lifetime employment in teasing out exactly what those limits are, we have no reason a priori to believe that the components of a system fully adapted to an unusual environment will behave "normally."

However, it remains an open question what time-scale is necessary for this sort of adaptation. Specifically, the notion of "Residual Normality" assumes a very strong version of mental modularity, a theoretical position at best controversial (Dunn & Kirsner, 2003; Juola & Plunkett 1998). Mental modules are assumed to be encapsulated to a degree that it is possible to damage a single module without affecting others at all. Not even an automobile displays this level of modularity – bad alignment of the front tires will immediately drop the fuel efficiency and will probably even result in the radio volume increasing. These are the result of a complex feedback system: The poor alignment makes increased tire noise, causing the human driver to turn up the radio to drown it out. The radio does not behave "normally" in the presence of faulty alignment.

This sort of feedback and recovery is explicit in the experimental setup of simulation 2; although a damaged module such as the indirect route may not directly interact with the direct route, the learning mechanism will compensate for errors by adjusting the weights in the undamaged module. However, the performance of the undamaged route will be "normal" only if there is no feedback at all in the system to alter performance. Three factors assure this in the connectionist system under discussion. First, the topology of the network makes it impossible for changes in indirect representation to affect direct representation, so that there is no immediate effect of a lesion to the indirect route on the weights of the direct route. Second, no further learning/change can occur after insult; in the case of the connectionist model, only "endstate" damage will result in normality, whereas in the case of a human, we assume that the neuroplasticity of an adult approximates zero. Third, the inflectional module itself is sufficiently encapsulated that it can be modeled in isolation from the rest of cognition.

Unfortunately, it is not clear and perhaps not even likely that any of these three conditions hold *in vivo.* The human brain appears to be deeply interconnected, to the point that the stress of hospitalization (Bates et al. 1987) or of unrelated cognitive processing (Blackwell & Bates 1995) can produce symptoms similar to the loss of a neural module. This dependence on cognitive resources provides an immediate pathway for cognitive change in unrelated modules immediately after insult and shows the dependence of these hypothesized "modules" on the rest of the brain.

Common sense and the new experimental findings both support the idea that the brain can compensate for its own failure to develop correct processes by altering other processes in the brain. Bates and colleagues (Bates et al. 1987) have shown that the correctness of processes in normal adults is subject to variation. Within this framework, the idea of mental modules as encapsulated "processes" is an increasingly difficult one to support. Our strongest evidence for modules within the system built by T&K-S is architectural: the topology of the network. Arguments from localization of brain function and neural mappings can produce the same evidence for architectural modularity in the brain, evidence that behavioral studies cannot duplicate. The issues raised by T&K-S are yet another reason to distrust behavioral evidence.

Did Residual Normality ever have a chance?

Susan C. Levine, Terry Regier, and Tracy L. Solomon

Psychology Department, University of Chicago, Chicago, IL 60637.
{s-levine; t-regier; tsolomon}@uchicago.edu
http://www.eci.uchicago.edu/faculty/S_Levine.html
http://www.psych.uchicago.edu/~regier

Abstract: Thomas & Karmiloff-Smith (T&K-S) show that the assumption of residual normality (RN) does not hold in connectionist simulations, and argue that RN has been inappropriately applied to childhood disorders. We agree. However, we suggest that the RN hypothesis may never have been fully viable, either empirically or computationally.

Thomas & Karmiloff-Smith (T&K-S) argue that the assumption of residual normality (RN) has been inappropriately applied to childhood disorders. The RN assumption is that "in the face of a selective developmental deficit, the rest of the system can nevertheless develop normally and independently of the deficit." (sect. 3.1). The authors discuss many reasons why this assumption, borrowed from the adult neuropsychological model, should not apply to developmental disorders. Most importantly, they argue that it ignores the role of the developmental process in producing out-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

comes. These views seem unassailable to us. Our only criticism, if any, is that they may be somewhat too unassailable, in the sense that a strict version of RN never really had a fighting chance, either empirically or computationally.

T&K-S review a set of studies on Specific Language Impairment (SLI) and Williams syndrome (WS) that assume RN. However, they acknowledge only briefly an area of research (i.e., pediatric neuropsychology) where RN is not supported. It has long been recognized that early brain lesions result in more subtle deficits than later lesions, although a broader range of functions is affected. For example, early left hemisphere lesions result in more subtle language deficits than later lesions (e.g., Basser 1962; Bates et al. 1999; Levine et al. 1987; Woods & Teuber 1978), but may affect spatial as well as verbal functions, resulting in lower IQs (e.g., Banich et al. 1990; Riva & Cazzaniga 1986). This pattern has been attributed to "crowding" – neural reorganization such that in addition to the functions that intact systems normally subserve, they become involved in functions that would normally be carried out by systems that have been damaged (e.g., Hebb 1949; Strauss & Verity 1983; Strauss et al. 1990). This clearly violates RN. This violation is a part of a larger picture in which the developmental process plays a role in shaping the outcome of early neurological disorders (Goldman-Rakic et al. 1983; Kolb 1995; Thal et al. 1991). For example, the average IQ of children with early focal brain injury is lower than that of nonlesioned children after but not before 7 years of age (Banich et al. 1990).

At the heart of T&K-S's paper are their computational simulations, using connectionist networks of a sort widely used in modeling both developmental and adult neurological damage. These simulations capture well, in broad outline, the effect of early damage on subsequent development. We have some minor concerns regarding the interpretation of these simulations. First, different amounts of damage were inflicted early and late in training, making it difficult to determine whether differences are attributable to time of damage, amount of damage, or some interaction. Second, the comparison of early and late simulated lesions with human performance is clouded by the fact that in the simulations, no training was provided following late lesions (see Marchman 1993). Brain-injured patients, in contrast, do have the opportunity to learn after an acquired injury. Indeed, it is suggested that residual deficits cannot be determined earlier than one year after injury (Spreen et al. 1995). A more comprehensive picture of the effects of early and later lesions might be provided by systematically crossing lesion timing with lesion size and amount of training.

Critically, T&K-S's central finding is that their networks do not exhibit RN during simulations. When one part of their network was damaged, the rest of the network was affected, rather than developing normally. This is a useful and relevant finding, particularly given the prominence of such networks in modeling human damage. Below we outline how the same general finding could have been arrived at analytically – although, of course, not all aspects of T&K-S's simulations could have been predicted in advance.

T&K-S use a connectionist network in which the input layer is fully connected to the output layer (the "direct" route), and is also fully connected to a hidden layer which is in turn fully connected to the output layer (the "indirect" route). The network's nodes have sigmoidal activation functions, and the network as a whole is trained under back-propagation, with a cross-entropy error function. Under these conditions, the weight update Δw_{rs} for a given weight w_{rs} on a connection that projects from sending node s to receiving node r in the output layer is:

$$\Delta w_{rs} = -\eta \frac{\partial E}{\partial w_{rs}} = -\frac{\eta}{\ln(2)}(a_r - d_r)a_s$$

Here E is the error measure, a_r and a_s are the activations of nodes r and s respectively, and d_r is the desired output for node r; η (the learning rate) and ln(2) are multiplicative constants. The critical point for our purposes is that this weight update rule is a linear function of a_r, the activation of the receiving output node.

Damage anywhere in the network (e.g., through lesioning) can be expected to affect the activation of at least one output node, a_r – because output nodes receive input from all input nodes and from all hidden nodes. This effect should be especially clear early in training, because output nodes will not yet be at asymptote. Given the above derivation, we can see that a change in that activation level, a_r, will affect the weight update on any connection that projects to node r – *whether or not that connection is a part of the lesioned section of the network.* For example, a lesion in the network's indirect route would change the activation of some output node – and this change in activation would in turn affect weight updates on all connections to that node, including connections that are a part of the direct route. Hence, in this very general sense at least, complete residual normality is an impossibility in networks of this sort.

Structural abnormality may not equal functional oddity

Yonata Levy

Psychology Department, The Hebrew University, Jerusalem, 91905, Israel.
msyonata@mscc.huji.ac.il

Abstract: This commentary questions the authors' claim that cognitive neuropsychology is defined by modularity and that other theoretical frames, that is, connectionism, are a priori rejected. It also points to the fact that whereas in genetic disorders there are developmental delay and asynchrony, there are few reports of deviant developmental trajectories that are never seen in typical development. It is suggested that the possibility that structure does *not* equal function in the developing brain, may be a viable option.

This commentary offers an objection, a commendation, and a question. The objection concerns the authors' use of the term cognitive neuropsychology. Cognitive neuropsychology is the name of the interdisciplinary area dedicated to the study of brain-cognition relationships. As such, it is theory-neutral and open to competing approaches. Among the most influential theoretical positions in this area of study is Fodor's (1983) modularity thesis, which has since been modified and elaborated by other people too (cf. Moscovic & Umiltà 1990). However, modularity does not define cognitive neuropsychology as a research field, as the authors seem to suggest. Rather, modularity is one possible (and highly debatable) framework for potential answers to the relevant issues. Another such framework is connectionism. By espousing the latter, one rejects modularity yet one does not reject cognitive neuropsychology.

Cognitive neuropsychology is unjustly blamed in the target article for other vices too. Contrary to what the authors say, cognitive neuropsychology has considered group studies along with single cases (cf. Cossu & Marshall 1990; Cossu et al. 1993), has searched for dissociations but also for associations (cf. Tager-Flusberg 2003), and has no objection in principle to dynamic, nonstatic models as befits developmental issues. Hence, I believe the following research programs, which the authors of the target article cite, are well within the domain of cognitive neuropsychology: Bishop (1997a), where a case is made for the nonmodularity of developmental disorders; Karmiloff-Smith (1997; 1998), where it is argued that modularity of cognitive functions is an emergent property that is a product of development, as well as the current work in which the effects of lesions on computer simulations of reading and past tense formation are demonstrated. Although none of this work espouses Fodorian modularity in any of its incarnations, and all assume dynamic changes that have consequences for the end-state, all deal with brain-cognition relationships, challenging other theoretical positions.

The praise relates to the authors' insistence that in cases of congenital disorders that affect the brain, behavioral impairments at

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

the endpoint of development are likely to be the outcome of an extended atypical developmental trajectory, determined by the initial structural anomalies in the system and by its atypical interaction with the environment (Karmiloff-Smith 1997; 1998). This is a plausible position in part because current genetic and biological understanding favors the view that in disorders of a genetic origin, the effects will be wide spread and nonspecific, and thus result in brain changes that are likely to affect learning and development. Developmental cognitive neuropsychology has not drawn a sufficiently clear distinction between acquired deficits and congenital ones, and has too often applied similar theoretical frameworks to both, disregarding the crucial effects of development.

The question that I would like to raise relates to the above position. The assumption that in cases of congenital disorders of genetic origin the child in fact learns with a different brain, predicts that disordered populations will more often than not exhibit an abnormal developmental course. However, it is rather rare to encounter reports of performance in disordered populations that is of a kind never seen in typically developing children. One typically sees reports of delays and asynchronous development, but rarely have truly deviant patterns been reported. A typical example concerns the notion of extended optional infinitive, describing asynchronous development in children with Specific Language Impairment (SLI) (Rice et al. 1998). Reports of developmental delays in neurodevelopmental disorders of genetic origin are too numerous to quote.

Especially problematic in this regard are claims based on group results when the comparisons are drawn between group means and prevalent tendencies. For example, consider the claim made by Cossu Rossini and Marshall (1993) that children with Down syndrome read without phonological awareness. This work might have qualified as an example of an atypical learning trajectory. However, the participants with Down syndrome did not perform at zero level. Rather, in some cases they scored 13 out of 21 correct, and there was just one child who scored 0 on three out of four tasks (Bryne 1993). Hence, whereas there were statistically significant differences between children with Down syndrome and typically developing children on phonological awareness tasks, I doubt whether one can attribute these differences to a different learning trajectory.

The authors cite work on face processing in Williams syndrome suggesting that, despite normal face recognition abilities, the underlying processes by which resolutions are arrived at is atypical (Deruelle et al. 1999). But are these processes of the kind never encountered in children with normal brains? In a similar vein, in all of the research on language in Williams syndrome there have been very few reports, if any at all, of errors that are never seen in normal children (but see Capirci et al. 1996). Most of the findings concern delayed acquisition, higher error rates, and either less (Karmiloff-Smith et al. 1997) or more (Clahsen & Almazan 2001) control over grammatical rules relative to memorized words. The conclusion from most of the work on Williams syndrome – and the authors' work is no exception – as well as from work on other developmental syndromes, is that there are set paths for language learning even under pathology, with variation in speed and onset of various components, yet with minimal options for deviation that are not time-related (Levy et al. 2000).

This is perhaps not surprising given the plasticity that is inherent in the young brain. Brain plasticity has been invoked mostly in cases of focal damage (cf. Bach 1990). However, if brain plasticity is assumed in congenital disorders of genetic origin, then the expectation is that despite genetic abnormalities that cause structural deviations from typicality, functional normalcy may be preserved throughout development. In fact, the question is to what extent structure equals function in the brain and, consequently, whether structural differences, such as are seen in genetic disorders, predict functional differences. The fact that there are healthy children who pursue developmental trajectories that characterize children with abnormal brains, suggests that these atypical developmental options are not a consequence of structural abnormality, and that they may operate in normalcy too. In conclusion, whether structure equals function in the brain is an empirical question that has to be addressed in reference to specific developmental domains. Theoretically, however, given that there is plasticity, the fact that there is structural abnormality does not in itself predict functional abnormality.

Coconstructed functionality instead of functional normality

Shu-Chen Li[a] and Ulman Lindenberger[b]

[a]*Center for Lifespan Psychology, Max Planck Institute for Human Development, D-14195 Berlin, Germany;* [b]*School of Psychology, Saarland University, D-66041 Saarbrücken, Germany.* **shuchen@mpib-berlin.mpg.de lindenberger@mx.uni-saar land.de http://www.mpib-berlin.mpg.de http://www.uni-saarland.de/fak5/lindenberger/members/ulman.htm**

Abstract: We agree with the critique of the Residual Normality assumption. Moreover, we challenge monolithic views of functional normality. Throughout life, development and adaptation require variations in cortical functional circuitry within and across individuals. We propose the principle of "coconstructed functionality" which maintains that brain-behavior functional correspondences are dynamically coproduced by neurobiological, experiential, and contextual processes.

Thomas & Karmiloff-Smith (T&K-S) argue against a canonical hypothesis of developmental cognitive neuropsychology. According to this hypothesis, similarities between domain-specific cognitive impairments observed after damage to the adult brain, on the one hand, and developmental disorders, on the other, point to similarities in underlying causes, thereby lending support to a modular organization of normal adult cognition. They challenge the validity of this inference by demonstrating that its underlying assumption (the "Residual Normality assumption") is untenable, primarily because developing systems generally compensate for selective deficits by modifying other parts of the system. We agree with this critique and would like to take it a step further. We call into question the notion of functional normality and advocate the principle of "coconstructed functionality" for investigating individual and life-span differences in brain-behavior mappings.

Coconstructed functionality. A monolithic divide between normality and abnormality is unproductive and misleading, if "normal" signifies invariance of functional correspondences between brain and behavior. Variations within and between individuals are fundamental to living organisms (Ford 1987; Mayr 1998). Rather than being one-to-one, functional mappings of brain structures and processes to behaviors tend to be many-to-many. At the behavioral level, research on the organization of intellectual abilities has long since identified substantial individual differences within the normal range of functioning (e.g., Deary 2001; Spearman 1904). Similarly, life-span intellectual development is conceptualized as dynamic, multifunctional, and multidirectional (e.g., Baltes et al. 1998; Cattell 1971; Engle 2002; Horn 1968). Among others (e.g., Siegler 1994), Francophone scholars emphasize the importance and ubiquity of multiple developmental pathways with vicariance (i.e., alternation among several processes fulfilling similar functions) being a source of within- and between-individual variations and development (Lautrey, in press; Reuchlin 1978; de Ribaupierre 1993). At the neurobiological level, intra-individual (e.g., Makeig et al. 2002), interindividual (e.g., Toga 2002), and life-span (e.g., Cabeza 2002; Johnson 2001; Raz 2000) differences in brain functioning are large and widespread in the absence of pathology (e.g., within the "normal range" of functioning). Cognition and behavior reflect the reciprocal interactions between neurobiological and experiential contexts of life that vary both within and across individuals. Therefore, instead of reaffirming the orthodox dichotomy between the normal

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

and the abnormal, "normality" is better viewed as the potential of individuals (and environments) to produce an adaptive range of behaviors in a variety of ways. Specifically, any attempt to define normality in terms of invariant brain-behavior mappings neglects the pivotal role of variation at all level and timescales of biological (Mayr 1998) and biocultural organization (Baltes & Singer 2001; Li 2003).

Life-span differences in brain-behavior mappings. In arguing against biological determinism, recent coconstructive theories (e.g., Baltes & Singer 2001; Cole 1999, Elman et al. 1996; Li 2003; Nelson 1996) have stressed the importance of interactive processes and developmental plasticity occurring across different timescales and levels in shaping brain-behavior relations. For example, face processing is less localized or specialized in infants than in adults. In infants, face processing involves both left and right ventral visual pathways; in adults, face processing primarily involves the right ventral visual pathway (de Haan et al. 2002). During senescence, cortical information processing in different brain areas becomes less specific or more integrated (see Cabeza 2002; Logan et al. 2002; Reuter-Lorenz 2002 for reviews). A recent cross-level neurocomputational theory suggests that senescent decrements in neuromodulation, which affect the distinctiveness of neuronal representations of environment-experience interactions, may account for some of the observed losses in cortical specialization (e.g., Li et al. 2001). It is unclear to what extent cortical reorganization in late adulthood directly reflects senescent changes in structural and neurochemical integrity, or compensatory adaptations to such changes. In either case, evidence from both child development and aging indicates that the modal functional circuitries brought to bear upon "nominal cognitive tasks" differ widely by age.

At the behavioral level, life-span changes in the structure of intelligence support our critical appraisal of functional normality. The differentiation-dedifferentiation hypothesis (e.g., Baltes et al. 1980) postulates that intellectual abilities are rather undifferentiated in childhood, undergo differentiation during maturation, leading to a multifaceted ability structure that stays largely invariant during adulthood, and again become more homogeneous (dedifferentiated) during aging. A recent study with a wide range of ages (6 to 89 years) and a comprehensive battery of intelligence tests and standard reaction-time tasks revealed that correlations among different aspects of intelligence and information-processing parameters are indeed higher at both ends of the life span (Li et al., in press). Again, such life-span transformations call for dynamic views on brain-behavior mappings.

Individual differences in brain-behavior mappings. Variability and plasticity of functional circuitry are not confined to the extreme ends of the life span; rather, they are general prerequisites for cumulative experiential tunings reflecting the provisions, demands, and constraints of specific sociocultural contexts. For example, posterior regions of the hippocampi, involved in processing spatial representation of the environment, have been reported to be significantly larger in London taxi drivers than in age-matched controls (Maguire et al. 2000). Similarly, there are marked differences between native speakers of English and Italian with respect to the brain areas involved in reading (Paulesu et al. 2000): English readers show greater activations in the left posterior inferior temporal gyrus and the anterior frontal gyrus, areas associated with word retrieval. Presumably, these differences reflect language-specific adaptations in the functional circuitry supporting grapheme-to-phoneme translation.

Conclusion. Life-span changes, interindividual differences, and intra-individual variability in brain-behavior mappings are not adequately described or explained in relation to a "gold standard" of normality. The principle of coconstructed functionality, with its emphasis on socioculturally embedded experiential tuning, seems more productive than the normal/abnormal dichotomy. To capture the dynamic and multivariate picture of brain-behavior relations, we need systematically to investigate biocultural coconstructive processes that lead to individual and life-span differences in cortical functioning. Undoubtedly, invariant properties of cortical organization will also emerge from such an approach, but they will be located at higher levels of abstraction, be it coherent patterns of neural assembly (e.g., Singer 1995), or general principles of human behavior (e.g., Nesselroade 1991).

ACKNOWLEDGMENT

We thank Paul B. Baltes for his comments on an early draft of the commentary.

What can developmental disorders tell us about modularity?

Gary F. Marcus

Psychology Department, New York University, New York, NY 10012.
gary.marcusl@nyu.edu **http://www.psych.nyu.edu/gary/**

Abstract: This commentary discusses the logic of inferring modularity or the lack of modularity from observed patterns of developmental disorders.

Developmental disorders are, as Thomas & Karmiloff-Smith (T&K-S) observe, not necessarily like cases of adult brain damage. Their simulations well illustrate why one cannot directly infer underlying mental organization from surface behavior, a problem that has long made the discipline of psychology challenging (Chomsky 1959; McClelland 1979). And T&K-S are quite right to highlight the perils of inferring modularity from a pairing of apparently normal behavior in some domains with impaired behaviors in other domains; seemingly normal behavior could always be the product of compensatory mechanisms. (Exactly the same worry hampers the study of second-language acquisition – when a nonnative speaker succeeds, it is difficult to know in advance whether she is relying on remnants of universal grammar or some kind of domain-general cognitive substitute.)

But the naive inference from impairment in multiple domains to a *lack* of modularity is equally riddled with danger. Although cognitive systems are often caricatured as being the product of either entirely modular or entirely shared resources, the reality is that every complex neural system – even ones that are modular in the sense of performing computations that are informationally encapsulated – probably relies on a mixture of domain-specific and domain-general resources. A system for language understanding, for example, probably relies on both shared resources like memory and symbol-manipulation machinery (Marcus 2001) and nonshared machinery that is specialized for linguistic representation (Hauser et al. 2002). A disorder that impaired the neural underpinnings of memory might well then impair both linguistic and nonlinguistic systems, even if the linguistic system included some specialized resources. Disorders that affect multiple domains hence may tell us little about questions about underlying modularity.

What is true at the psychological level is perhaps even more true at the genetic level. Even where two organs are specialized and physical distinct, they often depend in large part on overlapping sets of genes. A chicken's wing and its leg, for example, are for the most part built under the direction of the same set of genes. Out of the hundreds or thousands of genes involved in limb formation, only a handful style='color:black'>(Margulies et al. 2001), black'>. In a similar way, it is possible, perhaps even likely, that separate style='color:black'>on common sets of genes, just as separate limbs the case of the specialization of limbs, the specialization of neural function might rely on a handful of novel genes specific to that domain interacting with a larger number of conserved (shared) genes.

It follows that we must be extremely careful about inferring a lack of modularity from the distribution of disorders. We should expect that many disorders will affect multiple subsystems, yet keep our eyes open for deficits that are more focused. Even if, for

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

example, some variants of reading disorders turn out be consequences of genes (or environmental insults) that affect wide-ranging developmental processes, there may still turn out to be others that derive from rarer genes that have purely domain-specific consequences. In a recent study, Ramus and colleagues (Ramus et al., in press) identified a pure form of dyslexia in which only phonological knowledge is impaired (even after compensatory mechanisms are factored out). It would be foolish to rule out a priori the possibility that such a focused deficit might have a specific genetic basis. The same could be said for language disorders: truly domain-specific disorders might be rare in style='color:black'>, but their relative rarity should not lead us to rule out the possibility of modularity.

No simulation or theoretical analysis will save us from the careful empirical work ahead. T&K-S are right to advise caution, but we should be no quicker to dismiss modularity than we should be to accept it.

ACKNOWLEDGMENTS

This work is supported by the NIH, HFSP, the MacArthur Foundation, and the Center for Advanced Study in Behavioral Sciences.

Double dissociations never license simple inferences about underlying brain organization, especially in developmental cases

James L. McClelland[a] and Gary Lupyan[b]

[a]*Department of Psychology and the Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213;* [b]*Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213.*

jlm@cnbc.cmu.edu glupyan@cnbc.cmu.edu

http://www.cnbc.cmu.edu/~jlm http://www.cnbc.cmu.edu/~glupyan

Abstract: Different developmental anomalies produce contrasting deficits in a single, integrated system. In a network that inflects regular and exception verbs correctly, a disproportionate deficit with exceptions occurs if connections are deleted, whereas a disproportionate deficit with regulars occurs when an auditory deficit impairs perception of the regular inflection. In general, contrasting deficits do not license the inference of underlying modularity.

Thomas & Karmiloff-Smith (T&K-S) have done the field an important service by calling attention to the fact that lesions to a developing brain can have different consequences from lesions to adult brains. They are entirely correct in pointing out that adaptive learning processes will shape the acquired functions not only of those parts of the brain directly affected by the damage, but also other parts of the brain. In general, the article makes a point we heartily agree with, namely, that the standard logic of neuropsychological interpretation should not be applied uncritically to the interpretation of developmental disorders.

In our view (see also Plaut 1995), the standard logic of neuropsychological interpretation cannot be applied uncritically to the interpretation of any disorders, whether or not they are developmental disorders. By the phrase "the standard logic of neuropsychological interpretation" we mean the reliance on a double dissociation between performance on materials from two different experimenter-defined categories to infer that normal performance relies on separate modules specialized for processing the different categories of materials. This logic has repeatedly been used in both the adult neuropsychology literature and in the literature on developmental disorders discussed by T&K-S (for discussions, see Plaut 1995; Shallice 1988). For example, a double dissociation between living things and artifacts in picture naming and property verification has been used to argue for separate modules for different semantic categories (Warrington & McCarthy 1987), and a double dissociation in reading abstract versus concrete words has been used to argue for separate modules for abstract versus concrete nouns (Warrington 1981). Closer to the example used by T&K-S, a double dissociation in production of the past tenses of exception words versus regular words has been used to argue for separate brain mechanisms for words and rules (Pinker 1991; Pinker & Ullman 2002).

In all these cases, connectionist/parallel-distributed processing models (Rumelhart et al. 1986) have provided alternatives to the standard interpretations of these double dissociations (Farah & McClelland 1991; Joanisse & Seidenberg 1999; Plaut 1995). In general, these models take the following form: A single integrated and interactive system is used for processing items of both categories. Because of item characteristics that covary with category membership, performance on items from one category depends more on one part of the system, whereas performance on items from the other category depends more on another part of the system. For example, Plaut (1995) suggested that concrete and abstract words may differ in the number of semantic features. He trained a network with feed-forward and recurrent connections to map both concrete and abstract words from orthography to semantics. The concrete words contained more semantic features, so they used the recurrent connections more effectively and were less dependent on the feed-forward connections. Lesioning the feed-forward connections produced a relative deficit for abstract words, whereas lesioning the recurrent connections produced a relative deficit for concrete words. Standard neuropsychological reasoning would interpret this double dissociation as evidence of separate modules for abstract and concrete words, but there is no such modularization.

Similar problems arise in a developmental context. Different impairments imposed on the system at the beginning of training can differentially impact *learning* to perform correctly with items of different types. Again, standard neuropsychological reasoning would incorrectly imply that the underlying organization is modular.

To illustrate this point, we revisited Simulation One from T&K-S. We trained a single, three-layer, feed-forward network on both regular and exceptional past-tense forms using the same architecture and training patterns (kindly provided by Michael Thomas). Performance after training the intact network was 100% correct for items of both types, in line with the proposal (Rumelhart & McClelland 1986) that a single integrated system might underlie the processing of both regular and exceptional forms.[1] We replicated their "intact" condition, which produced 100% correct performance on both regulars and exceptions, and a fairly severe (80%) "starting point" lesion, which produced fairly good performance on regulars but poor performance on exceptions. We included another condition expected differentially to impair performance on regulars (following Hoeffner & McClelland 1993). This condition builds on the suggestion (Leonard 1998; Leonard et al. 1992) that the English regular past-tense inflection may be weakly represented in the speech signal and therefore difficult to perceive for children with certain forms of language impairment (see also Tallal 1995). Specifically, the nonsyllabic forms of the English past tense (/t/ and /d/, as in "liked" and "loved") involve very slight additions that agree in voicing with the preceding sound and can be very difficult to detect (Bird et al., in press), and the syllabic form (/^d/ as in "hated"‘) is unstressed. To simulate a deficit in perceiving these inflections, the units representing the past-tense inflection were sometimes set to 0 in the target past-tense pattern that the network is given as a model for what it should learn. This is based on the idea that children learn from what they hear and that the perceptual impairment makes the inflection sometimes imperceptible. This condition was otherwise identical to the intact condition. What we see in Table 1 is that in this new condition there is a disproportionate deficit in processing regular past tenses.[2]

Looking at Table 1, we clearly see a pattern of double dissociation. If we found two groups of children who exhibited the two patterns seen in the table and then employed standard neuropsy-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Table 1 (McClelland & Lupyan). *Performance inflecting regular and exception words after training under two different forms of developmental abnormality*

	Item Type	
Deficit Type	Regular	Exception
Intact Network	100%	100%
Sparse connectivity (80% of connections removed)	75%	22%
Perceptual Deficit (60% failure to perceive the regular inflection)	38%	97%

chological reasoning, we would conclude that the normal brain contains separate systems for processing regular items and exceptions. This inference would be incorrect, however, because in this case we know that intact performance is generated by a single system that processes both regular and exceptional forms correctly. We simply have two different deficits that differentially impair learning to process the different types of items. A lesion that produces sparse connectivity reduces the ability of the network to become sensitive to particular combinations of input phonemes that must be considered simultaneously to inflect an exception correctly. An impairment that impacts perception reduces the network's exposure to the information that indicates the correct pronunciation of the regular past tense.

In summary, T&K-S have sounded an important note of caution, indicating that standard neuropsychological reasoning cannot be applied uncritically to the interpretation of patterns of deficits seen in developmental disorders. We hope we have underscored their point by noting that this caution is important in interpreting adult as well as developmental cases. Our simulation suggests that developmental "double dissociations" can be especially misleading because developmental disorders can produce contrasting differential deficits in a single integrated mechanism, but the application of standard neuropsychological reasoning would interpret this pattern as evidence for a two-part system.

NOTES

1. The simulation reported by T&K-S showed relatively poor generalization to novel items. There are several possible reasons for this: (1) the training corpus (which is based on one used previously by Plunkett & Marchman 1993) employed a relatively large number of exceptions compared to regular forms; and (2) the patterns used for the variants of the regular past tense inflection do not reflect its phonological characteristics or its systematic relation to the phonological features of the stem. We would expect that a corpus that more realistically reflected the frequency structure of the language and the phonology of the regular inflection would produce a higher level of generalization.

2. We do not wish to defend the particulars of this simulation as an adequate model of the phonological impairment of any real children; specifically, we believe that in reality, such impairments affect perception of some aspects of exception items as well as regular items, so that the differential would not be as extreme.

Weak evidence for a strong case against modularity in developmental disorders

Ralph-Axel Müller

Department of Psychology, San Diego State University, San Diego, CA 92120. **amueller@sciences.sdsu.edu**
http://www.sci.sdsu.edu/~amueller/index.html

Abstract: Thomas & Karmiloff-Smith (T&K-S) provide evidence from computational modeling against modular assumptions of "Residual Normality" (RN) in developmental disorders. Even though I agree with their criticism, I find their choice of empirical evidence disappointing. Cognitive neuroscience cannot as yet provide a complete understanding of most developmental disorders, but what is known is more than enough to debunk the idea of RN.

The saddest part about this target article is its apparent necessity. It is surprising enough that some cognitive neuropsychologists discount plasticity in adults with acquired lesions. Brain reorganization in recovering aphasics has been shown in many studies, which contradicts Residual Normality (RN). Arguably, the precise links between language recovery and reorganization are under debate for adult acquired lesions (Rosen et al. 2000; Thulborn et al. 1999). In *developmental* populations, however, these links cannot be denied. Loss of the left hemisphere in childhood is often accompanied by good language outcome (Vargha-Khadem et al. 1997), and evidence suggests more pronounced interhemispheric reorganization in children than in adults (Müller et al. 1999). Even for comparatively well-defined structural lesions, the RN assumption is therefore inaccurate: The remaining neurofunctional system *changes* after damage. Can RN be appropriate for developmental disorders that are not even fully understood pathogenetically?

As Thomas & Karmiloff-Smith (T&K-S) argue convincingly, RN in developmental disorders requires strong assumptions that are probably untrue. They present evidence from neural network modeling, contending that cognitive neuroscience has no definitive answers. I disagree: First, computational models are informed by highly oversimplified properties of neural architecture and function, and results may not fully apply to biological brains. For example, the biological meaning of added noise or reduced discriminability in T&K-S's lesion models remains fuzzy. More importantly, neuroscience provides a wealth of evidence showing that RN cannot be expected in developmental disorders.

For a trivial start, genes do not code for cognitive modules (Gottlieb & Halpern 2002). As long as there is loose talk about "genetic double dissociations" (Pinker 1999), an innate "language acquisition device," and "the genetically determined component of the brain . . . that is dedicated to . . . language" (Chomsky 2002, pp. 83–85), ad nauseam repetition is unavoidable: The genome may be a "code," but one for proteins, not cognitive subsystems. The RN assumption could be reasonable if modular cognitive dysfunctions were linearly caused by aberrant genes. There are indeed developmental disorders with known genetic causes (e.g., phenylketonuria or fragile-X syndrome), but these affect multiple brain and body systems in pleiotropic ways. The same applies to language impairment in family KE, caused by a defective FOXP2 gene (Lai et al. 2001). Behavioral patterns and brain morphometry suggest a broad deficit spectrum with pronounced motor involvement, not a specific language impairment (Watkins et al. 2002a; 2002b).

I will discuss the limits of the RN assumption, taking autism as an exemplary developmental disorder. Modularity has been claimed regarding theory-of-mind (ToM) deficits in autism (Leslie 1992; Scholl & Leslie 2001), which are supported by behavioral and neuroimaging evidence (Baron-Cohen et al. 1999; Castelli et al. 2002; Rutherford et al. 2002). Neuroimaging and lesion studies have also identified brain sites in the temporal and frontal lobes that are normally involved in ToM (Calder et al. 2002; Stone et al. 1998). ToM in autism can therefore serve as a test case for the RN assumption in similar ways to past-tense formation in Williams syndrome and developmental language impairment, as discussed in T&K-S.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Many autistic symptoms (obsessive interests, repetitive behaviors, motor abnormalities, language deficits; Dawson et al. 2002) go beyond ToM. Theoretically, these could emerge from primary ToM dysfunction. However, this would contradict the RN assumption because it would imply developmental interaction between an impaired module and other cognitive domains. Structural imaging in autism implicates numerous brain loci, with additional evidence of abnormal gray and white matter growth curves (Akshoomoff et al. 2002). Functional neuroimaging shows atypical activation patterns in autism and Asperger's disorder for tasks assumed to tap into ToM (Baron-Cohen et al. 1999; Castelli et al. 2002; Happé et al. 1996). However, the sites of functional anomalies differ across studies, and overall numerous other regions have been implicated by functional neuroimaging (Rumsey & Ernst 2000). Hence, imaging does not support selective involvement of frontal or temporal regions normally involved in ToM.

Admittedly, evidence for diffuse brain anomalies, such as abnormal growth patterns, cannot provide explanation for the very specific (though variable) symptom profiles in autism. The modular RN approach is a shortcut that may generate simple models, but such models are counterproductive when neurodevelopmental evidence is much more complicated. Why is it so complicated? Because behaviorally defined developmental disorders, such as autism or developmental language impairments, are likely to require explanation in terms of numerous interacting genetic and environmental risk factors. Not only is this interaction complex at any given point in time, but it constantly changes throughout development (Gottlieb & Halpern 2002).

Consider serotonergic abnormalities in autism, as supported by neuroimaging (Chugani 2002), blood plasma, and pharmacological studies (Buitelaar & Willemsen-Swinkels 2000). Crucially, the role of serotonin *changes* during development. Whereas in the adult brain serotonin functions as a neurotransmitter, it plays neurotrophic roles during development impacting neuronal differentiation, myelination, synaptogenesis, and dendritic development (Whitaker-Azmitia 2001). Furthermore, critical periods for neurotrophic involvement of serotonin vary across brain regions. Although the evidence is inconclusive (Betancur et al. 2002), it is possible that serotonergic disturbances in autism are linked to specific genes. Even if one were to assume that serotonergic disturbances are necessary and sufficient causes of autism, their effects would be incompatible with modular RN assumptions because serotonin plays fundamentally different roles at different developmental stages. Effects of serotonergic disturbances during development will hence differ from those in adulthood. Furthermore, autistic pathogenesis is certainly more complex than an isolated defect of a single transmitter system. The same is true for developmental language impairments. This should highlight how much there is still to be learned from neuroscience for a true understanding of developmental disorders – and how little we can expect to learn from overly simplified modularity assumptions in cognitive neuropsychology.

The best is yet to come: The promise of models of developmental disorders

Yuko Munakata,[a] Jamie O. Edgin,[b] and Jennifer Merva Stedron[b]

[a]*Department of Psychology, University of Colorado, Boulder, Boulder, CO 80309-0345;* [b]*Department of Psychology, University of Denver, Denver, CO 80208.* **munakata@colorado.edu jogline@frontier.psy.du.edu jstedron@frontier.psy.du.edu http://psych.colorado.edu/~munakata/**

Abstract: The developmental modeling approach to investigating developmental disorders appears highly promising. In this commentary, we question the untapped potential of this approach for supporting insights into particular developmental disorders, developmental processes across the life span, and the viability of traditional theories of developmental disorders.

Thomas & Karmiloff-Smith (T&K-S) make a valuable contribution using neural network models to demonstrate the importance of developmental processes in developmental disorders. This approach should advance not only the study of developmental disorders (because neural network models provide a particularly well-suited tool for exploring the nonlinear, emergent processes likely to be at work in developmental disorders), but also the study of computational systems (because understanding developmental processes may be essential for understanding how such systems come to behave as they do). Our comments and questions below are aimed primarily at pushing this developmental modeling approach further to evaluate its potential to advance understanding of: (1) specifics of developmental disorders beyond general principles; (2) acquired deficits in adults, with developmental processes being relevant across the life span, and (3) the realistic viability of Residual Normality.

First, the current models serve as important existence proofs that the same damage to developing and mature systems have different effects on performance. As such, the models play an important role in clearly demonstrating the general point that developmental processes must be considered in investigating developmental disorders. However, one can ask what the models contribute toward advancing understanding of specific processes in specific developmental disorders. For example, what specific neural mechanisms are thought to be disrupted in Specific Language Impairment (SLI), Williams syndrome, autism, and so on? How do these neural mechanisms map onto various manipulations in network models, such as the removal of connections, the addition of noise in activation levels, and the alteration of discriminability in processing units? How do these specific neural mechanisms (and their computational analogs) interact with one another and with the developmental process to produce developmental outcomes that capture the profiles of specific developmental disorders? Direct mappings to specific neural mechanisms and specific developmental disorders will be critical for further advances within this developmental modeling approach (as argued elsewhere by one of the authors; Thomas 1999).

Second, the current approach seems to treat development as a process that ends at a particular point, such that developmental processes are relevant for understanding only developmental disorders and not acquired deficits in the mature system. This assumption seems unnecessary and unwarranted. Instead, consideration of developmental processes could shed light on deficits acquired in adulthood, and modeling could prove (and has proven) useful in this regard as well. For example, even in mature neural network models, the effects of lesions may be best understood by considering the developmental processes that created the mature system. In a reading model with the full set of orthography-phonology-semantic mappings as described in the target article, severe lesions to the orthography-phonology pathway led to semantic errors similar to those observed in deep dyslexia, despite the fact that the semantic system was undamaged (O'Reilly & Munakata 2000). This pattern could be understood in terms of the division of labor between the different pathways that arose during development. The semantic pathway developed while relying on contributions from the orthography-phonology pathway. As a result, when the orthography-phonology pathway was lesioned, the semantic pathway could no longer function in the way it had learned to with the orthography-phonology pathway intact. Hence, lesioning one pathway in the mature network did not simply reveal the intact functioning in the remaining pathway. In this way, a consideration of developmental processes made sense of how damage to the orthography-phonology pathway in a mature system could lead to semantic errors in the face of an intact semantic pathway.

Further, the current approach seems to treat cases of acquired deficits in the mature system as if no compensations or learning occur in response (in contrast to the developing system). This treatment of the mature system seems extreme, given that there is plasticity in the adult system. What is the evidence that devel-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

opmental processes do not apply to the adult case? Models could serve as useful tools for exploration of these issues as well. For example, following lesions to mature models, do the models perform similarly if learning is then turned off versus being left on? Or do even mature models show interesting compensations as well, if allowed to learn? Do such compensations differ from those observed in developmental cases? How do any of these profiles map onto those observed in people?

Third, in describing several ways to achieve Residual Normality, the paper conveys some skepticism about the likelihood that such conditions would be met in real systems; this point may be worth emphasizing and elaborating. As discussed in notes to the paper, models that might be viewed as meeting conditions of Residual Normality actually do not show pure specializations. For example, even with strong structure-function correspondences in the "what" and "where" pathways of an object processing model (Mareschal et al 1999), each pathway can still encode information relevant to the other channel. And even with an attempt to design an algorithm so that a system could learn to read only regular words, the system still learned to read some exception words (Coltheart et al. 1993). Hence, although one might be able to describe systems that in theory could achieve Residual Normality, it is unclear whether such systems could actually work in practice. Implemented computational models will therefore be essential for assessing the viability of theories positing Residual Normality. Such models would allow detailed testing of traditional theories and an assessment of their developmental and biological plausibility, comparable to the kind of evaluation that is now possible with models of other theories.

In sum, the developmental modeling approach provides an exciting advance in how to structure the study of developmental disorders. Further progress will likely depend upon greater specificity in the models to support insights into particular developmental disorders, a consideration of developmental processes across the life span, and the development of models to allow comparable evaluation of contrasting theories.

Residual Normality and the issue of language profiles in Williams syndrome

Csaba Pléh,[a] Ágnes Lukács,[b] and Mihály Racsmány,[c]

[a]*Center for Cognitive Science, Budapest University of Technology and Economics, Budapest, and Department of Neuropsychology and Psycholinguistics of the Hungarian Academy of Sciences, Műegyetem Rkp 9 R 203 H-111, Hungary;* [b]*Institute for Linguistics, Hungarian Academy of Sciences;* [c]*Department of Psychology, University of Bristol, and Department of Psychology, University of Szeged.* **pleh@itm.bme.hu alukacs@nytud.hu racsmany@edpsy.u-szeged.hu**

Abstract: One of the debated issues regarding Residual Normality (RN) is frequency sensitivity in Williams syndrome (WS). We present some data on frequency sensitivity in Hungarian WS subjects. Based on vocabulary measures, we suggest that instead of the across-the-board frequency insensitivity proposed by some, a higher frequency threshold characterizes these subjects' performance. Results from a category fluency task show that whereas frequency sensitivity in WS is in line with controls, error patterns imply a qualitatively distinct, looser categorical organization. Regarding the much-debated issue of morphological overgeneralizations, our data suggest that frequency sensitivity cuts across the divisions proposed by dual-process theories. In general, some of the frequency effects are the same as in typically developing populations, but with a delayed pattern. Frequency may be interpreted as supporting RN, but in WS it operates with higher thresholds that might be a general processing feature of WS individuals.

One of the key issues regarding the theoretical problem of postulating or questioning parallels between acquired neuropsychological disorders and developmental impairments of a neurogenetic origin, as Thomas & Karmiloff-Smith (T&K-S) duly emphasize, is the nature of some of the proposed developmental dissociative patterns. The goal of our commentary is to take up the issue of behavioral-cognitive patterns through presenting some of our results concerning lexical organization in Williams syndrome (WS).

WS is one of the favorite targets of dissociative developmental theories. This case is used as an example of impairment in spatial cognition with Residual Normality (RN) in language (for a historical survey, see Bellugi et al. 2000) and of specific within-language dissociations between an intact grammar or rule system (RN) and impaired lexical knowledge or associative memory, contrasted in a tone of double dissociation with Specific Language Impairment (SLI) subjects who are supposed to have an intact associative lexical system (RN) and an impaired grammatical/rule system (Clahsen 1999; Pinker 1991; 1999; Ullman & Gopnik 1999). T&K-S challenge this simplified identification of developmental patterns with adult dissociative profiles. We would like to present some data on the issue of frequency sensitivity in WS to indicate that the full complexity of dissociations arises as a result of development.

One certainly oversimplified interpretation of the within-language dissociation proposed by Pinker and Clahsen would be to suppose that in relation to or caused by the underdevelopment of their lexical-associative system, WS subjects are less sensitive to frequency across the board. In the frame of the Hungarian Williams Syndrome Project we obtained several types of evidence related to this issue. Table 1 summarizes some relevant measures in two lexicon-related tasks in our project. Fifteen subjects with WS (mean age 13.4) were compared to children matched in age (age control; AC) and Peabody receptive vocabulary scores (vocabulary control; VC) (mean age 7.1).

In the picture-naming vocabulary task, WS individuals showed a lower performance than either of their controls ($p < 0.05$), in line with the impaired lexical system conception. However, frequency effects in the WS group were stronger than in the AC group (interaction $p < 0.001$, although this difference might be due to a ceiling effect in the latter group) and were the same as frequency effects for the VC group ($p < 0.001$). Hence, WS persons are sensitive to frequency although their naming performance lags behind controls. Instead of their being insensitive to frequency, we might argue that they are more sensitive to frequency than controls, in the sense that they need more input to stabilize items in their memory system (for a more detailed presentation of results, see Lukács et al. 2001; in press).

On the other hand, interesting qualitative differences were observed between these groups in an oral category fluency task using eight categories (Food, Animals, Drinks, Musical Instruments, Occupations, Things to Read, Furniture, and Clothes). Remark-

Table 1 (Pléh et al.). *Some overall differences between the WS and control groups in lexical measures*

Task	Williams syndrome	Age control	Vocabulary control
Picture naming frequent/rare, % correct	91/65	97/93	96/79
Category fluency, average frequency of items	95.05		96.85
Category fluency, average number of rare items	16.02		9.58
Category fluency, noncategory members	3.82		1.81

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

ably, there were no differences in the number of items and in overall frequency effects between the WS and the control group. However, WS subjects produced more repetitions ($p < 0.001$), more noncategory member items ($p < 0.05$), and more members ($p < 0.01$) that are not listed in the normative adult data (Kónya & Pintér 1986). A more fine-grained analysis of the latter difference reveals that items listed by the WS subjects that are not listed in the normative database are often not unusual or rare but are pet names for food or animals, examples of which are not found in the control sample, in which children are past the age at which typically developing children use such names.

So WS individuals do not differ from their controls matched on receptive vocabulary scores in the frequency effects governing retrieval and production of lexical items belonging to a semantic category. Repetitions and the answers not obeying category restrictions are in agreement with the conclusions of earlier studies (Jarrold et al. 2000; Johnson & Carey 1998; Temple et al. 2002) – that instead of the lack of frequency effects, it is atypical organization or underspecification of conceptual knowledge that can explain peculiarities of language use in Williams syndrome.

The issue of rules versus items – and the implied double dissociation – is certainly crucial to the overall debate between dual-system (Clahsen 1999; Pinker 1999) and single-system approaches to the development and mental representation of language. As Thomas and Karmiloff-Smith (2002) point out, the much-discussed observations that WS subjects show strong morphological overgeneralizations (Clahsen & Almazan 1998) can be reinterpreted in the frame of single-route systems.

> While individuals with WS do show poorer performance on exception verbs than regular verbs, this appears to be in step with their delayed performance, because younger children also find exception verbs harder than regular verbs. There is no *specific* deficit for irregular verbs. (Thomas & Karmiloff-Smith 2002, p. 17)

The Hungarian nominal allomorphy system is especially relevant to test the rule issue because of the fact that there are subclasses of different generality, and saturation and regularity and frequency are much less tied there than in the English past tense (see Lukács & Pléh 1999). Earlier we reported data on a group of Hungarian WS subjects showing a rather strong tendency to overgeneralize, thus supporting a crude version of the dual-route system (Lukács et al. 2001). Recently, more elaborate analyses were performed on the same data set from the impaired group with a set of more properly matched controls. Table 2 shows correct percentages in an elicited morphological suffixation task in which subjects had to provide accusatives and plurals to nouns varying in both regularity and frequency.

The superiority of performance on regulars over irregulars was observed in WS ($p < 0.05$), whereas frequency effects were not significant. VCs seemed to follow a similar pattern in that they also had more errors on irregulars, but in their case overregularization of irregulars was clearly related to frequency. This may be interpreted as evidence for a slower language development in WS, as suggested by Thomas and colleagues (Thomas et al. 2001). However, there are still some signs to the effect that in WS subjects even frequent irregular items lead to overgeneralizations, and they seem to be delayed on regulars as well, arguing against a clear case of a selective impairment with irregulars.

Table 2 (Pléh et al.). *Correct performance of different groups with noun suffixation*

Noun type	Williams syndrome	Age control	Vocabulary control
Regular, frequent	89	100	99
Regular, rare	85	100	99
Irregular, frequent	74	100	92
Irregular, rare	66	99	64

Therefore, taken together our data on Hungarian WS subjects suggest that frequency does operate as a regulatory factor weighting connections in the memory system of this impaired population. In general, they seem to have relative difficulty stabilizing items in their memory, which shows up in different lexical tasks. But frequency does not disappear as a factor, it merely has a harder time. At the same time these subjects seem to have a looser semantic organization and a peculiar tendency to overgeneralize with irregulars, but they also err on regulars. These facts imply some peculiarities in their language representation, though not necessarily a two-process theory.

ACKNOWLEDGMENTS

Financial support for most of the research discussed in the commentary was provided by an OTKA grant (Hungarian National Science Foundation) T 029514 and an NKFP grant (5/0079/2002) to Csaba Pléh and by an NSF grant No. BCS01226151 to Ilona Kovács with Csaba Pléh as co-principal investigator. The empirical worked referred to was made possible by the Hungarian Williams Syndrome Association.

Evidence for a domain-specific deficit in developmental dyslexia

Franck Ramus

Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/ENS/CNRS), 75006 Paris, France. **ramus@lscp.ehess.fr**
http://www.lscp.net/persons/ramus/

Abstract: Thomas & Karmiloff-Smith (T&K-S) claim that "Residual Normality" is a priori unlikely, that is, that specific cognitive deficits should not exist in developmental disorders. Here I review evidence that a specific cognitive deficit is at the core of developmental dyslexia and I provide a possible neurological account thereof.

Thomas & Karmiloff-Smith's (T&K-S's) argument against Residual Normality (RN) is based on (1) ideological preconceptions concerning the development of cognitive modules; and (2) neural network simulations that suggest that RN is unlikely to occur. However, as they note, whether RN exists or not is an empirical issue, not one that will be solved by theoretical considerations or computer simulations. I propose to examine empirical data from developmental dyslexia. As this suggests that RN does sometimes occur, it follows that T&K-S's ideological preconceptions are mistaken and that their neural network simulations do not adequately model brain and cognitive development.

Developmental dyslexia is traditionally viewed as a specific developmental disorder: One that initially affects exclusively phonological representations and processes, with consequences for the acquisition of grapheme/phoneme correspondences and reading processes (Snowling 2000; Vellutino 1979). Because all other cognitive components seem entirely unaffected, dyslexia is usually seen as a case of RN.[1] However, a number of researchers point out that dyslexics' difficulties are not restricted to phonology and reading but include auditory (Tallal 1980), visual (Lovegrove et al. 1980), tactile (Stoodley et al. 2000), motor (Fawcett et al. 1996), learning (Nicolson & Fawcett 2000), and attention deficits (Hari & Renvall 2001). It is even suggested that dyslexia is a general sensorimotor syndrome (Stein & Walsh 1997). Obviously, all these findings undermine the claim for RN. The difficulty here is in seeing the forest for the trees.

Evidence for sensorimotor dysfunction in dyslexia comes from many studies reporting significant group differences between dyslexics and controls in sensory or motor tasks. However, when individual data is examined, it invariably shows that there is a large overlap in performance between the two populations and that the group effect is carried by a subgroup of dyslexics performing ab-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

normally (for a review, see Ramus 2003). This could simply reflect measurement error and task reliability. But many of the results reported in recent studies rely on several measures of the same psychophysical threshold in each individual, obtained using adaptive procedures. There are therefore good reasons to believe in individual data points. Hence, on any given sensory or motor task, a good proportion of dyslexics seems to have intact performance. This contrasts with the quasi-universality of the phonological deficit.

It could be that any single low-level deficit fails to account for all dyslexics, but that, taken together, they do: Perhaps one third of dyslexics have an auditory-based phonological deficit, one third have a visual deficit, and the rest have a cerebellar learning disorder; all of these would lead to reading difficulties. My collaborators and I have addressed this possibility by administering a comprehensive 10-hour test battery of psychometric, phonological, auditory, visual, and motor skills to 16 adult dyslexics and 16 controls. We found that all 16 dyslexics had a phonological deficit, 10 had an auditory deficit, 4 had a motor dysfunction, and 2 had visual problems (with partial overlap between the sensorimotor deficits). Most interestingly, 5 dyslexics seemed to have a pure phonological deficit, as their performance on the whole battery of auditory, visual, and motor/cerebellar tests was entirely normal (Ramus et al. 2003) – a clear demonstration of RN. These results are not an artifact of data collection cut short prematurely; on the contrary, we have adopted all the main state-of-the-art tasks used to reveal sensorimotor deficits. Yet, some dyslexics seem entirely untouched by these low-level deficits. Similar results have been obtained with dyslexic children (Kronbichler et al. 2002; White et al., submitted). Moreover, it is becoming increasingly clear that the auditory deficits observed in certain dyslexics have little causal relationship to the phonological deficit (Bishop et al. 1999a; see review by Ramus 2003).

Consistent with the currently known data, my view of dyslexia is that of a specific phonological deficit sometimes accompanied by a sensorimotor syndrome with variable manifestations and little additional impact on phonology and reading. According to T&K-S, such a disorder makes no neurological and developmental sense. Indeed, as the authors note, "Karmiloff-Smith and colleagues have argued that a priori, the effects of genetic abnormalities are likely to be widespread throughout the brain and unlikely to be isolated to single high-level cognitive modules" (target article, sect. 3.3). Yet I propose that there is a plausible neurodevelopmental model for this view of dyslexia.

Two types of brain anomalies that make cognitive sense have been observed in dyslexics' brains: (1) cortical layer ectopias and microgyri (focal anomalies of neural migration in the outer layer of the cortex) are located mainly in left perisylvian areas and can be seen as underlying the phonological deficit (Galaburda et al. 1985; Humphreys et al. 1990); (2) the magnocells of the lateral and medial geniculate nuclei of the thalamus are disordered, which is thought to underlie the visual and auditory deficits, respectively (Galaburda et al. 1994; Livingstone et al. 1991). Interestingly, the thalamic anomalies seem secondary to the cortical ones, rather than the other way around. Moreover, they arise only under certain conditions. Indeed, based on animal models, it seems that the thalamic anomalies arise from the conjunction of cortical anomalies and certain hormonal conditions, notably excessive testosterone concentrations during brain development (Fitch et al. 1997; Herman et al. 1997; Rosen et al. 1999).

These neurological data naturally suggest a model of developmental dyslexia in which cortical anomalies occurring in early brain development and located in cortical areas important for the development of phonology and reading are the main underlying cause of the phonological and reading impairments. In certain individuals only (under certain hormonal conditions), the disruption spreads to the thalamus and engenders sensory deficits.[2] This is consistent with behavioral-genetic studies showing that dyslexics' phonological deficit is highly heritable, whereas their auditory and visual deficits are not (Bishop et al. 1999a; Davis et al. 2001; Olson & Datta 2002).

Interestingly, the early focal cortical anomalies documented above are quite similar to the focal lesions seen in adult neurology. Of course, there is no doubt that focal cortical damage may have very different effects in a developing brain and in an adult brain (as already pointed out by Frith 1986). But this is a clear example of how the effects of genetic abnormalities can be isolated to certain *anatomical* modules.[3] Furthermore, the dependence of the spread of the disruption on hormonal factors demonstrates that "malformation in the cortex is not always associated with plasticity changes in connectionally related areas" (Galaburda 2001, pp. 456–57). Nevertheless, in order to explain the dyslexia data, our model requires that certain *anatomical* modules are unique in their ability to support certain *cognitive* modules; if they are disrupted, no amount of "plasticity" or "compensation" will allow the brain to relocate the function to another area. This seems to be a necessary conclusion from the observation that dyslexics never fully recover from their specific congenital phonological deficit. Note that the functional uniqueness of certain brain areas does not necessarily imply the existence of a detailed genetic blueprint; it may simply reside in the combination of their connections to certain inputs and outputs together with their representational and computational properties.

In summary, developmental dyslexia, when not accompanied by the sensorimotor syndrome, seems to be a prime example of residual normality, and there is a plausible neurological model supporting this view. I submit that such empirical work and neuroscience-based modeling should weigh more than ideological preconceptions and computer simulations. In fact, the incapacity of T&K-S's neural network models to generate or accommodate the sort of neurodevelopmental scenario proposed only highlights the fact that these models are too simplistic even to begin to capture the complexity and sophistication of neural and cognitive development. It therefore seems unwise to rely on them to draw very general conclusions about the brain, development, and cognition.

ACKNOWLEDGMENTS

I am grateful to Gary Marcus, Uta Frith, and Sarah White for comments on this paper.

NOTES

1. Note that this is a completely different issue from that of the phonological versus surface subtypes advocated by Castles and Coltheart (1993) and discussed at length by T&K-S. It is now quite clear to most dyslexia researchers that the phonological and orthographic reading routes cannot really develop independently and, therefore, that selective damage to one of them is unlikely in developmental dyslexia (e.g., Zabell & Everatt 2002). I therefore generally agree with T&K-S's criticism of this hypothesis. However, this is orthogonal to the question of RN outside dyslexics' phonological and reading systems.

2. According to Stein and Walsh (1997), the motor part of the syndrome can be explained from there by the massive inputs from the magnocellular system via the posterior parietal cortex to the cerebellum.

3. It is well established from studies of autoimmune mice that cortical ectopias are of genetic origin (Sherman et al. 1990).

Are developmental disabilities the same in children and adults?

Paula Tallal

Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102. **Tallal@axon.rutgers.edu**

Abstract: Thomas & Karmiloff-Smith (T&K-S) raise an issue of considerable theoretical importance: Are developmental disorders like cases of adult brain damage? However, a related question: Are developmental disabilities the same in children and adults? is rarely addressed. Failure to consider the cumulative and differing effects of aberrant development across the life span confounds the current literature on both developmental dyslexia and Specific Language Impairment.

The role of rapid auditory processing (RAP) constraints in the origin of language development and disorders is a hotly debated is-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

sue. Tallal and Piercy (1973a; 1973b) demonstrated that 6–8-year-old control children were able to respond correctly when two brief tones differing in frequency were presented rapidly in succession with only tens of milliseconds interstimulus intervals (ISIs). However, children with Specific Language Impairment (SLI) of the same age required several hundred milliseconds to respond at the same level of accuracy. They proposed that this deficit would constrain the ability of children with SLI to extract the rapidly successive acoustic changes that characterize the acoustic waveform of speech. Using computer-generated speech syllables, in which the temporospectral changes that occur within the formant transitions of speech could be controlled, they showed that children with SLI were significantly impaired in discriminating between speech syllables that were characterized by brief, rapidly successive acoustic changes (Stark & Tallal 1980; Tallal & Piercy 1974).

To demonstrate the specificity of this acoustic deficit, the duration of the intrasyllabic formant transition within the syllables /ba/ versus /da/ was computer synthesized to be 40 msec in one set of syllables and 80 msec in a second set. The results showed that whereas the children with SLI could not discriminate between the syllables with 40 msec duration formant transitions, they could when the duration of the formant transition was computer-extended to 80 msec (Tallal & Piercy 1975). This finding subsequently formed the basis of a new treatment for language learning impairments (Merzenich et al. 1996; Tallal et al. 1996).

Evidence from both longitudinal studies and family genetic studies demonstrates considerable overlapping etiology and co-occurrence between developmental oral language and reading impairments (Flax et al., in press; see Leonard 1998 for review). As such, Tallal (1980) investigated the relationship between deficits in phonological awareness (the ability to read nonsense words), a deficit that has been shown to characterize children with developmental dyslexia (Snow et al. 1998), and RAP. It was hypothesized that if RAP were directly related to reading impairment, then a high correlation would occur between the degree of RAP impairment and the degree of deficit in reading nonsense words. The results yielded a highly significant ($p < .001$) positive correlation coefficient ($rs = .81$). That is, the more difficulty a dyslexic child had in responding correctly to rapidly successive nonverbal auditory stimuli, the more difficulty they had in reading nonsense words. Although RAP was significantly correlated with measures of reading, it was not significantly correlated with performance IQ ($rs = .06$), demonstrating the specificity of this correlation (Tallal 1980).

Despite these striking relationships, the role of nonlinguistic processing deficits, specifically RAP, in the origin of SLI or dyslexia has become one of the most hotly debated topics in this field (Denenberg 1999; Mody et al. 1997). I would argue that at the heart of this theoretical debate are many of the same critical issues raised in Thomas & Karmiloff-Smith's (T&K-S's) target article. The entire theoretical premise that early constraints in RAP will lead to cascading effects on phonological analysis, finely tuned phonological representations, disordered phonological processing, deficits in the development of grammatical morphology, as well as phonological awareness, reading decoding, and reading comprehension deficit, *requires* a developmental account of experience-dependent Hebbian learning (Hebb 1949) in the nervous system (see Merzenich & Jenkins 1995 for review). Nonetheless, such a perspective is generally lacking in the search for the origins of developmental language-based learning impairments. Most problematic is a lack of concern for the age of subjects and age-appropriate methodologies, and specifically a growing trend of approaching issues of causality by studying adult subjects with a lifelong history of developmental language and/or reading impairments. Although such studies may be informative about the "endstate" of a lifetime of compensating for a developmental disability, it is not possible to approach issues of etiology, causality, or neurobiological substrates of a developmental disability by studying adults. Such issues are best informed by developmental studies.

Recent prospective longitudinal studies of infants born into families either with (FH+) or without (FH-), a family history of language learning impairment (LLI), illustrates the importance of taking a developmental perspective in evaluating the potential role of RAP thresholds on subsequent language development. RAP thresholds were obtained when infants were 7.5 months of age, and the sample was then followed through age 36 months. Results showed that significant differences in mean RAP thresholds were seen in infants born into families with a history of LLI as compared with controls (Benasich & Tallal 1996). Examination of relations between infant processing abilities and emerging language through 24 months of age revealed that RAP thresholds at 7.5 months were the single best predictor of language outcome, regardless of whether or not an infant did or did not have a family history of LLI. At age 3 years, RAP threshold and being male together predicted 39% through 41% of the variance in language outcome. Furthermore, these data accurately classified (91.4%) at infancy children who would be "impaired" (one SD below the mean) at age 3 on the verbal reasoning component of the Standford-Binet. Importantly, none of the infant variables proved capable of discriminating between 3-year-old children on the nonverbal portions of the Standford-Binet, demonstrating the specificity of this relationship between individual differences in infant RAP thresholds and subsequent individual differences in language development (Benasich & Tallal 2002).

Prospective longitudinal studies demonstrate how critically important taking a developmental perspective is in addressing issues pertaining to the causes and determinants of developmental disabilities, and how misleading studies in older children and adults with developmental disabilities may be in the search for developmental etiologies.

How connectionist simulations fail to account for developmental disorders in children

Christine Temple[a] and Harald Clahsen[b]

[a] *Developmental Neuropsychology Unit, University of Essex, Colchester CO4 3SQ, United Kingdom;* [b] *Department of Linguistics, University of Essex, Colchester CO4 3SQ, United Kingdom.* **tempc@essex.ac.uk harald@essex.ac.uk http://privatewww.essex.ac.uk/~harald**

Abstract: Using connectionist modelling, Thomas & Karmiloff-Smith (T&K-S) claim that developmental disorders in children are characterised by atypical trajectories and an ultimate functional architecture that is fundamentally different from normal. We argue that there is no empirical evidence for these claims in any developmental disorder and that the available evidence provides support for Residual Normality in both developmental and acquired disorders. We also refute the claim that modular accounts cannot encompass developmental trajectories in children with developmental disorders.

A fundamental debate concerning developmental disorders is the relationship between established skills and normal skills. An assumption of cognitive neuropsychology is that impaired performance reflects normal function minus impaired function, such that exhibited skills are part of the normal functional architecture (Ellis & Young 1988; Temple 1997). This has been termed subtractivity (Saffran 1982), transparency (Caramazza 1984) or Residual Normality (Thomas & Karmiloff-Smith). Evidence against subtractivity would come from cases in which the functional architecture of the system had altered with development of cognitive modules that do not exist in the normal brain. However, there is no such empirical evidence in relation to either adults or children.

A contrasting view to subtractivity is that performance in developmental disorders does not relate to normality but reflects abnormal development of the entire system. Thomas & Karmiloff-Smith (T&K-S) argue for a qualitatively different endstate to

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

normal, with "a different functional structure in children." On this basis, developmental disorders provide no insight into normal development, and child neuropsychology is erroneous in its method and conclusions. T&K-S argue from simulations, but there remains no empirical evidence in any developmental disorder that the ultimate functional architecture has fundamentally different organisation from normal, rather than merely lacking or having reduced development of components of normal functional architecture. Consequently, their simulations are not a valid reflection of developmental disorders in children.

In contrast to absence of empirical evidence supporting a different functional structure in abnormal development, there is extensive empirical evidence of Residual Normality in both developmental and acquired disorders in children. For example, in procedural dyscalculia, despite pervasive impairment in arithmetical procedures, knowledge of numerical facts can have 95% accuracy. In number fact dyscalculia, severely impaired knowledge of numerical facts coexists with excellent mastery of procedural algorithms (Sokol et al. 1994; Temple 1991; 1994). Similar arguments apply to developmental perceptual disorders that selectively affect perception of either movement or location (Ahmed & Dutton 1996; McCloskey et al. 1995). Another example is selective impairment in components of complex visual recognition with severely impaired face recognition but excellent visual word recognition in reading for Dr. S. (Temple 1992; 1997) and severely impaired visual recognition for irregular words yet intact visual memory for other complex material in surface dyslexia (Castles & Coltheart 1996). In each case, impaired skill coexists with excellent development of contrasting skill within the same domain. Such cognitive neuropsychological analyses highlight the focal modular impairments seen within developmental disorders.

T&K-S suggest that evidence for a distinct functional structure in a developmental disorder is found in face recognition in Williams syndrome (WS). Here it is argued that development is completely different from normal and cannot be explained in terms of a normal brain with "parts intact and parts impaired" (Karmiloff-Smith 1997). Early studies argued for good face processing in WS, but based only on face matching. Feature-based componential analysis of faces, dependent on local cues, is intact in WS, but global processing of the configuration of faces is impaired (Deruelle et al. 1999; Karmiloff-Smith 1997). However, normal face recognition begins with a featural componential strategy (Carey & Diamond 1977). This ability remains existent in the normal repertoire of skill. Normal children and adults can perform both local and configural analyses of faces successfully. Children with WS can perform only the former. In terms of the normal brain, componential processing is intact and configural processing impaired, a straightforward modular dissociation within face processing systems. There is no evidence for a new and different endstate.

A further example of modular impairment is the intact ability to form regular inflected words but selective impairment in irregular inflection in WS (Clahsen & Almazan 1998). T&K-S dismiss this study, referring to their own study reporting absence of such selective effects when scores are averaged across diverse ages and mental ages, rather than in age bands (Thomas et al. 2001). Clahsen and Temple (2003) offered a reanalysis of this data that remains compatible with a modular account and emphasises the importance of appropriate control matching within developmental levels. Furthermore, recent data from further cases of WS confirm the original finding of a selective (lexical) impairment with irregular inflection (Clahsen et al. 2003). T&K-S also ignore the comparable selective impairment in comparative-adjective formation (Clahsen & Temple 2003) and irregular plural formation (Clahsen & Almazan 2001) yet intact complex syntax in WS (Clahsen & Almazan 1998). They appear unable to account for subtle selective deficits of this kind.

A further debate concerns the nature of development itself. T&K-S discuss cognitive neuropsychology and its presupposition of modularity, and Residual Normality as if it necessarily ignores development. However, a modular account does not preclude the study of development or a developmental dimension to the functional architecture of normal systems. Indeed, the multiple-route cognitive neuropsychological models, for example, of literacy (e.g., Coltheart et al. 2001), have enabled more flexible interpretation of development than traditional post-Piagetian stage-models with one-dimensional rigidity and invariant sequence (e.g., Frith 1985). Another example comes from Specific Language Impairment (SLI). Longitudinal studies of English (Rice 1999) and German SLI children (Rothweiler & Clahsen 1994) demonstrate selective delayed onset of verb-finiteness marking the same developmental trajectory as in normal children, and persistent selective delay in verb-finiteness marking into late childhood. Hence, impairments in onset do not necessarily lead to atypical developmental trajectories.

T&K-S's simulation demonstrates development that proceeds differently in a damaged single-mechanism system. However, in a single-mechanism system, disorder can only be interpreted as delay or distortion because it does not have any other mechanism able to continue to operate normally despite initial damage. The introduction of multiple-route connectionist models has been an attempt to respond to the limitations of unitary systems in explaining neuropsychological phenomena (e.g., Plaut et al. 1996). T&K-S claim that even in a dual-mechanism system Residual Normality is absent. Their second simulation tries to demonstrate that if one route is damaged, the other will not develop normally. However, this does not imply that the impaired route works improperly for its domain. For example, damage to the indirect-route system in their dual-route simulation does not produce impairments for regular past-tense forms, and the endstate performance on regulars is intact in this damaged system.

An alternative is multiple-component cognitive neuropsychological models with potential to account for both individual variation in developmental sequence of skill acquisition and varied developmental disorders. Their success in modelling dynamic development can be seen in modelling of changes in language associated with progressive dementia (e.g., Hodges et al. 1992). The system changes with development, although the underlying functional architecture remains the same. Just as components of the functional architecture may degrade with dementia, so, in development, components of the architecture may unfold over time and become more complex in relation to the representations upon which they can act. A genetic blueprint may unfold, with different modules coming into play or maturing of capabilities within modules over time. In developmental disorders, components may fail to unfold, as in face processing in WS, or the representations upon which they act may be reduced or distorted, as in irregularly inflected word forms in WS. The success of cognitive neuropsychology in assessing dynamic change is also seen in studies charting progress of theory-driven remediation upon development (e.g., Brunsdon et al. 2002; Cardell & Chenery 1999). This approach provides a constructive starting point and framework for the study of development; and in the delineation of patterns of performance and disorder seen across development, the extent of modularity itself within development can be properly assessed.

In any developmental analysis, a systematic approach is needed to define the states that change. Combined with a longitudinal perspective, cognitive neuropsychology provides a method and framework for investigation and description of developmental change. T&K-S argue for the importance of development, but use empirical data averaged across the life span (e.g., Thomas et al. 2001) and provide no protocol with which to describe and measure empirical change in performance. There is no theoretical account of the mental representations that are changing within their postulated dynamic processes, and how these changes are to be quantified. Hence, from an empirical perspective, their model remains untestable.

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

The developmental cognitive neuroscience approach to the study of developmental disorders

Elise Temple

Department of Human Development, Cornell University, Ithaca, NY 14853.
et62@cornell.edu

Abstract: Functional magnetic resonance imaging studies of developmental disorders and normal cognition that include children are becoming increasingly common and represent part of a newly expanding field of developmental cognitive neuroscience. These studies have illustrated the importance of the *process of development* in understanding brain mechanisms underlying cognition and including children in the study of the etiology of developmental disorders.

Our current understanding of how the brain is organized owes a great deal to the study of brain lesions in adult patients. These studies have helped scientists gain insight into the brain mechanisms underlying language (Broca 1861), memory (Scoville & Milner 2000), attention (Posner et al. 1984), vision (Farah et al. 1989), and many other cognitive and perceptual systems. Thomas & Karmiloff-Smith (T&K-S) argue that it may not be appropriate to utilize the same adult patient model for the study of developmental disorders and the development of brain mechanisms for cognition. They make the case that it is imperative to take into account the *process of development* when studying developmental disorders. This argument has many implications, including the primary one emphasized by T&K-S, that one cannot equate developmental disorders with adult brain damage even if the symptoms are similar. Another implication of this argument is that it may be inappropriate to study the etiology of developmental disorders by studying adults who have the disorder.

Most functional brain imaging studies of the etiology of developmental disorders use adult subjects. In the early days of brain imaging, when positron emission tomography (PET) was the only choice for studying functional localization in humans' brains, this was a necessary experimental choice. PET requires the injection of radioactive isotopes and cannot be utilized in minors except in very specific circumstances. However, with the advent of functional magnetic resonance imaging (fMRI), a noninvasive technique for studying brain function and functional localization became available that can be used in children – as long as they are able to keep their movement restricted. Even with the advent of fMRI, many studies on developmental disorders have continued to use adult subjects. This practice reflects an assumption that the process of development is not important, and implies that the disrupted brain process associated with the disorder is isolated from and does not affect the rest of the brain.

fMRI studies of developmental disorders and normal cognition that include children are becoming increasingly common and represent part of a newly expanding field of developmental cognitive neuroscience. These studies have already illustrated the importance of studying children as opposed to adults. Some results have confirmed previous adult findings, but others have shown differences between the adult and child organization. For example, a study of amygdala response to fearful faces in children (Thomas et al. 2001) showed less amygdala response in children as compared to adults; interestingly, this was because of an increased response in the children to the neutral faces. In a developmental study of inhibitory control (Luna et al. 2001), differences were found not only between children and adults but also in adolescents. Some of the brain differences shown in adolescents were different from both the children and the adults, illustrating the need to study the whole developmental trajectory (even in the relatively simple task of controlling eye movements). In addition, a study of cognitive control (Bunge et al. 2002) showed differences between adults and children in the brain regions associated with effective cognitive control, suggesting that the frontal network adults use to suppress interference and inhibit responses is not fully developed in children.

In our own studies of developmental dyslexia, my colleagues and I have found both similarities and differences between children with dyslexia and adult studies of dyslexia. For example, we found that children with developmental dyslexia showed decreased activity in left hemisphere posterior language areas where adults with dyslexia had shown decreases in previous studies (Temple et al. 2001; 2003). This decrease in left temporo-parietal cortex has been shown in other studies of children (Shaywitz et al. 2002; Simos et al. 2002), suggesting that this disruption in temporo-parietal response seen in adults with dyslexia may be fundamental to the disorder (Temple 2002). However, in our studies we have also seen differences in the functional brain organization of children with dyslexia. We examined the brain response to rapid auditory stimuli in adults with dyslexia (Temple et al. 2000) and found that whereas normal reading adults showed left prefrontal responsivity to rapid auditory stimuli, the dyslexic reading adults did not show left prefrontal response to the same stimuli. In addition, two of three adult subjects who participated in a training program to improve their reading and rapid auditory processing ability showed increases in left prefrontal cortex after training.

When we studied the brain response to rapid auditory processing in normal and dyslexic reading children (Temple 2001), we found some similarities to and differences from the adult study. In normally reading children, left prefrontal cortex was responsive to rapid auditory stimuli, but other brain regions were also involved in a network of response that was more distributed than we had seen in adults. In children with dyslexia, we saw the same lack of prefrontal responsivity to rapid auditory stimuli that we had seen in adults, but we also saw lack of response in the whole larger network of brain areas responsive in normally reading children. After training we saw increased activity in not only the left prefrontal cortex of the children with dyslexia but across the larger network.

These results suggest to us that the neural response to rapid auditory stimuli undergoes developmental changes from childhood to adulthood, becoming more focused and involving an increasingly smaller network. In addition, this study suggests that this larger network is disrupted in developmental dyslexia and the results seen in adults with dyslexia represent only one aspect of the disrupted response seen in children with dyslexia.

These examples of recent studies of normal and abnormal developmental cognitive neuroscience illustrate the point that the brain mechanisms involved in cognitive processes do undergo developmental changes. In the case of developmental disorders, it may be that entire networks are disrupted. Taking these processes of development into account may be crucial to our understanding of the etiology of developmental disorders. I would argue that, when feasible, studies of the neurobiology underlying developmental disorders as well as normal cognitive functioning should be conducted with children and taking the process of development into account.

Models of atypical development must also be models of normal development

Gert Westermann and Denis Mareschal

Centre for Brain and Cognitive Development, School of Psychology, Birkbeck College, London WC1E 7HX, United Kingdom.
{g.westermann; d.mareschal}@bbk.ac.uk
http://www.cbcd.bbk.ac.uk/people/gert; denis/

Abstract: Connectionist models aiming to reveal the mechanisms of atypical development must in their undamaged form constitute plausible models of normal development and follow a developmental trajectory that matches empirical data. Constructivist models that adapt their structure to the learning task satisfy this demand. They are therefore more informative in the study of atypical development than the static models employed by Thomas & Karmiloff-Smith (T&K-S).

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

As demonstrated here by Thomas & Karmiloff-Smith (T&K-S), neural network models are a useful tool for assessing how damage at different stages of development affects the outcome of learning. However, if the specific results gained with models are to inform studies of atypical development in children, it is important that the models used to simulate atypical development should otherwise also constitute realistic models of normal development. Only if a model can give a good account of how development proceeds under normal circumstances, will damaging this model yield insights into the mechanisms of atypical development.

Both in reading and in past-tense acquisition, a plethora of data exists about the normal course of acquisition and the resulting adult processing system. Pervasive here are dissociations between regular and irregular forms in past-tense learning (Marcus et al. 1992), and between normal words and exception words in reading (Backman et al. 1984). It is important to establish that these developmental dissociations also exist in a model that is later used to explain atypical development by varying its parameters. This is not because the observed developmental profile is significant per se, but because it reveals important aspects of the underlying learning mechanisms, such as the internal reorganization of representations or the extraction of higher-order features from the learned data. Even if a model of learning that does not follow a realistic developmental course reaches a 100% success rate, it is unclear if it does so in ways that are comparable with children's learning, and hence, if damage to the model is comparable to damage to the child's learning system, that is, the brain. The analogy between atypical development and learning in a damaged model can thus only hold if the undamaged model can be shown to progress through a realistic course of acquisition, and preferably if the trained model displays dissociations comparable to those in adult subjects, for example, in psycholinguistic experiments or in acquired brain damage.

A class of models of cognitive development that have been shown to possess these properties is constructivist neural networks (Mareschal & Shultz 1996; Shultz et al. 1995; Westermann 2000). These are models that adapt their architecture by adding (and sometimes deleting) units and connections during the learning process in an experience-dependent way. As a result, based on interactions between initial constraints, experience with the environment, and the constructivist learning algorithm, these models end up with an architecture that is specifically adapted to the learning task. The motivation for this class of models comes from developmental theories that argue for a progressive increase in representational power (Mareschal & Shultz 1996), from learning theoretic considerations that show a structurally developing system to be fundamentally different from a static one (Quartz 1993), and from neurobiological evidence about experience-dependent brain development in infants and children (Johnson 1997; Quartz & Sejnowski 1997; Westermann 2000).

An important aspect of development that constructivist models can capture is the reorganization of representations and processing during development. As such they have been shown to model developmental stages in Piagetian tasks (for an overview, see Shultz et al. 1995) and functional specialization in the acquisition of verb inflections (Westermann 1998).

A reason why constructivist models are especially suited to modeling atypical development is that apart from the constraint alterations employed in the models discussed by T&K-S, here the developmental mechanism itself can be varied. For example, often in these models newly added structure takes on a specialized function, either to encode higher-level features or to process exceptional or difficult items. A variation of the rate at which new structure can be added would then be expected to lead to significant variations in the developmental outcome. Further, this type of variations is especially close to the effects of genetically based developmental disorders. For example, Williams syndrome is characterized by distinct deviations from normal brain structure, such as decreased cerebral volume, especially in the right occipital lobe (Reiss et al. 2000).

A specific constructivist model that has been employed in a task used by T&K-S is the CNN (Constructivist Neural Network) model (Westermann 1999; 2000). This model learned to inflect the German participle and, in doing so, it developed regions of specialization for regular and irregular participles, respectively. These regions could be selectively lesioned after training, leading to double dissociations that closely resembled those observed in subjects with acquired agrammatic aphasia (Penke et al. 1999). A comparison between a version of the model that developed its architecture based on experience with the learning task, and a static model that started out with the full architecture, showed that the developing but not the static model passed through a learning process that exhibited the same characteristic errors as children, and the developing model showed a clearer functional specialization in its subregions than the static model (Westermann 2000).

Taken together, these points suggest that constructivist models could give additional and more informative insights into the mechanisms of atypical development than the static models employed by T&K-S.

ACKNOWLEDGMENT

The writing of this commentary was supported by European Commission RTN grant HPRN-CT-2000-00065 to Gert Westermann.

Authors' Response

Residual normality: Friend or foe?

Michael Thomas[a] and Annette Karmiloff-Smith[b]

[a]*School of Psychology, Birkbeck College, University of London, London WC1E 7HX, United Kingdom;* [b]*Neurocognitive Development Unit, Institute of Child Health, London WC1N 1EH, United Kingdom.* **m.thomas@bbk.ac.uk**
a.karmiloff-smith@ich.ucl.ac.uk
http://www.ich.ucl.ac.uk/units/ncdu/NDU_homepage.htm

Abstract: In response to our target article, many of the commentators concentrated on our notion of *Residual Normality.* In our response, we focus on the questions raised by this idea. However, we also examine broader issues concerning the importance of incorporating a realistic theory of the process of development into explanations of developmental deficits.

One of the central theses of our target article and of our empirical research over the past 15 years is that one cannot take for granted that atypically developing brains present with a functional cerebral organisation composed of parts intact and parts impaired. We now refer to the idea that atypically developing brains are a selectively damaged version of the normal system, as the hypothesis of *Residual Normality* or RN. In our view, whether one challenges this hypothesis or blindly accepts its viability will exert a strong influence on the whole of one's empirical research programme. As a result, we view the issue of residual normality as central to the study of developmental disorders. To resummarise our position, (1) we are neutral as to whether RN is appropriate to any particular developmental disorder or cognitive domain within that disorder; (2) given what is known about the lack of specificity of genetic effects in brain development, and about the interactivity and compensation that characterise subsequent processes of cognitive development, we think that RN is an unlikely (but not impossible) hypothesis; the narrower and more domain-specific the proposed structural dissociation is, the less

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

likely we think it is; (3) given that RN is unlikely, it follows that convincing (developmental) empirical evidence must be presented to justify its use in explaining a developmental deficit.

We are therefore heartened to note that many of the commentaries on our target article are sympathetic to our arguments (e.g., **Christiansen et al.**, **Ferrari**, **Gerrans**, **Jackson & Coltheart**, **Juola**, **Levine et al.**, **Levy**, **Li & Lindenberger**, **McClelland & Lupyan**, **Müller**, **Munakata et al.**, **Pléh et al.**, **Tallal**, **E. Temple**, **Westermann & Mareschal**). A few of the remaining commentators disagree hotly with our thesis (e.g., **Friedmann & Gvion**, **Ramus**, **C. Temple & Clahsen**), whereas some of the other responses to the issue of RN stem, in our view, from a misinterpretation of our original statements (e.g., **Briscoe**, **Condray & Steinhauer**, **Cooper**, **Marcus**). We welcome the opportunity to clarify these issues and to respond to our critics on the many questions that our target article generated.

The rest of this response proceeds as follows. We first address the issue of an appropriate characterisation of adult cognitive neuropsychology. We then address commentaries on the theoretical and empirical grounding of RN, and its effect on data collection. Next, we consider claims that there is little evidence that developmental disorders ever produce atypical functional structure. We then deal with commentaries focusing on Williams syndrome (WS), before finishing with issues surrounding the use of computational modelling in theory development.

R1. Adult neuropsychology – a caricature?

In the target article, we suggested that the tendency to view atypically developing systems in terms of static modular systems was in part a consequence of attempting to extend uncritically the dissociation methodology (and models) of adult cognitive neuropsychology into the developmental realm. However, as **Levy** correctly points out, we used the terms "adult neuropsychology" and "cognitive neuropsychology" rather loosely in our endeavour to stress this contention. In doing so, we erroneously put into a single basket both the box-and-arrow approach of some adult neuropsychologists who argue for independently functioning modules (e.g., **C. Temple & Clahsen**), and the approach of those whose research focus is on the dynamics, interconnectivity, and plasticity of the adult brain (**Briscoe**, **Juola**, **Li & Lindenberger**, **Tallal**). But we disagree that cognitive neuropsychology is theory neutral (Levy) because the majority of research practice is based on the modularity assumption and the double dissociation method, as was stressed in several commentaries (particularly **McClelland & Lupyan**). Moreover, it is true that the term "endstate" is really a misnomer because cerebral development and plasticity continue under certain constraints throughout life (Briscoe, Li & Lindenberger, **Munakata et al.**). It nonetheless remains our belief that a substantial part of the research in adult neuropsychology is theoretically underpinned by a static view of a modular brain, and that this model has permeated much of the research on developmental disorders (Baron-Cohen 1998; Pinker 1999; Temple 1997). We fully agree that a two-way dialogue must exist between development models and adult endstate models (Li & Lindenberger, Munakata et al.) because, as Piaget always argued, what has to be attained by the child to reach the adult state is deeply affected by the way in which it is acquired over developmental time.

Does that mean that we think that adult models have no place in developmental studies? No. Clearly mature performance can be used to generate quite precise hypotheses about what the infant startstate might be like (**Jackson & Coltheart**, **Ramus**, **C. Temple & Clahsen**). However, these must always be treated as empirically testable hypotheses, not theoretical assumptions!

R2. Residual normality – theoretically grounded?

We agree that RN is more likely (but not guaranteed) to hold in cases of adult brain injury. This is because in the adult's previously normal brain, specialisation and localisation of function had already stabilised, so selective impairments might emerge.

What about the developmental case? A careful reading of our target article shows that, contrary to the interpretation by **Marcus** and **Friedmann & Gvion**, we do not dismiss a priori the possibility of residual normality in developmental disorders (although we deem it to be unlikely). What we do stress is that RN must be demonstrated empirically and must be couched within an explicit developmental theory that provides an account of how RN could hold, rather than merely assuming it. When we read (**C. Temple & Clahsen**) that "a genetic blueprint may unfold, with different modules coming into play or maturing of capabilities within modules over time. In developmental disorders, components may fail to unfold," it is unclear to us what "unfolding" means precisely, how it is timed (rigidly prespecified genetically?), or what mechanisms might cause an unfolding failure. In our view, none of the recent progress in understanding brain development (e.g., Anderson et al. 2001b; Bradshaw 2001; Huttenlocher 2002; Johnson 1997; Nelson & Luciana 2001) points to the passive maturational unfolding of the neural substrates of cognition. Moreover, as we show below, empirical studies of typically developing infants and toddlers also call into question the notion of passive maturational unfolding.

Bertenthal, Campos, and colleagues examined the relationship between the onset of self-locomotion and a subsequent developmental cascade affecting a wide range of both cognitive and social behaviours (Anderson et al. 2001a; Bertenthal et al. 1994; Campos et al. 1992). At approximately 8–10 months of age, typically developing infants begin to move about independently, by crawling, bottom-shuffling, and so on. The researchers studied two groups of 8.5-month-old infants: a locomotor group and a prelocomotor group. They found significant differences between these two groups across a range of psychological functions, from social referencing to solving the famous A/not B task. This suggests that the onset of self-locomotion is accompanied by a cascade of cognitive-social development. It could merely be argued that the same neural substrate that mediates the cognitive-social cascade also supports the onset of the motor skills required for self-locomotion, and that the two share a similar maturational timetable. However, the researchers also studied a third group of infants who, like the prelocomotor group above, had not yet begun to self-locomote, but who were provided with a "prosthesis,"

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

a manmade walker. It turns out that this group of children displayed the same cognitive-social developmental cascade typically associated with the onset of self-locomotion. This indicates that it is the experience of self-locomotion, and not simply a genetically specified maturational timetable, that triggers the developmental cascade. The same researcher carried out a cross-cultural study in China of children in families where self-locomotion was restricted. In this case, the cognitive-social developmental cascade was delayed for several months. Similar findings have been documented for healthy preterm infants (Matthews et al. 1996). All of these data suggest that experience can affect the timing of major developmental milestones, and if this is the case for normal development, it is likely to be so for atypical development, as well. Thus, the issue of maturation needs to be couched in a far more interactional framework than suggested in some of the commentaries.

Müller argues that evidence for diffuse brain anomalies, such as abnormal growth patterns, cannot provide explanation for the very specific symptom profiles in autism. We disagree. If one takes a truly developmental perspective, then low-level diffuse anomalies can give rise to quite specific higher-level deficits. We gave several empirical and computational examples of this in our target article. In fact, Müller's own excellent discussion of the changing role of serotonin during development could be used as a good example. But we entirely agree that widespread deficits in an atypical system don't necessarily rule out RN (**Marcus**, **Ramus**). However, they do require detailed explanation.

Likewise, claims of selective deficits need to be examined with caution. For example, when a clinical population displays equivalent behaviour to a control group in domain A but not in B, claims of RN and a single dissociation often abound. However, if this control group is matched on, say, Mental Age, this can mean that the control children are many years younger than the participants with the genetic disorder. So a more accurate account of such results would be that in the clinical group *both* domains are very delayed (with B more delayed than A), not that A is "normal/intact/preserved" and B selectively impaired! To dismiss delayed performance as irrelevant carries the assumption that the representational processes under investigation do not interact with others throughout developmental time, that is, it simply assumes RN. When exploring an atypically developing *system,* it is vital in our view to understand the difference between simple delay versus complex deviance and, without knowledge of the developmental trajectory in each case, we cannot establish the true sources of *both* differences and similarities amongst developmental disorders (Karmiloff-Smith et al. 2003b).

It is important to reiterate that we do not rule out modularity a priori (**Marcus**, **Ramus**) and have never done so in the past (Karmiloff-Smith 1992). Rather, we make two claims: (1) that, when interpreting empirical data, one must always keep in mind that modularity may be the *product* of a developmental process rather than its starting point, and (2) that one can never take modularity for granted, it must be demonstrated. We certainly do not think (as Marcus suggests) that the only alternative to modularity is "entirely shared resources" (see, e.g., Elman et al. 1996, as well as Karmiloff-Smith 1998 on how domain-relevant biases might *become* domain-specific over developmental time). However, we do agree with Marcus that disorders that affect multiple domains may be less informative about issues of modularity. Yet, when researchers did keep their "eyes open for deficits that are more focused," as Marcus suggests, they ended up making sweeping claims about innate modularity in, for example, Williams syndrome (e.g., Bellugi et al. 1994; Pinker 1999) that simply have not stood the test of time. The same has obtained for genetic specification. Early claims that specific genes directly contributed to specific cognitive outcomes in WS (Frangiskakis et al. 1996) were shown to be premature because of the complex interactions between genes (Karmiloff-Smith et al. 2003a; Tassabehji et al. 1999). So, the claim that a gene has purely domain-specific consequences from the observable deficit that it seems to cause in a single part of the system should be treated with great caution, as the following example nicely illustrates.

A report in the press heralded the discovery of a specific "gene for hearing." The article on which it was based, however, illustrates not only how indirect the effects of the gene are, but also how a gene can be widely expressed in an organism and yet only a single deficit will be observable (Lynch et al. 1997). Geneticists studying eight generations of a Costa Rican family found a 50% incidence of acquired deafness, with onset around age 10 and complete deafness by age 30. A single gene mutation was identified, with the last 52 amino acids in the gene's protein product misformed, and the first 1,213 amino acids formed correctly. This gene produces a protein that controls the assembly of actin. Actin organises the tiny fibres found in cell plasma that determine a cell's structural properties, such as rigidity. Because the genetic impairment is tiny and the protein functions sufficiently well to control the assembly of actin in most parts of the body, no other deficits are observable. However, it turns out that hair cells are especially sensitive to loss of rigidity, such that even this tiny impairment has a huge effect on them. The loss of rigidity in hair cells eventually results in deafness. In other words, what might look like a specialised gene for a complex trait like hearing is, on closer examination, very indirect: hearing is dependent on the interaction of huge numbers of genes, one of which affects the rigidity of hair cells and has cascading effects on the others. And this same gene is expressed in multiple areas of the organism, despite the fact that this is not obvious from the effects of this particular mutation. A "gene for hearing" might be a convenient shorthand but, as the work of Lynch and collaborators shows, it is a very misleading one, impeding a researcher from seeking to understand the probabilistic dynamics of the developmental outcome (Karmiloff-Smith 1998). We believe that a similar argument should be explored with respect to the acclaimed FOXP2 "gene for speech and language" (Lai et al. 2001; see Pinker 2001 commentary, heralding the decade of cognitive genetics).

R3. Residual normality – empirically grounded?

Some commentators argue that there is considerable evidence for selective deficits alongside intact modules, and no empirical evidence for the lack of RN (**Ramus**, **C. Temple & Clahsen**). First, it is clear that the notion of "intact" performance always depends on the sensitivity of the measurement scale. So, even the claim for RN in adults should be considered with caution (**McClelland & Lupyan**). Second, as we have argued repeatedly, all the new work on

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

brain development suggests that RN is unlikely, although clearly not impossible (see also **Müller**). Ramus claims that developmental dyslexia can exist without sensorimotor deficit. He provides the example of a subset of adult developmental dyslexics with a pure (i.e., specific) phonological deficit, which is heritable and seems to have a genetic cause (in terms of localised disorders of neural migration). Leaving aside for the moment the substantial theoretical leap from foetal brain ectopias to adult dyslexia, we welcome this example of an attempt to build an empirically testable developmental account to support a claim for RN. We believe that this interesting theory could gain a lot from emulating the detailed developmental work on FragileX syndrome, which does not leap from synapse to cognition (Cornish 2003; Cornish et al. 1998; Dykens et al. 1989; Fisch et al. 1999; Scerif et al., in press).

Ramus's work is precisely the type of endeavour we argue for in our target article, that is, that RN does not come for free; it must be demonstrated empirically and couched in a developmental perspective. We were interested, for example, that Ramus rejected RN between components of the reading system (in his footnote 1 he writes: "It is now quite clear to most dyslexia researchers that the phonological and orthographic reading routes cannot really develop independently, and, therefore, that selective damage to one of them is unlikely in developmental dyslexia") but that he argued for its plausibility at the broader level of phonological representations. This reflects our intuition that RN becomes a more plausible starting hypothesis for broader functional distinctions.

However, the **Ramus** example clearly requires further exploration. Explaining adult dyslexia through foetal brain anomalies without charting its developmental trajectory through to reading leaves aside many vital issues. For example, the account cannot generate a ready explanation for the high co-morbidity of dyslexia with dyscalculia (Ansari & Karmiloff-Smith 2002) without the additional assumption that spontaneous ectopias also occur in parietal cortex. Then the deficit becomes less specific, because it is possible that ectopias occur throughout such brains but only display their effects on some resulting cognitive domains. But these are empirically testable questions, a fact that we value.

Moreover, in our view there remain problems with **Ramus**'s approach. First, he appears to accept animal models uncritically, yet they present similar perils to the assumptions involved in going directly from adult to infant cognition (Karmiloff-Smith et al. 2002). Second, there are some methodological difficulties with taking subgroups of (behaviourally defined) dyslexics to justify pure deficits. It could be that this behavioural screen picks up just those individuals who happen to have above average visual or motor skills that allow them to compensate for the deficits that come with a wider view of developmental dyslexia, so that their phonological deficit is the only one to surface in the data. And there are indeed empirical and brain-imaging data showing that interindividual variation enables differing levels of compensation within developmental disorders (**E. Temple**; and see Thomas 2003).

To reiterate, third, and perhaps most seriously, **Ramus** leaps directly from foetal brain ectopias to adult reading deficits. How do ectopias affect the functional parameters of the developing system? Ramus presents dyslexia as a case of RN, yet it is merely assumed on the basis of the characteristics of the adult endstate that zero compensation has taken place. Although it may be possible to probe for the function of the history of compensation in the adult (**Jackson & Coltheart**), and even if the precise details of **Tallal**'s theory have been questioned, we strongly applaud her pioneering work using the prospective, longitudinal study of infants at risk to seek low-level impairments that cascade developmentally on the resulting high-level cognitive outcomes. This, in our view, is the way in which the Ramus enterprise should be taken forward, across a representative sample of at-risk infants, rather than a retrospective screening for selective deficits in adults.

R4. Effects of the RN assumption on empirical research programmes

How does one plan a research programme in the developmental field? For studies of normal development, researchers usually focus on age groups in which the domain of interest is still not acquired and trace its progressive acquisition until mastery. In atypical development, the main trend has been to concentrate on areas of deficit, with summary measures of areas of proficiency. Our research strategy has always been very different. In the case of normal development, we focused on age groups who already seemed to have behaviourally mastered a particular domain and then traced subsequent representational change (Karmiloff-Smith 1992). In the atypical domain, we took the unusual step of initially focusing on domains of reported proficiency (Karmiloff-Smith 1998), such as language and face processing in Williams syndrome. This was because of our discovery that in normal development, identical behaviour at two different ages could be sustained by very different underlying representations. We thus challenged the RN implications of merely accepting "scores in the normal range" and the use of IQ scores to undermine the fact that an adult has relatively fluent language (Karmiloff-Smith 1998).

By way of illustration, if one reported that an adult with an IQ of 58 has good language, this might sound surprising and interesting to the reader, suggestive that language is dissociated from general intelligence. If, however, one rephrased this and reported that an adult with a Mental Age of 9 years has good language, the claim seems far less impressive and less immediately suggestive of a modular dissociation – normal children age 9 have excellent language. A focus on IQ scores skips over the fact that IQ calculations are affected by chronological age, and that if an individual plateaus as a result of his or her disorder, then IQ will be lower, the older the chronological age.

Thus, one of the aims of our target article was to demonstrate the way in which the assumption of residual normality influences empirical research. If one has a proclivity to accept RN as a likely hypothesis, then the discovery of scores in the normal range in a particular domain is sufficient to support claims for selective preservation of certain functions, and that part of the research can be concluded. If, however, one questions the automatic assumption of RN, then one is motivated to pursue further probes as to the underlying cognitive processes sustaining the behaviour.

One commentary illustrates the perils of uncritically accepting RN. Take, first, the example of face processing in

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Williams syndrome. We claimed that WS face processing was atypical because, despite scores in the normal range, this clinical population tends to use predominantly componential processing, whereas normal controls use configural processing. **C. Temple & Clahsen** retort that these results merely illustrate a case of "*straightforward* modular dissociation within face processing systems" (emphasis added) between intact componential and impaired configural processing, implying that there is nothing more to be said. By contrast, our studies have shown that when individuals with WS process stimuli componentially, their brains are activated differently from controls. More interestingly still, when we carried out a cross-syndrome comparison of two clinical groups in which a predominance of componential processing has long been documented (autism and Williams syndrome), the brain functioning of these two populations was radically different again (Grice et al. 2001). So, the descriptor "componential" turns out to be nothing more than that – still a description of behaviour. It has at least three different meanings with respect to different populations, suggesting that further empirical probing is crucial. There is nothing "straightforward" about it! Had we simply accepted the original "normal" face processing scores as evidence for RN, we would never have gone on to explore these important within-syndrome and cross-syndrome differences.

A hair-trigger acceptance of RN also tends to nullify the need to trace developmental trajectories. By contrast, we deem it vital to trace developmental disorders back to their origins in infancy. This has alerted us to the importance of understanding how lower-level deficits can impact on higher-level cognitive outcomes. For example, in the visual domain, our studies have shown that simple eye movement saccade planning is very atypical in infants and toddlers with Williams syndrome, whereas it is sluggish in those with Down syndrome (Brown et al., in press). Such work has revealed a series of deficits in early development, giving rise to atypical learning trajectories. This has turned out to obtain even in domains where the adult outcome is rather proficient, particularly language development in Williams syndrome, for example, for which sweeping claims for intact modules have been made (e.g., Pinker 1999). Our studies have revealed considerable deviance in how infants with WS *learn* language, including deficits in early speech segmentation, in triadic attention, in word-referent mapping, and in brain lateralisation (Laing et al. 2002; Nazzi & Karmiloff-Smith 2002; Nazzi et al. 2003; see also, Singer Harris et al. 1997). Such deviant learning processes obtain in older people with WS, as well. In the domain of reading, and despite careful individual matching on word reading levels, our studies showed that the process by which the clinical group learned to read new words differed from controls (Laing et al. 2001). All of this work underlines the danger of simply accepting the RN assumption.

R5. A radically different brain?

Some commentators thought we were claiming that the atypical brain must be fundamentally different from the normal brain (**C. Temple & Clahsen**). This was not the case, because obviously at some level of description all human brains share some very general principles. What we do claim is that it is very unlikely that the specific learning mechanisms are the same in the normal and atypical brains. And we continue to challenge the view that the atypical brain is merely a normal brain with parts intact and parts impaired. In our view, as genotyping and phenotyping methodologies improve, subtle and often widespread differences will be identified in the brains of children with developmental disorders, without necessarily giving rise to overt deficits in behaviour (viz., our earlier example of the differential effects of an actin deficit; see Karmiloff-Smith & Thomas, in press). Where to draw the line between subtle and radical differences in brain structure may be a matter of taste, and, as **E. Temple** has stressed, differences between normal and atypical brain function often lie at the level of interacting networks rather than independently functioning modules (C. Temple & Clahsen). We maintain that the brain must be considered as a whole, interacting organism within a developmental framework, and that, if RN were to obtain, it must be explained within that development perspective.

We agree, however, with **Li & Lindenberger** that it may well be unproductive to draw a monolithic divide between normality and abnormality, because variations within and between individuals are fundamental to living organisms. Although normality may well be better viewed as the potential of individuals and environments to produce an adaptive range of behaviours, as they submit, there are of course constraints on that process, and it is these constraints and how they might differ in atypical development that is of concern to us here. In challenging the blind assumption of RN, we believe that there are many indications in the literature that atypical brains do not present with a neat division between intact and impaired parts, and that there is indeed evidence for atypical structure. For example, the brains of older children and adults with Down syndrome are more active than normal brains, due perhaps to a failure to modularise during development (i.e., a lack of progressive specialisation and localisation). It has also been argued that autism may be characterised by a process of rapid modularisation such that circuits become dedicated too rapidly to overly specialised functions (Oliver et al. 2000).

FragileX (FraX) also presents with an atypical brain in which development plays a crucial role. The silencing of the FMR1 gene, whose associated protein plays a crucial role in activity-dependent synaptic plasticity, has wide-spread effects on neurone structure and function across the cortex (Churchill et al. 2002). Despite these generalised effects, the brain in FraX is characterised by decreased volumes of the cerebellum and enlarged hippocampus. Yet, these regional differences cannot be simply accounted by regional differences in FMR1 expression, suggesting a crucial role of development in determining the adult outcome. Early and continuously treated phenylketonuria (PKU) is characterised by specific effects on the dopaminergic content of dorso-lateral prefrontal cortex (Diamond 2001; Diamond et al. 1997). However, differences in the developmental timing of the exposure to elevated phenylalanine result in subtly different effects on the executive profile in PKU (Antshel & Waisbren, in press; Koch et al. 1999). In all of these cases, ontogenetic development has played a crucial and necessary role, and the resulting brain is not "merely lacking or having reduced development of components of normal functional architecture" (**C. Temple & Clahsen**). In our view, developmental disorders are potentially very informative about normal development, not in terms of

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

which modules are intact or impaired, but in terms of constraints on plasticity of a developing brain (Huttenlocher 2002; Thomas 2003).

Some theorists, however, might choose to dissociate atypicalities in the neural substrate from functional structure assessed at a behavioural level. Following this line of argument, if behavioural data in a disorder appear normal, then underlying abnormalities in brain structure or processes cannot have functional significance. Take the case of a young child who had suffered damage to the left hemisphere, and subsequently shown a recovery of language into the normal range via reorganising cerebral function to the right hemisphere. A proponent of this position would argue that a normal functional structure for language processing had emerged. If one ignores evidence from neuroscience, what evidence is there that atypical modular structure is to be found in developmental disorders? There is a methodological problem here, in that standardised test batteries are designed only to assess the presence or absence of normal functions. The requirement would seem to be for genuinely novel behaviours or genuinely novel error patterns. We do believe such evidence exists, for example, in date calculation in autism, in the cross-modality experiences of individuals with synaesthesia, and in the errors made by children with WS during language acquisition (Capirci et al. 1996). Indeed, evidence of "crowding effects" in the later development of children with significant early damage to one hemisphere (Anderson et al. 2001b) belie the claim that all atypical development can be couched as graded acquisition of a normal module set (**Levine et al.**).

R6. Williams syndrome – myths and facts

We already discussed above the issue of face processing in individuals with Williams syndrome, but it is their proficient language that has tantalised many thinkers, leading to claims about modularity and genetic specification (e.g., Pinker 1999). However, as **Pléh et al.** convincingly show, finer-grained analyses and cross-linguistic constraints challenge many of the assumptions about WS language.

In our target article we raised questions about the claims of Clahsen and Almazan (1998) regarding the purported selective impairment in irregular inflection in WS. There are, in our view, various problems with the Essex University data and the interpretation thereof. First, unfortunately the normal data against which the WS data were contrasted do not replicate, even by researchers of a similar theoretical persuasion (van der Lely & Ullman 2001; and see Thomas et al. 2001). Could it be that the normal Essex data were mistakenly entered into the wrong column? We set out to replicate and extend Clahsen and colleagues' original finding on the four children with WS and were surprised when the basic effect was not present in our much larger population. If an effect is robust with four participants, it should have been more than evident with 21 participants. Indeed, we have concerns about the representativeness of the four participants with WS used in most of the Essex studies, because five of the findings based on this group (on past tense formation, pluralisation, noun compounding, semantic fluency, and speeded naming) have not replicated with other larger samples of individuals with WS (past tense: Thomas et al. 2001; Zukowski 2001; pluralisation and compounding: Zukowski 2001; semantic fluency: Jarrold et al. 2000; Scott et al. 1995; speeded naming: Thomas et al. 2003). Second, our data for WS point to deficits in *both* the irregular and the regular inflections, once language level is partialled out, suggesting no selective deficit (even Clahsen and Temple's [2003] reanalysis of our data reveals deficits for both regular and irregular inflections compared to controls, with no evidence of an interaction, although their focus is restricted to the irregular deficit). Third, even if the Essex data for the WS and controls were to hold (and we welcome more data on this point), the interpretation they offer lacks a developmental perspective. We of course agree with **C. Temple & Clahsen** and with **Tallal** that longitudinal data can be more informative than cross-sectional data. But in our view, cross-sectional data can be treated within a developmental framework by building developmental trajectories (Thomas et al. 2001; see also Ansari et al., in press).

Perhaps the main difference between our view and that of **C. Temple & Clahsen** in relation to this particular debate revolves around our preference for an explanation focusing on the developmental process itself rather than in terms of deficits to static structures (a specific impairment in accessing subnodes of lexical entries: Clahsen & Almazan 1998). Thus, in recent, detailed computational modelling of the past tense formation in WS (Thomas & Karmiloff-Smith, in press), we have shown that were a selective deficit in irregular inflection to be characteristic of some individuals with WS, this could be explained as the developmental consequence of various atypicalities in lexical-semantics. This prediction then prompts further empirical work exploring the development of lexical-semantics in this population. This brings us to our final section.

R7. Computational models as a tool exploring RN

A number of commentators raise issues about the modelling work in our target article, either with regard to specifics of the models employed (**Cooper**, **Juola**, **Levine et al.**, **Levy**, **McClelland & Lupyan**, **Müller**, **Westermann & Mareschal**), or to the general enterprise of using computational models to study developmental disorders (**Christiansen et al.**, **Gerrans, McClelland & Lupyan**, **Munakata et al.**). To address these issues, and to clarify the role we see our particular simulations playing, we think it useful to draw a distinction between two different ways in which computational models can be used.

First, models can be employed at a fairly abstract level, where one only seeks to capture general characteristics of the problem domain. The aim here is to explore the patterns of behavioural data that can be generated by models embodying particular principles of processing. The end product is an expansion of the range of candidate inferences that can be drawn from patterns of human behavioural data to underlying structure. We will call this type of model *Abstract.* Abstract models tend to be useful in the earlier stages of theory development in a given field, because they are an engine for conceptual clarification.

Second, models can be used in an attempt to simulate detailed patterns of empirical data from a target cognitive domain, while incorporating as many empirically motivated psychological (and perhaps neural) constraints as possible. We will call this type of model *Specific.* The aim of Specific

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

models is to evaluate the viability of particular hypotheses and generate new testable predictions. Such models are characteristic of more theoretically developed fields of enquiry, which are supported by a rich body of empirical data. The simulations presented in this target article fall within the category of Abstract models. But we have also carried out simulations that explore in detail the acquisition of English past tense formation in Williams syndrome (Thomas & Karmiloff-Smith, in press), that fall within the category of Specific models.

With this distinction in mind, we can now address some of the modelling issues raised. First, we of course agree with **Munakata et al.** that Abstract models of developmental disorders are insufficient on their own, and that subsequent work needs to explore more direct mappings to particular neural mechanisms and particular developmental disorders (see also Thomas 2000). However, as we found in our work with modelling data from individuals with Williams syndrome, the translation of sometimes vague theoretical claims (e.g., that individuals with WS have "poorly specified semantic representations") into precise computational manipulations is not always straightforward. Importantly, the modelling process forces clarification in these instances.

Even if a Specific model is successful in its goal of simulation, there is still some way to go to assess the model's utility for understanding a disorder. **Christiansen et al.** provide an excellent set of criteria for such an assessment, that may in the future prove a benchmark against which models are compared. Christiansen et al. point out that even if a model is successful in simulating an atypical developmental trajectory following a parameter manipulation, there remains the question of whether this was the only possible parameter manipulation that would lead to a fit of the data. Many issues are raised by the proposal that parameter variations to "normal" computational models can explain atypical data, among them the possibility of multiple cognitive causes for behaviourally defined disorders and the relation of individual differences to atypical development. We do not have space to consider these issues here (see Thomas, in press; Thomas & Karmiloff-Smith 2002), save to note that Christiansen et al. raise an important point, and that in our modelling of WS we indeed went to some lengths to explore both the background flexibility of our model in capturing different possible patterns of developmental deficits, and the model's potential to capture individual variability as well as the group trend of atypical development.

Simulations 1 and 2 in our target article fall under the category of Abstract models. Our aim was to explore the kinds of inferences that might be made from behavioural deficits in developmental disorders to the underlying functional structure, as well as the developmental trajectory that may have led to the adult state. Furthermore, we wished to advance our own thinking about the kinds of developmental commitments that were being made in arguing for RN, that is, that one could strike out a single module of an adult system to explain a selective deficit in a developmental disorder. To do this, we utilised a widely employed modelling architecture, feedforward backpropagation networks.

Several commentators extend the ethos of our simulations. **McClelland & Lupyan** use our model to demonstrate that a single undifferentiated system could show a developmental double dissociation under different startstate manipulations. The model thus widens the possible inferences that might be made from behavioural double dissociation data, suggesting here that developmental double dissociations need not imply independent underlying structures. In the same spirit, **Gerrans** examines Pinker's claim that WS and SLI form a "genetic" double dissociation in the domain of inflectional morphology necessarily implies innate modularity (see Thomas & Karmiloff-Smith, in press, for further discussion of this issue).

Other commentators additionally focus on our particular choice of model type, and ask whether this choice limits the lessons that can be drawn concerning RN. It was suggested that the choice of the connectionist paradigm implies a rejection of modular accounts of development (**Levy**). Several commentators question the choice of backpropagation on the grounds of biological plausibility (**Cooper**, **Müller**), or on the grounds that in Simulation 2, it biased the system against showing RN because an error produced by damage in one route was necessarily fed back to drive changes into the other route (**Juola**, **Levine et al.**). The choice of feedforward backpropagation networks is also questioned on the basis that their relation to the neural substrate is unclear (**Müller**); that perhaps constructivist networks, with representational power that increases during training, are better models of development (**Westermann & Mareschal**); or simply that a wider range of model types needs to be considered before drawing conclusions about the likelihood that RN holds in the developing cognitive system (**Cooper**).

Before we respond to these comments, two specific points of clarification. **Levine et al.** suggest that in Simulation 1, different amounts of damage were applied early and late, making comparison difficult. This is incorrect. With the exception of startstate lesions, where performance was at ceiling for the range of lesions applied in the endstate, damage levels were *identical* across startstate and endstate. Levine et al. are correct that we did not include retraining after damage to the endstate (see also **Juola**). This is also true of many models of acquired deficits that we cited. This "frozen" endstate was indeed a simplification for the purposes at hand, but we fully agree that it is unrealistic. Our current work is exploring patterns of recovery, and, in particular, critical period effects in the progressive emergence of modularity. Second, **C. Temple & Clahsen** suggest that in Simulation 2, early damage to the Indirect route (normally preferentially involved in learning irregular patterns) had no effect on the learning of regular patterns (normally preferentially carried out in the Direct route). They therefore argue that the model shows RN. However, this is incorrect. They have not looked at the results closely enough. The greater the damage to the Indirect route, the more the Direct route has to take on the function of learning irregular patterns. This causes a subtle deficit to its normal acquisition of regular patterns. After a 60% initial lesion to the Indirect route, regular performance drops from 100% to 99%. After a 90% initial lesion to the Indirect route, regular performance drops to 95%. RN does not hold here, but the evidence for this is subtle. The model therefore nicely illustrates the point that issues of RN will not be settled without detailed examination of empirical data.

Now to the more general points. First, we need to be clear about why we chose feedforward backpropagation networks, and about the theoretical framework within which we are operating. Feedforward backpropagation networks were chosen in order to make contact with the large number of existing models of cognitive development that

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

use this architecture. To evaluate the effect of the developmental process on early damage, and to investigate the likelihood of RN holding in development, it seemed reasonable to start by using an architecture frequently used over the last fifteen years to model development. Similarly, we chose our parameter manipulations based on those widely employed in existing models (see sect. 7.1). In addition, we viewed feedforward networks as sufficient to help advance our thinking on the conditions under which RN might hold (which we laid out in sect. 9). Even if feedforward networks are not the right kind of architecture, we would know what features to seek in other architectures to ascertain whether they display RN.

With regard to the theoretical framework within which we are operating, the objective is to put the developmental process at the heart of explanations of developmental disorders (Karmiloff-Smith 1998). Given the current understanding of mechanisms of change in developmental psychology, we do not view this framework as currently committed to a particular learning mechanism. Nor is it committed to the idea that the startstate parameters (in which we include the learning algorithm as well as issues of architecture, representations, and so forth; see Karmiloff-Smith & Thomas 2003; Thomas & Karmiloff-Smith 2002) must be frozen throughout development. It is currently an open question whether parameters such as "plasticity," "memory capacity," or "representational power" change across development and what factors may drive such "constructivist" changes. The theoretical framework is, however, committed to the idea that differences in the computational constraints that shape development are the appropriate terms in which to explain developmental deficits, rather than in terms of a list of "intact" and "impaired" adult modules.

First, **Levy** claims that the use of connectionist models implies an a priori rejection of modularity. We must disagree with this. Connectionism per se implies no position on modularity; indeed many of the individual models put forward to explain cognitive development in particular domains could be viewed as cognitive modules. Furthermore, Simulation 2 in the target article explicitly examined the emergence of modular function in a connectionist network, albeit of a more graded form than in the classical definition.

Next, in response to the comment that the relationship between feedforward networks and brain processes is unclear (**Müller**), we would stress that our first aim is to make contact with the clinical data that capture deficits in developmental disorders, and that our focus is thus at the level of cognitive modelling. We would agree that a future aim should be to include more neural constraints in such models.

The use of backpropagation often raises questions regarding biological plausibility (**Cooper**). We take backpropagation as representative of gradient descent algorithms, themselves a method of achieving error-driven learning. O'Reilly (1998) argues that error-driven learning is one of two main types of learning occurring in the brain, the other being self-organisation/unsupervised learning. The brain learns abstract representations of the world and then learns the optimal mappings between them. We do not have a particular commitment to backpropagation but rather to error-driven learning. One of our current research issues is to explore the implications of developmental deficits in systems using other error-driven algorithms and in self-organising systems. Nevertheless, we believe that backpropagation has served a useful theoretical purpose in highlighting the issues at stake in explaining the causes of developmental deficits.

Levine et al. argue that one can show analytically that the dual-route system trained by backpropagation must fail to show RN. The analytical approach is an important source of clarity, although we have tended to find that the analytical account follows exploratory modelling work. For example, our current work is exploring the factors that tend to drive emergent specialisation in distributed systems, and the way in which plasticity, competition, structure-function correspondences, computational resources, and frequency all interact across training to determine whether a system produces specialisation or redundancy. We believe that an analytical account of this interaction is still some distance away at the current time.

Nevertheless, it is entirely valid to ask whether backpropagation has something unusual about it that makes it a poor choice to explore RN. This really concerns the wider issue of *generality.* Should not we really be using constructivist networks? ask **Westermann & Mareschal**. What of the range of other learning architectures available, exemplified in the domain of past tense formation? asks **Cooper**. What of symbolic learning algorithms? Would our conclusions concerning RN hold for these other models? As we stated above, use of these other developmental computational models would be quite consistent with our theoretical framework. But questions of generality are legitimate ones. Our modelling work with past tense formation in Williams syndrome forced us to address this issue (Thomas & Karmiloff-Smith, in press). Using a feedforward backpropagation network, we identified startstate parameters manipulations that would deflect the developmental trajectory of a normal model of past tense formation to capture the WS profile. We then considered whether these manipulations would have the same result in four other competing models, including a constructivist architecture and one based on ACT-R. As far as we could tell, the manipulations would have broadly the same effect – the results were not backpropagation specific. Moreover, several of the competing models had dual-route architectures. Our conclusion was that none of the architectures of these Specific models would show RN.

Let us be clear, however, that we do not think our modelling work somehow "disproves" RN. It explores the constraints that would need to be met for RN to obtain. Whether RN holds is ultimately an empirical issue to be resolved on a domain-by-domain basis (so long as the right sort of data is collected). Our modelling work has, however, led us to be sceptical concerning the viability of RN. As **Munakata et al.** spotted in our Notes 12 and 13, we reviewed a case where even a computational system in which RN seems likely suggested possible routes of compensation under closer analysis; and a case that demonstrated that achieving RN in a multi-component learning system is hard even if one deliberately tries to engineer it.

Lastly, **Jackson & Coltheart** embrace the utility of computational modelling, but have reservations about using developmental rather than static (hand-wired) models. They do not find it helpful to map hypothetical problems with the acquisition of knowledge (in terms of theories concerning processing deficits, impoverished environments, and so forth) onto manipulations of trainable connectionist mod-

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

els as a way of exploring possible accounts of developmental deficits. For them, "doing so requires leaps of faith" which they "prefer not to take." We find this a slightly strange position. The point of modelling is to build constraints into the model based on empirical evidence, and then evaluate the model's performance in capturing behavioural data patterns. If one is not building in constraints, then one might as well use a look-up table to simulate data. It is not clear to us why Jackson & Coltheart view the constraints that they build into static models such as the DRC (Coltheart et al. 2001) as somehow safe, but those based on dynamic principles as too risky. We view this position as overly conservative. It is true that sometimes it can be hard to derive a unique mapping from a psychological constraint to a particular feature of the model, but this difficulty is no greater for the developmental (or dynamic) case than for the adult (or static) case. Apart from the computational constraints we have investigated in the target article and elsewhere, we believe there is much to be gained from exploring the implications of, for example, different environments by altering input sets, or (in the case of reading), different types of explicit tuition by altering training regimes. One day static and developmental models must join together. This won't happen unless we begin building from both sides of the river.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.

Ahmed, A. & Dutton, G. N. (1996) Cognitive visual dysfunction in a child with cerebral damage. *Developmental Medicine and Child Neurology* 38:736–43. [CT]

Akshoomoff, N., Pierce, K. & Courchesne, E. (2002) The neurobiological basis of autism from a developmental perspective. *Development and Psychopathology* 14:613–34. [R-AM]

Anderson, D. I., Campos, J. J., Anderson, D. E., Thomas, T. D., Witherington, D. C., Uchiyama, I. & Barbu-Roth, M. A. (2001a) The flip side of perception-action coupling: Locomotor experience and the ontogeny of visual-postural coupling. *Human Movement Science* 20(4–5):461–87. [rMT]

Anderson, V., Northam, E., Hendy, J. & Wrennall, J. (2001b) *Developmental neuropsychology: A clinical approach.*. Psychology Press. [arMT]

Ansari, D., Donlan, C., Thomas, M. S. C., Ewing, S., Peen, T. & Karmiloff-Smith, A. (2003) How children with Williams syndrome understand why counting counts. *Journal of Experimental Child Psychology* 85:50–62 [rMT]

Ansari, D. & Karmiloff-Smith, A. (2002) A typical trajectories of number development. *Trends in the Cognitive Sciences* 6(12):511–16. [rMT]

Antshel, K. M. & Waisbren, S. E. (in press) Timing is everything: Executive functions in children exposed to elevated levels of phenylalanine. *Neuropsychology.* [rMT]

Bach, R. P. (1990) Brain plasticity as a basis for recovery of function in humans. *Neuropsychologia* 28:547–54. [YL]

Backman, J., Bruck, M., Hebert, M. & Seidenberg, M. (1984) Acquisition and use of spelling sound correspondences in reading. *Journal of Experimental Child Psychology* 38:114–33. [GW]

Baddeley, A. D. (2000) The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 11:417–23. [JB]

Baltes, P. B, Cornelius, S. W., Spiro, A., Nesselroade, J. R. & Willis, S. L. (1980) Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology* 6:625–35. [S-CL]

Baltes, P. B., Lindenberger, U. & Staudinger, U. M. (1998) Life-span theory in developmental psychology. In: *Handbook of child psychology, vol. 1: Theoretical models of human development,* ed. R. M. Lerner. [Editor-in-chief: W. Damon.] Wiley. [S-CL]

Baltes, P. B. & Singer, T. (2001) Plasticity and the aging mind: An exemplar of the biocultural orchestration of brain and behaviour. *European Review* 9:59–76. [S-CL]

Banich, M. T., Levine, S. C., Kim, H. & Huttenlocher, P. (1990) The effects of developmental factors on IQ in hemiplegic children. *Neuropsychologia* 28:35–47. [SCL]

Barbeau, D., Liang, J. J., Robitaille, Y., Quirion, R. & Srivastava, L. K. (1995) Decreased expression of the embryonic form of the neural cell adhesion molecule in schizophrenic brains. *Proceedings of the National Academy of Sciences USA* 92:2785–89. [RC]

Baron-Cohen, S. (1998) Modularity in developmental cognitive neuropsychology: Evidence from autism and Gilles de la Tourette syndrome. In: *Handbook of mental retardation and development,* ed. J. A. Burack, R. M. Hodapp & E. Zigler. Cambridge University Press. [arMT]

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A. & Williams, S. C. (1999) Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience* 11:1891–98. [R-AM]

Basser, L. S. (1962) Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. *Brain* 85:427–60. [SCL]

Bates, E., Friederici, A. & Wulfeck, B. (1987) Comprehension in aphasia: A cross-linguistic study. *Brain and Language* 32:19–37. [PJ]

Bates, E., Vicari, S. & Trauner, D. (1999) Neural mediation of language development: Perspectives from lesion studies of infants and children. In: *Neurodevelopmental disorders. Developmental cognitive neuroscience,* ed. H. Tager-Flusberg. MIT Press. [SCL]

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z. & St. George, M. (2000) The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience* 12(1):7–30. [CP]

Bellugi, U., Wang, P. & Jernigan, T. L. (1994) Williams syndrome: An unusual neuropsychological profile. In: *Atypical cognitive deficits in developmental disorders: Implications for brain function,* ed. S. Broman & J. Grafman. Erlbaum. [arMT]

Benasich, A. A. & Tallal, P. (1996) Auditory temporal processing thresholds, habituation, and recognition memory over the first year. *Infant Behavior and Development* 19:339–57. [PT]

(2002) Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research* 136:31–49. [PT]

Bertenthal, B. I., Campos, J. J. & Kermoian, R. (1994) An epigenetic perspective on the development of self-produced locomotion and its consequences. *Current Directions in Psychological Science* 3(5):140–45. [rMT]

Betancur, C., Corbex, M., Spielewoy, C., Philippe, A., Laplanche, J. L., Launay, J. M., Gillberg, C., Mouren-Simeoni, M. C., Hamon, M., Giros, B., Nosten-Bertrand, M. & Leboyer, M. (2002) Serotonin transporter gene polymorphisms and hyperserotonemia in autistic disorder. *Molecular Psychiatry* 7:67–71. [R-AM]

Bird, H., Lambon Ralph, M. A., Seidenberg, M. S., McClelland, J. L. & Patterson, K. (in press) Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language.* [aMT, JM]

Bishop, D. V. M. (1990) Handedness, clumsiness and developmental language disorders. *Neuropsychologia* 28:681–90. [aMT]

(1997a) Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *Quarterly Journal of Experimental Psychology* 50A:899–923. [aMT, YL]

(1997b) *Uncommon understanding: Development and disorders of language comprehension in children.* Psychology Press. [JB]

Bishop, D. V. M., Bishop, S. J., Bright, P., James, C., Delaney, T. & Tallal, P. (1999a) Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research* 42:155–68. [FR]

Bishop, D. V. M., Carlyon, R. P., Deeks, J. M. & Bishop, S. J. (1999b) Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language and Hearing Research* 42:1295–310. [aMT]

Blackwell, A. & Bates, E. (1995) Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience* 7:228–57. [PJ]

Bradshaw, J. L. (2001) *Developmental disorders of the frontostriatal system: Neuropsychological, neuropsychiatric, and evolutionary perspectives.* Psychology Press. [rMT]

Bray, N. W., Reilly, K. D. & Grupe, L. A. (1997) Neural network models and mechanisms of strategy development. *Developmental Review* 17:525–66. [MHC]

Bretherton, I. (1996) Internal working models of attachment relationships as related to resilient coping. In: *Development and vulnerability in close relationships,* ed. G. G. Noam & K. W. Fischer. Erlbaum. [MF]

Briscoe, J., Bishop, D. V. M. & Norbury, C. (2001a) Phonological processing, language and literacy: A comparison of children with mild to moderate hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 42. [JB]

Briscoe, J., Gathercole, S. E. & Marlow, N. (2001b) Everyday memory and

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

cognitive ability in children born very prematurely. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 42(6):749–54. [JB]
Broca, P. (1861) Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). *Bulletin de la Société Anatomique* 6:330–57. [ET]
Brown, G. D. A. (1997) Connectionism, phonology, reading, and regularity in developmental dyslexia. *Brain and Language* 59:207–35. [MHC, aMT]
Brown, G. D. A. & Loosemore, R. (1995) A computational approach to dyslexic reading and spelling. In: *Developmental and acquired dyslexia,* ed. C. K. Leong & R. M. Joshi. Kluwer Academic. [MHC]
Brown, J. H., Johnson, M. H., Paterson, S. J., Gilmore, R., Longhi, E. & Karmiloff-Smith, A. (2003) Spatial representation and attention in toddlers with Williams syndrome and Down syndrome. *Neuropsychologia* 41(8):1037–46. [rMT]
Bruner, J. (2001) Self-making and world-making. In: *Narrative and identity: Studies in autobiography, self, and culture,* ed. J. Brockmeier & D. Carbaugh. John Benjamins. [MF]
Brunsdon, R. K., Hannan, T. J., Nickels, L. & Coltheart, M. (2002) Successful treatment of sublexial reading deficits in a child with dyslexia of the mixed type. *Neuropsychological Rehabilitation* 12:199–229. [CT]
Bryne, B. (1993) Learning to read in the absences of phonemic awareness? A comment on Cossu, Rossini and Marshall (1993). *Cognition* 48:285–88. [YL]
Buitelaar, J. K. & Willemsen-Swinkels, S. H. (2000) Medication treatment in subjects with autistic spectrum disorders. *European Child and Adolescent Psychiatry* 9(1):I85–I97. [R-AM]
Bullinaria, J. A. (1997) Modelling reading, spelling, and past tense learning with artificial neural networks. *Brain and Language* 59:236–66. [aMT]
Bullinaria, J. A. & Chater, N. (1995) Connectionist modelling: Implications for cognitive neuropsychology. *Language and Cognitive Processes* 10:227–64. [aMT]
Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J. & Gabrieli, J. D. (2002) Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron* 33:301–11. [ET]
Cabeza, R. (2002) Hemispheric asymmetry reduction in older adults. *Psychology and Aging* 17:85–100. [S-CL]
Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I. & Young, A. W. (2002) Reading the mind from eye gaze. *Neuropsychologia* 40:1129–38. [R-AM]
Campos, J. J., Bertenthal, B. I. & Kermoian, R. (1992) Early experience and emotional development: The emergence of wariness of heights. *Psychological Science* 3(1):61–64. [rMT]
Capirci, O., Sabbadini, L. & Volterra, V. (1996) Language development in Williams syndrome: A case study. *Cognitive Neuropsychology* 13(7):1017–39. [YL, rMT]
Caramazza, A. (1984) The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language* 21:9–20. [CT]
Cardell, E. A. & Chenery, H. J. (1999) A cognitive neuropsychological approach to the assessment and remediation of acquired dysgraphia. *Language Testing* 16:353–88. [CT]
Carey, S. & Diamond, R. (1977) From piecemeal to configurational representation of faces. *Science* 195:312–14. [CT]
Castelli, F., Frith, C., Happé, F. & Frith, U. (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125:1839–49. [R-AM]
Castles, A. & Coltheart, M. (1993) Varieties of developmental dyslexia. *Cognition* 47:149–80. [JB, FR, aMT]
(1996) Cognitive correlates of developmental surface dyslexia: A single case study. *Cognitive Neuropsychology* 13:25–50. [CT]
Castles, A., Datta, H., Gayan, J. & Olson, R. (1999) Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology* 72:73–94. [NEJ]
Cattell, R. B. (1971) *Abilities: Their structure, growth and action.* Houghton Mifflin. [S-CL]
Chapman, L. J. & Chapman, J. P. (2001) Commentary on two articles concerning generalized and specific cognitive deficits. *Journal of Abnormal Psychology* 110:31–39. [RC]
Chomsky, N. (1959) A review of *Verbal behavior,* by B. F. Skinner. *Language* 35:26–58. [GFM]
(2002) *On Nature and language.* Cambridge University Press. [R-AM]
Christiansen, M. H. & Chater, N. (2001) Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences* 5:82–88. [MHC]
Chugani, D. C. (2002) Role of altered brain serotonin mechanisms in autism. *Molecular Psychiatry* 7(2):S16–S17. [R-AM]
Churchill, J. D., Grossman, A. W., Irwin, S. A., Galvez, R., Klintsova, A. Y., Weiler, I. J. & Greenough, W. T. (2002) A converging methods approach to FragileX Syndrome. *Developmental Science* 40:323–38. [rMT]
Clahsen, H. (1999) Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22:991–1060. [CP]
Clahsen, H. & Almazan, M. (1998) Syntax and morphology in children with Williams syndrome. *Cognition* 68:167–98. [CP, CT, arMT]
(2001) Compounding and inflection in language impairment: Evidence from Williams syndrome (and SLI). *Lingua* 111:729–57. [YL, CT]
Clahsen, H., Ring, M. & Temple, C. (2003) Lexical and morphological skills in English children with Williams syndrome. In: *Language competence and development in Williams syndrome,* ed. S. Bartke & J. Siegmueller. Benjamins. [CT]
Clahsen, H. & Temple, C. (2002) Words and rules in Williams Syndrome. In: *Towards a definition of specific language impairment in children,* ed. Y. Levy & J. Schaeffer. Erlbaum. [aMT]
(2003) Words and rules in children with Williams Syndrome. In: *Language competence across populations,* ed. Y. Levy & J. Schaeffer. Erlbaum. [CT, rMT]
Cohen, I. L. (1998) Neural network analysis of learning in autism. In: *Neural networks and psychopathology,* ed. D. J. Stein & J. Ludik. Cambridge University Press. [MHC, aMT]
Cohen, J. D. & Servan-Schreiber, D. (1992) Context, cortex, and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychological Review* 99:45–77. [RC, aMT]
Cole, M. (1999) Culture in development. In: *Developmental psychology: An advanced textbook, 4th edition,* ed. M. H. Bornstein & M. E. Lamb. Erlbaum. [S-CL]
Coltheart, M. (1999) Modularity and cognition. *Trends in Cognitive Sciences* 3:115–20. [aMT]
Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993) Models of reading aloud: Dual route and parallel-distributed-processing approaches. *Psychological Review* 100:589–608. [YM]
Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. C. (2001) DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108:204–56. [NEJ, CT, arMT]
Condray, R., Steinhauer, S. R., van Kammen, D. P. & Kasparek, A. (in press) The language system in schizophrenia: Effects of capacity and linguistic structure. *Schizophrenia Bulletin.* [RC]
Conway, C. M., Ellefson, M. R., Dale, R., Christiansen, M. H. & Schaper, K. (in preparation) *Connectionist models of developmental disorders: A critical appraisal.* [MHC]
Cornish, K. M. (in press) *Deconstructing the attention deficit in Fragile X syndrome: A developmental neuropsychological approach* (submitted). *Journal of Child Psychology and Psychiatry* [rMT]
Cornish, K. M., Munir, F. & Cross, G. (1998) The nature of the spatial deficit in young females with Fragile X syndrome: A neuropsychological and molecular perspective. *Neuropsychologia* 36(11):1239–46. [rMT]
Cossu, G. & Marshall, J. C. (1990) Are cognitive skills prerequisite for learning to read and write? *Cognitive Neuropsychology* 7:21–40. [YL]
Cossu, G., Rossini, F. & Marshall, J. C. (1993) When reading is acquired but phonemic awareness is not. *Cognition* 46:129–38. [YL]
Crick, F. (1989) The recent excitement about neural networks. *Nature* 337:129–32. [RPC]
Crombie, A. C. (1994) *Styles of scientific thinking in the European tradition: The history of argument and explanation especially in the mathematical and biomedical sciences, vols. 1, 2, and 3.* Duckworth. [MF]
Davis, C. J., Gayan, J., Knopik, V. S., Smith, S. D., Cardon, L. R., Pennington, B. F., Olson, R. K. & DeFries, J. C. (2001) Etiology of reading difficulties and rapid naming: The Colorado Twin Study of Reading Disability. *Behav Genet* 31(6): 625–35. [FR]
Dawson, G., Webb, S., Schellenberg, G. D., Dager, S., Friedman, S., Aylward, E. & Richards, T. (2002) Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology* 14:581–611. [R-AM]
Deary, I. J. (2001) Human intelligence differences. A recent history. *Trends in Cognitive Science* 5:127–30. [S-CL]
De Haan, M., Pascalis, O. & Johnson, M. H. (2002) Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience* 14:199–209. [S-CL]
Denenberg, V. H. (1999) A critique of Mody, Studdert-Kennedy and Brady's "Speech perception deficits in poor readers: Auditory processing of phonological coding?" (1997) *Journal of Learning Disabilities* 32:379–83. [PT]
De Ribaupierre, A. (1993). Structural invariants and individual differences: On the difficulty of dissociating developmental and differential processes. In: *The new structuralism in cognitive development: Theory and research on individual pathways,* ed. R. Case & W. Edelstein. Karger. [S-CL]
Deruelle, C., Mancini, J., Livet, M. O., Cassé-Perrot, C. & de Schonen, S. (1999) Configural and local processing of faces in children with Williams syndrome. *Brain and Cognition* 41:276–98. [YL, CT, aMT]
Diamond, A. (2001) A model system for studying the role of dopamine in prefrontal cortex during early development in humans. In: *Handbook of*

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

developmental cognitive neuroscience, ed. C. Nelson & M. Luciana. MIT Press. [rMT]

Diamond, A., Prevor, M., Callender, G. & Druin, D. P. (1997) Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development (Monograph No. 252)* 62(4):1–207. [rMT]

Donnai, D. & Karmiloff-Smith, A. (2000) Williams syndrome: From genotype through to cognitive phenotype. *American Journal of Medical Genetics: Seminars in Medical Genetics* 97:164–71. [PG]

Dunn, J. & Kirsner, K. (2003) What can we infer from double dissociations? *Cortex* 39(1):1–7. [PJ]

Dykens, E. M., Heap, R. M., Ort, S., Finucane, B., Shapiro, L. & Leckman, J. F. (1989) The trajectory of cognitive development in males with FragileX syndrome. *Journal of American Academy of Child and Adolescent Psychiatry* 28(3):422–26. [rMT]

Ellis, A. W. (1993) *Reading, writing, and dyslexia: A cognitive analysis, 2nd edition.* Erlbaum. [NF]

Ellis, A. W. & Young, A. W. (1996) *Human cognitive neuropsychology: A textbook with readings.* Psychology Press. [NF]

(1988) *Human cognitive neuropsychology.* Erlbaum. [CT]

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: A connectionist perspective on development.* MIT Press. [PG, S-CL, arMT]

Engle, R. W. (2002) Working memory capacity is executive attention. *Current Directions in Psychological Science* 11:19–23. [S-CL]

Farah, M. J., Hammond, K. M., Mehta, Z. & Ratcliff, G. (1989) Category-specificity and modality-specificity in semantic memory. *Neuropsychologia* 27:193–200. [ET]

Farah, M. J. & McClelland, J. L. (1991) A computational model of semantic memory impairment: Modality specificity and emergent category-specificity. *Journal of Experimental Psychology: General* 120:339–57. [JLM]

Farah, M. J., O'Reilly, R. C. & Vecera, S. P. (1993) Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review* 100:571–88. [aMT]

Fawcett, A. J., Nicolson, R. I. & Dean, P. (1996) Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia* 46:259–83. [FR, aMT]

Fisch, G. S., Carpenter, N., Holden, J. J., Howard-Peebles, P. N., Maddalena, A., Borghgraef, M., Steyaert, J. & Fryns, J. P. (1999) Longitudinal changes in cognitive and adaptive behaviour in Fragile X females: A prospective multicenter analysis. *American Journal of Medical Genetics* 83(4):308–12. [rMT]

Fitch, R. H., Brown, C. P., Tallal, P. & Rosen, G. D. (1997) Effects of sex and MK-801 on auditory-processing deficits associated with developmental microgyric lesions in rats. *Behavioral Neuroscience* 111:404–12. [FR]

Flax, J. F., Realpe-Bonilla, T., Hirsch, L. S., Brzustowicz, L., Bartlett, C. & Tallal, P. (in press) Specific language impairments in families: Evidence for co-occurrence with reading impairments. *Journal of Speech, Language, and Hearing Research.* [PT]

Fodor, J. A. (1983) *The modularity of mind.* MIT Press. [YL, aMT]

(2000) *The mind doesn't work that way: The scope and limits of computational psychology.* MIT Press. [RC, aMT]

Ford, D. H. (1987) *Human as self-constructing living systems. A developmental perspective on behavior and personality.* Erlbaum. [S-CL]

Frangiskakis, J. M., Ewart, A. K., Morris, A. C., Mervis, C. B., Bertrand, J., Robinson, B. F., Klein, B. P., Ensing, G. J., Everett, L. A., Green, D., Proschel, C., Cutowski, N. J., Noble, M., Atkinson, D. L., Odelberg, S. J. & Keating, M. T. (1996) LIM-kinase1 hemizygosity implicated in impaired visuospatial constructive cognition. *Cell* 86:59–69. [rMT]

Frawley, W. (2002) Control and cross-domain mental computation: Evidence from language breakdown. *Computational Intelligence* 18:1–28. [aMT]

Friedmann, N. & Gvion, A. (2001) Letter position dyslexia. *Cognitive Neuropsychology* 18:673–96. [NF]

(2002) Letter position dyslexia following left occipito-parietal lesion. Paper presented at the Science of Aphasia: Functional Neuroimaging Studies of Language and its Impairment Conference, Maratea, Italy, June 2002. [NF]

Friedmann, N. & Nachman-Katz, I. (in press) Neglect dyslexia in a Hebrew-reading child. *Cortex.* [NF]

Friedmann, N. & Novogrodsky, R. (2003) Syntactic movement in Hebrew-speaking children with G-SLI. Paper presented at EUCLIDES meeting, Wales, UK, May 2003. [NF]

Friedmann, N. & Rahamim, E. (2002) Developmental letter position dyslexia: The effect of diacritics and comparison of treatment methods. Paper presented at the SCRIPT conference, Tel Aviv University, Tel Aviv, Israel, December 2002. [NF]

Frith, U. (1985) Beneath the surface of developmental dyslexia. In: *Surface dyslexia,* ed. K. E. Patterson, J. C. Marshall & M. Coltheart. Erlbaum. [CT]

(1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia* 36: 69–81. [FR]

(1995) Dyslexia: Can we have a shared theoretical framework? *Educational and Child Psychology* 12:6–17. [aMT]

(in press) Nativism and neuroconstructivism in the explanation of Williams syndrome. *Biology and Philosophy.* [PG]

Galaburda, A. (2001) Brain and sounds: Lessons from "dyslexic" rodents. In: *Language, brain and cognitive development: Essays in honor of Jacques Mehler,* ed. E. Dupoux. MIT Press. [FR]

Galaburda, A. M., Menard, M. T. & Rosen, G. D. (1994) Evidence for aberrant auditory anatomy in developmental dyslexia. *Proceedings of the National Academy of Science USA* 91:8010–13. [FR]

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F. & Geschwind, N. (1985) Developmental dyslexia: Four consecutive patients with cortical anomalies. *Annals of Neurology* 18:222–33. [FR]

Gerrans, P. (2002) Theory of mind in evolutionary psychology. *Biology and Philosophy* 17:305–21. [PG]

(2003) Nativism and neuroconstructivism in the explanation of Williams syndrome. *Biology and Philosophy* 18:1–11.

Gerrans, P. & McGeer, V. (2003) Theory of mind in autism and schizophrenia. In: *Individual differences in theory of mind: Influences on typical and atypical development,* ed. B. Repacholi & V. Slaughter. Psychology Press. [PG]

Goldman-Rakic, P., Isseroff, A., Schwartz, M. & Bugbee, N. (1983) The neurobiology of cognitive development. In: *Handbook of child psychology, vol. 2: Infancy and developmental psychobiology,* ed. M. Haith & J. Campos. Wiley. [SCL]

Gooding, D. C. & Iacono, W. G. (1995) Schizophrenia through the lens of a developmental psychopathology perspective. In: *Developmental psychopathology: Vol. II. Risk, disorder, and adaptation,* ed. D. Cicchetti & D. J. Cohen. Wiley. [RC]

Gottlieb, G. & Halpern, C. T. (2002) A relational view of causality in normal and abnormal development. *Development and Psychopathology* 14:421–35. [R-AM]

Grant, J., Karmiloff-Smith, A., Gathercole, S. A., Paterson, S., Howlin, P., Davies, M. & Udwin, O. (1997) Phonological short-term memory and its relationship to language in Williams syndrome. *Cognitive Neuropsychiatry* 2:81–99. [aMT]

Grice, S., Spratling, M. W., Karmiloff-Smith, A., Halit, H., Csibra, G., de Haan, M. & Johnson, M. H. (2001) Disordered visual processing and oscillatory brain activity in autism and Williams Syndrome. *NeuroReport* 12:2697–2700. [arMT]

Hacking, I. (1998) *Mad travelers: Reflections on the reality of transient mental illnesses.* University of Virginia Press. [MF]

(2002) *Historical ontology.* Harvard University Press. [MF]

Halliday, L. & Briscoe, J. (2002) Reading strategies in children with mild to moderate hearing loss. Unpublished raw data. [JB]

Happé, F., Ehlers, S., Fletcher, P. C., Frith, U., Johansson, M., Gillberg, C., Dolan, R. J., Frackowiak, R. S. J. & Frith, C. D. (1996) "Theory of mind" in the brain: Evidence from a PET scan study of Asperger syndrome. *NeuroReport* 8:197–201. [R-AM]

Hari, R. & Renvall, H. (2001) Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Science* 5:525–32. [FR]

Harm, M. W. & Seidenberg, M. S. (1999) Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review* 106:491–528. [MHC, NEJ, aMT]

Harm, M. W. & Seidenberg, M. S. (2001) Are there orthographic impairments in phonological dyslexia? *Cognitive Neuropsychology* 18(1):71–92.

Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–79. [GFM]

Hebb, D. O. (1949) *The organization of behavior: A neuropsychological theory.* Wiley. [SCL, PT]

Herman, A. E., Galaburda, A. M., Fitch, R. H., Carter, A. R. & Rosen, G. D. (1997) Cerebral microgyria, thalamic cell size and auditory temporal processing in male and female rats. *Cerebral Cortex* 7:453–64. [FR]

Hill, E. (1998) A dyspraxic deficit in specific language impairment and developmental co-ordination disorder? Evidence from hand and arm movements. *Developmental Medicine and Child Neurology* 40:388–95. [aMT]

Hinton, G. (1989) Connectionist learning procedures. *Artificial Intelligence* 40:185–234. [aMT]

Hinton, G. E. & Sejnowski, T. J. (1986) Learning and relearning in Boltzmann machines. In: *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1: Foundations,* ed. D. E. Rumelhart, J. L. McClelland & the PDP Research Group. MIT Press. [aMT]

Hinton, G. E. & Shallice, T. (1991) Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review* 98:74–95. [aMT]

Hodges, J. R., Patterson, K., Oxbury, S. & Funnell, E. (1992) Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain* 115:1783–1806. [CT]

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Hoeffner, J. (1992) Are rules a thing of the past? The acquisition of verbal morphology by an attractor network. *Proceedings of the 14th Annual Meeting of the Cognitive Science Society.* Erlbaum. [aMT]

Hoeffner, J. H. & McClelland, J. L. (1993) Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. *Proceedings of the 25th Child Language Research Forum,* ed. E. V. Clark. Stanford University Press. [MHC, JLM, aMT]

Hoffman, R. E. (1996) Neural networks, cortical connectivity and schizophrenic psychosis. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Hoffman, R. E. & McGlashan, T. H. (1997) Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. *American Journal of Psychiatry* 154:1683–89. [MHC]

Horn, J. L. (1968) Organization of abilities and the development of intelligence. *Psychological Review* 75:242–59. [S-CL]

Horrobin, D. (1998) The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophrenia Research* 30:193–208. [RC]

Humphreys, K., Ewing, S. & Karmiloff-Smith, A. (2002) *Face processing in Williams syndrome: Infant precursors in developmental disorders.* Presentation to the International Conference on Infant Studies, Toronto, April 2002. [aMT]

Humphreys, P., Kaufmann, W. E. & Galaburda, A. M. (1990) Developmental dyslexia in women: Neuropathological findings in three patients. *Annals of Neurology* 28:727–38. [FR]

Huttenlocher, P. R. (2002) *Neural plasticity: The effects of the environment on the development of the cerebral cortex.* Harvard University Press. [rMT]

Isaacs, E. B., Lucas, A., Chong, W. K., Wood, S. J., Johnson, C. L., Marshall, C., Vargha-Khadem, F. & Gadian, D. G. (2000) Hippocampal volume and everyday memory in children of very low birthweight. *Pediatric Research* 47:73–720. [JB]

Jackson, N. E. & Coltheart, M. (2001) *Routes to reading success and failure: Toward an integrated cognitive psychology of atypical reading.* Psychology Press. [NEJ, aMT]

Jacobs, R. A. (1997) Nature, nurture, and the development of functional specializations: A computational approach. *Psychonomic Bulletin Review* 4:229–309. [aMT]

(1999) Computational studies of the development of functionally specialized neural modules. *Trends in Cognitive Science* 3:31–38. [aMT]

Jacobs, R. A. & Jordan, M. I. (1992) Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* 6:181–214. [aMT]

Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hinton, G. E. (1991) Adaptive mixtures of local experts. *Neural Computation* 3:79–87. [aMT]

Jarrold, C., Hartley, S. J., Phillips, C. & Baddeley, A. D. (2000) Word fluency in Williams syndrome: Evidence for unusual semantic organisation? *Cognitive Neuropsychiatry* 5:293–319. [CP, rMT]

Joanisse, M. F. (2000) Connectionist phonology. Ph.D. dissertation, University of Southern California. [aMT]

Joanisse, M. F. & Seidenberg, M. S. (1998) Specific language impairment in children: A deficit in grammar or processing? *Trends in Cognitive Sciences* 2:240–47. [aMT]

(1999) Impairments in verb morphology following brain injury: A connectionist model. *Proceedings of the National Academy of Sciences USA* 96:7592–97. [JLM, aMT]

Johnson, M. H. (1997) *Developmental cognitive neuroscience.* Blackwell. [rMT, GW]

(1999) Cortical plasticity in normal and abnormal cognitive development: Evidence and working hypothesis. *Development and Psychopathology* 11:419–37. [aMT]

(2001) Functional brain development in humans. *Nature Review Neuroscience* 2:475–83. [S-CL]

Johnson, M. H. & Karmiloff-Smith, A. (1992) Can neural selectionism be applied to cognitive development and its disorders? *New Ideas in Psychology* 10(1):35–46. [aMT]

Johnson, S. C. & Carey, S. (1998) Knowledge enrichment and conceptual change in folkbiology: Evidence from Williams syndrome. *Cognitive Psychology* 37:156–200. [CP]

Juola, P. & Plunkett, K. (1998) Why double dissociations don't mean much. In: *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society, Orlando, Florida, March 1998,* pp. 561–66, ed. H. MacKenzie. Erlbaum. [PJ]

Karmiloff-Smith, A. (1992) *Beyond modularity: A developmental perspective on cognitive science.* MIT Press/Bradford Books. [arMT]

(1997) Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology* 13:513–24. [YL, aMT, CT]

(1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences* 2:389–98. [PG, YL, arMT]

Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P. & Udwin, O. (1997) Language and Williams syndrome: How intact is "intact"? *Child Development* 68:246–62. [YL, aMT]

Karmiloff-Smith, A., Grant, J., Ewing, S., Carette, M. J., Metcalfe, K., Donnai, D., Read, A. P. & Tassabehji, M. (2003a) Using case study comparisons to explore genotype/phenotype correlations in Williams Syndrome. *Journal of Medical Genetics* 40:136–40. [rMT]

Karmiloff-Smith, A., Scerif, G. & Ansari, D. (2003b) Double dissociations in developmental disorders: Theoretically misconceived, empirically dubious. *Cortex* 39:161–63. [rMT]

Karmiloff-Smith, A., Scerif, G. & Thomas, M. S. C. (2002) Different approaches to relating genotype to phenotype in developmental disorders. *Developmental Psychobiology* 40:311–22. [arMT]

Karmiloff-Smith, A. & Thomas, M. S. C. (2003) Developmental disorders. In: *The handbook of brain theory and neural networks, 2nd edition,* ed. M. A. Arbib. MIT Press. [rMT]

(in press) What can developmental disorders tell us about the neurocomputational constraints that shape development? The case of Williams syndrome. *Development and Psychopathology.* [rMT]

Karmiloff-Smith, A., Tyler, L. K., Voice, K. Sims, K., Udwin, O., Howlin, P. & Davies, M. (1998) Linguistic dissociations in Williams syndrome: Evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia* 36: 343–51. [aMT]

Keating, D. P. & Miller, F. K. (1999) Individual pathways in competence and coping: From regulatory systems to habits of mind. In: *Developmental health and the wealth of nations: Social, biological, and educational dynamics,* ed. D. P. Keating & C. Hertzman. Guilford Press. [MF]

Knight, R. A. & Silverstein, S. M. (2001) A process-oriented approach for averting confounds resulting from general performance deficiencies in schizophrenia. *Journal of Abnormal Psychology* 110:15–30. [RC]

Koch, R., Moseley, K., Ning, J., Romstad, A., Guldberg, P. & Guttler, F. (1999) Long-term beneficial effects of the phenylalanine-restricted diet in late-diagnosed individuals with phenylketonuria. *Molecular Genetics and Metabolism* 67:148–55. [rMT]

Kolb, B. (1995) *Brain plasticity and behavior.* Erlbaum. [SCL]

Kónya, A. & Pintér, G. (1986) Kategória norma a verbális emlékezet vizsgálatához. (Category norms for testing verbal memory). *Magyar Pszichológiai Szemle* 20:93–111. [CP]

Kronbichler, M., Hutzler, F. & Wimmer, H. (2002) Dyslexia: Verbal impairments in the absence of magnocellular impairments. *NeuroReport* 13:617–20. [FR]

Ladd, G. T. (1892) Psychology as a so-called "natural science." *Philosophical Review* 1:24–53. [MF]

Lai, S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–23. [R-AM, rMT]

Laing, E., Butterworth, G., Ansari, D., Gsödl, M., Longhi, E. Panagiotaki, G., Paterson, S. & Karmiloff-Smith, A. (2002) Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science* 5:233–46. [PG, rMT]

Laing, E., Hulme, C., Grant, J. & Karmiloff-Smith, A. (2001) Learning to read in Williams syndrome: Looking beneath the surface of atypical reading development. *Journal of Child Psychology and Psychiatry* 42(6):729–39. [rMT]

Lautrey, J. (in press) A pluralistic approach to cognitive differentiation and development. In: *Models of intelligence for the new millennium,* ed. R. J. Sternberg, J. Lautrey & T. I. Lubart. American Psychological Association. [S-CL]

Leonard, L. B. (1998) *Children with specific language impairment.* MIT Press. [JLM, PT]

Leonard, L. B., McGregor, K. K. & Allen, G. D. (1992) Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research* 35:1076–85. [JLM]

Leslie, A. M. (1992) Pretense, autism, and the theory-of-mind module. *Current Directions in Psychological Science* 1:18–21. [R-AM, aMT]

Levine, D. S. (1996) Modeling dysfunction of the prefrontal executive system. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Levine, S. C., Huttenlocher, P. R., Banich, M. T. & Duda, E. (1987) Factors affecting cognitive functioning in hemiplegic children. *Developmental Medical Child Neurology* 29:27–35. [SCL]

Levy, Y., Tennebaum, A. & Ornoy, A. (2000) Spontaneous language of children with specific neurological syndromes. *Journal of Speech, Language and Hearing Research* 43:351–65. [YL]

Li, S.-C. (2003) Biocultural orchestration of developmental plasticity across levels: The interplay of biology and culture in shaping the mind and behaviour across the lifespan. *Psychological Bulletin* 129:171–94. [S-CL]

Li, S -C. & Lindenberger, U. (1999) Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

the dedifferentiation of cognitive abilities in old age. In: *Cognitive neuroscience of memory,* ed. L.-G. Nilsson & H. Markowitsch. Hogrefe and Huber. [aMT]

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W. & Baltes, P. B. (in press) Lifespan developmental transformations in the couplings of mental abilities and underlying cognitive processes. *Psychological Science.* [S-CL]

Li, S.-C., Lindenberger, U. & Sikström, S. (2001) Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences* 5:479–86. [S-CL]

Ling, C. X. & Marinov, M. (1993) Answering the connectionist challenge: A symbolic model of learning the past tenses of English verbs. *Cognition* 49:235–90. [aMT]

Livingstone, M., Rosen, G., Drislane, F. & Galaburda, A. (1991) Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Science USA* 88:7943–47. [FR]

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C. & Buckner, R. L. (2002) Under-recruitment and non-selective recruitment: Dissociable neural mechanisms associated with aging. *Neuron* 33:827–40. [S-CL]

Lovegrove, W. J., Bowling, A., Badcock, B. & Blackwood, M. (1980) Specific reading disability: Differences in contrast sensitivity as a function of spatial frequency. *Science* 210:439–40. [FR]

Lukács, Á. & Pléh, C. (1999) Hungarian cross-modal priming and treatment of nonsense words supports the dual process hypothesis. Commentary on Harald Clahsen's "Lexical entries and rules of language." *Behavioral and Brain Sciences* 22:6. [CP]

Lukács, A., Pléh, C. & Racsmány, M. (in press) Language in Hungarian Williams syndrome children. [CP]

Lukács, Á., Racsmány, M. & Pléh, C. (2001) Vocabulary and morphological patterns in Hungarian children with Williams syndrome: A preliminary report. *Acta Linguistica* 48:243–69. [CP]

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., Keshavan, M. S., Genovese, C. R., Eddy, W. F. & Sweeney, J. A. (2001) Maturation of widely distributed brain function subserves cognitive development. *Neuroimage* 13:786–93. [ET]

Lutz, A. (2002) Toward a neurophenomenology as an account of generative passage: A first empirical case study. *Phenomenology and the Cognitive Sciences* 1:133–67. [MF]

Lynch, E. D., Lee, M. K., Morrow, J. E., Welcsh, P. L., Leon, P. E. & King, M. C. (1997) Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the Drosophila gene diaphanous. *Science* 278:1315–18. [rMT]

MacWhinney, B. & Leinbach, J. (1991) Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 40:121–57. [aMT]

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J. & Frith, C. D. (2000) Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences USA* 97:4398–403. [S-CL]

Majerus, S., Palmisano, I., van der Linden, M., Barisnikov, K. & Poncelet, M. (2001) An investigation of phonological processing in Williams syndrome. *Journal of the International Society* 7(2):153. [aMT]

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E. & Sejnowski, T. J. (2002) Dynamic brain sources of visual evoked responses. *Science* 295:690–94. [S-CL]

Manis, F., Seidenberg, M. S., Doi, L., McBride-Chang, C. & Petersen, A. (1996) On the basis of two subtypes of developmental dyslexia. *Cognition* 58:157–95. [aMT]

Marchman, V. A. (1993) Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience* 5:215–34. [SCL]

Marcus, G. F. (2001) *The algebraic mind: Integrating connectionism and cognitive science.* MIT Press. [GFM]

(2003) *The birth of the mind: The genetic origins of our mental life.* Basic Books. [GFM]

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. & Xu, F. (1992) Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 4 (entire volume). [GW]

Marenco, S. & Weinberger, D. R. (2000) The neurodevelopmental hypothesis of schizophrenia: Following a trail of evidence from cradle to grave. *Development and Psychopathology* 12:501–27. [RC]

Mareschal, D., Plunkett, K. & Harris, P. (1999) A computational and neuropsychological account of object-oriented behaviours in infancy. *Developmental Science* 2:306–17. [YM, aMT]

Mareschal, D. & Shultz, T. R. (1996) Generative connectionist networks and constructivist cognitive development. *Cognitive Development* 11:571–603. [GW]

Margulies, E. H., Kardia, S. L. & Innis, J. W. (2001) A comparative molecular analysis of developing mouse forelimbs and hindlimbs using serial analysis of gene expression (SAGE). *Genome Research* 11:1686–98. [GFM]

Marshall, J. C. & Newcombe, F. (1973) Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research* 2:179–99. [NF]

Matthews, A., Ellis, A. E. & Nelson, C. A. (1996) Development of preterm and full-term infant ability on AB, recall memory, transparent barrier detour, and means-end tasks. *Child Development* 67(6):2658–76. [rMT]

Mayall, K. & Humphreys, G. W. (1996) Covert recognition in a connectionist model of pure alexia. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Mayr, E. (1998) *Evolutionary synthesis: Perspectives on the unification of biology.* Harvard University Press. [S-CL]

McCarthy, R. & Warrington, E. (1990) *Cognitive neuropsychology.* Academic Press. [aMT]

McClelland, J. L. (1979) On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review* 86:287–330. [GFM]

McCloskey, M., Rapp, B, Yantis, S. et al. (1995) A developmental deficit in localising objects from vision. *Psychological Science* 6:112–17. [CT]

McLeod, P., Plunkett, K. & Rolls, E. T. (1998) *Introduction to connectionist modelling of cognitive processes.* Oxford University Press. [aMT]

Merzenich, M. M. & Jenkins, W. M. (1995) Cortical plasticity, learning and learning dysfunction. In: *Maturational windows and adult cortical plasticity,* ed. B. Julesz & I. Kovacs. Addison-Wesley. [PT]

Merzenich, M., Jenkins, W., Johnston, P., Schreiner, C., Miller, S. & Tallal, P. (1996) Temporal processing deficits of language-learning children ameliorated by training. *Science* 271:77–81. [JB, PT]

Miller, K. D. & Erwin, E. (2001) Effects of monocular deprivation and reverse suture on orientation maps can be explained by activity-instructed development of geniculocortical connections. *Visual Neuroscience* 18:821–34. [aMT]

Mills, D. L., Alvarez, D., St. George, M., Appelbaum, G., Bellugi, U. & Neville, H. (2000) Electrophysiological studies of face processing in Williams syndrome. *Journal of Cognitive Neuroscience* 12(S):47–64. [aMT]

Mody, M., Studdert-Kennedy, M. & Brady, S. (1997) Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology* 64:199–231. [PT]

Morris, D. B. (1998) *Illness and culture in the postmodern age.* University of California Press. [MF]

Moscovic, M. & Umiltà, C. (1990) Modularity and neuropsychology: Modules and central processes in attention and memory. In: *Modular deficits in Alzheimer-type dementia,* ed. M. Schwarz. MIT Press. [YL]

Müller, R.-A., Rothermel, R. D., Behen, M. E., Muzik, O., Chakraborty, P. K. & Chugani, H. T. (1999) Language organization in patients with early and late left hemisphere lesion: A PET study. *Neuropsychologia* 37:545–57. [R-AM]

Nagarajan, S, Mahncke, H., Salz, T., Tallal, P., Roberts, T. & Merzenich, M. (1999) Cortical auditory signal processing in poor readers. *Proceedings of the National Academy of Sciences USA* 96:6483–648. [PG]

Nazzi, T. & Karmiloff-Smith, A. (2002) Categorization abilities in infants and toddlers with Williams syndrome. *NeuroReport* 13:1259–62. [rMT]

Nazzi, T., Paterson, S. & Karmiloff-Smith, A. (2003) Early word segmentation by infants and toddlers with Williams syndrome. *Infancy* 4:257–71. [rMT]

Nelson, C. A. & Luciana, M. (2001) *Handbook of developmental cognitive neuroscience.* MIT Press. [rMT]

Nelson, K. (1993) The psychological and social origins of autobiographical memory. *Psychological Science* 4:7–14. [MF]

(1996) *Language in cognitive development.* Cambridge University Press. [S-CL]

Nesselroade, J. R. (1991) The warp and the woof of developmental fabric. In: *Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohlwill,* ed. R. M. Downs, L. S. Liben & D. S. Palermo. Erlbaum. [S-CL]

Neville, J. J., Mills, D. L. & Bellugi, U. (1994) Effects of altered auditory sensitivity and age of language acquisition on the development of language-related neural systems: Preliminary studies of Williams syndrome. In: *Atypical cognitive deficits in developmental disorders: Implications for brain function,* ed. S. Broman & J. Grafman. Erblaum. [aMT]

Newcombe, F. & Marshall, J. C. (1981) On psycholinguistic classifications of the acquired dyslexias. *Bulletin of the Orton Society* 31:29–46. [NF]

Nichols, G. & Prigogine, I. (1989) *Exploring complexity: An introduction.* W. H. Freeman. [PJ]

Nicolson, R. I. & Fawcett, A. J. (2000) Long-term learning in dyslexic children. *European Journal of Cognitive Psychology* 12:357–93. [FR]

Novogrodsky, R. & Friedmann, N. (2002) Relative clause comprehension in Hebrew-speaking school-age children with G-SLI. Paper presented at the Euresco Syntax of Normal and Impaired Language Conference, Corinth, Greece, June 2002. [NF]

O'Laughlin, C. & Thagard, P. (2000) Autism and coherence: A computational model. *Mind and Language* 15:375–92. [MHC]

Oliver, A., Johnson, M. H., Karmiloff-Smith, A. & Pennington, B. (2000)

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Deviations in the emergence of representations: A neuroconstructivist framework for analysing developmental disorders. *Developmental Science* 3:1–23. [arMT]

Olson, R. & Datta, H. (2002). Visual-temporal processing in reading-disabled and normal twins. *Reading and Writing* 15(1–2):127–149. [FR]

O'Reilly, R. C. (1998) Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Sciences* 2(11):455–62. [rMT]

O'Reilly, R. C. & Munakata, Y. (2000) *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain.* MIT Press. [YM]

Patterson, K. E. (1981) Neuropsychological approaches to the study of reading. *British Journal of Psychology* 72:151–74. [NF]

(1990) Alexia and neural nets. *Japanese Journal of Neuropsychology* 6:90–99. [aMT]

Patterson, K, Plaut, D. C., McClelland, J. L., Seidenberg, M. S., Behrmann, M. & Hodges, J. R. (1996) Connections and disconnections: A connectionist account of surface dyslexia. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Patterson, K., Seidenberg, M. S. & McClelland, J. L. (1989) Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In: *Parallel distributed processing: Implications for psychology and neuroscience,* ed. R. G. M. Morris. Oxford University Press. [aMT]

Patterson, K. & Shewell, C. (1987) Speak and spell: Dissociations and word-class effects. In: *The cognitive neuropsychology of language,* ed. M. Coltheart, G. Sartori & R. Job. Erlbaum . [aMT]

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., Pesenti, S., Gallagher, A., Perani, D., Price, C., Frith, C. D. & Frith, U. (2000) A cultural effect on brain function. *Nature Neuroscience* 3:91–96. [S-CL]

Penke, M., Janssen, U. & Krause, M. (1999) The representation of inflectional morphology: Evidence from Broca's aphasia. *Brain and Language* 68:225–32. [GW]

Piaget, J. (1971) *Biology and knowledge.* University of Chicago Press. [aMT]

Pinker, S. (1991) Rules of language. *Science* 253:530–35. [JLM, CP, aMT]

(1999) *Words and rules.* Basic Books/Weidenfeld and Nicolson. [PG, CP, R-AM, arMT]

(2001) Talk of genetics and vice versa. *Nature* 413:465–66. [rMT]

Pinker, S. & Ullman, M. T. (2002) The past and future of the past tense. *Trends in Cognitive Sciences* 6:456–63. [JLM]

Plaut, D. C. (1995) Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology* 17:291–321. [JLM]

(2002) Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology* 19(7):603–39. [aMT]

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. (1996) Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103:56–115. [aMT, CT]

Plaut, D. C. & Shallice, T. (1993) Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology* 10:377–500. [aMT]

Plunkett, K., Bandelow, S. & Juola, P. (2001) *Stochastic approaches to the breakdown of inflectional morphology in a connectionist network.* (in preparation). [aMT]

Plunkett, K. & Juola, P. (1999) A connectionist model of English past tense and plural morphology. *Cognitive Science* 23:463–90. [aMT]

Plunkett, K. & Marchman, V. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38:1–60. [MHC, aMT]

(1993) From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48:21–69. [JLM, aMT]

(1996) Learning from a connectionist model of the English past tense. *Cognition* 61:299–308. [aMT]

Posner, M. I., Walker, J. A., Friedrich, F. J. & Rafal, R. D. (1984) Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience* 4:1863–74. [ET]

Precel, K. & Friedmann, N. (2002) The relationship between attention deficits in ADHD and in specific developmental dyslexia. Paper presented at the Israeli Association for Literacy 17th SCRIPT conference, Bar-Ilan University, Ramat Gan, Israel, June 2002. [NF]

Quartz, S. (1993) Neural networks, nativism, and the plausibility of constructivism. *Cognition* 48:223–42. [GW]

(in press) Innateness and the brain. *Biology and Philosophy.* [PG]

Quartz, S. R. & Sejnowski, T. J. (1997) The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences* 20:537–96. [PG, GW]

Ramus, F. (2001) Dyslexia – talk of two theories. *Nature* 412:393–95. [FR]

(2003) Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology* 13. [FR]

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S. & Frith, U. (2003) Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain* 126:841–65. [GFM, FR]

Rastle, K. & Coltheart, M. (1999) Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance* 25:482–503. [NEJ]

Raz, N. (2000) Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In: *Handbook of aging and cognition,* ed. F. I. M. Craik & T. A. Salthouse. Erlbaum. [S-CL]

Reggia, J., Goodall, S., Chen, Y., Ruppin, E. & Whitney, C. (1996a) Modeling post-stroke cortical map reorganization. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Reggia, J. A., Ruppin, E. & Berndt, R. (1996b) *Neural modeling of brain and cognitive disorders.* World Scientific. [aMT]

Reiss, A. L., Eliez, S., Schmitt, J. E., Straus, E., Lai, Z., Jones, W. & Bellugi, U. (2000) Neuroanatomy of Williams syndrome: A high-resolution MRI study. *Journal of Cognitive Neuroscience* 12(1):65–73. [GW]

Reuchlin, M. (1978) Processus vicariants et différences individuelles. *Journal de Psychologie* 2:133–45. [S-CL]

Reuter-Lorenz, P. A. (2002) New visions of the aging mind and brain. *Trends in Cognitive Sciences* 6:394–400. [S-CL]

Rice, M. (1999) Specific grammatical limitations in children with Specific Language Impairment. In: *Neurodevelopmental disorders,* ed. H. Tager-Flusberg. MIT Press. [CT]

Rice, M. L., Wexler, K. & Hershberger, S. (1998) Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research* 41:1412–31. [YL]

Riva, D. & Cazzaniga, L. (1986) Effects of brain lesion sustained before and after age one. *Neuropsychologia* 24:423–28. [SCL]

Rosen, G. D., Herman, A. E. & Galaburda, A. M. (1999) Sex differences in the effects of early neocortical injury on neuronal size distribution of the medial geniculate nucleus in the rat are mediated by perinatal gonadal steroids. *Cerebral Cortex* 9:27–34. [FR]

Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., Dromerick, A. W., Fiez, J. A. & Corbetta, M. D. (2000) Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology* 55:1883–94. [R-AM]

Rossen, M., Klima, E. S., Bellugi, U., Bihrle, A. & Jones, W. (1996) Interaction between language and cognition: Evidence from Williams syndrome. In: *Language learning and behaviour,* ed. J. H. Beitchman, N. J. Cohen, M. M. Konstantareas & R. Tannock. Cambridge University Press. [aMT]

Rothweiler, M. & Clahsen, H. (1994) Dissociations in SLI children's inflectional systems: A study of participle inflection and subject-verb agreement. *Journal of Logopedics and Phoniatrics* 18:169–79. [CT]

Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tense of English verbs. In: *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models,* ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [JLM, aMT]

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (1986) *Parallel distributed processing: Explorations in the microstructure of cognition, vols. 1 and 2.* MIT Press. [JLM, aMT]

Rumsey, J. M. & Ernst, M. (2000) Functional neuroimaging of autistic disorders. *Mental Retardation and Developmental Disabilities Research Reviews* 6:171–79. [R-AM]

Ruppin, E., Horn, D., Levy, N. & Reggia, J. A. (1996) Computational studies of synaptic alterations in Alzheimer's disease. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]

Rutherford, M. D., Baron-Cohen, S. & Wheelwright, S. (2002) Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders* 32:189–94. [R-AM]

Saffran, E. M. (1982) Neuropsychological approaches to the study of language. *British Journal of Psychology* 73:317–37. [CT]

Scerif, G., Cornish, K., Wilding, J., Driver, J. & Karmiloff-Smith, A. (in press) Visual selective attention in typically developing toddlers and toddlers with Fragile X and Williams syndrome. *Developmental Science.* [rMT]

Scholl, B. J. & Leslie, A. M. (2001) Minds, modules, and meta-analysis. *Child Development* 72:696–701. [R-AM]

Scott, P., Mervis, C. B., Bertrand, J., Klein, B. P., Armstrong, S. C. & Ford, A. L. (1995) Semantic organization and word fluency in 9- and 10-year-old children with Williams syndrome. [Special Issue.] *Genetic Counseling* 6:172–73. [rMT]

Scoville, W. B. & Milner, B. (2000) Loss of recent memory after bilateral hippocampal lesions (1957). (Reprint of 1957 article). *Journal of Neuropsychiatry and Clinical Neurosciences* 12:103–13. [ET]

Seidenberg, M. S. & McClelland, J. L. (1989) A distributed, developmental model of word recognition and naming. *Psychological Review* 96:523–68. [aMT]

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

Shallice, T. (1988) *From neuropsychology to mental structure.* Cambridge University Press. [JLM, aMT]
Shallice, T., Warrington, E. K. & McCarthy, R. (1983) Reading without semantics. *Quarterly Journal of Experimental Psychology* 35A:111–38. [aMT]
Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., Constable, R. T., Marchione, K. E., Fletcher, J. M., Lyon, G. R. & Gore, J. C. (2002) Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry* 52:101–10. [ET]
Sherman, G. F., Morrison, L., Rosen, G. D., Behan, P. O. & Galaburda, A. M. (1990) Brain abnormalities in immune defective mice. *Brain Research* 532:25–33. [FR]
Shrager, J. & Johnson, M. H. (1996) Dynamic plasticity influences the emergence of function in a simple cortical array. *Neural Networks* 9:1119–29. [aMT]
Shultz, T. R., Schmidt, W. C., Buckingham, D. & Mareschal, D. (1995) Modeling cognitive development with a generative connectionist algorithm. In: *Developing cognitive competence: New approaches to process modeling,* ed. T. Simon & G. Halford. Erlbaum. [GW]
Siegler, R. S. (1994) Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science* 3:1–5. [S-CL]
Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., Davis, R. N., Fitzgerald, M. & Papanicolaou, A. C. (2002) Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology* 58:1203–13. [ET]
Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270:758–75. [S-CL]
Singer Harris, N. G., Bellugi, U., Bates, E., Jones, W. & Rossen, M. (1997) Contrasting profiles of language development in children with Williams and Down syndromes. *Developmental Neuropsychology* 13:345–70. [rMT]
Snow, C. E., Burns, S. M. & Griffin, P. (1998) *Preventing reading difficulties in young children.* National Academy Press. [PT]
Snowling, M. J. (2000) *Dyslexia, 2nd edition.* Blackwell. [FR]
Sokol, S. M., Macaruso, P. & Gollan, T. H. (1994) Developmental dyscalculia and cognitive neuropsychology. *Developmental Neuropsychology* 10:413–41. [CT]
Spearman, C. (1904) General intelligence, objectively determined and measured. *American Journal of Psychology* 15:201–93. [S-CL]
Spitzer, M. (1996) Phantom limbs, self-organizing feature maps, and noise-driven neuroplasticity. In: *Neural modeling of brain and cognitive disorders,* ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]
Spreen, O., Risser, A. H. & Edgell, D. (1995) *Developmental neuropsychology.* Oxford University Press. [SCL]
Stark, R. & Tallal, P. (1981) Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America* 69:568–74. [PT]
Stein, D. J. & Ludik, J. (1998) *Neural networks and psychopathology.* Cambridge University Press. [aMT]
Stein, J. & Walsh, V. (1997) To see but not to read; The magnocellular theory of dyslexia. *Trends in Neurosciences* 20:147–52. [FR]
Stone, V. E., Baron-Cohen, S. & Knight, R. T. (1998) Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience* 10:640–56. [R-AM]
Stoodley, C. J., Talcott, J. B., Carter, E. L., Witton, C. & Stein, J. F. (2000) Selective deficits of vibrotactile sensitivity in dyslexic readers. *Neuroscience Letters* 295:13–16. [FR]
Strauss, E., Satz, P. & Wada, J. A. (1990) An examination of the crowding hypothesis in epileptic patients who have undergone carotid amytal test. *Neuropsychologia* 28:1221–27. [SCL]
Strauss, E. & Verty, C. (1983) Effects of hemispherectomy in infantile hemiplegia. *Brain and Language* 20:1–11. [SCL]
Strauss, M. E. (2001) Demonstrating specific cognitive deficits: A psychometric perspective. *Journal of Abnormal Psychology* 110:6–14. [RC]
Stuart, M. & Howard, D. (1995) KJ: A developmental deep dyslexic. *Cognitive Neuropsychology* 12:793–824. [NF]
Taatgen, N. A. & Anderson, J. R. (2002) Why do children learn to say "broke"? A model of learning the past tense without feedback. *Cognition* 86(2):123–55. [aMT]
Tager-Flusberg, H. (2000) Differences between neurodevelopmental disorders and acquired lesions. *Developmental Science* 3:33–34. [aMT]
(2003) Language impairment in children with complex neurodevelopmental disorders: The case of autism. In: *Language competence across populations: Toward a definition of specific language impairment,* ed. J. Levy, & J. Schaeffer. Erlbaum. [YL]
Takeuchi, J. K., Koshiba-Takeuchi, K., Matsumoto, K., Vogel-Hopker, A., Naitoh-Matsuo, M., Ogura, K., Takahashi, N., Yasuda, K. & Ogura, T. (1999) Tbx5 and Tbx4 genes determine the wing/leg identity of limb buds. *Nature* 398:810–14. [GFM]
Tallal, P. (1980) Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language* 9:182–98. [FR, PT]
(1995) Development and disorders of speech and language: Implications for neural and behavioral plasticity. In: *Maturational windows and adult cortical plasticity,* ed. B. Julesz & I. Kovacs. Addison-Wesley. [JLM]
Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M. & Merzenich, M. M. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271:81–84. [PT]
Tallal, P. & Piercy, M. (1973a) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241:468–69. [PT]
(1973b) Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia* 11:389–98. [PT]
(1974) Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia* 12:83–93. [PT]
(1975) Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia* 13:69–74. [PT]
Tallal, P. & Stark, R. (1981) Speech acoustic cue discrimination abilities of normally developing and language impaired children. *Journal of the Acoustical Society of America* 69:568–74. [aMT]
Tassabehji, M., Metcalfe, K., Karmiloff-Smith, A., Carette, M. J., Grant, J., Dennis, N., Reardon, W., Splitt, M., Read, A. P. & Donnai D. (1999) Williams syndrome: Use of chromosomal microdeletions as a tool to dissect cognitive and physical phenotypes. *American Journal of Human Genetics* 64:118–25. [rMT]
Temple, C. M. (1984) New approaches to the developmental dyslexias. In: *Advances in neurology, 42: Progress in aphasiology,* ed. F. C. Rose. Raven Press. [NF]
(1991) Procedural dyscalculia and number fact dyscalculia: Double dissociation in developmental dyscalculia. *Cognitive Neuropsychology* 8:155–76. [CT]
(1992) Developmental memory impairment: Faces and patterns. In: *Mental lives: Case studies in cognition,* ed. R. Campbell. Blackwell. [CT]
(1994) The cognitive neuropsychology of the developmental dyscalculias. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition* 133:351–70. [CT]
(1997) *Developmental cognitive neuropsychology.* Psychology Press. [NF, CT, arMT]
(2001) Magnetic resonance imaging of reading disorders: Neural correlates of reading difficulty in adults and children. Unpublished doctoral dissertation, Stanford University Medical School. [ET]
(2002) Brain mechanisms in normal and dyslexic readers. *Current Opinion in Neurobiology* 12:178–83. [ET]
Temple, C. M., Almazan, M. & Sherwood, S. (2002). Lexical skills in Williams syndrome: A cognitive neuropsychological analysis. *Journal of Neurolinguistics* 15(6):463–95. [CP, aMT]
Temple, C. M. & Marshall, J. C. (1983) A case study of developmental phonological dyslexia. *British Journal of Psychology* 74:517–33. [NF]
Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. E. (2003) Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Science USA* 100:2860–65. [ET]
Temple, E., Poldrack, R. A., Protopapas, A., Nagarajan, S., Salz, T., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. (2000) Disruption of the neural response to rapid acoustic stimuli in dyslexia: Evidence from functional MRI. *Proceedings of the National Academy of Science USA* 97:13907–12. [ET]
Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. (2001) Disrupted neural responses to phonological and orthographic processing in dyslexic children: An fMRI study. *NeuroReport* 12:299–307. [ET]
Thal, D., Marchman, V. A., Stiles, J., Aram, D., Trauner, D., Nass, R. & Bates, E. (1991) Early lexical development in children with focal brain injury. *Brain and Language* 40:491–527. [SCL]
Thomas, K. M., Drevets, W. C., Whalen, P. J., Eccard, C. H., Dahl, R. E., Ryan, N. D. & Casey, B. J. (2001) Amygdala response to facial expressions in children and adults. *Biological Psychiatry* 49:309–16. [ET]
Thomas, M. S. C. (2000) Neuroconstructivism's promise. *Developmental Science* 3(1):35–37. [YM, rMT]
(2003) Limits on plasticity. *Journal of Cognition and Development* 4(1):95–121. [arMT]
(in press) Multiple causality in developmental disorders: Methodological implications from computational modelling. *Developmental Science.* [rMT]
Thomas, M. S. C., Grant, J., Barham, Z., Gsödl, M., Laing, E., Lakusta, L. Tyler, L. K., Grice, S., Paterson, S. & Karmiloff-Smith, A. (2001) Past tense formation in Williams syndrome. *Language and Cognitive Processes* 16:143–76. [CP, CT]
Thomas, M. S. C. & Karmiloff-Smith, A. (1999) Quo vadis modularity in the 1990s? *Learning and Individual Differences* 10:245–50. [aMT]
(2002) Modelling typical and atypical cognitive development: *Handbook of childhood development,* ed. U. Goswami. Blackwell. [CP, arMT]
(2003) Connectionist models of development, developmental disorders and

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press

individual differences. In: *Models of intelligence: International perspectives,* ed. R. J. Sternberg, J. Lautrey & T. Lubart. American Psychological Association. [arMT]

(in press) Modeling language acquisition in atypical phenotypes. *Psychological Review.* [MHC, arMT]

(in preparation) Critical periods and plasticity in feedforward connectionist networks: Implications for cognitive development and language acquisition. [aMT]

Thomas, M. S. C., Dockrell, J. E., Messer, D., Parmigiani, C., Ansari, D. & Karmiloff-Smith, A. (2003) *Naming in Williams syndrome.* (submitted). [rMT]

Thomas, M. S. C., Grant, J., Barham, Z., Gsödl, M., Laing, E., Lakusta, L., Tyler, L. K., Grice, S., Paterson, S. & Karmiloff-Smith, A. (2001) Past tense formation in Williams syndrome. *Language and Cognitive Processes* 16:143–76. [arMT]

Thompson, G. B. & Johnston, R. (2000). Are nonword and other phonological deficits indicative of a failed reading process? *Reading and Writing* 12:63–97. [NEJ]

Thulborn, K. R., Carpenter, P. A. & Just, M. A. (1999) Plasticity of language-related brain function during recovery from stroke. *Stroke* 30:749–54. [R-AM]

Toga, A. W. (2002) Neuroimage databases: The good, the bad, and the ugly. *Nature Reviews Neuroscience* 3:302–309. [S-CL]

Tyler, L. K., de Mornay Davies, P., Anokhina, R., Longworth, C., Randall, B. & Marslen-Wilson, W. D. (2002a) Dissociations in processing past tense morphology: Neuropathology and behavioural studies. *Journal of Cognitive Neuroscience* 14:79–94. [aMT]

Tyler, L. K., Randall, B. & Marslen-Wilson, W. D. (2002b) Phonology and neuropsychology of the English past tense. *Neuropsychologia* 40:1154–66. [aMT]

Udwin, O. & Yule, W. (1991) A cognitive and behavioural phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology* 13:232–44. [aMT]

Ullman, M. T. (in press) Evidence that lexical memory is part of temporal lobe declarative memory, and that grammatical rules are processed by the frontal/basal-ganglia procedural system. *Brain and Language.* [aMT]

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J. & Pinker, S. (1997) A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* 9:266–76. [aMT]

Ullman, M. T. & Gopnik, M. (1999) Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics* 20:51–117. [CP, aMT]

Ullman, M. T., Izvorksi, R., Love, T., Yee, E., Swinney, D. & Hickok, G. (in press). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgement of inflection in aphasia. *Brain and Language.* [aMT]

Van der Lely, H. K., Rosen, S. & McClelland, A. (1998) Evidence for a grammar-specific deficit in children. *Current Biology* 8:1253–58. [GFM]

Van der Lely, H. K. J. & Ullman, M. T. (2001) Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes* 16:177–218. [arMT]

Vargha-Khadem, F., Gadian, D. G., Watkins, K. R., Connelly, A., Van-Paesschen, W. & Mishkin, M. (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277:376–80. [JB]

Vargha-Khadem, F. & Mishkin, M. (1997) Speech and language outcome after hemispherectomy in childhood. In: *Paediatric epilepsy syndromes and their surgical treatment,* ed. I. Tuxhorn, H. Holthausen & H. E. Boenigk. John Libbey. [R-AM]

Vellutino, F. R. (1979) *Dyslexia: Research and theory.* MIT Press. [FR]

Vicari, S., Brizzolara, D., Carlesimo, G., Pezzini, G. & Volterra, V. (1996a) Memory abilities in children with Williams syndrome. *Cortex* 32:503–14. [aMT]

Vicari, S., Carlesimo, G., Brizzolara, D. & Pezzini, G. (1996b) Short-term memory in children with Williams syndrome: A reduced contribution of lexical-semantic knowledge to word span. *Neuropsychologia* 34:919–25. [aMT]

Volterra, V., Capirci, O. & Caselli, M. C. (2001) What atypical populations can reveal about language development: The contrast between deafness and Williams syndrome. *Language and Cognitive Processes* 16:219–39. [aMT]

Vuletic, L. & Ferrari, M. (in press) *Perspectives on unusualness: A transfer boy.* Jessica Kingsley Press. [MF]

Warrington, E. K. (1981) Concrete word dyslexia. *British Journal of Psychology* 72:175–96. [JLM]

Warrington, E. K. & McCarthy, R. (1987) Categories of knowledge: Further fractionation and an attempted integration. *Brain* 110:1273–96. [JLM]

Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. (2002a) Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain* 125(3):452–64. [R-AM]

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., Frackowiak, R. S., Mishkin, M. & Gadian, D. G. (2002b) MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain* 125(3):465–78. [R-AM]

Weinberger, D. R. (1996) On the plausibility of "the neurodevelopmental hypothesis" of schizophrenia. *Neuropsychopharmacology* 14:1S–11S. [RC]

Westermann, G. (1998) Emergent modularity and U-shaped learning in a constructivist neural network learning the English past tense. In: *Proceedings of the 20th Annual Meeting of the Cognitive Science Society,* ed. M. A. Gemsbacher & S. J. Derry. Erlbaum. http://www.cogsci.ed.ac.uk/~gert/publications/cogsci98.pdf [aMT, GW]

(1999) Single mechanism but not single route: Learning verb inflections in constructivist neural networks. *Behavioral and Brain Sciences* 22:1042–43. http://www.cogsci.ed.ac.uk/~gert/publications/bbs-clahsen.pdf [GW]

(2000) Constructivist neural network models of cognitive development. Ph.D. thesis, Division of Informatics, University of Edinburgh. http://www.cogsci.ed.ac.uk/~gert/publications/thesis.pdf [GW]

Whitaker-Azmitia, P. M. (2001) Serotonin and brain development: Role in human developmental diseases. *Brain Research Bulletin* 56:479–85. [R-AM]

White, S., Milne, E., Rosen, S., Hansen, P. C., Swettenham, J., Frith, U. & Ramus, F. (submitted) The role of sensorimotor processing in dyslexia: A multiple case study of dyslexic children. [FR]

Woods, B. & Teuber, H-L. (1978) Early onset of complementary specialization of the cerebral hemispheres in man. *Transactions of the American Neurological Association* 98:113–17. [SCL]

Wright, B. A., Buonomano, D. V., Mahncke, H. W. & Merzenich, M. M. (1997a) Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience* 17:3956–63. [PG]

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M. & Merzenich, M. M. (1997b) Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature* 387:176–78. [PG]

Zabell, C. & Everatt, J. (2002) Surface and phonological subtypes of adult developmental dyslexia. *Dyslexia* 8:160–77. [FR]

Zorzi, M., Houghton, G. & Butterworth, B. (1998) Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance* 24:1131–61. [aMT]

Zukowski, A. (2001) Uncovering grammatical competence in children with Williams syndrome. Unpublished doctoral dissertation, Boston University. [rMT]

https://doi.org/10.1017/S0140525X02320132 Published online by Cambridge University Press