13

QUO VADIS MODULARITY IN THE 2020s?

Michael S.C. Thomas and Daniel Brady

Both of us had the privilege of being postdoctoral fellows for Annette: one of us (MT) when Annette set up her own research unit in 1998 after leaving MRC Cognitive Development Unit, the other (DB) in Annette's final project, an investigation of infant development in Down syndrome (DS) and its link with lifelong elevated risk for Alzheimer's disease in the syndrome (Karmiloff-Smith et al., 2016). In this chapter, we consider the idea of *modularity* because it figured in both our collaborations with Annette (henceforth, AKS) – but between the two collaborations, the usage of the concept had changed.

For MT, his first publication with AKS was a commentary that challenged the use of the Fodorian sense of modularity (Fodor, 1983) to explain deficits in developmental disorders, as it featured in four papers then appearing in the journal *Learning and Individual Differences*. We entitled our commentary, 'Quo vadis modularity in the 1990s?' (or, for non-Latin speakers, 'modularity, where are you going?'; Thomas & Karmiloff-Smith, 1999). For DB, modularity featured in his use of Graph Theoretic methods to analyse time series correlations between the electroencephalogram (EEG) data recorded from electrodes spread across the scalps of infants with DS as they watched a video. Graph Theory was deployed to understand more about the functional connectivity of these infants' brains - which regions appeared to be working with other regions - and the relation of functional connectivity both to the infants' chronological age and to their cognitive ability.

For this Festschrift volume, as we enter the third decade of the 21st century, we can ask this question once again: *quo vadis, modularity?* In this chapter, we will see that over the decades and with the advance of neuroscience, the notion of modularity has been recast: from an *a priori* design principle which it would make good sense for potential cognitive systems to employ, to a data-driven concept based on how the brain is actually working.

Modularity as an advantageous *a priori* design principle

In 1998, when MT joined AKS's new research group, the Neurocognitive Development Unit, based at London's Institute of Child Health, AKS set him three goals. The first was to advance research on neurodevelopmental disorders, pursing ideas in her seminal paper *The Key to Developmental Disorders Is Development Itself* (Karmiloff-Smith, 1998). This involved investigating the genetic syndrome Williams syndrome (WS), characterised by overall learning disability, but relative strengths in language and face recognition, and weaknesses in visuospatial cognition. But, AKS asked, was language truly a strength? Was language developing typically, independently of the rest of cognition? How come it was not at chronological age level? How did language fit with the overall pattern of learning disability? This line of work was reasonably successful: a set of studies suggested that language skills in WS were broadly in line with overall mental age, but with subtle patterns of unevenness – that is, atypicality – such as in semantic compared to phonological skills (Thomas et al., 2001; Laing et al., 2005; Ansari et al., 2003; Thomas & Karmiloff-Smith, 2005; Thomas, Forrester, & Richardson, 2006; Thomas, 2006). A similar pattern was found in detailed investigations of face processing, although the behavioural strength here was more marked, and there was greater evidence for atypical underlying cognitive processes (Annaz, Karmiloff-Smith, Johnson, & Thomas, 2009; Grice et al., 2001; Karmiloff-Smith et al., 2004).

The second goal AKS set for MT was to begin a programme of computational modelling. Since there existed at the time reasonable computational models of typical cognitive development, based on artificial neural networks, was it possible to build computational models of developmental disorders, such as of language development in WS? This work was reasonably successful as well. MT built a model of the acquisition of morphosyntax in WS, where different types of atypical constraint (arising from the structure of semantic representations, phonology representations, or computational properties learning to mediate the relationship between language representations) could shape developmental trajectories in ways that did or did not reflect those observed in individuals with WS (Thomas & Karmiloff-Smith, 2003a). The model made more concrete the theoretical notion that deficits in developmental disorders were the result of atypical constraints acting on the developmental process by specifying that process in neurocomputational terms.

The third goal proved more challenging. AKS wanted MT to build a computational model of *representational redescription* (RR). RR was AKS's account of how children could develop cognitive flexibility. It was based on the proposal that there was a general phase of cognitive development lying beyond behavioural mastery, where the knowledge underlying mastery of a given ability becomes both more accessible to conscious awareness (or more 'explicit') and also more open to flexible application (Karmiloff-Smith, 1992). For MT, this work did not go so well. He ran into a stumbling block when attempting to think about RR in computational terms. Was RR, on the one hand, a process *intrinsic* to a cognitive system that was acquiring some behaviour, such that even when the system had achieved behavioural mastery, it somehow went on to reorganise its knowledge into a form that could be 'offered up' to other cognitive systems, such as those involved in cognitive control, verbal report, or metacognition? What would drive this reorganisation, if mastery was already achieved? Or was RR, on the other hand, a process concerned with the enabling of latent, pre-existing *connectivity between* different cognitive systems, so that once the internal representations of the target system began to exhibit interesting internal structures or covariations relevant to other systems (cognitive control, verbal report, etc.), these latent connections would begin to strengthen? Could RR involve both intrinsic and connectivity elements? The dilemma was hard to resolve, but theoretical clarification was necessary prior to proceeding with implementation in a computational model.

Meanwhile, the attention of both AKS and MT was focused on atypical development, and RR moved to the backburner. In her final years, AKS was perhaps on the verge of returning to RR: she noted the possibility that brain imaging might test the RR hypothesis through studying emerging hierarchical network structure in 'resting state' functional brain imaging data (Karmiloff-Smith, 2010); and viewed Dehaene's proposal of a global workspace for flexible use of knowledge (Dehaene et al., 2014) as a possible brain basis for the concept of RR (Karmiloff-Smith, 2015). It is fantastic to see McClelland and his team (this volume) taking RR ideas forward in a neurocomputational framework, as well as Hughes and Edgin (this volume), considering a link between RR and sleep, and Plunkett (this volume) consider RR in relation to vocabulary development. No competing theory has emerged in the meantime to explain the developmental emergence of explicit knowledge and cognitive flexibility in children.

However, behind all three of these goals lay a pervasive idea that was key to AKS's thinking at the time – that of *modularity*. There were, in the 1990s, explanations of uneven cognitive profiles in developmental disorders that made reference to innate modules. For AKS, this was the wrong way to go. For her, modules were an outcome of development, not a precursor to it (Karmiloff-Smith, 1992, 1994, 1998, 2009, 2015). Multiple convergent methods within cognitive neuroscience, including computational modelling and functional brain imaging, would be necessary to specify the nature of developmental processes. And even when modularity had emerged in a given domain, somehow development could move beyond it, so that children could begin to show insight into their own behaviour.

Let us take a moment to recapitulate the original, Fodorian sense of what a module was (Fodor, 1983), around which this debate revolved. The theory was based on a cognitivist approach that sought to cleave *cognition* from *perception*, and consider each as a decomposable into a set of specialised sub-processes. Fodor constructed his arguments by appealing to what would constitute a sensible way to design a cognitive system, grounded by (limited) empirical reference to intuitive phenomena such as visual illusions, and supported by thought experiments. The key idea was *encapsulation of information*.[1] Specialised and dedicated processing systems would contain their own dedicated information; these systems would be impervious to information from outside in the rest of the cognitive system. Intuitively, it is evident that we continue to see visual illusions even when we know they are illusions. For Fodor, this is because low-level vision is a module – fast, automatic, self-contained, independently delivering its results to the rest of cognition. Indeed, it would be a poor design for a perceptual system to rely on background knowledge: if you see a tiger leaping at you while out shopping, you need to duck, not think about whether it's likely you'll encounter a tiger in the shopping mall. Fodor distinguished perceptual modules from the central cognitive system. The central system is involved in thought and reasoning, where in his view background knowledge is crucial to its operation. Should you go to the shops today? A vast range of factors might be relevant (including the chance of tigers). Fodor suggested that somewhere in between perception and the central system, there might be *cognitive modules* for self-contained cognitive domains such as language, or perhaps syntax. Along with properties of fast processing, automatic operation and encapsulated information, these cognitive modules would likely be innately specified, hardwired (that is, neurally specific), and not assembled from other cognitive components. Modularity amounted to a *cluster* concept, a set of properties that Fodor thought would likely hang together within the architecture of the cognitive system.

The idea of cognitive modules fitted with what was observed in cases of specific loss of abilities following acquired focal brain damage in adults – as if a bit of the cognitive system were implemented by a bit of the brain, and damage to that bit of the brain could produce a specific deficit in the corresponding ability. But in 1990s, some researchers were trying to apply this adult-based framework to uneven cognitive profiles in developmental disorders, such as autism, developmental language disorder (then called Specific Language Impairment), dyslexia, and prosopagnosia. For AKS, however, Fodor's framework left out a key aspect: development. The increasing understanding of functional brain development, as articulated in *Rethinking Innateness* (Elman et al., 1996; see Johnson, this volume) suggested that although modularity may be characteristic of the adult state, it was likely a product of developmental processes from an initially less differentiated and specialised system. Therefore, the static adult framework was inappropriate to characterise developmental disorders. AKS and MT spent some time making this argument more concrete by comparing the different profiles of deficits in computational models of developmental deficits versus models of acquired deficits (Thomas, & Karmiloff-Smith, 2002; Thomas, & Richardson, 2006; Thomas *et al.* 2006).

In one sense, Fodor's notion of modularity is hard to argue against – and that sense is *falsification*. The six proposed properties of a module were not necessary and sufficient conditions, just properties likely to cluster together, based on what would be a sensible way to design a cognitive system (and some supporting intuitive phenomena). The absence of one or more feature from the cluster for some target behaviour would not be enough to reject the idea. (For example, visual word recognition is learned, not innate, but shows other characteristics of modular function – it is fast, automatic, can be specifically lost after acquired brain damage in alexia – but its lack of innateness would not serve to falsify the cluster concept of modularity; maybe some modules can be learned, Fodor would say.) So the idea of modularity was never rejected.

The main flaw with the original concept of modularity, and indeed of the cognitivist approach of the 1980s and early 1990s more generally, was that it dealt in ways that cognitive systems *could* work – hence the close contemporary affiliation of cognitive theory with artificial intelligence research. However, it did not necessarily deal with the way the cognitive system *does work*, which is constrained by its actual biology and the attendant evolutionary history that shaped the biology. The actual way the cognitive system works has become more apparent with progress in cognitive neuroscience over the intervening years. The brain operates via content-specific systems and modulatory systems (both focal and diffuse) which coordinate their operation; it operates via distributed networks of interacting systems which are configured dynamically according to task and context; hierarchical systems learn to extract ever higher-order invariances from sensory input or to generate motor output; cortical systems are continuously interacting with limbic systems whose structures reflect the deeper evolutionary goals of the species through the emotions. The increasing influence of cognitive neuroscience led some commentators to revise how the notion of modularity should usefully be deployed. Already by the end of the century, Coltheart (1999) was arguing that modularity should be ascribed to a cognitive system solely on the basis of the system being *domain specific*. This did not even preserve the notion of information encapsulation that was the motivating factor in Fodor's original conception.

Modularity as a data-driven description of brain structure and function

Let us pause and let some time pass. AKS has pursued her neuroconstructivist agenda, using cross-syndrome comparisons in an attempt to triangulate the developmental constraints that operate on cognition, in domains such as language and face recognition, and has continued to integrate wider sets of developmental cognitive neuroscience methods. MT has left to broaden his approach to thinking about different types of cognitive variation beyond disorders, such as intelligence and giftedness, and how to link cognitive neuroscience findings to education. Less is heard of modularity, and less still of RR. We come to Annette's final project.

The London Down Syndrome (LonDownS) Consortium was launched in 2013, funded by a Wellcome Trust Strategic grant. Its aim was to investigate the link between Alzheimer's disease and DS. Adults with DS are at greater risk of developing Alzheimer's disease (AD) compared to the general population, and the onset of AD tends to occur at an earlier age (Head, Powell, Gold, & Schmitt, 2012). One reason is that one of the key genes implicated in AD, the amyloid beta precursor protein (*APP*) gene, lies on chromosome 21; there are three copies of chromosome 21 in DS, potentially disrupting the gene dosage for all the genes on this chromosome. The consortium took a lifespan perspective, both seeking to understand the degenerative disease of AD and to predict as early as in infancy which individuals with DS are most at risk of the disease in adulthood. AKS worked with the infant and child cohorts, investigating whether risk factors for AD would be reflected in early developmental trajectories (the answer: sometimes, sometimes not; D'Souza et al., 2020; Thomas et al., 2020 submitted). She also addressed the difficulties of developing cognitive tests for infants and children with DS and how to match these with the tests being developed for the mouse models of the syndrome (see D'Souza et al., Chapter 16).

The LonDownS project was a large affair, five years in length including longitudinal follow-ups of children and adults, and a multidisciplinary team set to investigate dementia in adults, mouse models of DS, genetics of AD risk, in vitro modelling of DS neurons cultured from induced pleuripotent stem cells, and of course, AKS's study of infancy in DS, with the largest UK cohort that she could manage to put together (over 100, in the end, with around half of those being followed up longitudinally) (see Karmiloff-Smith et al., 2016). AKS embraced the complexity of development, instead of attempting to simplify, seeking to measure her cohort across multiple behaviours, with multiple methods (behavioural, eye-gaze monitoring, brain imaging, genetics), including measures of parent-child interaction and sleep. DB, the second author of this chapter, joined her team to focus on the analysis of brain imaging data for the infants with DS and typically developing controls.

It is here that the notion of modularity resurfaces, in a different guise. A new technique has emerged to analyse both structural and functional brain imaging data, one that focuses on networks, construed within the mathematical framework of *Graph Theory*. We will go into a little more detail in the following paragraphs to convey the technical innovations of Graph Theory and the potential link of imaging data to cognition. Graph Theory takes data summarising the correlations between many variables and translates these into graphs, where each variable is a node in the network, and the patterns of correlations are reflected in the links joining the nodes. For brain imaging data, the nodes may be parcellated brain regions (structural analyses), signals from optical or electrical sensors spread over the scalp, or blood oxygenation levels from specific brain regions (functional analyses) (see Sporns, 2014). In structural analyses, the key question is whether two regions are connected, while for functional analyses, the key question is whether the activity of pairs of nodes is correlated over time, and therefore presumably involved in similar processes.

Once a network has been generated from a set of correlational data, a range of metrics can be generated describing the graph. Three of the most widely used measures are whether a network is characterised by *hubs*, the *average path length* between any two nodes, and the level of *modularity* (Vertes & Bullmore, 2015). Hubs are nodes that have a number of links that greatly exceeds the average, suggesting they are key locations for communication between less connected areas. The average path length is a measure of efficiency, indicating how many links have to be traversed to move from any node to any other node. Short average path length indicates a network architecture that facilitates fast communication. Modularity in graph theory is a measure that assesses the division of a network into groups (also called *clusters* or *communities*): networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules (Newman, 2006). In functional terms, the activity of a set of nodes (call it, group A) may be highly correlated among each other, but uncorrelated to another group (B), while the nodes in B in turn show high correlations among each other. Figure 13.1 shows examples of networks exhibiting low modularity and high modularity, respectively. In brain terms, therefore, a module now comes to mean sets of brain areas which appear to be strongly physically connected among themselves but not to other regions (structural modularity), or sets of regions whose activity seems to be highly correlated to each other but not to other regions (functional modularity). Modularity in this sense indeed appears to be a property of biological systems (Newman, 2006; Sporns, 2014).

Analyses of brain imaging data employing graph theoretic techniques have found that graph metrics appear to correspond to interesting cognitive dimensions. When Crossley, Mechelli, and Vertes (2013) examined co-activated regions across a large set of functional imaging studies, they found a coactivation network that was modular, with occipital, central, and default-mode modules predominantly coactivated by specific cognitive domains including, respectively, perception, action, and emotion. Networks appear to alter across development and are sensitive to differences in atypical development (Morgan, White, Bullmore, & Vertes, 2018; Vertes & Bullmore, 2015; Whitaker et al., 2016). When Hilger, Ekman, Fiebach, and Basten (2017) correlated measures of general intelligence to networks

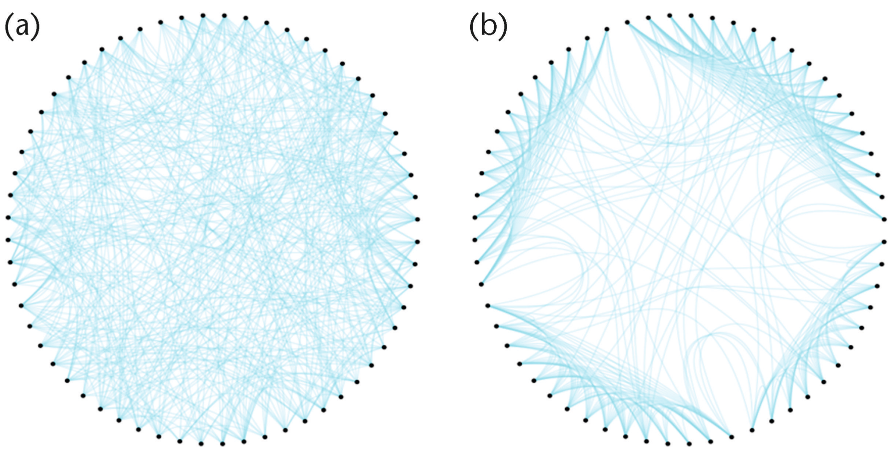

FIGURE 13.1 Illustrative network analysis results. Both networks have the same number of nodes and vertices, but network (a) has low modularity while network (b) has high modularity

constructed from resting-state functional magnetic resonance imaging data in a sample of over 300 individuals, they found brain networks characterised by substantial modularity. Intelligence was not associated with global features of modularity, such as the number or size of modules. However, intelligence was associated with node-specific measures of within- and between-module connectivity, particularly in frontal and parietal brain regions that have previously been linked to intelligence.

AKS, DB and their team acquired resting-state EEG data from the cohort of infants and toddlers with DS as they watched a short 2-minute video. The prospect was that this kind of sensitive measure might detect differences in underlying brain processes reflecting later risk/resilience to AD even before they are manifest in behaviour, on the basis that in typical development, genes conferring additional risk for AD have been observed to influence white matter growth as early as 2–25 months of age (Dean, Jerskey, & Chen, 2014). The analyses of these data are still on-going, and we present here some of the provisional results to illustrate the method, and also some of its limitations.

In EEG analyses, the small voltages changes produced at the scalp by brain activity are measured in the millisecond range. This activity can be broken down into the amount of energy produced (power) and phase in different frequency bands, typically delta (0.1–3 Hz), theta (4–7 Hz), alpha (8–15 Hz), beta (16–31 Hz), and, gamma (32–100 Hz). Comparing the phase of the activity between nodes (in the case of EEG this is usually sensors) can be used as a way of quantifying the degree of connectedness between the nodes, and undertaking this for all combinations of nodes allows creation of a network or graph. Specifically, the degree of correlation between nodes was determined by the weighted Phase lag index (https://www.nbtwiki.net/doku.php?id=tutorial:phase_lag_index), so the graph measures were not built from the overall energy at a particular frequency band (the power) but instead the similarity between phases of EEG activity at the nodes.

The goal here was to explore whether individual differences in the properties of these networks correlated to measures of cognition. Figure 13.2 shows scatter plots linking each network's cluster coefficient in each frequency band to the children's receptive language ability, as measured by the Mullen test (Mullen, 1995). It is evident here that no relationship was present for any frequency band, despite extensive variation in language ability. Figure 13.3 focuses on just one band, beta, and examines the change in five different network metrics across chronological age.[2] Cluster coefficient, global and local efficiency each reduced with age, while mean shortest path length increased, and modularity showed no change. Figure 13.4 contrasts the correlation of modularity and mean shortest path length with overall mental age on the Mullen: there was no relationship with modularity, but an increase in shortest path length with increasing mental age. These results replicate the developmental sensitivity of network metrics and the absence of a correlation of overall modularity with intelligence observed by Hilger et al. (2017). The increase in shortest path length both with age and with mental age is

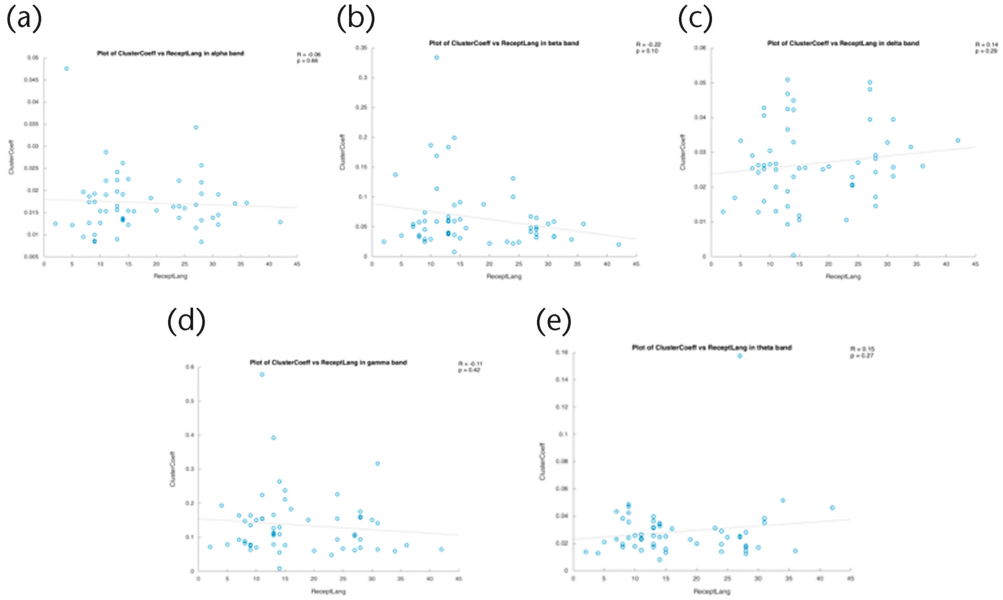

FIGURE 13.2 Example network analysis results from EEG data as infants and toddlers with DS watched a short video. Graphs were generated on 40 seconds of data for around 40 children for 5 frequency bands (delta, theta, alpha, beta, and gamma). Graphs were proportionally thresholded so 20% of strongest connections remain. Data show the relationship between the cluster coefficient (the degree to which nodes in a graph tend to cluster together) and receptive language ability at different frequency band: (a) alpha (b) beta (c) delta (d) gamma, and (e) theta

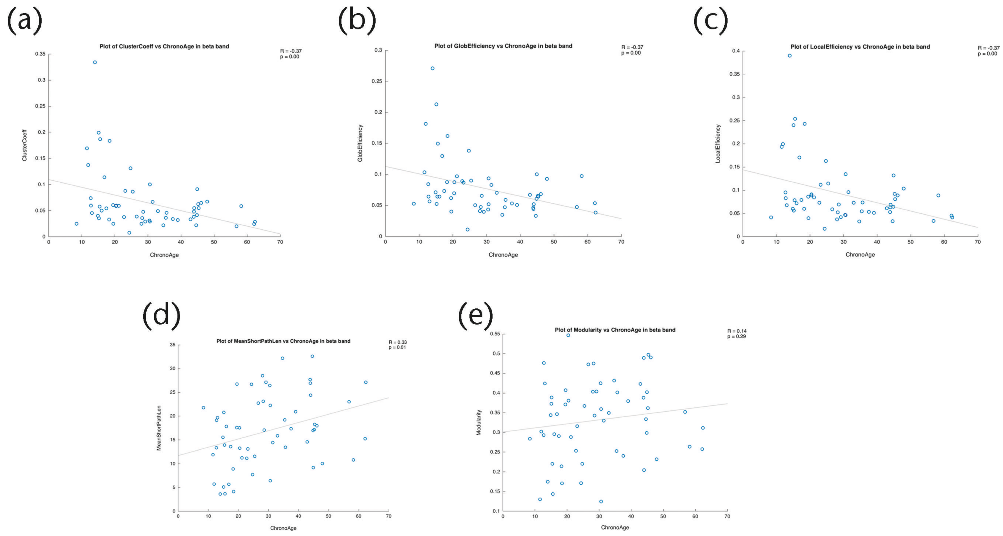

FIGURE 13.3 Example network analysis results from EEG data as infants and toddlers with DS watched a short video. Data show the relationship between different graph metrics and chronological age in the beta frequency band. (a) Cluster coefficient (b) global efficiency (c) local efficiency (d) mean shortest path length, and (e) modularity

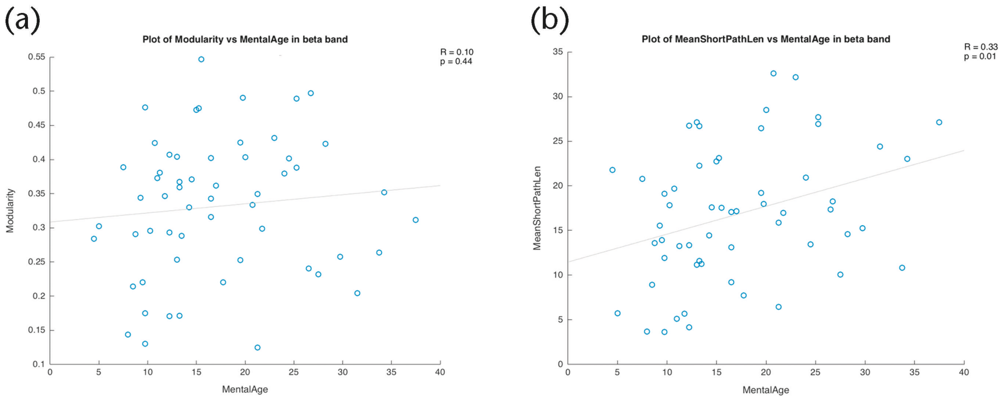

FIGURE 13.4 Relationship between (a) the modularity metric and mental age in the beta frequency band; and (b) the mean shortest path length and mental age in the beta band

suggestive of worsening efficiency in the network, which is consistent with the emerging developmental delay observed in DS across early childhood (D'Souza et al., in preparation).

However, these data remain provisional because they need to be compared to the patterns observed in typical development (currently underway) and because of their sensitivity to decisions made in constructing the graphs. For instance, EEG data are noisy: decisions need to be made in the quality of the signal that is included into the analyses; in unweighted graphs, links are either present or absent: decisions need to be made on strength of the correlation between any two nodes required before a link is generated. Moreover, as with any neuroscience data generating a large number of comparisons, there is the challenge of how to correct for elevated risk of Type 1 statistical errors (false alarms). Many of the other comparisons in other frequency bands were null. In these analyses, we have 5 different measures of cognition from the Mullen test plus chronological age, 5 different graph measures, and 5 different frequency bands (150 comparisons). Without an *a priori* hypothesis, the appropriate significance level (alpha) would be .0003 rather than the standard of .05, which represents a stiff test for noisy data. The sense is that such analyses might give fleeting glimpses into emerging architectures, but would stand in need of replication in other samples before being viewed as robust.

Final thoughts

We see here, then, the notion of modularity recast. It is now a data-driven concept based on how the brain is actually working, rather than an *a priori* design principle, one that would seem to make sense for potential cognitive systems to employ. While functional brain imaging presumes some degree of domain-specificity of local neural processing, overall function emerges from the interaction of many regions. Moreover, these regions may not necessarily align with the modules or component processes articulated in cognitive theories (Price & Friston, 2005). Since modularity is now data driven, it is an empirical issue whether it is a property that alters across development or varies with intellectual ability.

In our view, the routine use of brain imaging data to complement cognitive theories of development makes it unlikely we will return to the 1980s Fodorian notion of modularity, as research proceeds into the 2020s. For the mammalian brain, evolution does not appear to have followed the principle of componential design that might have yielded the benefits Fodor had in mind. Developmental studies suggest that specialisation is in part an emergent property influenced either by experience or internally generated neural activity, so called *interactive specialisation* (Johnson, this volume). The on-line interactivity of different brain regions characterised by graph theory does not seem consistent with the key idea of information encapsulation championed by Fodor.

However, this reality does not negate Fodor's key insight – we still need to characterise how, in brain terms, fast perception is different to slow thinking, why optical illusions are resistant to knowledge, or why the arcane knowledge underlying the processing of syntax appears unconscious and specific to language. To take one contemporary puzzle: when teenagers spend hundreds of hours playing action video games, they change their brains and enhance their visual perception skills, particularly for peripheral attention. Yet, despite this massive effort, and the observed behavioural change, brain imaging suggests that no enhancement has occurred to 'bottom-up', automatically triggered, exogenous attention, only to voluntarily exerted top-down attention (Altarelli, Green, & Bavelier, 2020). Why should that be? What is different about exogenous attention? Altarelli et al. report that it remains largely unknown why initial orienting under the control of exogenous attention is not altered by intensive action video game play, despite it being advantageous to improve this skill for game performance. They argue one possible explanation is that orienting develops very early during development and is in a large part mediated by sub-cortical structures, such as the nucleus basalis or the pulvinar; these structures may be less plastic later in life than cortical brain regions that support top-down attentional mechanisms. This account seems close to the kind of dimension of cognition that Fodor had in mind: a fast, automatic, early developing, modular perceptual function, to be contrasted with a slower, more strategic and informationally rich system exerting top-down control. Similarly, we are yet to construct convincing explanations for the phenomena that led AKS to propose representational redescription in the early part of her research career – that in many cognitive domains (but not all), expertise can be accompanied by explicit knowledge and greater flexibility in a phase of development extending beyond behavioural mastery of the task itself.

These challenges, posed by theoretical adversaries such as AKS and Fodor, remain for the future. Yet it is our belief that the multi-disciplinary approach advocated by AKS, with development at its core, is best positioned to yield the answers we seek.

Acknowledgements

We would like to acknowledge all of Annette's team on the LonDownS project who contributed to this research including George Ball, Joanne Ball, Amelia Channon, Charlotte Dennison, Dean D'Souza, Hana D'Souza, Maria Eriksson, Jennifer Glennon, Kate Hughes, Amanda Lathan, Esha Massand, Beejal Mehta, Lidija Nikolova, Jessica Schulz, Laura Simmons, Margaux Trombert, and Kate Whitaker. The scripts for collection of the resting state EEG data were kindly shared by the BASIS project.

The LonDownS project is supported by a Wellcome Trust Strategic Award [098330/Z/12/Z], the Baily Thomas Charitable Fund, and the Jerome LeJeune Foundation. The London Down Syndrome (LonDownS) Consortium principal investigators are Andre Strydom (chief investigator), Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK, and Division of Psychiatry, University College London, London, UK; Elizabeth Fisher, Department of Neurodegenerative Disease, Institute of Neurology, University College London, London, UK; Dean Nizetic, Blizard Institute, Barts and the London School of Medicine, Queen Mary University of London, London, UK, and Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore; John Hardy, Reta Lila Weston Institute, Institute of Neurology, University College London, London, UK, and UK Dementia Research Institute at UCL, London, UK; Victor Tybulewicz, Francis Crick Institute, London, UK, and Department of Medicine, Imperial College London, London, UK; Annette Karmiloff-Smith, Birkbeck University of London, London, UK (deceased); Michael Thomas, Birkbeck University of London, London, UK; and Denis Mareschal, Birkbeck University of London, London, UK.

Notes

1 Fodor (1985): 'A **module** is (inter alia) an informationally encapsulated computational system – an inference-making mechanism whose access to background information is constrained by general features of cognitive architecture, hence relatively rigidly and relatively permanently constrained' (p. 3).

2 Graph Theory network metric definitions: *Cluster coefficient*: A measure of the degree to which nodes in a graph tend to cluster together. To what extent are the nodes to which a given node is connected, connected to each other? *Global efficiency*: The efficiency of a network measures how efficiently it exchanges information. Global efficiency quantifies the exchange of information across the whole network where information is concurrently exchanged. *Local efficiency*: The local efficiency quantifies a network's resistance to failure on a small scale. The local efficiency of a node reflects how well information is exchanged by its neighbours when it is removed. *Mean shortest path length*: The average number of steps along the shortest paths for all possible pairs of network nodes, which indexes the information flow across a network. *Modularity*: **Modularity** measures the strength of division of a network into groups. Networks with high modularity have dense connections between the nodes within modules, but sparse connections between nodes in different modules.

References

Altarelli, Green C. S., & Bavelier, D. (2020). Action video games: from effects on cognition and the brain to potential educational applications. In M. S. C. Thomas, D. Mareschal & I. Dumontheil (Eds.), *Educational Neuroscience: Development Across the Lifespan*. London: Psychology Press.

Annaz, D., Karmiloff-Smith, A., Johnson, M. H., & Thomas, M. S. C. (2009). A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome and Williams syndrome. *Journal of Experimental Child Psychology*, *102*, 456–486.

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.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, *3*(*3*), 115–120.

Crossley, N. A., Mechelli, A., Vertes, P. E., et al. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences USA*, *110*(28), 11583–11588.

Dehaene, S., Charles, L., King, J. R., & Marti, S. (2014). Toward a computational theory of conscious processing. *CurrOp Neurobio*, *25*, 76–84.

D'Souza, H. D., Mason, L., Mok, K. Y. , Hardy, J., Karmiloff-Smith, A.Mareschal, D. the LonDownS Consortium Strydom, A., & Thomas, M. S. C. (2020), Differential associations of apolipoprotein E ε4 genotype with attentional abilities across the life span of individuals with down syndrome, *JAMA Network Open*, *3*(9), e2018221. doi: 10.1001/jamanetworkopen.2020.18221

D'Souza, H., Karmiloff-Smith, A., Mareschal, D., & Thomas, M. S. C. (2019). *The Down syndrome profile emerges gradually over the first years of life*. Manuscript submitted for publication.

Dean, D. C. 3rd, Jerskey, B. A., Chen, K., et al. (2014). Brain differences in infants at differential genetic risk for late-onset Alzheimer disease: A cross-sectional imaging study. *JAMA Neurology*, *71*(1), 11–22.

Elman, J., Karmiloff-Smith, A., Bates, E., Johnson, M., Parisi, D., & Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. A. (1985). Precis of the modularity of mind. *Behavioral and Brain Sciences*, *8*, 1–42.

Grice, S. J., Spratling, M. W., Karmiloff-Smith, A., Halit, H., Csibra, G., de Haan, M., et al. (2001). Disordered visual processing and oscillatory brain activity in autism and Williams syndrome. *Neuroreport*, *12*, 2697–2700.

Head, E., Powell, D., Gold, B. T., & Schmitt, F. A. (2012). Alzheimer's Disease in Down Syndrome. *European Journal of Neurodegenerative Disease*, *1*(3), 353–364.

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017). Intelligence is associated with the modular structure of intrinsic brain networks. *Scientific Reports*, 7, 16088. http://doi.org/10.1038/s41598-017-15795-7

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press/Bradford Books.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398. http://doi.org/10.1016/S1364-6613(98)01230-3

Karmiloff-Smith, A. (2010). Neuroimaging of the Developing Brain: Taking "Developing" Seriously. *Human Brain Mapping*, *31*, 934–941.

Karmiloff-Smith, A. (2013). Challenging the use of adult neuropsychological models for explaining neurodevelopmental disorders: Developed versus developing brains. *The Quarterly Journal of Experimental Psychology*, *66*, 1–14.

Karmiloff-Smith, A. (2015). An alternative to domain-general or domain-specific frameworks for theorizing about human evolution and ontogenesis. *AIMS Neuroscience*, *2*(2), 91–104.

Karmiloff-Smith, A., (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, *45*(1), 56–63.

Karmiloff-Smith, A., Al-Janabi, T., D'Souza, H., Groet, J., Massand, E., Mok, K., … Strydom, A. (2016). The importance of understanding individual differences in Down syndrome. *F1000Research*, *5*, 389. http://doi.org/10.12688/f1000research.7506.1

Karmiloff-Smith, A., Thomas, M. S. C., Annaz, D., Humphreys, K., Ewing, S., Grice, S., Brace, N., Van Duuren, M., Pike, G., & Campbell, R. (2004). Exploring the Williams syndrome face processing debate: The importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(7), 1258–1274.

Laing, E., Grant, J., Thomas, M. S. C., Parmigiani, C., Ewing, S., & Karmiloff-Smith A. (2005). Love is… an abstract word: The influence of lexical semantics on verbal short-term memory in Williams syndrome. *Cortex*, *41*(2), 169–179.

Morgan, S. E., White, S. R., Bullmore, E. T., & Vertes, P. E. (2018). A network neuroscience approach to typical and atypical brain development. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(9), 754–766.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8577–8696.

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service Inc.

Price, C. J., & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, *22*(3), 262–275. http://doi.org/10.1080/02643290442000095

Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, *17*, 652–660. http://doi.org/10.1038/nn.3690

Thomas, M. S. C. (2006). Williams syndrome: Fractionations all the way down? *Cortex*, *42*, 1053–1057.

Thomas, M. S. C. & Karmiloff-Smith, A. (1999). Quo vadis modularity in the 1990s? *Learning and Individual Differences*, *10*(*3*), 245–250.

Thomas, M. S. C. & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioral and Brain Sciences*, *25*(6), 727–788.

Thomas, M. S. C. & Karmiloff-Smith, A. (2003a). Modelling language acquisition in atypical phenotypes. *Psychological Review*, *110*(4), 647–682.

Thomas, M. S. C. & Karmiloff-Smith, A. (2003b). Connectionist models of development, developmental disorders and individual differences. In R. J. Sternberg, J. Lautrey, & T. Lubart (Eds.), *Models of Intelligence: International Perspectives*, (pp. 133–150). American Psychological Association.

Thomas, M. S. C., & Johnson, M. H. (2018). Karmiloff-Smith from Piaget to neuroconstructivism. In: A. Karmiloff-Smith, M. S. C. Thomas & M. H. Johnson, M. H. (Eds.), *Thinking Developmentally from Constructivism to Neuroconstructivism: Selected Works of Annette Karmiloff-Smith*. London: Routledge.

Thomas, M. S. C., & Karmiloff-Smith, A. (2005). Can developmental disorders reveal the component parts of the human language faculty? *Language Learning and Development*, *1*(1), 65–92.

Thomas, M. S. C., & Richardson, F. (2006). Atypical representational change: Conditions for the emergence of atypical modularity. In Y. Munakata, & M. Johnson (Eds.), *Attention & Performance XXI: Processes of Change in Brain and Cognitive Development* (pp. 315–347). Oxford: Oxford University Press.

Thomas, M. S. C., Dockrell, J. E., Messer, D., Parmigiani, C., Ansari, D., & Karmiloff-Smith, A. (2006). Speeded naming, frequency and the development of the lexicon in Williams syndrome. *Language and Cognitive Processes*, *21*(6), 721–759.

Thomas, M. S. C., Forrester, N. A., & Richardson, F. M. (2006). What is modularity good for? In Proceedings of The 28th Annual Conference of the Cognitive Science Society (pp. 2240–2245), July 26-29, Vancouver, BC, Canada.

Thomas, M. S. C., Grant, J., Barham, Z., Gsodl, 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* (2/3), 143–176.

Thomas, M. S. C., Ojinaga Alfageme, O., D'Souza, H., Patkee, P. A., Rutherford, M. A., Mok, K. Y., Hardy, J., the LonDownS Consortium, & Karmiloff-Smith, A. (2020). A multi-level developmental approach to exploring individual differences in Down syndrome: genes, brain, behaviour, and environment, *Research in Developmental Disabilities*, *104*, 103638. doi: 10.1016/j.ridd.2020.103638. Epub 2020 Jul 10. PMID: 32653761; PMCID: PMC7438975.

Thomas, M. S. C., Purser, H. R. M., & Richardson, F. M. (2013). Modularity and developmental disorders. In: P. D. Zelazo (Ed), *Oxford Handbook of Developmental Psychology*. Oxford: Oxford University Press.

Vertes, P. E., & Bullmore, E. T. (2015). Annual Research Review: Growth connectomics–the organization and reorganization of brain networks during normal and abnormal development. *Journal of Child Psychology and Psychiatry*, *56*(3), 299–320.

Whitaker, K. J., Vertes, P. E., et al. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proceedings of the National Academy of Sciences USA*, *113*(32), 9105–9110. DOI: 10.1073/pnas.1601745113