Educational neuroscience and learning

Andrew Tolmie, Psychology and Human Development, UCL Institute of Education

Biographical details

Andrew Tolmie is Chair of Psychology and Human Development at the UCL Institute of Education, University College London. He is also Deputy Director of the UCL/Birkbeck Centre for Educational Neuroscience, and was Editor of the *British Journal of Educational Psychology* from 2007-12. He is a developmental psychologist with longstanding interests in the growth of children's conceptual representations and behavioural skills, and the relationships between these, particularly in the elementary school age range. Most of his work has focused on educationally-relevant topics and settings, with a substantial emphasis on science learning, but also on the acquisition of road-crossing skills among children. He was recently a member of a Royal Society working group reporting on science and mathematics education 5-14, inputting to this in particular on the role of teacher-learner interaction and peer collaboration in effecting representational growth.

Abstract

Advances in neuroscience have had a profound impact on psychological understanding of learning, and evidence continues to be accumulated in relation to the complex characteristics which it exhibits. Educational neuroscience attempts to coordinate evidence from behavioural and neuroimaging studies in order to obtain a more complete understanding of learning that can then be used to specify the pedagogical approaches and educational systems that will support these most effectively. This chapter offers a realistic portrayal of the ways that such evidence might influence teaching now and in the future. The translational nature of the field poses challenges because of its requirement for collaboration between researchers and educational practitioners. However, research in literacy, number development, science learning and executive function illustrates the potential of the field to explain both typical and atypical learning in a coherent fashion and to identify novel pedagogical strategies that fully address individual variation in capability.

Keywords

Neuroscience, learning, individual variability, literacy, number development, science learning, executive function

Introduction: cognitive neuroscience and its tools

Cognitive neuroscience (as opposed to cellular or biological neuroscience) focuses on the relationship between cognitive function and neural systems, especially within humans. Both biological and cognitive neuroscience have informed understanding of the neural basis of learning, but given the complex interconnected nature of the learning which results from formal education, it is cognitive neuroscience which typically provides us with insights that have educational relevance. Cognitive neuroscience has distinct origins from educational psychology, however, and the lens it provides on educational phenomena has therefore been an external one, until very recently at least. This has created a range of challenges in defining its utility and how it might be appropriately applied. It is on these issues that the present chapter will focus.

The immediate precursor of cognitive neuroscience, neuropsychology, relied on fine-grained case studies of impairments in cognitive function in patients who had suffered brain injury. Triangulation of careful descriptions of anatomical damage with variation in the capacity to process written text, for example, made it possible to some extent to ascertain the relationship between regions of the brain and function (see e.g., Wilson, 1999). Work of this kind helped establish the role of frontal areas of the brain in high-level cognitive activity – planned and deliberate thinking of the type most relevant to education. However, many of its conclusions look crude from modern perspectives. In particular, injuries to specific regions created a false sense of localisation of functions (e.g., language abilities in the left hemisphere of the brain and spatial in the right), when in fact they were more commonly blocks in wider networks of neural activity – the leading characteristic of neural organisation from a contemporary viewpoint.

As technology developed in medical contexts became more accessible, though, new imaging techniques emerged which could finally support in vivo studies of neural activity, making it possible to examine what happens at the brain level during mental processing of information and coordination of behavioural responses. There is now a suite of such techniques providing us with different types of data, each with its own particular utility (see e.g., Dick et al., 2014).

The most important of these is functional Magnetic Resonance Imaging (fMRI). This is a method of measuring areas of increased blood flow in the brain, using magnetic fields to detect concentrations of oxygenated red blood cells, which exhibit different magnetic properties to unoxygenated ones. By comparing blood flow for one task relative to another (the Blood Oxygen Level Dependent or BOLD contrast), the areas of the brain which exhibit greater neural activity when participants are engaged in specific types of cognitive processing can be inferred, and the impact of variations in task assessed. Since the brain is effectively transparent to fMRI, it has proved a successful means of mapping the apparent organization of function throughout the cortical structure. The use of visual displays to show which areas 'light up' also lends an appealing immediacy to communication of the results. Its strength lies in mapping *where* activity is occurring, however, and it is not sensitive to *when* it happens, making it hard to judge the sequence of activation of different regions.

Electroencephalograms (EEG), and particularly their use to record Event-Related Potentials (ERP), provide a contrasting time- but not location-sensitive technique. EEGs detect small fluctuations in electrical activity within the brain via electrodes placed on the scalp. By indexing changes against the moment of presentation of some specific stimulus (e.g., a particular word pair), it is possible to examine how this impacts on the pattern of electrical activity (assumed to reflect the firing of neurons) over the ensuing second or so. Since that activity is affected by a multitude of simultaneous processes, detecting the effects of an event from a single trial is usually impossible. However, by repeating recording over a number of identical trials (as many as 100), the 'noise' created by other neural activity can be averaged out, enabling the event-related firing to stand out. The relative lack of locational sensitivity (clearer activations may be captured by electrodes on one part of the scalp, giving some clues) makes it harder to judge what particular patterns of electrical activity might mean. However, a mapping between event characteristics and spikes in activity ('components') at different post-event times has gradually emerged from compositing of data (see e.g., Luck & Kappenman, 2012) which makes it possible to infer the likely nature of a reaction (for instance, the N400 – a negative spike at about 400 ms post-event – is associated with perceived semantic violations). As a result, EEG has grown in popularity, especially given its low overhead on equipment and administration costs relative to fMRI.

Both fMRI and EEG suffer from problems with contamination of data by movement artefacts – instigation and control of physical movement creates its own activity within the brain, masking the target activity in which researchers are interested. This means that physical movement has to be severely restricted, with behavioural responses to stimuli typically limited to small finger movements on dedicated key pads. This means that while data can be gathered in vivo, they are not necessarily ecologically valid. It also means that there are age limitations on the gathering of data, with the period between 2 and 6 years of age – crucial in developmental terms – almost unknown territory as far as brain imaging is concerned, because it is extremely difficult to get young children to stay still enough. Near Infra-Red Spectroscopy (NIRS) has been developed as an attempt to reduce these problems, and to create something that is usable in real-world contexts such as classrooms. Like fMRI, this measures areas of blood flow in the brain, but via associated changes in temperature as detected by an EEG-like net of detectors worn on the scalp. NIRS has the advantage of being both location- and time-sensitive, but it can only detect activity in the surface layer of the brain, and as yet its resolution is poor relative to fMRI (Dick et al., 2014).

Structural MRI provides a different type of resource, effectively using variations in the concentration and alignment of water molecules in the brain to infer the location of physical cell structure. This makes it possible to map changes in connectivity within the brain – the white matter linking neurons to each other - over extended periods. This has led to understanding of how changes in cortical thickness relate to specialisation of brain function and broader development of intellectual capability (see e.g. Ramsden et al., 2011). Transcranial Magnetic Stimulation (TMS) and Transcranial Direct Stimulation (TDS) provide yet other techniques, where rather than looking for traces of activation to relate to cognitive outcome, the activation is instigated either magnetically or via application of minute current to look at cognitive consequences.

Taken together, the tools available to contemporary cognitive neuroscience permit coordination of different types of evidence into a sophisticated picture of neural activity as it actually occurs. However, there is one key limitation that is common to all the techniques outlined above: we can only interpret the neural data by correlating it with behavioural data which indicates the *content* of cognitive activity. None of them – as yet, anyway – are therefore capable of telling us anything about learning on their own. The significance of this point for educational neuroscience will become clearer in what follows.

Why educational neuroscience?

At root the brain is a simple organ, comprised mostly of one cell type, the neuron, linked via branches called dendrites to synaptic junctions with other neurons, allowing small electrical currents to be passed indirectly from neuron to neuron. In simple brain structures, neural activation starts with sensory input – the firing of receptors in the eyes, ears, nose, mouth or skin brought about by different stimuli – and ends with behavioural output – the firing of nerves to direct muscle movement. Neural activity in humans and other mammals is much more complex than this, but the basic division of activity into input, processing and output holds.

The complexity of the brain derives from the number of neurons involved – as many as a trillion in humans, each linked to up to ten thousand others (Nauta & Feirtag, 1986). These are organised to some extent into structures with specific functions, such as the visual cortex, responsible for processing visual input. However, many are unspecialised – particularly the frontal cortex – and most are capable of making new connections, providing the foundation for learning by building up intricate networks which perform detailed information processing.

The principal developmental changes in neural organisation that occur between gestation and adulthood are now well understood (see e.g., Gogtay et al., 2004;

Thomas, 2012), and provide a context for finer-grained changes that derive from learning. Large scale neural structures and connectivity between them develop prenatally, and typically-developing infants are born with primary capabilities intact. However, the cortex (the surface layer of the brain) is characterised by plasticity (the capacity to make novel connections) across the lifespan, with its micro-structure determined by local activity (i.e., experience). As the adage goes, ‘what fires together, wires together’, leading to progressive specialisation, effectively ‘tuning’ the cortex according to need (cf. Johnson, 2005, on interactive specialisation).

This process is amplified by profuse growth during infancy of new synaptic connections (synaptogenesis), creating massive potential for novel activations. This is followed by an extended period, up to puberty, of synaptic pruning, during which unused or under-used connections are shed, and the remaining connections are strengthened, focusing neural pathways on these. Activity increasingly comes to rest on networks involving multiple brain regions, as processing becomes more sophisticated. Structural and functional development continues in adulthood, with grey matter (cell bodies) thinning till age 30, but white matter (connections) increasing until at least 60.

These changes in structure are accompanied by changes in functionality. Between 2 and 5 years, children acquire near adult-like basic sensory abilities. Complex processing (e.g., of faces) takes longer to develop, continuing up to puberty. Basic understanding of the social world is good in early childhood, but accurate prediction of intentions and goals emerges later; and language processing shifts similarly from the simple in early childhood to the complex by age 10 or so. Memory systems and ability to control impulses and emotions emerge more slowly alongside increasing activity in the prefrontal cortex (PFC). Puberty sees a further spurt of synapse formation and pruning in the PFC as executive functions (selective attention, working memory, problem solving, multi-tasking) improve steadily, though risk-taking behaviour also increases alongside changes in the evaluation-of-reward network. Despite these major shifts, though, there is no evidence to support the existence of *critical* periods during which learning of a specific type has to occur, although there do appear to be sensitive periods – it is easier to learn a language before 7 years, for instance (Thomas, 2012).

Why is any of this relevant to education? There is a common concern (see e.g., Varma et al., 2008) that if neuroscience data is only interpretable via coordination with behavioural output, it can never do more than identify correlated activity. How then does it add to our understanding to know what changes in neural activity and structure are associated with learning and cognitive growth? There are two related answers to this question.

The first concerns theoretical completeness. A full understanding of learning processes and the constraints upon them, and optimal coordination of this understanding with teaching practices, are core concerns for educators as well as researchers: incomplete understanding will always leave open the possibility of educational failure. However, the diversity of past theorising about learning (see e.g., Marx & Goodson, 1976) underscores the complexity involved: since most theories have supporting evidence, it cannot be that they are wrong but rather that each provides a small window on a bigger picture. Full understanding of learning is therefore not possible without integrating these diverse strands of work.

One key consideration within this is that learning rests on a physical system and basic processes (i.e., the brain and nervous system) whose basic operation pre-date human culture. Few past theorists have attempted to address the interface between culture and this physical system in any direct fashion, though Vygotsky is a notable exception. He characterised the brain as providing a biological substrate supporting basic psychological operations, as in other species, which then became transformed by the use of external tools, especially the sign functions inherent in language, into 'higher psychological processes' i.e., deliberate, controlled cognition, which is fundamental to education (Vygotsky, 1978).

Vygotsky's account is theoretical, however, and whilst this notion of transformation has attracted many adherents (e.g., Cole, 1996; Tomasello, 1999), the *extent* of the changes it produces are unclear even if the account is essentially accurate. Gazzaniga (1998), for instance, argues that much neural activity demonstrably has nothing to do with deliberate cognition, so the latter might be better thought of as a surface overlay. Understanding the nature of this relationship is central to an understanding of the nature of conscious, deliberate cognition and how it arises. Even correlational data is relevant here, especially if it reveals patterns of activation that are more differentiated than behavioural responses, suggesting pre-conscious or unconscious activity that nevertheless impinges on outcomes.

The second answer is that, as understanding has progressed, it has become apparent that there are multiple ways in which cognitive systems *could* potentially function (see e.g., Thomas, 2012). We need to understand which actually hold in the context of different types of cognitive activity if we are to properly grasp the nature how these are instantiated. Again, even correlational data can be important for distinguishing between possibilities, especially if they run counter to expectation.

Identification of regularities in the way in which neural activity is organised has in fact led to important reconceptualizations of human cognition. The most striking example is recognition of the inadequacies of information processing models derived from computing, which had been the guiding framework for cognitive psychology and cognitive science since the 1960s. It has become evident that the brain does not engage in narrow sequential processing of items of information, but operates instead through large-scale parallel processing and pattern extraction (Rumelhart & McClelland, 1986; Karamanis & Thomas, 2011). One key implication is that learning is inherently domain and even context specific – connections are made in the service of particular activities, and, unlike computers, there are no inbuilt general functions or 'modules' serving cognitive processing across a variety of contexts, though these may emerge over time and experience. Despite the assumption of students' ability to derive broadly applicable knowledge from specific experiences which is inherent in educational curricula from early years onwards, transfer and generalisation may in fact often be hard won and in need of deliberate support strategies (Brown, 1990). This signals how important the contribution of educational neuroscience might potentially be in reshaping how we conceive of learning.

What is educational neuroscience?

The term 'educational neuroscience' implies a concern solely with explaining education through neuroscientific principles. In fact, although there is no consensual definition, there is general agreement that educational neuroscience is not reductionist, involves a range of methodologies, and is concerned with the coordination of evidence of *many* types, with neuroscience research providing just one strand (Butterworth & Tolmie, 2014; Thomas, 2013). It attempts to draw attention to the relevance of the neural dimension, but not to the exclusion of other levels.

The assumed model of the learner (see Mareschal et al., 2014) is a multilevel system operating at neural, cognitive and social/environmental levels, with complex and multi-directional interactions between these. These lead to substantial individual variation in presentation during schooling and educational outcomes, and sometimes unexpected, even puzzling constraints on progress. This variation is masked to some extent by equifinality – achievement of ostensibly the same outcome via different routes – which creates an impression of uniformity across learners that can be catered for by one-to-many delivery within classes using standardised approaches to teaching. However, as learners' backgrounds become more diverse, especially within urban populations, and the criteria for achievement become more exacting, assumptions of uniformity become increasingly hard to sustain.

To address this variation properly, what is needed is organised effort to a) map and understand the cross-level interactions, seeking regularities and typologies wherever possible; b) derive implications for practice that are sensitive to individual – and topic to topic – variation in what is effective; and c) consider methods of translating these into actual practice, taking pragmatic constraints into account. These objectives present substantial challenges, however, especially with respect to building sufficient shared understanding to enable the necessary communication and coordination of activity between researchers and practitioners.

The difficulties have been recognized since Bruer (1997) argued the endeavour was 'a bridge too far', because the connections between education and neuroscience were too remote and based too often on misunderstandings or overgeneralizations. As illustration, he cited the claim – common then among educators – that synaptogenesis represents a critical period for promoting learning. This derived from mistaken beliefs about the nature of synaptogenesis and the functions it supports, stemming from limited contact with relevant literature. In fact, the learning of explicit procedures and content that is focal to education continues long after the early synaptic spurt has given way to pruning and consolidation, suggesting synaptogenesis has little bearing.

Other forms of ‘neuromyth’ still persist (e.g., left and right brain learning – as already noted, most cognitive functions rest on neural systems distributed throughout the brain). Other ideas that have taken root have more credibility but little systematic evidence in their favour (e.g., the value of training key functions like working memory – while training produces improvement on standard measures, transfer of gains to classroom performance is typically limited). The misunderstandings are not restricted to educators, though. Hruby (2012) details instances of the failure of researchers to engage with expertise regarding practice, which is essential if translational goals are to be met. Common failures include glossing over the conclusions of practice-based research, or attributing perspectives to practitioners that they do not actually hold.

The real potential of educational neuroscience will only become clear when these problems are adequately addressed. Bruer (1997) argued that progress required discrete linkages between educators and instructional psychologists, and between psychologists and neuroscientists. These boundaries have since become blurred, however: many educational researchers have acquired expertise in neuroscience methods, and better lines of communication have been established between teachers and researchers of different types. Precise prescriptions as to productive linkages are therefore probably outdated. Nevertheless, it remains plain that if educational neuroscience has any real future, it will demand even wider collaboration between researchers from different backgrounds and practitioners of different types. It will also need support from government and policy makers because of the implications it carries for the design of educational systems.

Educational neuroscience may therefore be best thought of structurally as akin to public health and its translation of biological and environmental science into practical action to control disease and promote wellbeing. This implies a requirement to build teams of individuals representing different strands of activity, focused on mutually identified areas of need and methods of addressing these. To be effective, there would need to be an unbiased consensus across key players with different professional backgrounds, based on (within bounds) shared knowledge of the relevant science.

This is a complex balancing act, made harder by the fact that educational neuroscience research to date is piecemeal and unevenly developed. There is much work on dyslexia which has informed models of atypical and typical development and thence remedial and mainstream teaching of literacy (see e.g., Hulme & Snowling, 2009). However, few other areas approach this level of activity, and some (e.g., science learning, creativity) have only been addressed by a handful of researchers. Moreover, there is no unifying theoretical framework; to the extent that there is a consensus across researchers, this is based on a shared belief that full understanding of learning demands consideration of neural processes, not what form the resulting models should take. A coherent approach to translation is unlikely to possible without a more coordinated research base than we have at present, covering typical and atypical learning in key areas.

Do we really need educational neuroscience?

Given the challenges involved, is educational neuroscience actually worth the effort? I will address this question by summarising work in the three areas focused on by the OECD Programme for International Student Assessment (PISA): literacy development; number and mathematics; and science learning. I will also consider research on executive function and working memory, as ostensibly domain-general cognitive functions, before concluding with a note on other significant areas of activity.

Literacy development

Learning to decode a written language rests on the implicit sense of its organisation provided by earlier acquisition of its oral forms (Hulme & Snowling, 2009). This includes the division of sounds into meaningful word units or morphemes, the basic vocabulary encoded in this way, and acceptable ways of putting this into sequences to create extended meaning – the grammar of the language. There is good evidence that children have an innate sensitivity to these characteristics as they manifest in their local language community, and that imitative learning plays an important role in acquiring their use (Hulme & Snowling, 2009; Saxton, 2010).

This prior knowledge makes it possible for children to map the orthographic forms they encounter onto recognised elements of language (Hulme & Snowling, 2009). This happens initially at whole word level and in rote fashion (the pre-alphabetic stage): children simply learn the association between written symbols and spoken words that are significant to them. Rote learning of this kind is inefficient, however, and where the language uses alphabetic orthography, more detailed mapping is possible at the grapheme-phoneme (i.e., letter-sound) level. Once children have begun to grasp these correspondences, they use this mapping to learn relationships between letter strings and their pronunciation (the alphabetic stage). At first, they are better able to work with larger units than single letters and phonemes, such as syllables ('crust') and onset-rime structures ('cr' and 'ust'). These mark important points of transition to full phonological strategies which make use of their developing knowledge of the ways in which phonemes are represented by letters. These strategies are gradually automated as they become familiar, the child starts to extract the spelling patterns they regularly encounter (the orthographic stage), and they return to whole-word recognition – but based on explicit knowledge about word composition, and strategies for decoding unknown words.

Phonological awareness – the ability to identify specific sounds in speech, and to decompose spoken words into these – is a key precursor of learning to read, since it underpins the shift to the alphabetic stage, and makes it possible to bind orthography into pre-existing neurocognitive structures that link phonology and meaning. There has consequently been a shift in the UK towards the use of phonics teaching to promote awareness of phonemes in the early years of primary/elementary school, though this has involved a complex mix of politics (Wyse & Goswami, 2008; cf. the point about the role of government). In English, this approach needs to be accompanied by work on whole-word recognition because the number of irregular spellings limits the value of grapheme-phoneme conversion as a method of learning to read. Children need to recognise where these mappings do not apply and words are not pronounced as they are written (e.g., ‘yacht’).

Studies using fMRI have identified the main neural pathways involved in reading and the anomalies exhibited by children with dyslexia (Demonet et al., 2004), for whom the main problem is a deficit in the emergence of phonological awareness. Although there are competing theories about the nature of this deficit, one argument is that the key step in learning to read efficiently is the development of amodal processing structures in the temperoparietal and left ventral cortex, which integrate phonology, orthography and semantics within the Visual Word Form Area (VWFA; McCandliss et al., 2003). Metaphonological awareness – the capacity to deliberately utilise knowledge of word sounds – is strongly implicated in developing these structures (Kovelman et al., 2012). This work has demonstrated that differences separating non-impaired from impaired readers also predict individual variation in typically developing children, implying that dyslexia is not a qualitatively different syndrome but a more pronounced deficit, marked by poor connectivity and anomalies in neurotransmitter profile in the VWFA.

The impact of this deficit is affected by the letter-sound relationships in the child’s native language. Reading disorders are less apparent where there is a consistent relationship between letters and phonemes, as in Spanish, or in writing systems that do not require segmentation into phonemes, such as Chinese (Hulme & Snowling, 2009). This suggests that the crucial factor is children’s sensitivity to complexity in orthographic mapping onto phonology.

Computer modelling of brain-like processing structures also indicates that regularity determines how difficult it is to build up stable representations of letter-sound relationships (Harm et al., 2003). Such modelling has explored how the underlying deficit might be corrected for in an irregular language like English, as well as identifying the stage at which remediation of different types might be successful (Ziegler et al., 2007). Since studies using EEG measures have made it possible to identify children at risk of dyslexia before they begin learning to read (Lyytinen et al., 2001), remediation can begin early. Phoneme practice has emerged as a key intervention, and evaluation of brain activation changes in children who respond positively indicates normalisation is achievable in the majority of cases (Griffiths & Stuart, 2013). Candidate genes that may affect brain development have also now been found (Schulte-Körne et al., 2007), making it possible to contemplate yet earlier forms of remediation.

Number and mathematics

Most research on mathematics learning has focused on the growth of children's understanding of number and arithmetic in early years and primary/elementary school, since (as with word reading) this basic level underpins later achievement. There is good evidence that infants (in common with other primates) possess two innate perceptual systems for recognising number. One enables them to distinguish small exact quantities up to 3 without counting – what is termed subitization. The other allows them to judge relative differences in large quantities (e.g., 16 versus 32) (Xu & Spelke, 2000). It has been argued that these systems provide the foundation for verbal counting skills in the same way that phonological awareness provides a basis for reading.

The link between perception of number and counting is less exact, however, since children have to grasp new concepts that only apply to the latter (Gelman & Gallistel, 1978): the one-to-one principle (each object to be counted only gets one count word), the stable order principle (count words get used in fixed order), the order irrelevance principle (the order in which objects are counted does not affect the total), and the cardinality principle (the last count word used represents the total number). Current evidence suggests grasp of these principles requires prolonged practice, and children move more easily from the innate systems to estimation (Gilmore et al., 2007).

Basic counting ability is usually in place by the time children go to school, providing the foundation for arithmetic skills (Hulme & Snowling, 2009). These are more disparate than those involved in reading, and include: single- and then multiple digit addition, subtraction, multiplication and division, along with varying strategies for carrying out these computations; understanding of tens and units, and the role of place value in distinguishing between these; translation between arithmetical problems presented in concrete, verbal and numerical formats; knowledge of number facts (e.g., any number multiplied by 10 equals the original number shifted a place leftwards followed by 0); and derived fact strategies (e.g., use of the multiplication by 10 fact to decompose sums into sub-problems).

There appears to be no consistent order in the emergence of these skills, or any strong relationship between them as they are acquired (Dowker & Sigley, 2010), so children show wide individual variation and sometimes surprising gaps. Dyscalculia, a deficit in the ability to learn arithmetic was identified over thirty years ago, but research on innate number awareness has transformed understanding. The neural pathways involved in typical numerical capacities are now relatively well-mapped (Castelli et al., 2006), and dyscalculia appears to be due to a deficit in number awareness associated with abnormalities in a region of the parietal lobe specialized for enumeration. This is reflected in poor performance on tasks such as comparing dots in two visual displays to say which has more (Landerl et al., 2004). Less is known about the nature of this deficit than is the case for dyslexia, but interventions aimed at strengthening number sense via practice with the kinds of displays used to diagnose problems appear to be effective (Butterworth & Yeo, 2004). These have greater impact than traditional interventions for children with number problems, which focus on simple repetition of number-object correspondences: it is not counting as such that is the problem.

The relationship to innate abilities appears to be weaker than it is for reading, though. Achievement in arithmetic has been found to be more strongly related to knowledge of number facts and working memory ability (Cowan & Powell, 2014; Soltesz et al., 2010). Given the role of attention and executive control in counting and computation, and the way in which knowledge impacts on the strategies used for solving problems, this is unsurprising. It makes it harder to tell what kinds of teaching strategy might be most effective, however, especially for mainstream learners.

Science learning

The cognitive capacities underpinning science understanding also originate in the pre-school period, and again key aspects appear to be innate. There are three aspects to science learning: factual knowledge (evidence), grasp of procedure (generation of evidence), and conceptual grasp (understanding of phenomena). Despite the emphasis in much teaching on the first, and in research on the development of scientific thinking on the second, it is concepts that are arguably central, especially during initial learning, since they provide unifying ideas about causal processes that help learners make sense of subsequent experience and provide the basis for meaningful scientific testing.

There is good evidence that infants are innately sensitive to causal information, so that, for example, when one object strikes another and the second object moves, this is perceived as the first causing the movement (Leslie & Keeble, 1987). Neuroscience evidence supports this claim, indicating a neural foundation for causal perceptions (Satpute et al., 2005). By 2 years, children are capable of employing retrospective assessment of associations between cause and outcome to anticipate the outcome of specific events, and do so with considerable statistical accuracy (Schulz et al., 2007).

However, this early understanding is not typically the resource for science learning it might be. Innately organised perceptions are based on tacit sensitivity to covariation between events, and are not automatically verbally accessible – being able to catch a ball because its trajectory has been accurately anticipated is not the same as being able to describe the forces at work. Explicit concepts, which can be broken down into elements that can be described in language and subjected to analysis and discussion, appear commonly to have separate origin in conversation about everyday experiences, to relate to causal perceptions in unpredictable and often inaccurate fashion, and to be resistant to change (Howe et al., 2012). EEG studies with adults reveal distinct neural activations associated with accurate tacit perceptions and inaccurate explicit concepts (Kallai & Reiner, 2010), indicating failure to integrate them is a commonplace outcome.

The crucial educational step is therefore the accurate mapping of explicit concepts onto tacit perceptions and consequent calibration between observable regularities and inferred causal relationships. However, this mapping is constrained in various ways. In adults, fMRI studies show that coordination of observational data with pre-existing concepts is strongly associated with activity in the dorsolateral prefrontal cortex (DLPFC), a region linked with executive function (Fugelsang & Dunbar, 2005). Executive control is known to be relatively late maturing (Zelazo et al., 2008), suggesting that management of cognitive load may substantially limit tacit-explicit mapping before age 10. Although direct evidence is as yet scant, individual variation in attentional and working memory abilities may further affect pick-up and coordination of perceptual information, and there may even be so far unrecognised dyscalculic-like deficits in sensitivity which impose additional constraints for some learners. There is also clear evidence of the importance for conceptual progression of learners being provided with explicit vocabulary and explanatory constructs as part of events within which language and observation are brought together (Philips & Tolmie, 2007). Such experiences vary from child to child to a far greater extent than exposure to count words and counting experiences.

In terms of educational approaches, this analysis suggests primary/elementary school teachers need to provide children with experiences that a) involve manipulation of causal events, b) draw attention to the exact sequence within these (e.g., what happens when an object that barely floats is dropped into a tank of water?), c) provide descriptions to help make this sequence explicit, and d) connect these descriptions to explanatory constructs. In fact, collaborative group work in science typically embodies all these features, and has been shown to produce robust improvements in conceptual understanding (Tolmie et al., 2010). Its success has generally been attributed to its capacity to generate explanatory dialogue, but the educational neuroscience approach suggests it is the combination of observation and dialogue which is critical. The group context may further assist learning by providing good reason to refer to observations explicitly, and making informational load more manageable by distributing it between individuals.

Executive function and working memory

Encoding, storage and retrieval of information using long- and short-term memory are central to learning. Long-term memory provides a cumulative store that can be used in myriad ways, even though the information it contains is typically inexact because the process of encoding extracts and organises material in meaningful units or chunks, mapping it onto existing knowledge. Short-term memory retains high levels of detail (encoding is more exact), but for much shorter durations, and was seen by early research simply as a 'loading platform' for long-term memory – a transitional stage during which items were rehearsed in preparation for transfer. Incoming information was stored very briefly in precise detail in an iconic memory buffer (cf. pre-conscious processing), before salient aspects were extracted for rehearsal (Atkinson & Shiffrin, 1968; Murdock, 1962).

Current theories assign short-term memory the much more crucial role of principal mental workspace, reflected in the term working memory (Baddeley, 1986; Baddeley & Hitch, 1974). Working memory consists of three interrelated mechanisms: the phonological loop, which deals with auditory material and rehearsal; the visuo-spatial scratchpad, which provides a processing space for manipulating visual information; and the central executive, which directs operations. Together these provide the core system for conscious awareness of and attention to information, where we act on it deliberately, decomposing tasks into sub-problems and sub-goals, central features of human problem-solving (Newell & Simon, 1972).

The operation of working memory both influences and is influenced by learning. For instance, in decomposing a multiplication task into sub-problems, it helps substantially to know in advance which components are likely to be easiest to compute and in what sequence, in order to avoid taxing the scratchpad and the phonological loop. This makes the central executive critical: where possible, it retrieves past examples of similar problems and the strategies used to solve these from long-term memory, and then enacts these. Where there has been substantial relevant experience, the process is near-automatic. Where past experience provides no direct help, the executive seeks ways to proceed, often using analogies to provide clues (Gentner & Jeziorski, 1989). When the solution has been arrived at, it checks the steps are sound, and that the end product looks plausible. The process of monitoring effectiveness is central to learning, ensuring mistakes are avoided the next time a similar problem is encountered, and highlights the key role of the central executive as the mechanism underpinning metacognition. It does not just organise the manipulation of information, it watches its own operation and corrects itself when it goes wrong. As part of this, its activity includes inhibition of inappropriate operations, especially when these have become automated (for instance when the rules governing a well-practiced task are changed).

The working memory system is seen by contemporary theories as central to human cognitive functioning, directing mental activity, drawing on existing knowledge and creating new knowledge as part of its operation; and becoming self-regulating as children develop, taking a central role in planning activity, monitoring feedback and ultimately directing further learning (Blair & Diamond, 2008). As a result, the role of executive function in particular has unsurprisingly become well-documented in the contexts of a) reading, where it is crucial for directing the uptake and processing of information across assemblages of text (Hulme & Snowling, 2009); b) arithmetic, where it is central to the manipulation of numerical information (Cowan & Powell, 2014); and c) scientific thinking, where it is required to coordinate evidence with concepts (Fugelsang & Dunbar, 2005).

There is also growing evidence that its inhibitory dimension plays a key role in the development of scientific reasoning, where incorrect intuitive beliefs (e.g., elephant cells are bigger than mouse cells) must be suppressed in favour of correct but counterintuitive ones (they are actually the same size). Indeed, the ability to suppress intuitive beliefs may be a crucial difference between scientific novices and experts (Masson et al., 2014). It may play a similar role in areas of arithmetic such as fractions, where children and even adults find it difficult to grasp that increases in denominator size entail a decrease in overall value.

Given its importance, there has been much interest in the trainability of executive control and working memory. Up to a point research in this area has been successful. Various studies have shown that practice on tasks with high working memory loads leads to improved performance (e.g., Shipstead et al., 2012). Houdé et al. (2000) have shown that training in inhibitory control leads to activation shifting from posterior to frontal areas of the brain, including the DLPFC, consistent with executive function being promoted. However, this success has been widely coupled with failures of transfer to untrained contexts (e.g., Thorell et al., 2009).

It seems likely that these failures result from misapplication of the computing analogy noted earlier; functions develop on the basis of connections serving specific activities, not as general modules, and therefore have to be promoted in the context in which they are required. Current research is beginning to test this hypothesis, which carries considerable implications regarding generalisation in learning. To date, this has been poorly researched, despite the fact that most conceptions of the curriculum depend on it as a given. Neuroscience evidence suggests we actually have no adequate models of generalisation, and that the extended connections inherent in the notion may often only occur when deliberately engendered.

Other areas of work

Although I have focused here on literacy development, number, science learning and executive function, educational neuroscience covers a considerably wider range of work. For example, there is growing research on socio-emotional development from infancy through adolescence, and how this is influenced by the organization and refinement of neural structures relating to processing of facial information, mentalizing (the attribution of mental states to others), and sensitivity to acceptance and rejection by others (Blakemore et al., 2014). The apparent importance of refinements during adolescence in particular suggests that forms of social education have a potential value in high school curricula which is presently largely unrecognized. Another strand of this research has helped identify abnormalities in the amygdala which are associated with impaired emotional awareness and callous-unemotional forms of anti-social behaviour, suggesting alternative approaches to educational intervention with the most intransigent disruptive students (Jones et al., 2009). Research in the area of second language learning has repeatedly demonstrated an important predictive role for low-level sensitivity to perceptual regularities in the language being learned (Frost et al., 2013), underscoring the importance of immersive pedagogies. Other pertinent work concerns shifts in adolescent sleep patterns and their implications for the timing of the school day (Lockely & Foster, 2012).

Implications

The preceding section illustrates how educational neuroscience may be capable of making a distinctive contribution both to our understanding of key areas of learning and to the development of evidence-based pedagogical practices that will improve outcomes for typical and atypical learners. It is important to note that, across the different areas considered, the best work rests on a combination of behavioural and neuroscience methodologies including direct brain imaging and computational modelling. At present, effective interventions are based largely on behavioural work (much of this drawing on teacher expertise in its inception). However, by combining this with neuroscience, it has become possible to explain better why these work, and build models of individual variation at a level of sensitivity that behavioural studies alone are not able to achieve (Thomas, 2013). In time, wholly novel interventions which address this level of detail may begin to emerge.

There are other important implications, however. First, it is apparent that literacy, number and science learning call upon a relatively unique constellation of functions and processes, with even the types of tacit-explicit mapping involved in each being distinctive in character. This may mean there are few general principles underpinning successful learning across different areas – although the apparent importance of symbolic mapping processes of different types in each area (see e.g., Grabner et al., 2011) suggests that there may be some common thread of this nature (cf. the issue of the culture/brain interface). We need to keep looking for such principles as they may be beyond current conception, and more coordinated research activity is crucial for this reason. Even then, research in each area will probably remain at best only loosely interconnected for the immediate future.

These points may also mean that there are unlikely to be any effective overarching pedagogies – if different areas of education rest on different functions and skills, these will typically require different specific methods of support, which may in turn differ according to individual need. The key objective at this stage is to map the skills necessary to progress in focal areas of the curriculum, the main steps involved in developing these, and the types of clinical and non-clinical variation that can arise. There is a particular need for more emphasis on mainstream development in order to extend the potential benefits of educational neuroscience to broad populations, rather than just those with specific deficits. These mappings can then be used to derive evidence-based strategies for promoting learning in each area and managing individual variability.

This suggests it may also be more appropriate to move away from assessment processes based on uniform standards to ones based on progress against expected individual trajectories, as projected from data on a range of neural, cognitive and environmental indices at school entry. The techniques to support an enhanced value-added approach of this kind already exist, and the profiling of educational input it would promote is entirely consistent with developments in medicine and public health.

To facilitate such changes, the training of teachers and provision for support of their continuing development would need to shift emphasis to promoting understanding of the range of evidence-based models applicable to their working context, placing them in a position somewhat akin to that of a medical general practitioner in terms of the type of skill-set they require. This would be a radical change, and given the scale of the new knowledge to be understood and deployed, it may entail a further need to strike a different balance between school-based and out-of-school education, with the former concentrating on provision of finely honed development of key skills, which are then elaborated according to personal choice in a range of other settings. If the evidence-based approach at the core of educational neuroscience suggests progress requires shifts of this nature, then ultimately governments and wider society will need to decide whether these are desirable objectives for education, and how far to embrace – and resource – them.

References

Atkinson, R.C. and Shiffrin, R.M. (1968). Human memory: a proposed system and its control. In K.W. Spence and R.T. Spence (Eds.), *The psychology of learning and motivation, Vol. 2*, pp 89-105. New York: Academic Press.

Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A.D. and Hitch, G.J. (1974). Working memory. In G.Bower (Ed.), *The psychology of learning and motivation: advances in research and theory*, pp 47-90. New York: Academic Press.

Blair, C. & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology,* 20, 899-91.

Blakemore, S-J., Cohen Kadosh, K., Sebastian, C.L., Grossmann, T. and Johnson, M.H. (2014). Social development. In D. Mareschal, B. Butterworth and A. Tolmie (Eds.), *Educational Neuroscience* pp. 268-296. Chichester: Wiley-Blackwell.

Brown, A.L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, 14, 107-133.

Bruer, J. T. (1997). Education and the brain: A bridge too far. *Educational Researcher*, 26(8), 4–16.

Butterworth, B. & Tolmie, A. (2014). Introduction. In D. Mareschal, B. Butterworth & A. Tolmie (Eds.), *Educational neuroscience* pp 1-12. Chichester: Wiley-Blackwell.

Butterworth, B. and Yeo, D. (2004). Dyscalculia Guidance. London: NFER-Nelson.

Castelli, F., Glaser, D.E. and Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. *Proceedings of the National Academy of Science*, 103, 4693-4698.

Cole, M. (1996). *Cultural psychology*. Cambridge, MA: Belknap/Harvard University Press.

Cowan, R. and Powell, D. (2014). The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology*, 106, 214-229.

Demonet, J.-F., Taylor, M.J. and Chaix, Y. (2004). Developmental dyslexia. *Lancet*, 363, 1451-1460.

Dick, F., Lloyd-Fox, S., Blasi, A., Elwell, C. and Mills, D. (2014). Neuroimaging methods. In D. Mareschal, B. Butterworth and A. Tolmie (Eds.), *Educational Neuroscience* pp. 13-45. Chichester: Wiley-Blackwell.

Dowker, A. and Sigley, G. (2010). Targeted interventions for children with arithmetical difficulties. *British Journal of Educational Psychology Monograph Series II Volume 7: Number Development and Difficulty*, 65-81.

Frost, R., Siegelman, N., Narkiss, A. and Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, 24, 1243-1252.

Fugelsang, J.A. and Dunbar, K.N. (2005). Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, 43, 1204-1213.

Gazzaniga, M.S. (1998). *The mind's past*. Berkeley, CA: University of California Press.

Johnson, M. H. (2005). Sensitive periods in functional brain development: problems and prospects. *Developmental Psychobiology*, 46, 287–292.

Gelman, R. and Gallistel, C.R. (1978). *The child's understanding of number*. Cambridge MA: Harvard University Press.

Gentner, D. & Jeziorski, M. (1989). Historical shifts in the use of analogy in science. In B. Gholson, W. Shadish, R. Neimeyer & A. Houts (Eds.), *Psychology of science: contributions to metascience.* Cambridge: Cambridge University Press.

Gilmore, C.K., McCarthy, S.E. and Spelke, E. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, 447, 589-591.

Gogtay, N., Giedd, J.N., Lusk, L. et al (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, 101 (21), 8174-8179.

Grabner, R.H., Reishofer, G. & Koschutnig, K. (2011). Brain correlates of mathematical competence in processing mathematical representations. Frontiers in Human Neuroscience, 5:130 (accessed 20.02.15).

Griffiths, Y. & Stuart, M. (2013). Reviewing evidence-based practice for pupils with dyslexia and literacy difficulties. *Journal of Research in Reading*, 36, 96-116.

Harm, M.W., McCandliss, B.D. and Seidenberg, M.S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, 7, 155-182.

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B. and Tzourio-Mazoyer, N. (2000). Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training. *Journal of Cognitive Neuroscience*, 12, 721-728.

Howe, C., Tavares, J.T. and Devine, A. (2012). Everyday conceptions of object fall: Explicit and tacit understanding during middle childhood. *Journal of Experimental Child Psychology*, 111, 351-366.

Hruby, G.G. (2012). Three requirements for justifying an educational neuroscience. *British Journal of Educational Psychology*, 82, 1-23.

Hulme, C. and Snowling, M.J. (2009). *Developmental disorders of language learning and cognition*. Chichester: Wiley-Blackwell.

Jones, A.P., Laurens, K.R., Herba, C., Barker, G. and Viding, E. (2009). Amygdala hypo-reactivity to fearful faces in boys with conduct problems and callous–unemotional traits. *American Journal of Psychiatry*, 166, 95-102.

Kallai, A. Y. and Reiner, M. (2010). The source of misconceptions in physics: When event-related potential components N400 and P600 disagree. European Association for Research on Learning and Instruction Neuroscience and Education Special Interest Group Biennial Meeting, Zurich Switzerland.

Karaminis, T. N. and Thomas, M. S. C. (2011). Connectionism. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning*, Part 3, pp 767-771. Berlin: Springer.

Kovelman, I., Norton, E.S., Christodoulou, J.A. et al. (2012). Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. *Cerebral Cortex*, 22(4), 754-764.

Landerl, K., Bevan, A. and Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8-9 year old students. *Cognition*, 93, 99-125.

Leslie, A. M. and Keeble, S. (1987). Do six-month old infants perceive causality? *Cognition* 25, 265–288.

Lockley, S.W. and Foster, R.G. (2012). *Sleep: A very short introduction*. Oxford University Press, Oxford UK.

Luck, S.J. and Kappenman, E.S. (2012). *The Oxford handbook of event-related potential components*. Oxford: Oxford University Press.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T.K., Laakso, M.-L., Leinonen, S., Leppanen, P.H.T., Lyytinen, P., Poikkeus, A.M., Puolakanaho, A., et al. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology*, 20, 535-554.

Mareschal, D., Butterworth, B. and Tolmie, A. (Eds.) (2014). *Educational neuroscience*. Chichester: Wiley-Blackwell.

Mark, M.H. and Goodson, F.E. (1976). *Theories in contemporary psychology, 2nd Edition*. New York: Macmillan.

Masson, S., Potvin, P., Riopel, M. and Brault Foisy, L-M. (2014). Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. Mind, Brain and Education, 8, 44-55.

McCandliss, B.D., Cohen, L. & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293-299.

Murdock, B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64, 482-488.

Nauta, W.J.H. and Feirtag, M. (1986). *Fundamental neuroanatomy*. New York: Freeman.

Newell, A. & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Philips, S. and Tolmie, A. (2007). Children's performance on and understanding of the Balance Scale problem: The effects of parental support. *Infant and Child Development*, 16, 95-117.

Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S. C., Ellis, C., Shakeshaft, C., Seghier, M. L. and Price, C. J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature*, 479, 113-116.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.

Satpute, A.B., Fenker, D.B., Waldmann, M.R., Tabibnia, G., Holyoak, K.J. and Lieberman, M.D. (2005). An fMRI study of causal judgments. *European Journal of Neuroscience*, 22, 1233-1238.

Saxton, M. (2010). *Child language: acquisition and development*. London: Sage.

Schulte-Körne, G., Ziegler, A., Deimel, W., Schumacher, J., Plume, E., Bachmann, C., Kleensang, A., Propping, P., Nöthen, M., Warnke, A., et al. (2007). Interrelationship and familiality of dyslexia related quantitative measures. *Annals of Human Genetics*, 71, 160-175.

Schulz, L.E., Gopnik, A. and Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, 10, 322-332.

Shipstead, Z., Hicks, K.L. and Engle, R.W. (2012). Cogmed working memory training: Does the evidence support the claims? Journal of Applied Research in Memory and Cognition, 1, 185-193.

Soltesz, F., Szucs, D. and Szucs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: A developmental study. *Behavioral and Brain Functions*, 6:13 (accessed 03.06.13.).

Thomas, M. S. C. (2012). Brain plasticity and education. *British Journal of Educational Psychology - Monograph Series II Volume 8: Educational Neuroscience*, 142-156.

Thomas, M.S.C. (2013). Educational neuroscience in the near and far future: Predictions from the analogy with the history of medicine. *Trends in Neuroscience and Education*, 2, 23-26.

Thorell, L.B., Lindqvist, S., Bergman Nutley, S., Bohlin G. & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 12, 106-113.

Tolmie, A., Topping, K., Christie, D., Donaldson, C., Howe, C., Jessiman, E., Livingston, K. and Thurston, A. (2010). Social effects of collaborative learning in primary schools. *Learning and Instruction*, 20, 177-191.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Varma, S., McCandliss, B.D. and Schwartz, D.L. (2008). Scientific and pragmatic challenges for bridging education and neuroscience. *Educational Researcher*, 37, 140-152.

Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wilson, B.A. (1999). *Case studies in neuropsychological rehabilitation*. Oxford: Oxford University Press.

Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal,* 34(6), 691-710.

Xu, F. and Spelke, E. (2000). Large number discrimination in 6-month old infants. *Cognition*, 74, B1-B11.

Zelazo, P.D., Carlson, S.M. and Kesek, A. (2008). The development of executive function in childhood. In C.A. Nelson and M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience*, Second edition. Cambridge MA: MIT Press.

Ziegler, J., Castela, C., Pech-George, C., Georgel, F., Alarioa, F.-X. and Perry, C. (2007). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition*, 107, 151-178.