Principles of Learning, Implications for Teaching: A Cognitive Neuroscience Perspective

USHA GOSWAMI

Cognitive neuroscience aims to improve our understanding of aspects of human learning and performance by combining data acquired with the new brain imaging technologies with data acquired in cognitive psychology paradigms. Both neuroscience and psychology use the philosophical assumptions underpinning the natural sciences, namely the scientific method, whereby hypotheses are proposed and tested using quantitative approaches. The relevance of 'brain science' for the classroom has proved controversial with some educators, perhaps because of distrust of the applicability of so-called 'medical models' to education. Nevertheless, the brain is the main organ of learning, and so a deeper understanding of the brain would appear highly relevant to education. Modern science is revealing the crucial role of biology in every aspect of human experience and performance. This does not mean that biology determines outcomes. Rather, there is a complex interplay between biology and environments. Improved knowledge about how the brain learns should assist educators in creating optimal learning environments. Neuroscience can also identify 'biomarkers' of educational risk, and provide new methodologies to test the effects of educational interventions.

INTRODUCTION

Neuroscience is a large field, and all aspects of neuroscience have some relevance to education. For example, some neuroscientists study how cells grow in the foetal brain, and some neuroscientists study the chemical 'neurotransmitters' that cells use to transmit information to each other. Neither area of study appears *a priori* relevant to education, and yet better knowledge of foetal brain development offers important insights into why children born to alcohol-addicted mothers have particular learning difficulties, for example in numeracy (Kopera-Frye, Dehaene

and Streissguth, 1996). Knowledge about neurotransmitters has led to the marketing of drugs that enhance learning ('cognitive enhancers'). As students can take cognitive enhancers before they sit their exams, this aspect of neuroscience has a potentially dramatic impact on educational outcomes (Turner and Sahakian, 2006). Cognitive neuroscience is one sub-field within neuroscience. Cognitive neuroscience takes psychological theories about the mind (e.g. that short-term and long-term memory are distinct systems) or symbolic descriptions of mental processes (e.g. that we can think using images versus 'inner speech') and explores them by measuring electro-chemical activity in the brain. Interpretations of neural activity are constrained by using experimental paradigms drawn from cognitive psychology (see Szücs and Goswami, 2007).

Cognitive neuroscience involves both the direct measurement of electro-chemical activity, and connectionist modelling. Direct measurement of the brain shows the patterns of activity across large networks of neurons (called cell assemblies) that correspond to mental states such as remembering a telephone number. Although this is essentially correlational information, brain imaging also reveals the time course of the activity (e.g. which neural structures are activated, in which order) and interactions and feedback processes within these large networks. Modelling enables in-principle understanding of how synchronised neuronal activity within cell assemblies results in learning and development. I will argue that all of these kinds of information are useful for education, even though the field of cognitive neuroscience is still in its early stages.

COGNITIVE NEUROSCIENCE AND RESEARCH QUESTIONS RELEVANT TO EDUCATION

At birth, considerable brain development has already taken place. Most of the neurons (brain cells) that will comprise the mature brain are already present, and have migrated to the appropriate neural areas. Neural structures such as the temporal cortex (audition) and the occipital cortex (vision) have formed, but will become progressively specialised as the infant and young child experiences environmental stimulation. Neural specialisation depends on the growth of fibre connections between brain cells both within and between different neural structures (called 'synaptogenesis'). Some fibre connections reflect 'experience expectant' processes. These connections are usually in the sensory system, and reflect abundant early growth in response to classes of environmental stimulation (such as visual field information) that the brain 'expects' (via evolution) to receive. Connections that are not used frequently are then aggressively pruned. Experience-expectant plasticity is biologically pre-programmed. Other fibre connections are 'experience dependent'. Here the brain is growing connections in order to encode unique information that is experienced by the individual. Every person has a distinctive environment, even children growing up within the same family, and so experience-

dependent connections are the ones that make each brain subtly different. Experience-dependent connections include the connections formed by education. Experience-dependent synaptogenesis enables life-long plasticity with respect to new learning.

The specialisation of neural structures occurs within developmental trajectories that are constrained by both biology and environment. We can study these developmental trajectories by asking questions about structure and function, and some of these questions are relevant to education. Such questions include which neural structures are important for learning different educational inputs (e.g. reading versus arithmetic), which interconnections between structures are important for different educational inputs, what the temporal sequencing looks like, and what can go wrong. We can also try and distinguish cause from effect.

Research Question 1: Neural Structures for Learning

A very active area of cognitive neuroscience is the study of which neural structures are active as the brain learns different inputs or performs different tasks. The method most frequently used is fMRI, or functional magnetic resonance imaging, which measures changes in blood flow in the brain. Blood will flow to neural structures that are active, hence this kind of imaging reveals which parts of the brain are most involved in certain tasks. Such research is essentially correlational, but it is still of value for education.

For example, researchers have measured the main neural structures that are active in novice readers as they perform tasks with print. These studies have found that the neural structures for *spoken* language are those that are most active (Turkeltaub *et al.*, 2003). Such data would appear to rule out theories of reading acquisition that argue that children first learn to read using 'logographic' strategies, going directly from print to meaning without recoding from print to sound (e.g. Frith, 1985). As children are exposed to more and more printed words, brain imaging studies show that a structure in the visual cortex becomes increasingly active (Cohen and Dehaene, 2004). Dubbed the 'visual word form area', this structure appears to store information about letter patterns for words and chunks of words and their connections with sound. It is experience-dependent, storing the learning that results as children are exposed to more and more printed words and recode them into sound. It will therefore respond to 'nonsense words' that have never been seen before, because chunks of these nonsense words will be familiar from prior learning experiences (e.g. 'brate' is a nonsense word, but it is analogous to real words like 'crate' and 'brake').

Therefore, although the brain did not evolve for reading, it is able to develop fibre connections to support reading that encode print experience into the nervous system. This is achieved by recruiting neural structures that already perform very similar functions, such as object recognition. For example, the visual word form area is adjacent to the area in the visual

system that is most active during picture naming. One prominent educational neuroscientist calls this feature of experience-dependent plasticity the 'neuronal recycling hypothesis' (Dehaene, 2008). Dehaene argues that our evolutionary history and genetic organisation constrain new cultural acquisitions to some extent, as new learning must be encoded by a brain architecture that evolved to encode at least partially similar functions over primate evolution.

Nevertheless, studies using fMRI have revealed that unexpected neural structures can be involved in educational performance. For example, studies of arithmetic and number processing with adults have shown that an area in parietal cortex is particularly active whenever numerical magnitude must be accessed (Dehaene, 1997). Closely adjacent areas of parietal cortex are activated when judgements about size or weight are required. This has led Dehaene to argue that the parietal cortex is the location of an approximate, analogue magnitude representation in the human brain. However, when Dehaene and his colleagues compared brain activation for two arithmetic tasks, one involving exact addition (e.g. $4+5=9$) and one involving approximate addition (e.g. $4+5=8$), they found that the parietal structures were only most active in the second case (Dehaene *et al.*, 1999). During exact calculation the greatest relative activation occurred in a left-lateralised area in the inferior frontal lobe, which is traditionally regarded as a language area. Dehaene *et al*. argued that this was because exact arithmetic requires retrieval of over-learned 'number facts', which are stored in the language areas. This implies that part of school mathematical learning is linguistic. The multiplication tables and 'number facts' that children are taught seem to be learned as verbal routines, like the months of the year, rather than as numerical concepts that are coded by the analogue magnitude representation. *A priori*, one might not have expected the brain to develop fibre connections in the language system in order to encode classroom activities assumed to develop neural structures for mathematics.

Research Question 2: The Interconnections Between Neural Structures

Analyses that reveal the interconnections between different neural structures can provide a variety of information relevant to education. Connectivity analyses reveal the white matter tracts formed by bundles of fibre connections, called by neuroscientists 'information highways' in the brain. Connectivity analyses can be informative with respect both to how a child is achieving a particular level of behavioural performance, and with respect to which developmental pathways were important in attaining that particular level of performance. For example, connectivity analyses carried out with children at risk for reading difficulties have revealed two types of neural system supporting reading (Shaywitz and Shaywitz, 2005). One group of children studied longitudinally by Shaywitz and Shaywitz, called persistently poor readers (PPR), had met criteria for poor reading both at the beginning of the study in 2nd/3rd Grade, and also later in the

study in 9th/10th Grade. Another group of children, called accuracy-improved poor readers (AIR), had met the criteria for poor reading in 2nd/3rd Grade but no longer met these criteria in 9th/10th Grade. Connectivity analyses suggested that reading achievement depended on rote memory for the PPR group, whereas the AIR group were utilising the same neural networks as a control group of children who had never had a reading impairment. Nevertheless, the AIR group were not activating these neural networks as efficiently as the control children. Hence apparently identical performance (in this case, the number of words read correctly) depended on *different* interconnections. The finding that the PPR group had developed fibre connections to right hemisphere systems supporting working memory and memory retrieval rather than between the traditional left-hemisphere language systems that typically support reading is suggestive of different developmental pathways. This kind of information can be useful for education, for example in suggesting the optimal focus for further remediation.

Research Question 3: The Time Course of Neural Activation

Children who are at risk for reading difficulties also show differences in the time course of neural activation in comparison to typically-developing children. One useful imaging method for gaining precise information about neural timing is EEG (electroencephalography). EEG measures the extremely low voltage changes caused by the electro-chemical activity of brain cells by using highly sensitive electrodes that are placed on the scalp. This method yields information precise to the millisecond. EEG studies suggest that the brain has decided whether it is reading a real word or a nonsense word within 160–180 ms of presentation, for children as well as for adults (e.g. Csepe and Szücs, 2003; Sauseng, Bergmann and Wimmer, 2004). When children who are at risk of reading difficulties are shown printed nonsense words (like 'lan') in kindergarten, they are significantly slower to show neural activity (Simos *et al.*, 2005; this study used a variant on EEG that combines information on electrical activity with information on blood flow, called MSI or magnetic source imaging). The researchers found this slowed neural activation even when the children were shown single letters. The at-risk group required almost 100 ms more for neural activity to begin in comparison to kindergartners who were not at risk for reading difficulties. Simos *et al.* (2005) reported that delayed activation actually increased over time for the at-risk group. For example, when providing the sounds for single letters at the end of Grade 1, processing time had increased by around 200 ms for the at-risk children. No increase was found for the children who were not at risk. Such information is important, as it is evidence that the child is not simply disengaged from the task or 'not trying'. Rather, the child has a specific neural processing difficulty.

The time course of neural activation can also be used to test inferences drawn from fMRI studies about the role of different neural structures. For

example, in their study of exact and approximate addition, Dehaene and his colleagues tracked the precise time course of brain activation using EEG. They found that the electrical signals to exact versus approximate trial blocks were already different by 400 ms. This was before the possible answers to the additions were displayed. Dehaene *et al*. argued that the time course data supported the idea that exact calculation and approximate calculation relied on different neural structures. If the two types of calculation had relied on the same neural structure, then the time course of activation should not have differed *before* the answers were displayed.

Research Question 4: Neural Correlation versus Causation

Given the current state-of-the-art in brain imaging, most neuroimaging data are correlational and do not provide information about causation. As in all scientific enquiry, therefore, experimental design is crucial to how useful the data will be for contributing to research questions. For example, it is important to control for other factors that might be important for any correlations that are found, and to use control groups. As in all developmental science, longitudinal studies are the most informative, and the studies discussed above taken from the reading literature are a good example (Simos *et al*., 2005; Shaywitz and Shaywitz, 2005). Intervention studies are also required when reliable correlations are found, in order to manipulate the variables in any association. This enables the study of 'dose-response relationships': if a particular factor is having a particular effect, then receiving more of that factor should increase the effect. A good example of intervention studies in education comes from the extensive literature supporting the importance of phonological awareness for reading development. Interventions to boost phonology show average effect sizes of 0.70, showing that experiencing an intervention makes a large difference to later attainment in reading (Bus and van IJzendoorn, 1999).

AVOIDING THE SEEDING OF NEUROMYTHS

When evaluating neuroscience research, it is important to be vigilant: correlations are still correlations, even when they involve physiological measures. Yet many correlational findings that reach the popular media are given causal interpretations. A good example comes from data sets that have been interpreted to show that fatty acids such as fish oils play a potentially causal role in learning. Unsaturated fatty acids are important in brain development and in neural signal transduction, which has led to the belief that omega-3 and omega-6 highly unsaturated fatty acids may be good for the brain. For example, in a recent paper, Cyhlarova *et al*. claimed that 'the omega-3/omega-6 balance is particularly relevant to dyslexia' (Cyhlarova *et al*., 2007, p. 116). This claim was based on a study measuring the lipid fatty acid composition of red blood cell membranes in 52 participants, 32 dyslexic adults and 20 control adults. No differences

between dyslexics and controls were found in any of the 21 different measures of membrane fatty acid levels taken by the researchers. However, a *correlation* was found between a total measure of omega-3 concentration and overall reading in the whole sample. This correlation by itself does not show that fatty acids have anything to do with dyslexia. The correlation depends on the whole group, potentially relevant variables such as I.Q. have not been controlled, and there is no intervention to test whether the association is a causal one. No plausible mechanism was proposed by the experimenters to explain why omega-3 concentration should have specific effects on reading, rather than general effects on any culturally-acquired skill (for example, arithmetic).

Nevertheless, when physiological variables such as changes in brain activation are involved, it is easy to suspend one's critical faculties. This has been demonstrated empirically by Weisberg and her colleagues (Weisberg *et al.*, 2008). These researchers gave adults bad explanations of psychological phenomena, either with or without accompanying neuroscientific information. The neuroscientific details were completely irrelevant to the explanations given, and yet the adults rated the explanations as far more satisfying when such details were present. The researchers concluded that the neuroscience details were very seductive. These details suggested to the participants that the explanations given were part of a larger explanatory system based on physiology. These seductive details then interfered with participants' ability to judge the quality of the explanations that they were being given. Weisberg *et al.* point out that this propensity to accept explanations that allude to neuroscience makes it all the more important for neuroscientists to think carefully about how neuroscience information is viewed and used outside the laboratory.

PRINCIPLES OF LEARNING FROM COGNITIVE NEUROSCIENCE

Although the field of educational neuroscience is still relatively new, there are a number of principles of learning demonstrated by empirical studies that can safely be incorporated into education and teaching. Some of these are now discussed.

A. Learning is Incremental and Experience Based

Although this may seem almost too trivial to note, the fact that the brain develops fibre connections to encode each experience that we have into our nervous systems is fundamental when considering education. The growth of interconnected networks of simple cells distributed across the entire brain eventually results in complex cognitive structures such as 'language' or 'causal knowledge'. The complexity that can be achieved by simple incremental learning has been demonstrated by connectionist modelling, which has shown a number of important 'in principle' effects. For example, complex 'rules' or principles such as the syntactic 'rules' of

language can be learned incrementally. There is no need for a specialised 'language acquisition device' (see Chomsky, 1957) that pre-encodes 'innate' knowledge about the general rules that all languages obey, along with innate knowledge of permitted variations. Simple incremental learning also yields 'critical period' effects, originally argued to show that some kinds of learning are particularly effective during a given time window. Connectionist modelling shows that these critical time windows are a natural part of the learning trajectory when learning is incremental. Similarly, incremental learning processes can explain apparent 'gaps' in learning exhibited by children. Learning is distributed across large networks of neurons, and so factors like the number of relevant neurons firing, their firing rates, the coherence of the firing patterns, and how 'clean' they are for signalling the appropriate information will all vary depending on how the current environmental input activates the existing network (see Munakata, 2001). As the fibre connections growing in response to received inputs are strengthened over time, it can also become difficult to reorganise the system when a new learning environment is experienced. This offers one potential mechanism for explaining why it is more difficult to learn a second language later in life (Munakata and McClelland, 2003). The fibre connections representing the phonemes of the first language are already entrenched, so (for example) Japanese speakers who have always encoded 'l' and 'r' as the same sound find it difficult to now encode them as different sounds (the same neural network will be activated by new tokens of both 'l' and 'r'). However, some people can learn to be fluent in a second language later in life, and so entrenchment is not irreversible. In the case of the 'l'/'r' distinction, multi-sensory input appears to be particularly effective (seeing as well as hearing other speakers produce the target sounds, Hardison, 2003).

These demonstrations of the importance of incremental environmental input show that the learning environments created in schools by teachers and other professionals will have important cumulative effects. Clearly, it is important to avoid creating learning environments that support the acquisition of maladaptive connections, for example environments that feel unsafe or stressful. Further, deep understanding of a given educational domain is required in order to present the cumulative information in the optimal sequence for the novice learner (this reflects classic educational concerns for consolidation of learning coupled with progression in curricula). The growth of new fibre connections in the brain always occurs in response to new inputs, and so claims that brain-based learning packages enable 'neuroplasticity' is a redescription of what always occurs for any new learning experience. Perhaps most challengingly, the biological necessity for learning to be incremental questions the notion that we can ever engender 'conceptual change'. Any neural network develops over time, and cannot suddenly be 'restructured' by one learning experience. On the other hand, certain experiences may result in previously distinct parts of the network becoming connected, or in inefficient connections that were impeding understanding being pruned away. Eventually, connectionist models may be able to show how and

when this is achieved in a given learning trajectory. Current connectionist modelling is limited by available statistical and mathematical alogorithms, and so the 'in principle' effects discussed above should be taken as suggestive rather than as established fact. Advances in the mathematical modelling of complex networks may reveal other principles of incremental learning.

B. Learning is Multi-Sensory

Different neural structures are specialised to encode different kinds of information, with sensory information being the most obvious example (e.g. visual information is encoded primarily by fibre growth in the visual cortex, auditory information is encoded primarily by fibre growth in the auditory cortex). However, most environmental experiences are multi-sensory, and therefore fibre connections between modalities are ubiquitous. Furthermore, because learning is encoded cumulatively by large networks of neurons, cell assemblies that have been connected because of prior experiences will continue to be activated even when a *particular* aspect of sensory information in a *particular* experience is *absent*. This ability of the brain to respond to *abstracted dependencies* of particular sensory constellations of stimuli enables, for example, a missed word to be filled in when someone coughs across another speaker. Even though our brain received the sensory information about the cough rather than the phonemes in the missed word, prior learning of the statistical regularities between words in connected discourse enables the brain to 'fill in' the missing information (e.g. Pitt and McQueen, 1998, for a related example).

   Although this example comes from within a modality (audition), the brain does the same thing across our senses (Noesselt *et al.*, 2007). Again, this principle implies that if children are taught new information using a variety of their senses, learning will be stronger (that is, learning will be represented across a greater network of neurons connecting a greater number of different neural structures, and accessible via a greater number of modalities). A nice example comes from a study by James (2007), who used fMRI to track the neural networks that developed as preschool children learned to recognise letters. Before taking part in the learning activities, the children's brains were imaged as they looked at letters and other familiar visual stimuli (such as cartoon hearts). As might be expected, significant activation of visual neural structures was found. As part of their initial reading activities, the children were then taught to recognise and write the letters. For example, visual letter knowledge was increased by helping children to recognise the target letters in story books, and to pick out target letters from 4 alternative possibilities (including reversed letters). The children were also taught to form and write the letters, thereby using another modality (kinaesthetic) in conjunction with the visual and auditory modalities (recognising and naming the letters). Following the writing training, the children's brains were again imaged

while they looked at letters and other familiar visual stimuli. This time, significant activation in motor areas was found for the letters, even though the children were *looking* at the letters and were not making any writing movements. Because their multi-sensory learning experiences had led fibre connections to develop both within the visual system and between the visual system and the motor system, the motor parts of the network were activated even though the current environmental experience was purely visual. Whether this multi-sensory representation can actually be interpreted as showing stronger learning requires further empirical work. Nevertheless, it is clear that information stored in multiple modalities is being activated despite the fact that sensory stimulation is only occurring in one modality (here, vision). This kind of empirical paradigm offers a way of investigating whether children really can be said to have different learning 'styles', for example being 'visual' or 'kinaesthetic' learners. Given the principles of how the brain learns, this seems *a priori* unlikely.

C. Brain Mechanisms of Learning Extract Structure from Input

The fact that incremental learning yields abstracted dependencies is a powerful mechanism for cognitive development (Goswami, 2008a). As the brain experiences particular sensory constellations of stimuli over multiple times, what is common across all these experiences will naturally be represented more strongly than what is different. This is because the fibre connections that encode what is common will become stronger than the fibre connections that encode the novel details. This mechanism effectively yields our 'basic level concepts', such as 'cat', 'dog', 'tree' and 'car' (Rosch, 1978). After 100 'cat' experiences, the strongest fibre connections will represent what has been common across all experienced instances, such as '4 legs', 'whiskers', 'tail', and so on. Therefore, the brain will have developed a generic 'prototype' representation of a cat. Sensory constellations of stimuli are also dynamic in space and time, and so simply by processing features of the input, and correlations and dependencies between these features, the brain will be learning about dynamic spatio-temporal structure and therefore about causal relations (Goswami, 2008c). This means that the child's brain can in principle construct detailed conceptual frameworks from watching and listening to the world. As we learn language and attach labels to concepts, the neural networks become more complex, and as we learn new information via language, fibre connections will form in response that encode more abstract information and therefore more abstract concepts.

I discuss many examples of this in my book (Goswami, 2008a), but to take an example pertinent to education, these learning mechanisms mean that the brain will extract and represent structure that is present in the input *even when it is not taught directly*. An example is the higher-order consistencies in the spelling system of English that I have previously described as 'rhyme analogies' (Goswami, 1986). Spelling-to-sound relations in English can often be more reliable at the larger 'grain size'

of the rhyme than at the smaller 'grain size' of the phoneme (Treiman *et al.*, 1995). For example, the pronunciation of a single letter like 'a' differs in words like 'walk' and 'car' from its pronunciation in words like 'cat' and 'cap'. The pronunciation in 'walk' or 'car' can be described as irregular, but it is quite consistent across other rhyming words (like 'talk' and 'star'). One way of exploring whether the brain is sensitive to these higher-level consistencies in letter patterns is to see how children read aloud novel 'nonsense' words that they have not encountered before. For example, children can be asked to read nonsense words matched for pronunciation like *daik*, *dake*, *loffi* and *loffee*. Only the rhyme spelling patterns in items like *dake* and *loffee* will have been learned from prior experiences with analogous real words (like 'cake' and 'toffee'). Hence only these 'chunks' of print and their connections with sound should be stored in the neural Visual Word Form Area (VWFA). There are no real English words with letter chunks like '*aik*'. English children indeed show a reliable advantage for reading aloud such analogous nonsense words, despite the fact that the individual letter-sound correspondences in the non-analogous items (e.g. '*d*', '*ai*', '*k*' in *daik*) were matched for orthographic familiarity to those in the analogous items (like *dake*, see Goswami *et al.*, 2003). Such data suggest that orthographic learning, presumably in the VWFA, reflects these higher-level consistencies, even though 'rhyme analogy' reading strategies have not been taught directly to these children (see Ziegler and Goswami, 2005, for converging evidence using other paradigms).

Whether learning (i.e. reading performance) would be even stronger if such strategies *were* taught directly to children remains an open question. In fact, this is an important general question for education. Although it is often noted that learning is 'embedded' in the experiences of the individual, one goal of education is to help all individuals to extract the higher-order structure (or 'principles' or 'rules') that underpin a given body of knowledge. It is generally felt that a combination of 'discovery led' and directly transmitted knowledge provides the best way of doing this, but there are many disagreements over the optimal balance between such teaching methods in different domains and in different views of pedagogy. A deeper understanding of how the brain uses incremental experience to extract underlying structure may help to inform such debates.

D. Learning is Social

We have social brains. The wealth of studies of infant and animal cognition are showing more and more clearly that the complex mammalian brain evolved to flourish in complex social environments. For example, there appear to be specialised neural structures in the human brain for encoding information about agents and their goal-directed actions (the so-called mirror neuron system—see for example Iacoboni *et al.*, 2005). Neurons in this system will respond to a biological agent

performing a certain action (e.g. a person lifting a cup), but not to a robot performing an identical action (Tai *et al*., 2004). In infancy, the attribution of intentionality to the goal-directed actions of others occurs surprisingly early in development. For example, Meltzoff has shown that 14-month-old infants will imitate actions that they have never witnessed, but which they infer to have been intended through watching a particular goal-directed action by a biological agent. If a baby watches an adult carrying a string of beads towards a cylinder, missing the opening and dropping the beads, the watching baby inserts the beads straight into the cylinder (see Meltzoff, 1995). Babies who watch a robot hand modelling an action, even a completed action, do not imitate. Gergely, Bekkering and Kiraly (2002) demonstrated that infants of the same age make very sophisticated inferences about intentional states. They used a different imitation paradigm devised by Meltzoff, in which an experimenter activates a light panel by leaning forwards and pressing it with her forehead. Sixty-nine percent of babies who watched this event also switched on the light panel by using their foreheads. However, in another version of the event, the experimenter said that she felt cold, and she held a blanket around her shoulders with her hands while she illuminated the light panel with her head. Now 79% of watching 14-month-olds used their hand to illuminate the light panel. They appeared to infer that in this latter scenario the experimenter used her forehead because her hands were constrained. She was not using her forehead intentionally because it was necessary to achieve her goal. Hence they no longer copied the 'forehead' action. As Carpenter, Ahktar and Tomasello (1998) have pointed out, infants who selectively imitate only the intentional acts of others will thereby acquire many significant cultural skills.

The social nature of human learning means that learning with others is usually more effective than learning alone, and that language and communication are central to this social process. This was recognised a long time ago in the theory of cognitive development proposed by Vygotsky (1978). Vygotsky argued that cognitive development did not just happen in the brain of the individual child. It also depended on interactions between the child and the cultural tools available for mediating knowledge. A primary cultural tool was language. As well as providing a symbolic system for communication with others, language enabled children to reflect upon and change their own cognitive functioning (in the terminology of modern psychology, by developing 'metacognitive' and 'executive function' skills). Furthermore, Vygotsky proposed the notion of the 'zone of proximal development' (ZPD). In contrast to independent problem solving, the ZPD was the larger area of potential development that was created when learning was supported by others. Most ZPDs are created by the social and cultural contexts created by parents and teachers (although Vygotsky argued that play with other children also creates important ZPDs). This theoretical perspective is reflected in education in socio-cultural theory. Csibra and Gergely (2006) revisited these theoretical ideas, recasting them in the language of brain science. They proposed that human brains have adapted to transfer

relevant cultural knowledge to conspecifics, and to *fast-learn* the contents of such teaching via a species-specific social learning system. They called this learning system 'pedagogy'. On my reading, their definition of pedagogy is essentially cultural knowledge transfer via collaborative learning in the zone of proximal development. Gergely, Egyed and Kiraly (2007) have further pointed out that such a system requires a default assumption about other agents, which is that they are trustworthy and benevolent sources of universally shared cultural knowledge.

E. Cortical Learning can be Modulated by Phylogenetically Older Systems

Traditionally, cognitive psychology has separated the study of cognition (thinking and reasoning) from the emotions. Neuroscience has shown that cognitive and emotional processes are integrated in the brain at multiple levels. For example, cortical structures such as orbito-frontal cortex integrate cognitive and emotional information during learning via interactions with phylogenetically older structures that are primarily involved in emotional processing, such as the amygdala. The amygdala controls our response to threat (as in the 'fight-flight' response). A particularly active area of neuroscience with respect to the emotional modulation of cognition is decision making ('neuroeconomics'; Coricelli, Dolan and Sirigu, 2007). Economics as a discipline has learned that human behaviour cannot be explained solely in rational cognitive terms. Learning based on cumulative emotional experience plays an important role in anticipating the possible future consequences of economic choices, and economic models incorporating emotional measures such as regret appear to be more efficient at modelling human behaviour. Cumulative emotional experience must also play a role in the efficiency of learning. This suggests that models of classroom learning must incorporate the emotions in order to better understand the behaviour of learners.

There is certainly some relevant data from both human and animal cognitive neuroscience. For example, learning in adults is impaired when the impact of punishment relevant to reward is enhanced, and the neurotransmitters that control this process can be identified (Cools *et al.*, 2005). The effect of anxiety on impeding learning can also be increased or reduced pharmacologically (i.e. with drugs). Animal models show that when previously innocuous sensory signals are processed by the amygdala as emotionally salient and aversive, this affects learning in rats (Stutzmann and LeDoux, 1999). The emotional information is prioritised by the brain, and receives privileged access to attention. When experiences are aversive, emotional responding blocks learning. The prioritising of emotional information was presumably evolutionarily adaptive, as emotional stimuli lead to enhanced sensory processing by the brain, enabling better behavioural responding. Such studies may provide a neural framework for explaining why children who are anxious do not learn efficiently. However, the data are suggestive rather than conclusive. Furthermore, there are no comparable animal models exploring the effects of highly

rewarding events on attention and learning (rather than aversive events). Emotions seem likely to produce both enhancement and costs to efficient learning, but how this might operate in learning environments like classrooms is currently not well-understood, at least by neuroscience.

F. Learning Shows Life-Long Plasticity and Compensation

Studies of the brain also suggest that it is never too late to learn. Some neural structures are still developing in the mid-twenties (e.g. the frontal cortex), and experience-dependent plasticity means that fibre connections continue to form to represent new learning throughout adulthood. Learning in adulthood often enlarges the relevant neural networks, as in the case of skilled adult pianists, who have enlarged cortical representations in auditory cortex specific to piano tones (Pantev *et al.*, 1998). In London taxi drivers, who need a good 'mental map' of London (called 'The Knowledge'), a small brain area involved in spatial representation and navigation called the hippocampus becomes enlarged. Hippocampal volume is correlated with the amount of time spent as a taxi driver (Maguire *et al.*, 2000). Similarly, the cortical representation of piano tones was found by Pantev and colleagues to be correlated with the amount of time that different individuals spent in piano practice. Greater synaptic density is associated with more learning. When neural networks are lesioned or partially destroyed (for example, by a stroke), connections can re-form. There appear to be differences between neural structures in how effectively this can occur, nevertheless these compensatory mechanisms (which can lead to significant recovery of pre-stroke function) again show that plasticity continues into late adult life. In terms of the brain mechanisms for learning, therefore, there is considerable continuity. This principle of learning provides empirical support for the efficacy of lifelong access to education ('lifelong learning').

Finally, I consider two other aspects of cognitive neuroscience research that are relevant to education.

BIOMARKERS OF LEARNING EFFICIENCY

Neuroscience also offers the promise of biomarkers or neural markers for learning. Cognitive processes are difficult to study directly, because they are theoretical rather than observable. Even processes such as memory have to be inferred from behaviour–we cannot yet pinpoint our memories as linked to particular brain cells, for example, even though the main cell assemblies or neural structures supporting memory are known. Nevertheless, when systematic experimentation enables a meaningful relationship to be identified between hypothetical cognitive processes and neural variables, the neural variables can become physiological markers or biomarkers of these cognitive processes. Therefore, the neural variables can be used to identify those who might be at educational risk. For example, a child may be at risk because aspects of sensory processing are

impaired, and biomarkers could show the presence of this processing impairment before any behavioural symptoms have appeared (Goswami, 2008b). The identification of neural markers would thereby enable very early intervention, before a potential learning difficulty has had time to become entrenched. Such interventions are most likely to be environmental, and thereby the provenance of education. Other potential biomarkers include genotypes and blood, plasma and spinal fluid markers, for example based on protein changes.

   While the latter biomarkers have been explored most in relation to diseases of the brain such as Alzheimer's disease, neural variables are being explored in relation to learning difficulties. Developmental dyslexia, developmental dyscalculia, specific language impairment, autism and attention deficit hyperactivity disorder are all active areas of research. For example, aspects of language function can be measured very early using EEG. Particular components of the waveform evoked in response to particular aspects of language input (evoked response potentials or ERPs) indicate the integrity of phonological (sound-based) versus semantic (meaning-based) aspects of language processing. These ERP components are already proving to have predictive value in determining who might be at risk of a language impairment (Friedrich and Friederici, 2006). As biomarkers like these can be measured in sleeping babies, they offer opportunities for very early identification of educational risk, without requiring attentive responding. As children with learning difficulties often have attentional difficulties, these methods offer a marked improvement over behavioural indices. However, they do raise potential ethical concerns. For example, early identification may result in benefit, but may also result in harm if the child is then placed in a category that carries stigma. There is also the question of who has the responsibility for detecting and monitoring these biomarkers. The assumption here is that benefits will outweigh harms, as the interventions that can follow early identification should improve learning trajectories over the whole life course. Nevertheless, there is a responsibility to ensure that children do not experience negative discrimination (for example, lowered expectations of their learning abilities) as a result.

MAXIMISING THE LEARNING EFFICIENCY OF THE BRAIN

The brain is an organ of the body, and in terms of efficiency of function it is not distinct from other organs such as the heart with respect to some of the basic factors that affect its ability to work effectively. Nutrition and diet are important for effective function, as are sleep and exercise. The obvious roles of nutrition and exercise have led to various 'brain based learning' claims about particular types of nutrition (e.g. fish oils, water) and particular types of exercise (e.g. Brain Gym[R], see Howard-Jones, 2007, for a useful discussion) and their particular effectiveness for learning by children. When evaluating such claims, it is of course critical to check the quality of the science cited in their support, and to distinguish

correlational data from causal data. In terms of underlying physiological mechanisms, there have been some pharmacological insights with respect to sleep, which appears to have an interesting role in consolidating learning. During sleep, there is behavioural inactivity accompanied by distinct electrophysiological changes in brain activity. These changes seem to affect memory, which appears to be consolidated during slow-wave and rapid-eye-movement sleep via the actions of certain neuro-transmitters. Newly encoded memories are stabilised, and are integrated with pre-existing (long term) memories via the action of neurotransmitters like acetylcholine. Low cholinergic activity is associated with slow-wave sleep, and when cholinergic activity is blocked in the brain, there is enhanced consolidation of memory (see Marshall and Born, 2007, for a summary).

One possibility is that we lose consciousness during sleep because the brain needs to use the same neural networks that support conscious activity for the processing and long-term storage of recently-acquired information. Indeed, empirical studies of learning finger tapping sequences show that there is improvement in the learned skill after sleep, with performance levels after sleep that are significantly higher than performance levels at the end of initial training, before sleep occurred (e.g. Fischer *et al.*, 2002). However, sleep has to occur within a certain time window after training in order for benefits to accrue (this time window is approximately within 16 hours of the learning period). Conversely, severe insomnia is associated with decrements in learning. The growing evidence base with respect to sleep shows the importance of these basic aspects of human behaviour to education.

CONCLUSION

Cognitive neuroscience is important for education because it enables a principled understanding of the mechanisms of learning and of the basic components of human performance. It also enables componential understanding of the complex cognitive skills taught by education. Many of the principles of learning uncovered by cognitive neuroscience might appear to support what teachers knew already. For example, aspects of pedagogy such as the value of multi-sensory teaching approaches or of creating safe and secure environments for learning are highly familiar. Nevertheless, cognitive neuroscience offers an empirical foundation for supporting certain insights already present in pedagogy and disputing others. The evidence from neuroscience is not just interesting scientifi-cally. It enables an evidence base for education in which mechanisms of learning can be precisely understood.

Nevertheless, the evidence base that it offers is a challenging one. An interesting analogy is provided by Clark, who was discussing language, but whose analogy also works for the entire cognitive system. Clark's argument is that we can conceptualise the brain as a 'loose-knit, distributed representational economy' (Clark, 2006, p. 373). Some elements in the

economy might conflict with other elements in the economy, but this is inevitable, as there is no 'homunculus' or single central overseer who determines learning. Rather, there are many interacting parts of the overall reasoning machinery that the brain is maintaining at the same time. The activity of all of these parts is what the child brings to the classroom, and different parts are more or less affected by different cognitive or emotional experiences. The child brings a 'vast parallel coalition of more-or-less influential forces whose unfolding makes each of us the thinking beings that we are' (ibid.). To borrow from another insightful commentator on the potential of cognitive neuroscience for cognitive development (Diamond, 2007), the truly ambitious goal for education is to cross and integrate the disciplinary boundaries of biology, culture, cognition, emotion, perception and action. Biological, sensory and neurological influences on learning must become equal partners with social, emotional and cultural influences if we are to have a truly effective discipline of education.

Correspondence: Usha Goswami, Centre for Neuroscience in Education, Faculty of Education, University of Cambridge, 184 Hills Rd, Cambridge CB2 8PQ, UK.
E-mail: ucg10@cam.ac.uk.

REFERENCES

Bus, A. G. and van IJzendoorn, M. H. (1999) Phonological Awareness and Early Reading: A Meta-Analysis of Experimental Training Studies, *Journal of Educational Psychology*, 91.3, pp. 403–414.

Carpenter, M., Akhtar, N. and Tomasello, M. (1998) Fourteen- through 18-month-old Infants Differentially Imitate Intentional and Accidental Actions, *Infant Behavior and Development*, 21, pp. 315–330.

Chomsky, N. (1957) *Syntactic Structures* (The Hague/Paris, Mouton).

Clark, A. (2006) Language, Embodiment, and the Cognitive Niche, *Trends in Cognitive Sciences*, 10.8, pp. 370–374.

Cohen, L. and Dehaene, S. (2004) Specialization within the Ventral Stream: The Case for the Visual Word Form Area, *NeuroImage*, 22, pp. 466–476.

Cools, R., Roberts, A. C. and Robbins, T. W. (2005) Serotoninergic Regulation of Emotional and Behavioural Control Processes, *Trends in Cognitive Sciences*, 12.1, pp. 31–40.

Coricelli, G., Dolan, R. J. and Sirigu, A. (2007) Brain, Emotion and Decision Making: The Paradigmatic Example of Regret, *Trends in Cognitive Sciences*, 11.6, pp. 258–265.

Csibra, G. and Gergely, G. (2006) Social Learning and Social Cognition: The Case for Pedagogy, in: Y. Munakata and M. H. Johnson (eds) *Processes of Change in Brain and Cognitive Development. Attention and Performance XX1* (Oxford, Oxford University Press), pp. 249–274.

Csepe, V. and Szücs, D. (2003) Number Word Reading as a Challenging Task in Dyslexia? An ERP Study, *International Journal of Psychophysiology*, 51, pp. 69–83.

Cyhlarova, E., Bell, J., Dick, J., Mackinlay, E., Stein, J. and Richardson, A. (2007) Membrane Fatty Acids, Reading and Spelling in Dyslexic and Non-Dyslexic Adults, *European Neuropsychopharmacology*, 17, pp. 116–121.

Dehaene, S. (1997) *The Number Sense: How the Mind Creates Mathematics* (New York, Oxford University Press).

Dehaene, S. (2008) Cerebral Constraints in Reading and Arithmetic: Education as a 'Neuronal Recycling' Hypothesis, in: A. M. Battro, K. W. Fischer and P. J. Lena (eds) *The Educated*

Brain: Essays in Neuroeducation (Cambridge, Pontifical Academy of Sciences and Cambridge University Press), pp. 232–247.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R. and Tsivkin, S. (1999) Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence, *Science*, 284, pp. 970–974.

Diamond, A. (2007) Interrelated and Interdependent, *Developmental Science*, 10.1, pp. 152–158.

Fischer, S., Hallschmid, M., Elsner, A. and Born, J. (2002) Sleep Forms Memory for Finger Skills, *Proceedings of the National Academy of Sciences*, 99, pp. 11987–11991.

Friedrich, M. and Friederici, A. D. (2006) Early N400 Development and Later Language Acquisition, *Psychophysiology*, 43, pp. 1–12.

Frith, U. (1985) Beneath the Surface of Developmental Dyslexia, in: K. E. Patterson, J. C. Marshall and M. Coltheart (eds) *Surface Dyslexia* (London, Lawrence Erlbaum Associates), pp. 301–330.

Gergely, G., Bekkering, H. and Király, I. (2002) Rational Imitation in Preverbal Infants, *Nature*, 415, p. 755.

Gergely, G., Egyed, K. and Király, I. (2007) On Pedagogy, *Developmental Science*, 10.1, pp. 139–146.

Goswami, U. (1986) Children's Use of Analogy in Learning to Read: A Developmental Study, *Journal of Experimental Child Psychology*, 42, pp. 73–83.

Goswami, U. (2008a) *Cognitive Development: The Learning Brain* (Philadelphia, PA, Psychology Press of Taylor and Francis).

Goswami, U. (2008b) Reading, Complexity and the Brain, *Literacy*, 42.2, pp. 67–72.

Goswami, U. (2008c) Analogy and the Brain: A New Perspective on Relational Primacy, *Behavioural and Brain Sciences*, 31, pp. 387–388.

Goswami, U., Ziegler, J., Dalton, L. and Schneider, W. (2003) Nonword Reading Across Orthographies: How Flexible is the Choice of Reading Units?, *Applied Psycholinguistics*, 24.2, pp. 235–47.

Hardison, D. (2003) Acquisition of second–language speech: effects of visual cues, context and talker variability, *Applied Psycholinguistics*, 24, pp. 495–522.

Howard-Jones, P. (2007) *Neuroscience and Education: Issues and Opportunities*. Commentary by the Teaching and Learning Research Programme (London, TLRP). Available online at: http://www.tlrp.org/pub/commentaries.html

Iacoboni, M., Molnar–Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C. and Rizzolatti, G. (2005) Grasping the Intentions of Others with One's Own Mirror Neuron System, *PLOS Biology*, 3, pp. 529–35.

James, K. H. (2007) *Perceptual-Motor Interactions in Letter Recognition: fMRI Evidence*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Boston, MA, 29 March–1 April 2007.

Kopera-Frye, K., Dehaene, S. and Streissguth, A. P. (1996) Impairments of Number Processing Induced by Prenatal Alcohol Exposure, *Neuropsychologia*, 34, pp. 1187–1196.

Marshall, L. and Born, J. 2007 The Contribution of Sleep to Hippocampus-Dependent Memory Consolidation, *Trends in Cognitive Science*, 11, pp. 442–450.

Maguire, E. A., Gadian, D., Johnsrude, I., Good, C., Ashburner, J., Frackowiak, R. and Frith, C. (2000) Navigation-Related Structural Change in the Hippocampi of Taxi Drivers, *Proceedings of the National Academy of Sciences*, 97.8, pp. 4398–4403.

Meltzoff, A. N. (1995) Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children, *Developmental Psychology*, 31, pp. 838–850.

Munakata, Y. (2001) Graded Representations in Behavioral Dissociations, *Trends in Cognitive Sciences*, 5.7, pp. 309–315.

Munakata, Y. and McClelland, J. L. (2003) Connectionist Models of Development, *Developmental Science*, 6, pp. 413–429.

Noesselt, T., Rieger, J., Schoenfeld, M., Kanowski, M., Hinrichs, H. and Driver, J. (2007) Audiovisual Temporal Correspondence Modulates Human Multisensory Superior Temporal Sulcus, Plus Primary Sensory Cortices, *Journal of Neuroscience*, 27.42, pp. 11431–41.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E. and Hike, M. (1998) Increased Auditory Cortical Representation in Musicians, *Nature*, 393, pp. 811–814.

Pitt, M. A. and McQueen, J. M. (1998) Is Compensation for Coarticulation Mediated by the Lexicon? *Journal of Memory and Language*, 39, pp. 347–370.

Price, C. J. and McCrory, E. (2005) Functional Brain Imaging Studies Of Skilled Reading And Developmental Dyslexia, in: M. J. Snowling and C. Hulme (eds) *The Science of Reading: A Handbook* (Oxford, Blackwell Publishing), pp. 473–496.

Rosch, E. (1978) Principles of Categorisation, in: E. Rosch and B. B. Lloyd (eds). *Cognition and Categorisation* (Hillsdale, NJ, Erlbaum).

Sauseng, P., Bergmann, J. and Wimmer, H. (2004) When Does the Brain Register Deviances from Standard Word Spellings? An ERP Study, *Cognitive Brain Research*, 20, pp. 529–532.

Shaywitz, S. E. and Shaywitz, B. A. (2005) Dyslexia (Specific Reading Disability), *Biological Psychiatry*, 57, pp. 1301–1309.

Simos, P. G., Fletcher, J., Sarkari, S., Billingslet, R., Castillo, E., Pataraia, E., Papanicolaou, A., Francis, D. and Denton, C. (2005) Early Development of Neurophysiological Processes Involved in Normal Reading and Reading Disability: A Magnetic Source Imaging Study, *Neuropsychology*, 19.6, pp 787–798.

Snowling, M. J. (2000) *Dyslexia* (Oxford, Blackwell).

Stutzmann, G. E. and LeDoux, J. E. (1999) GABAergic Antagonists Block the Inhibitory Effects of Serotonin in the Lateral Amygdala: A Mechanism for Modulation of Sensory Inputs Related to Fear Conditioning, *Journal of Neuroscience*, 19.RC8, pp. 1–4.

Szücs, D. and Goswami, U. (2007) Educational Neuroscience: Defining a New Discipline for the Study of Mental Representations, *Mind, Brain and Education*, 1, pp. 114–127.

Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N. and Castiello, U. (2004) The Human Premotor Cortex is 'Mirror' Only for Biological Actions, *Current Biology*, 14, pp. 117–120.

Treiman, R., Mullennix, J., Bijeljac-Babic, R. and Richmond-Welty, E. D. (1995) The Special Role of Rimes in the Description, Use, and Acquisition of English Orthography, *Journal of Experimental Psychology: General*, 124, pp. 107–136.

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A. and Eden, G. F. (2003) Development of Neural Mechanisms for Reading, *Nature Neuroscience*, 6.6, pp. 767–773.

Turner, D. C. and Sahakian, B. J. (2006) Neuroethics of Cognitive Enhancement, *BioSocieties*, 1, pp. 113–123.

Vygotsky, L. (1978) *Mind in Society* (Cambridge, MA, Harvard University Press).

Weisberg, D. S., Keil, F. C., Goodstein, E. R. and Gray, J. R. (2008) The Seductive Allure of Neuroscience Explanations, *Journal of Cognitive Neuroscience*, 20.3, pp. 470–477.

Ziegler, J. C. and Goswami, U. (2005) Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory, *Psychological Bulletin*, 131.1, pp. 3–29.