Chapter 3

How the ‘brain’ learns

At an atomic level we would not be able to see any difference between what is going on in the materiality of a school desk and the materiality of a schoolchild sitting at the desk. Both would show a buzzing world of electrons, protons, mesons, and so on. If we expanded our level of focus outwards, at what point would a distinctly educational reality emerge? We could say, with some certainty, that the materiality of schooling coalesces out of atomic soup with desks, chairs, windows, books, writing equipment and all the other elementary artefacts of the classroom. With the physical body of a student, the smallest material unit of analysis would be something very different, something like a neuron interacting with other neurons in a specific way. At a microscopic level, a neuronal network is an alive process, humming with charges and changes; completely different from the dead microfibres of a piece of paper overlapping each other to give a surface. Neurons actively change their state and connections when learning happens; paper fibres absorb and break when written upon, they do not dynamically adapt and change themselves in an organised way. Note that the educational imagination is jumping from a tight focus on the smallest unit of analysis in the material dimension of the physical school to the smallest unit of analysis in the material dimension of the physical body of a learner. On the one hand it is the smallest of moves, from a material desk to the materiality of a learner occupying the desk. But on the other hand it is a massive jump from something that is not alive to something that is alive; and more than just alive, something that is self-aware.

It is important, at this point, not to lose our educational focus by suddenly plunging into the intricacies of neuroscience and biology. Our question is not about
how the human body and brain function, but what the assorted smallest and largest
structures are in the field of education; and the smallest physical unit of learning
in us is something like a neuron interacting with other neurons in a specific way.
I use the word specific because we are not looking at how the brain functions in
general, or how neurons function specifically, but what happens to students when
involved in school learning. We have to be clear that the interaction we study is
an educational reality, not just brain activity. This book is not about neuroscience.
It is about education in all its heights and depths, which touches, at its finest level
of focus, on the way neuronal functioning helps us understand the process of
learning; not just any learning that happens as life teaches us its hard and beautiful
lessons, but the way we learn within educational frameworks. There are a growing
number of texts that combine neurobiology with education. One that I have found
particularly useful in getting this focus right is The Unified Learning Model: How
Motivational, Cognitive and Neurobiological Sciences Inform Best Teaching Practices
by Shell et al. (2010). I have used it to structure the first third of this chapter.

Episodic and semantic knowledge

We can start our search for what is distinctly educational in the brain by looking
for a distinction similar to one between everyday experience and formal school
knowledge. Such a distinction can be found in the difference between episodic
knowledge and semantic knowledge. Episodic knowledge develops in the process
of living our life and is immersed in the tastes, smells, feelings and details of
particular daily experiences. We don’t normally have to force ourselves to attend
to these experiences: they happen and we are in the happening, participating in its
situated richness. The episodic nature of this knowledge results in specific detail
fading quickly from long-term memory while habits and dispositions grow from
the repeated nature of daily events surrounding eating, playing and sleeping.

For example, allow me to ask you how you spent your evening precisely
30 days ago. The evening will not spring readily to mind. You could do some
heavy reconstructive work to get there, by trying to track something specific
that happened on that day, or the previous day, and work from the memory
outwards, but as time periods get longer and longer you land up with trends: like
Thursday night is burger night, Friday night is getting out night and Sunday night
is for romance, or should have been if we weren’t both so bloody tired. It’s the
repeated harmonies in the ebb and flow of ordinary life that combine into scripts
that govern how we conduct ourselves; into long-term memories of what our
bedroom looked like; and into routines that allow us to function automatically
and do things without paying much attention to them. The massive experiential
richness of the actual event dwindles into the comfortable trace of the usual.

Semantic knowledge is memory that is specifically not about our own life: it
does not come automatically and easily from the personal process of experiencing
life (Shell et al., 2010, p. 38). It has a context, although this is not the rich intimacy of our own lived context, but one that has to be built up, step by painful step. Although semantic knowledge starts flush in the thick of lived experience, it has to construct a context that is delocalised, one that specifically relates to the knowledge event rather than to our own lives. Semantic knowledge will always be located within an episodic context as we are always learning in our bodies here and now, but it will slowly separate off from the lived context towards its own specific contextual set. That is why it is really important to ensure the lived experience of learning at school is a positive one because semantic knowledge starts out in the episodic experience of that classroom on that day and resonates with previous experiences and events. But as semantic knowledge grows it has to separate itself from the particular experience of everyday life and develop the internal logic of what is being actively learnt in its own terms. It formalises itself.

The key image to hold in mind when thinking about the difference between episodic and semantic knowledge is that episodic knowledge starts off rich and full of the tastes, sounds, touches, sights and feelings of life and reduces itself to a routine script of what is common to all the variations; whereas semantic knowledge starts off small and impoverished and slowly grows, link by artificial link. When a child starts to learn how to read you can see this fruitful tension at its clearest. On the one hand the teacher will labour to keep an experiential episodic richness and play in the classroom whilst at the same time beginning to introduce the completely arbitrary logic of letters and their combinations. Slowly, but surely, the child begins to learn explicitly something that builds on itself and creates formal links with concepts that are higher than it, or examples that are different from it, creating an organised semantic web (figure 3.1).

![Hierarchical semantic knowledge structure](image-url)

*Figure 3.1 Hierarchical semantic knowledge structure*
Note how this resonates with the educational imagination travelling up to more abstract concepts (canary – bird – animal) and across to equivalent levels (bird/fish; canary/ostrich/shark/salmon). But also note that a child who has had a rich episodic set of experiences with canaries, ostriches, sharks and salmon will make far more sense of the formal structure. The episodic experiences, even though they have reduced to traces, provide a fertile soil. It’s not an either/or with episodic and semantic knowledge; it’s and/and. The key point I want to make, however, is that semantic memory needs directed attention and motivation for a sustained period. It does not come automatically as a part of living, eating, exploring your surroundings and sleeping. It is hard work no matter how well integrated into tastes, sounds and actions; hard work because we have to direct and hold attention on something that is arbitrary and formal, and because, as we all know, attention is limited.

*Neuron basics*

And it is the way neurons work when we pay attention that gets us closer to our educational unit of analysis rather than the intricate biology of the neuron in its own terms. It’s how neurons function when we learn that is the focus, not its internal mechanisms; but even so we need a bare minimum of mechanical detail.

A neuron has two ends, one that receives impulses (dendrite) and another that releases impulses (axon). To release an impulse the neuron has to receive impulses that push it beyond its threshold. Only when this limit is broached does the neuron fire. In other words, the neuron gets excited from all the impulses coming its way. At a certain liminal point, if all these impulses sweetly hit the same spot, the neuron just can’t help itself and has to fire away. I can relate. Unlike me, the more a neuron fires, the easier it is for it to fire again. The firing ability of a neuron changes through use. But it’s not only that the firing threshold drops with use: the connections with other neurons involved in stimulating or receiving the firing are strengthened as well. If you are connected together and fire together then the connections are strengthened and it’s easier to do it all again next time round. This gets us to a simple elementary definition of learning: ‘Learning occurs when the firing ability of a neuron is changed’ (Shell et al., 2010, p. 8).

*The vital role of working memory*

The problem is that there is an enormous amount of sensory stimulation pouring into us and circulating around us at any given moment and we can only attend to some of what is entering and circulating. The space where this happens is working memory and it is here that we hit the interconnection between learning and the functioning of our brain head on. If we had unlimited amounts of working memory and could pay attention to everything at once then learning would be
something very different from that of human experience. The memory span of young adults when pushed to its maximum is around seven items, as discovered by Miller (1956) in his famous paper ‘The magical number seven, plus or minus two’. We know this from the way we struggle to remember someone’s cell phone number unless it has a sweet pattern.

However, most of our learning tasks do not involve simple bits, but of deciding what to do with a bit: is it right or wrong, up or down, inside or outside, here or there? Notice that this means we shall remember a lot less than seven items because our working memory has some of its space taken up with decision-making processes. This has resulted in differences between popular accounts of the capacity of working memory, varying all the way from 7+/−2 items to one item on which a decision is being made. The more complex the decision to be made, the more working memory is taken up with the options presented by the decision. More recent research points to four slots of working memory being more likely than 7+/−2, but the point holds – the more complex the action, the less bits you can work with.

It is the constricted way that working memory functions which fundamentally structures the learning process of the human species. The massive number of firing connections swim desperately to a restricting attention channel they have to flow through to get into long-term memory on the other side. Only some make it, but when they do we have the fertile production of knowledge. So what is this cervix of the brain called working memory and how does it work?

First, not everything you experience gets into working memory. There is way too much going on around and inside you to be dealt with by four slots. Your biological and cognitive make-up tends to allow what is novel or conspicuous into working memory. Second, if all slots of your working memory are taken up with a process, then other things, especially if they are not novel or conspicuous, are just ignored. This means that we are not really paying much attention to most of what is going on in our environment, as is shown by all sorts of experiments designed to make most of us look like idiots. One of the most famous is the passing a basketball video designed by Daniel Simons (you can try it yourself. Spoiler alert: if you want the video to work its magic, then don’t read what follows until you try it.)

The instruction you are asked to follow is to ‘count how many times the players wearing white pass the basketball’. You then see six students, three in white and three in black passing two basketballs to each other. I concentrated hard on counting the passes and got it wrong, somehow counting sixteen instead of fifteen. What I did not see was a gorilla walk right into the middle of the group, beat his chest and then walk out. My working memory was taken up by counting the white players passing and trying to ignore the black players, so I ignored the black gorilla strolling in and out of the scene, because he moved at the same pace as the others and was also dressed in black.
My students experience something similar on teaching practice when I chat to them afterwards about their lessons. They are so busy trying to remember what they needed to say and do that they show a remarkable lack of awareness about what the students are doing. When I ask them afterwards about key events in the class, they have hardly any memory of them.

We all function partly like this: to concentrate on something, we withdraw our attention from other things. When I am concentrating hard I do not hear what is going on around me. My mother used to have to come right up to me and shout to break my concentration when I was listening to Jet Jungle and Squad Cars on the radio; and even then I experienced her voice as coming from a distance. What I concentrated on I tended to remember, especially if I thought about it afterwards. It’s obvious on one level: learning requires attention on the one hand and repetition on the other. To do this properly you need to focus on the object at hand for a sustained period of time and then go over it again in some way. Attention is what keeps something in working memory from slipping away after twenty or so seconds; and repetition is what enables transfer from temporary memory to long-term memory. We have to be careful with repetition, as some pedagogues could read its importance as meaning rote learning and drill. Two points need to be made here: first, a distinction between shallow learning and rote learning; and, second, the need to transform what is being repeated for it to stick in long-term memory, especially with more complex areas of knowledge.

We cannot avoid shallow learning: it’s a natural part of learning something new where initially you don’t tend to get the full picture, or understand how the different parts hang together and the intricate nuances. But by sticking to the task, long-term memory develops out of a learner paying attention and making meaning. Rote learning is devoid of meaning: all that is focused on is getting the form of things into the mind through drilling. Just because you initially learn in a shallow way does not mean that you need to be drilled with rote memorising strategies.

You do need to repeat and go over what you are learning to embed it in long-term memory, but the repetition does not have to be in exactly the same form each time. Exactly the opposite, actually: it’s in making meaning of the element – of coming back to it from different angles and placing it within larger networks that give it a frame – that a shift from temporary to long-term memory happens. As a temporary memory shifts into long-term memory it transforms; and the reason why it changes is that it has to be inserted into already existing networks in long-term memory. You don’t just plonk something in long-term memory: it gets there in a more sustained and memorable way when you first get its essential meaning and second link it to other things you already know and understand. Getting the essential meaning and placing it in a larger network are really part and parcel of the same activity – making sense of something – because it is in seeing how the particular elements relate to other elements that the attachment of meaning
happens. This means that the temporary memory has to be transformed into something that makes sense through connecting it to a relevant network. The larger and more dense the already existing set of connections, the more able you are to place an element in a bigger picture and see how it holds there, enabling its shift into long-term memory. As the gospel of Matthew noted, ‘Whoever has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him’ (Matthew 13:12, New International Version of the Bible).

The real danger of pointing at automatisation is that teachers think the take home pedagogic message is that drilling their pupils is key. The problem with drilling is that students lose focus and motivation, and if they are not actually attending to the task at hand when learning then its effectiveness is lost. You need to attend to the topic and the best way to work with attention is to make the topic interesting, worthwhile and relevant, not necessarily to everyday experiences, but to the growing semantic network of knowledge in long-term memory. And the more the student actually knows how to do this herself, the more she gets control of hir own meta-cognitive processes, then the more effective learning becomes. There are many reasons why we celebrate the independent learner who takes control of hir own learning, but what we don’t often see is that active learners (rather than drilled learners) are supported by the working of our brains.

*The many become one and are increased by one: the wonderful world of chunking*

What is interesting about connecting things into networks is that the element now holds as a moment in a bigger picture. If you understand the big picture, then you have the elements of the picture contained within it. This is a vital point regarding the intersection of pedagogy with the four slots of working memory, probably the most important point of this chapter. Why?

The reason is that the big picture with all its elements holds as one thing, needing only one slot of working memory. It is really hard, almost impossible, to increase the number of slots you have in working memory. It is far easier to develop your understanding of the connections between elements and build them up into a bigger pattern that holds, and become acquainted with how the pattern works, until a point is reached where the whole pattern is one functioning unit. This is what experts do. They don’t have more slots in their working memory than novices: it’s partly that they work with bigger chunks and this makes it look as if they have really quick, sharp minds. There is a classic test that reveals this involving chess grandmasters, A-level (mid-range) players, and novices (De Groot, 1965). Each set of characters were given two to ten seconds to memorise a chessboard position with 25 pieces. Grandmasters were astounding, recreating positions from scratch with 93% accuracy. That’s basically getting only two of twenty-five pieces in the wrong place. A-level players were able to get
approximately half the pieces in the right place and novices only a third. It would seem that grandmasters have photographic memories. I say ‘seem’ because if you take the same 25 chess pieces and put them randomly on the chessboard, then suddenly the grandmaster’s memory is only as good as the novice next to him. He was not remembering 25 individual pieces and their places; he was remembering patterns of pieces (Chase and Simon, 1973). The thousands of hours spent playing and studying chess meant that the grandmaster was not playing with individual pieces, but with patterns of pieces, and patterns of patterns, each of which holds as one chunk.

You can see this if you watch the grandmaster closely as he reconstructs the board from memory (figure 3.2). First, he thinks a little and then places around six pieces down quickly. Then he thinks again before placing around another six down, and so on with the third and fourth sets. The setting down of the pieces occurs in bursts of patterns, not of individual pieces. The novice works with individual pieces rather than with patterns and so is really limited in this task.

I experienced this at first hand when playing a visiting chess grandmaster at my school, only he did not just play against me but against 29 other students as well, all of us arranged in a circle of desks in the hall. He walked from board to board taking a couple of seconds before making a move, with me, the novice, agonising about what move to make. I was working with individual pieces and trying to think ahead with something like the following thought process: if I move this
piece here, then he can move this piece or this piece or this piece, which means I could then either move this piece or that piece, which means … where was I again? And then he was back at my table, taking a couple more seconds before moving decisively. It was like a vice slowly strangling me: my options became more and more limited, I became pinned down and exchanges of pieces resulted in the loss of more of mine. Then, after going around the tables and playing with 29 other players, he came to my table, looked at the positions, made one move and it was checkmate. It felt like I was playing against a god, some kind of superhuman being. What I was really playing against was a person who had put many of the patterns and moves into long-term memory and was able to recall and manipulate them in working memory. As for me, I had no long-term memory network around chess, except how the pieces worked and two opening moves. I had to rely on working memory for individual moves, not for overarching patterns, and what I held in my head thirty seconds before just melted away.

The implications of this point for teaching and learning are clear, as numerous educators have pointed out (Kirschner, Sweller and Clark, 2006). I would like to put one implication forward in two complementary ways: the need for novice learners faced with novel information to have explicit instruction; and recognition of how crucial it is to build up dense complexes of integrated information in long-term memory as a key component of developing problem solving skills.

If you are a novice on X, then you are going to need clear and explicit guidance on X so that your long-term memory networks around X consolidate. If you are left to try to discover what is going on at this point, then all you have to rely on is working memory (which fades all the time) and arbitrary resonances in long-term memory (that might not be that useful.) This is the fundamental insight of cognitive load theorists such as Sweller, Kirschner and Clark. A novice can work on problem solving for a very long time and come away learning hardly anything about the problem. A worked example, on the other hand, in which every step is explicitly shown and the answer clearly given, results in novices grasping the process because they only have to comprehend the process, not discover it. As working memory focuses on these essential relations rather than thrashing around, the relations are laid down in long-term memory, enabling students to think progressively (sic) on their feet when it comes to solving a problem as they have the resources in long-term memory to help. It then becomes crucial to move away from explicit worked examples to more open problem solving and discovery type questions, as you want the learners to work with their own long-term memories, not follow step-for-step what is on the page. It is the dense collection of integrated information and strategies stored in long-term memory that enable effective problem solving, not some generic problem solving course that is supposed to work across all sorts of different domains.

Studies have worked out how many patterned chunks grandmasters carry in their long-term memory, and it is around fifty thousand (Chase and Simon,
1973). That is astonishing. But you are carrying thousands of chunks as well; only they are not chess patterns but chunks of letters called words. There is much debate about how many words a well-educated person actually knows (D’Anna, Zechmeister and Hall, 1991, pp. 109–122). But the argument is that what we consider exceptional in a grandmaster is similar to what we all do with language. Shakespeare, by the way, knew 66,534 words, if statisticians like Efron and Thisted (1976) are to be believed. The point is that schooling is about developing such extended sets of chunks in subject specific semantic knowledge networks.

Two key features stand out from the above discussion. First, to shift something from short-term to long-term memory it is important to place the item into an existing network and this means transforming the item so that it fits within it. The denser and more developed the network, the easier it is to locate the item in some part of it. Second, to increase working memory capacity it is almost impossible to increase the number of its slots, so the best route forward is to increase the size of the chunks. Notice that these two features are intimately related. Increased networks allow for improved transfer of an element into long-term memory and increased networks result in chunking, where what is worked with is not a single element but patterns of elements in a network. It is a virtuous circle, especially when these long-term memories are accessible to working memory.

Let’s now go back to the previous discussion around episodic and semantic knowledge because the distinction between them is the difference between two kinds of ways networks can function. Episodic knowledge comes easily from the personal process of experiencing life. Semantic knowledge is built up, step by painful step, chunk by chunk, from elements that have a specific set of meanings and relationships that are not based on the stream of life, but on logical and defined associations. The larger and denser this defined network of relations, the more readily a semantic element finds a place in the network as a transformed part or moment of the network. This is because working memory looks for similar patterns in long-term memory to the element on which it is focusing. It’s the foundational act of intelligence, to predict what is going to happen based on what has gone before. We all do this with episodic knowledge. Based on our experiences of the world, we predict what will happen based on past similar experiences. That is why we respect and listen to people with life experience as they are able to negotiate current occurrences intelligently based on a wealth of past associations.

The issue for education is how we do this for semantic forms of knowledge as well as for episodic knowledge. The more these semantic elements can be built up into meaningful chunks, and the chunks are organised into dense patterns or schema, the more working memory is able to cope with larger loads in quicker ways. Crucially, working memory is also able to find matches in long-term memory for what it is temporarily holding. The pedagogic struggle is to shift the matches from purely episodic similarities, to matches that have both semantically dense and episodically rich resonances. A student needs to build up systematic
networks of knowledge that allow for informed placement and chunking. This takes work. Unlike episodic knowledge that comes from experiencing life, semantic knowledge has to be built up through an artificial process that has non-local logics that shift away from processes of everyday existence. But once semantic knowledge networks build in long-term memory, an expansion of working memory capacity results, not only in terms of its ability to carry larger chunks, but also, with repetition, of its ability to work with increasing speed because it has access and assistance from long-term memory. And that means that more stuff gets into long-term memory more quickly; resulting in a further expansion of both working and long-term memory, and so on, in a virtuous cycle.

What this makes clear is that a non-negotiable goal of school learning should be an expansion of long-term memory, especially of the semantic knowledge type. Each time a chunk is retrieved from long-term memory, the knot of neurons at the base of the memory fire. If you are working with an extended complex semantic network in long-term memory, then you are activating massive parts of the network each time you retrieve an element of it. Crucially, each time the knot of neurons fire, the easier it becomes to fire the next time round. Simply put, the more you repeat something and the more you retrieve it from long-term memory, the quicker and easier the access to that memory next time round, and the less strain you put on working memory. And the more these memories are chunked, then the larger the meaning base that is activated.

**Automatisation**

Continuous practice eventually results in automatisation and although it is unclear how the neuronal base of this works, it is clear that automaticity frees up working memory for other activities. When an action or practice becomes automatic, what has happened is that all the possible wrong ways of doing something have been ironed out, and what is left is the successful groove with a limited set of actions that are well oiled. This makes the activity easier and faster. This easy speeding up is accentuated by the fact that neuronal firing speeds increase with continued repetition, resulting in very quick reaction time. When this is combined with the ability of your mind to focus on other issues that demand attention, you can attend to a complex problem space with more intelligence and openness. Earlier, I used the example of student teachers struggling to attend to everything happening in their classes because they were focused either on getting the information of the lesson correct, or control of the class. Either classroom control or the content of the lesson suffer, often both. An expert teacher, on the other hand, is able to teach the lesson and attend to all the emotional and managerial issues at the same time because much of what she is doing has become automatic, freeing working memory to attend to the many complexities of the classroom dynamic. As an aside for teacher educators, this means we should have different definitions of
excellence for a novice and an expert. Excellence for a student teacher should not be defined in terms of what an expert teacher does, but on what a novice teacher can do well, given the fact that she is a novice.

From neuroscience to education

So if we start with a single neuron firing, and expand outwards towards what in this neuronal story resonates with the educational process, then we find it lies in the relationship of working memory to long-term memory, with a specific focus on the intersection with increasing growth of semantic knowledge and its dynamic interaction with episodic knowledge. Notice that we have moved away from the intricacies of the molecular functioning of the brain to more general functions of the mind (like working and long-term memory) and its intersection with knowledge types that are very hard to see in any meaningful way in the chemical reactions of the brain. The more we dig into the micro-architecture of the brain, the less relevant the application to education.

This is a key lesson for the educational imagination: how to negotiate the boundary where education ends and other worlds take over. It’s an exceptionally hard judgement to make and the crossover line is blurred, with networks running outwards and inwards and across the line. A rule of thumb is always to hold in mind a pedagogic act and ask if what is being stretched towards still has educational relevance. Allow me to try to demonstrate this with the borderland between neuroscience and cognitive science.

Neuroscience is about neurons and how they function. Cognitive science is a more interdisciplinary focus on how minds work. We have to be really careful about how we apply neuroscience to education (Willingham, 2009). Education has a set of foci and goals that are very different from neuroscience. We focus on learning, a complex space involving a learner, teacher, knowledge forms and pedagogic processes with goals of character development, social and cultural activity, economic improvement and political engagement that spin outwards and upwards, as this book shows. Neuroscience focuses on neurons in the brain in a descriptive and analytic way. We can use this as educators, but only with care. Cognitive science has clearer applications to education because its focus is on how the mind/brain works, although we would immediately add that the learner is more than a mind and the learning process is more than about cognition. Closer to education would be cognitive psychology with its focus on perception, memory, language and mental representation; and even closer still would be developmental psychology with its focus on how humans develop over their life span. Closest to education would be educational psychology, until we get to pure education, whatever that is. Is there a way we can get these different levels clear to assist us in negotiating a space that goes from the smallest element of neuronal activity upwards to brain to mind to person to learning how to live productively in our modern world?
The problem is that we are not simply working with levels of scale or complexity, but with jumps across ontological and functional categories, like brain and mind, individual and society; jumps in focus from neuronal interactions to learning processes; and jumps across different disciplinary fields with different types of research. If we simply take a scale approach that moves from smallest to largest, then we go from a synapse to an individual neuron, to a neuronal network, to a cortical sub-region, to the brain, to the central nervous system, to the body.

This zooming in and out of the body does important things. It gets us out of a tendency to think that we are really just our brains and the rest of our bodies are mere containers for its precious processes. Where it does not help is in the jump from the functioning of the material brain to the way our minds work, as this is a jump not in scale but in ontological type. Nor does it help us jump from a singular focus on the brain to the applied field of teaching and learning. The danger of developing an educational imagination stretching all over the show is that one refuses to discriminate between the radically differing realities lying very close to each other in the actual processes of teaching and learning. An educational imagination has to learn how to recognise and negotiate boundaries, not just indiscriminately jump around between them. Obviously all these levels and layers intersect and play out together, so it is often hard and hazardous to separate them out, but with practice and informed study, both the separation out of levels and their intersection become easier to see.

Neural and behavioural levels of analysis

One way to help imagine the connections between levels at this specific point is to separate a neural level of analysis from a behavioural level of analysis and show how they relate. Willingham and Lloyd (2007, p. 141) provide us with a useful diagram (figure 3.3) that does just this.

This chapter’s point of focus is at the level of a cognitive construct, specifically that of working memory and long-term memory, for it is here, we have argued, that the smallest useful unit of analysis for educational purposes can be found. Note that a cognitive construct is not a neural construct. It uses much of the information about neural levels of analysis work, but it has shifted ontological levels to focus on mind rather than electrified meat. When we have gone into lower levels of analysis (like how neurons and neural networks function), this has been simply to give a grounding picture of working and long-term memory. Our major point of interest has been how working and long-term memory intersect with semantic knowledge and how this plays out in educational terms. This is a behavioural level of analysis located between cognitive and educational constructs. We have been careful not to focus on the millions of neuronal
synapses between neurons or on our internal representations of the physical, social and personal world going on in our minds. This is too fine-grained a level of analysis for the educational imagination, not because it is invalid in its own terms, but because it sits outside the education boundary. One of the failures of the educational imagination is to pretend that it is allowed to go wherever it wants, because it is the imagination. All this results in is dilettantism and misapplication. There is a rigour to the educational imagination that comes from a defining of boundaries that gives it the limitations from which creativity can deepen and play. So at this point we are focusing on cognitive constructs in the mind, as these are the building block units of educational constructs that we will get to later in this chapter (such as cognitive load theory, reading theory or multiple intelligence pedagogy.)

<table>
<thead>
<tr>
<th>NEURAL LEVEL OF ANALYSIS</th>
<th>BEHAVIORAL LEVEL OF ANALYSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central nervous system</td>
<td>School</td>
</tr>
<tr>
<td>Gross anatomic structure</td>
<td>Classroom</td>
</tr>
<tr>
<td>Nucleus, cortical subregion</td>
<td>Individual mind</td>
</tr>
<tr>
<td>Neural network</td>
<td>Educational construct</td>
</tr>
<tr>
<td>Individual neuron</td>
<td>Cognitive construct</td>
</tr>
<tr>
<td>Synapse</td>
<td>Internal representation or process</td>
</tr>
</tbody>
</table>

Figure 3.3 Neural and behavioural levels of analysis

From molecular neuroscience to educational psychology

Kalra and O’Keefe (2011) attempt to illuminate these levels in a different way by contrasting how textbooks in educational psychology, developmental psychology, cognitive psychology, cognitive neuroscience and molecular neuroscience work differently, and it is fascinating to see how distinctive they are. Let’s take the extremes of educational psychology (closest to education) and molecular neuroscience (furthest from education) and compare their overall perspective, research methods and specific focus.

The overall perspective of educational psychology is on the application of psychology to education for teachers; its research methods involve qualitative and quantitative methods about how children think and learn; and its focus is on how effective this research is to inform pedagogic techniques and curriculum construction.

The overall perspective of molecular neuroscience is intensely focused on how synapses function. Its research methods involve single cell recording, genetic manipulation and experiments with the impact on the synapse of assorted drugs;
and its focus is to work out the distinctive properties of neurons in relation to other cell types.

There is no doubt that molecular neuroscience has implications for education, but a massive amount of reworking is needed before any sense can be made of the highly technical research. As educators, we have to rely on the recontextualisation of molecular neuroscience to education that takes into account the massive difference in overall perspective, research methods and focus.

There is a danger in assuming that because molecular neuroscience has the smallest and most tangible focus area that it is possible to make educational recommendations based on its findings. There are some areas where this is the case, like dyslexia, where neurological disorders affect learning (Willingham and Lloyd, 2007), but for the most part molecular neuroscience is too fine-grained a level of analysis. This becomes obvious when we compare the different disciplinary research methods.

It should be clear that single cell recording of non-human animal studies does not have much to say about the pedagogy best suited to children studying science at grade 6 level in a poor context. It’s as crazy as nuclear physicists waxing lyrical about the basis of our ethical and moral positions drawing on the functioning of the sub-atomic. That said, this should not stop the project to combine mind, brain and education into an interdisciplinary region that asks how research at the level of the mind and brain can inform educational policy and practice. It is just that the enormous difficulty of working between these levels has to be clearly seen and negotiated. We have started the process by using working memory and long-term memory as a focus point, as it gives us core cognitive constructs that exist at the intersection of cognitive neuroscience and cognitive psychology with powerful implications for teaching and learning, as seen in the Unified Learning Model of Shell et al. (2010) and the work of cognitive load theorists such as Sweller and van Merriënboer. It is in their work that you can find clear and precise accounts of how cognitive science intersects with pedagogy and instructional design.

Cognitive load theory

The stepping-off point for the intersection of pedagogy and our cognitive architecture is the simple point that our working memory is limited. If we are just remembering stuff, then we can hold around seven bits in our minds. But as soon as we start to process information through making sense of it by contrasting it, placing it and comparing it to what we already know, then we can only hold around two or three bits at the same time. Crucially, this information only stays in our working memories for around twenty seconds. Pedagogic practices have to take the limitations of working memory very seriously, for this provides the gateway to deep learning: every piece of knowledge has to pass through its gates, gates that have a limited carrying and holding capacity.
A curious feature of working memory is that it has two different channels that receive information, visual and auditory, and when both channels are used together, the capacity of working memory increases (Paivio, 1986, p. 86). If you listen to your teacher while at the same time seeing a demonstration on the board, you will understand more than when focusing on either on its own. I can testify to this effect from my own mathematical studies in the Khan Academy. Salman Khan has made around 4,000 videos that he has downloaded onto the Internet for students to learn maths from addition and subtraction all the way through to calculus. If you watch a video you will not see him talking, but will see the maths problems being written out. You listen to him talking while seeing the problem unfold.

This simple method, used by teachers across the world, makes maths easier to understand than simply working from the textbook or listening to an explanation. Videos that actually film a lecturer talking waste your working memory capacity. You don’t need to see someone talking; you need to hear him or her (hir). But if at the same time as hearing them you watch examples, diagrams, figures and key phrases, then your capacity to understand and remember increases. This dual coding of information is also what makes power point presentations that simply put up what is being said boring. Your visual channel works at a different speed to your audio channel: as a result you read the slide in a couple of seconds and then have to sit through the laborious verbal account of exactly the same thing. Don’t put up your words on power point; put up supporting images.

Take a look at figure 3.4, based on Cooper (1998), to catch what my words are describing.

Even if you are using both audio and visual channels to increase the capacity of working memory, it still has a very limited carrying capacity. Earlier in this chapter we saw how chunking provided the key technique to overcome this limitation. It might be that you can only carry around four elements at a time in your working memory, but the size of these chunks can vary from a single bit of information to a massive networked process. Suddenly, what seemed to be a terribly limited capacity can be expanded almost to infinity, not by increasing the number of slots in working memory, but by increasing the size and complexity of the chunk. It is the networked schemas you carry in your long-term memory that give your working memory real power. If you do not have these complex schemas in your long-term memory, then you are doomed to work with around four tiny elements at a time, rather than with four huge processes. You can see this in figure 3.4 with long-term memory. Either you work with the tiny elements at the bottom or with the massive chunks at the top that include the elements underneath them within their processes.

The pedagogic question then arises: how do we increase schema construction in long-term memory? It becomes pedagogically imperative to work on getting information into long-term memory in an organised, schematic, way.
But even if you have these complex schemas in your head, working memory can still be limited by having to pay attention to the way they work. With increased practice, many of the schemas start to function automatically. You don’t have to think about it, it just comes spontaneously in habitual form. The more automated the process, the more working memory you will have to attend to the problem at hand. It’s not only pedagogically imperative to build complex schemas, but also to automate them.

The place you build schemas is in working memory, so pedagogy must attend to the cognitive load expected of the student and ensure that working memory is not overwhelmed or starved. This is a tricky process because the more experienced the student is in the topic under exploration, the more developed his schemas will be, meaning that more can be done, more quickly. If the student is a novice then care must be taken with the cognitive load as the carrying capacity of working memory will be limited. So, what pedagogic techniques can be used to reduce cognitive load by focusing on what is intrinsically necessary and eliminating elements that are extraneous? What can be cut away to make the load lighter, but
still keep what is crucial for understanding, meaning making and the construction of schemas? Sweller, van Merriënboer and Paas (1998) identify six pedagogic techniques that reduce extraneous cognitive load for novices:

1. **Goal free effect**: Don’t give the long-term goal of the problem under exploration at the same time as the problem itself. Rather focus on the problem and allow the goal to emerge once the problem has been understood (figure 3.5). If you reveal the goal before tackling the problem, then the students have to try to work out the relationship between the problem and the goal before they have understood the problem.

   ![Figure 3.5 Goal free effect](image)

2. **Worked example effect**: Don’t just give novice students a problem to solve. Rather start with a worked example (figure 3.6) that shows the steps of the problem and how to solve it in its simplest form.

   ![Figure 3.6 Worked example effect](image)

3. **Completion problem effect**: Once the students have been given a worked example provide a problem that has some of the steps already learned and get them to complete it (figure 3.7). By reducing the size of the problem space to one or two steps that need completion you reduce extraneous load.

   ![Figure 3.7 Completion problem effect](image)
4. **Split attention effect**: Be careful of multiple sources of information that expect students to integrate the different bits (figure 3.8). Rather provide one integrated source of information on which they can focus.

5. **Modality effect**: If you are going to use different sources, then make sure that you combine auditory with visual channels. For example, replace a visual combination of a picture and a written explanation with a picture and an auditory explanation.

6. **Redundancy effect**: Be careful of multiple sources of information that all do something similar (figure 3.9). Rather have one source that does it all properly.
The key reason why you want to free up some space for working memory is that it needs space to make sense of the problem at hand, struggle with it and make meaning of it, so that the element of knowledge can be placed within an ever-growing schematic set developing in long-term memory. This process of meaning making also increases cognitive load, but it is a germane cognitive load that results in the crucial shifting of information from the limited world of working memory into knowledge networked within the infinite world of long-term memory. It is this germane cognitive load that must be increased by cutting away extraneous elements.

In effect, the total memory capacity we have inside ourselves when engaged in a task is taken up by three different kinds of load (figure 3.10): the intrinsic load of the task itself; the extraneous load of instructional choices around how to learn the task; and the germane load that comes with thinking about the task and making meaning from it (see http://kellymorganscience.com/why-designing-experiments-is-so-hard-for-students-what-we-can-do-to-help/).

For a person skilling herself in pedagogic imagination, this level of focus on working memory provides precise pedagogic variables that assist in both the descriptive analysis of learning and proscriptive advice about how to improve learning.

1. **Intrinsic load**: some tasks are more complex than others due to their inner make up and what they already expect the learner to be able to understand and do. Doing a calculus problem expects you already to have mastered algebra. If you do not have a good working knowledge of algebra, you will really struggle with a calculus problem even if it is carefully worked out for you. So one way to make the intrinsic load more manageable is to ensure that the basics needed to do the task are already well understood. A student with a working understanding of calculus will find a calculus problem easier than a novice calculus student.
2. **Extraneous load**: involves instructional strategies to teach and learn a task. You cannot avoid extraneous load, but you can reduce it by becoming skilled in those pedagogic choices that work to focus attention on the problem at hand, rather than distract attention with showy side effects. The six effects listed above provide a good beginner’s guide. I say beginner because the six effects work best for novice learners of a task, not experts. Experts enjoy cracking a difficult, obscure, multiple task, but don’t assume that because experts like it a novice will like it as well. In all likelihood, the novice will collapse with cognitive overload.

3. **Germane load**: the work done by thinking, reflecting and making meaning of a task. It is vital that space is left in working memory for this process otherwise the academic work done does not make it into the bigger schematic structures of long-term memory. Even worse, without meaning making, schematic structures in long-term memory stay thin and weak.

It is at this level of focus that the educational imagination can feel the powerful contribution cognitive science makes to pedagogy. We have not gone into ever-finer details about the chemical composition of the synapse, or the micro-architecture of the brain, important as these are. That is left for specialists in the neuro and cognitive sciences. What we have found in the work of cognitive load theory is a micro-level of focus that has direct pedagogic purchase. If we keep this combined focus on our cognitive architecture and its educational implications, but expand one level upwards towards a more general picture of the developing mind, what do we find?

### A general model of the developing mind

A general model of the architecture of the developing mind would not be a picture of the brain, or a neural map of brain functioning, as it would need to work with a variety of cognitive constructs in the mind and hold them together in a developmental system that speaks to education. Something like figure 3.11 illustrates how to do a model of the mind rather than a picture of the brain (Demetriou, Spanoudis and Mouyi, 2010).

What makes the model so elegant is the way it represents four different features of our minds in one diagram. The first is stages of development, represented as four different stages going upwards from sensorimotor to abstract (four stages of general development levels). The second is the core capacities that run as the spine through the model. We have partly dealt with this core in terms of working memory and shown how to increase the speed, span and control of core capacities through automaticity, chunking and motivated attention. The third feature is the specialised capacity spheres that point to the working of our minds with different aspects of existence in specially adapted ways.
Finally, we have the hyper-cognitive system that self-reflexively builds up a systematic set of schemas and models. It’s a key process as it catches the way students take control of their own learning by observing what happens to them when they learn and then learning how they learn. It’s partly what is developed when working with the germane load of a task. It gets students to take control of their own processes, make them meaningful and place them within bigger schemas. So we have a basic core; different domains of thought that arise from the various types of information, relations and problems we encounter in the world (specialised capacity spheres); different capacity spheres going through general development levels; and finally that part of our minds that self-reflexively makes sense of it all.

Figure 3.11 General model of the architecture of the developing mind integrating concepts from Demetriou and Case
The diagram is a prime example of a sophisticated educational imagination at work. It has shifted from a literal image of the brain to a combination of abstract elements held together in one image. But it does more than this as I hope to show in the following pages: it breaks specific limiting ways in which we have historically thought about human development and education. The first is to think that we go through general stages of development without taking into account that different specialised capacities develop at different speeds in different ways. The key breakaway point is to realise that we have different domains of thought that do different things, rather than just one overall path of development.

**Domains of thought**

A key operating principle of the education imagination is that it needs to move up and down all the levels that are educationally relevant. Once you get used to this climbing idea it becomes really important to grasp the different domains you need to master when climbing. For example, when climbing through levels of development in thinking, you have to recognise that there are different types of thinking, each of which needs its own climbing requirements. It’s not only that the educational imagination must travel through levels; it has to understand the different domains it travels through as well.

There is some dispute about how many different domains of thought we have and how they relate to each other. But rather than get into a debate around how many, I would prefer to give a brief account of six of them and how they develop from elementary to more complex structures so that you can see how different domains work with similar levels. Each domain starts with an elementary way of dealing with specific aspects of the world and then develops over time into more complex ways.

**Categorical reasoning**

We all simplify the complexity of the world by predicting what will happen based on what we have already experienced. We expect similar events to produce similar outcomes and do this almost from birth. This domain gives us ‘the seeds of inductive inference’ (Demetriou, Spanoudis and Mouyi, 2010, p. 15) that build into understandings of similarities and differences between objects. These different objects are classified and ordered, giving us means to organise the complexity of the world in predictable ways.

**Quantitative reasoning**

The world is continually in a state of change. Things increase and decrease, break into parts or come together. From a very young age we have the ability to work with increase and decrease. As we get older this ability improves and we develop ways of counting, sharing, splitting up or keeping things for ourselves (Demetriou, Spanoudis and Mouyi, 2010, p. 15).
Spatial reasoning
The world has objects of different sizes and shapes that are near or far away from us. To negotiate this aspect of the world we develop ways to track their movement and mentally rotate what they look like (Demetriou, Spanoudis and Mouyi, 2010, p. 15). This helps us recognise objects from different angles and work between them.

Causal reasoning
When we push something it moves. Our basic interactions with the world show that we have an effect on the world: our cry produces a mother; a thrown object moves away from us. These basic experiences develop into a causal understanding of the world where we begin to experiment with what things work and what things don’t, eventually developing systematic ways to isolate cause and effect relations.

Social reasoning
From a very young age we recognise human beings as different from other species and respond to smiles with a smile and a growl with a cry. This develops into an ability to work out what is going on inside other human beings based on their actions and appearance and to respond accordingly, eventually resulting in decentring, where we can take on the role of the other.

Verbal reasoning
Verbal reasoning to some extent works across the above domains as it enables social interaction, guides actions and combines insights from different domains into coherent inferences. Basic reasoning such as ‘if this then that’ and ‘either this or that’ develop an ability to work on relationships between things or processes and this develops into truth and validity relations where the honesty and reliability of the statements are evaluated.

Combining domains of thought with levels of development

Although these domains are separated off for ease of presentation, this in no way means that they do not intersect and overlap. The danger here is that once you grasp that there are different domains, you then imagine that they have to be separate. The best way to show how to get beyond a silo way of thinking is graphically. In figure 3.12 you can see that each domain is distinct, but there are places of connection (Rose and Fischer, 2009, p. 414).

Nor is it necessary that all the domains develop at equal times in equivalent ways. There are different developmental trajectories (figure 3.13) that can result in different levels of development at different times (Rose and Fischer, 2009, p. 416). Some domains do not necessarily become as highly developed as others.
Note that I am demonstrating this spreading and inter-related pattern of hierarchical development because it assists the educational imagination to discriminate between different domains or streams flowing through the same set of levels. This book is all about climbing through different kinds of levels. But even when one stays within one terrain – levels of intellectual development – it is possible to find that there are different streams running through the levels. This is a key discriminatory ability the educational imagination must master. And it must be able to imagine the different streams working at different speeds and in different ways, as can be seen in figure 3.13. To represent the same point using Demetriou’s model we can show one domain reaching the fourth level and another only reaching the third (figure 3.14).
We can see in figures 3.13 and 3.14 a personal developmental trajectory that has different domains at different levels of development. This can be contrasted to another personal development trajectory (figure 3.15) that has managed to develop the different domains up to the same level (Rose and Fischer, 2009, p. 415).
Figure 3.15 is crucial because it gets you to think back to a synthetic possibility of convergence once you have grasped the possibility of divergence. Just because the different domains diverge does not mean that one cannot get them to converge. Is it not worthy to aim at a balanced and harmonious set of domains as an educational ideal, rather than overdevelopment of some and underdevelopment of others?

Getting beyond simple and inaccurate models of the mind

What I like about the above five figures is how they challenge three overly simplified and restricting models of the way our minds work. Figures 3.11–3.15 enable the educational imagination to work in supple and lithe ways across levels of development and domains of thought, and shine a clarifying light on overly simple images of levels of development, IQ and multiple intelligences.

Ladder of development

The first out-dated model (figure 3.16) is a simple ladder of development view where we go through set stages at set times.

![Figure 3.16 Basic stages of development based on Piaget](image)

There is nothing wrong with initially learning about stages of development in this way. As climbers through an educational landscape we need to start with baby steps and the simplified view of Piaget’s stages of development does teach us the basic rules of climbing.

But Piaget was far more than this. He was a master of the climbing device and we have used his reaching for the heights to climb even further. We don’t simply have stages of development; we have different domains of thought developing at different rates to different levels. We also have the dawning recognition that if the domains have places of connection that work together, then it is risky to push for a speeding up of one domain at the expense of the others, for although separate they also inter-relate.
Finding the g-spot: intelligence quotient

The second out-dated model (figure 3.17) is one that assumes that behind all these different domains of thought is one general form of intelligence that can predict how we perform across all domains. This has been a blight on the educational landscape for over a century.

This assumption is that no matter what the domain, a single factor determines performance across all domains; something called ‘g’ for general intelligence. If we did four very different types of tests (a,b,c,d) that worked with different domains we would tend to perform at the same level in each because general intelligence ability is at the heart of how we perform in all four tests.

Multiple intelligences

With the realisation that we have different domains of thought, it became clear that it was possible to perform differently in different tests depending on the domain and the test. But this recognition can result in a third oversimplified model suggesting that we have lots of different types of intelligences and nothing holding them together. In figure 3.18, there are six different types of intelligence each measured by two different tests. For example, verbal intelligence (1) could have a spelling test (a) and a grammar test (b): what determines how well you do in the tests has everything to do with your verbal intelligence and not much to do with any other type of intelligence.

This is close to Gardner’s multiple intelligence theory (MIT) view that has become so popular in schools across the world. At the intuitive heart of MIT lies a simple recognition that individuals can be strong or weak in certain skills, but not necessarily in others. Gardner formulated this position as a critique of the general intellectual potential view. Surely some individuals are better at some things than others and this cannot be condensed into one general number? Gardner

---

**Figure 3.17 Single factor ‘g’ determines performance across domains ‘a’, ‘b’, ‘c’ and ‘d’**
developed seven types of intelligence, recently updated to nine. It breaks open the educational imagination to the possibility of teaching in different ways for different intelligences. The problem with this position is that it does not accurately represent how the core capacities central to the functioning of our minds have an impact on the different types of intelligence. To put it crudely, we have both a general intelligence function and multiple intelligences. If there are four tests (a, b, c and d as in figure 3.19) that work with two separate domains of thought (1 and 2), there are strong correlations both to specific domains of thought and to general intelligence (g). Willingham (2004) provides an excellent review of MIT and I have used his critique to format this discussion and the figures 3.17–3.19. I strongly recommend his blog found at www.danielwillingham.com.

General mental ability tends to predict how strong or weak performance will be in different domains of thought; and the domains of thought tend to predict performance in the two related tests in contrast to the other tests. Don’t think that because you have broken away from a single restricting view of intelligence into multiple intelligences that you have successfully extended the educational imagination. It’s only half the story. Take a look at how Demetriou depicts (figure
3.20) the hierarchical relationship between the six domains of thought and a more general mental ability (Demetriou, Spanoudis and Mouyi, 2010, p. 14). For our purposes the factor analysis is not important. Rather, it is the way the above analysis catches how the six domains stand as both separate and related to each other through a more general mental ability. It is as much a part of the educational imagination to work with convergence as well as divergence.

![Hierarchical relationship between six domains of thought and more general mental ability](figure3.20.png)

What gives Demetriou’s model of the mind its synthetic elegance is how it shows all three out-dated models (basic stages, general intelligence and multiple intelligences) to have insightful elements, but to be severely limited on their own. The core capacities of speed, span and control refer to a general intelligence capacity, but show how this core relates to more specific domains of thought that have their own functioning principles. The domains of thought catch Gardner’s insight into multiple intelligences, but show how these also work with a central
core; and the general development levels highlight how stages of development work, but relate this to different domains that can develop at different times and speeds. All of these factors have to take serious account of the hyper-cognitive system that gives us independence and control over our own intellectual functioning. Demetriou’s model is a triumph of the educational imagination in terms of levels, domains, core functioning and the sweet wrapping of self-reflexivity.

To end off this chapter I would like to stay with the contrast between Gardner’s multiple intelligences and Demetriou’s model because it illustrates the seductiveness of thinking that diversity is the acme of imagination. The seduction runs something like this. The more intelligences we have the more creative possibilities for teaching and learning open out, demanding a pedagogy where we teach different children differently because of their unique combinations of multiple intelligences. Is that not the height of the educational imagination?

No, it is not, although when Howard Gardner’s Frames of Mind: The Theory of Multiple Intelligences was published over thirty years ago in 1983 it did seem so. How is it, you might ask, that a theory pointing to a wide range of intelligences rather than one general Intelligence Quotient (IQ) could be bad for the educational imagination? Surely knowing that we are made up of different sorts of intelligences must be a good thing, educationally speaking? If each of us has a different intelligence profile just like our own unique fingerprint, and if understanding what this unique combination is gives a better ability to work out how to teach and learn, then we had better implement multiple intelligence pedagogies in all our classrooms as soon as possible. It should result in a massive improvement in learning as each student learns in a way best suited to her profile. Many teachers have had the personal experience of struggling to get a concept across in a particular fashion, and then, after shifting to a completely different way of doing it, finding their learners grasping it. There is no doubt that different ways of teaching a concept can have a radical impact on understanding and that some learners tend to grasp a concept better when using a particular type of representation. Teachers and researchers have known this long before multiple intelligences came along. The reason why learners start to understand the concept is not because it resonates with their particular type of intelligence profile, but because the concept has been better represented or because the multiple angle of view has increased understanding. The risk with teachers who use MIT to structure their lessons is that they might confuse the best ways to represent a topic with different intelligences of students (Willingham, 2004). And there lies the danger, for by trying to represent the concept in ways that best fit the intelligence profile of the student, the teacher can forget to represent the concept in ways that best fit the concept.

I was a history teacher in a past life and when covering the Anglo-Boer War, I used photographs, maps and diaries, not in an attempt to work with the different
learning styles of my students, but to best represent the Anglo-Boer War. When doing Hitler I played some of his speeches, not to resonate with the auditory learning style of some of my learners, but because this presentation method best caught the hypnotic power of Hitler. I worked with different modalities in presenting my lessons, not because I was taking into account the different multiple intelligence profiles of my students, but because these were good ways to present the topic. There is far too little time in a classroom to waste it trying out different modalities of learning that are not really suited to the topic at hand, but resonate with the intelligence profile of the student. To get learners to construct letters out of twigs because they are kinaesthetically oriented is to waste time because we have excellent research on the best ways to teach letters to children. That is what is important; not that they spend half a day arranging twigs into patterns because it provides a feel for letters. Research on learning the alphabet does point to the importance of using different modalities, but this is based on systematic ways of getting all the children in the class to experience and make meaning of the letters, not on trying to accommodate different intelligence profiles. If you want to improve learning, then focus on what needs to be learned and how this can be done in deep and meaningful ways that bring out the essence of the concept.

What makes the adoption of MIT in classrooms even more worrying is the lack of empirical proof for separate multiple intelligences. We saw earlier in this chapter that recent work in neuroscience and cognitive development points to a general processing efficiency of the brain that carries through to different domains of intelligence (Demetriou, Spanoudis and Mouyi, 2010). If you have a really good working memory and can process information quickly and reflectively, this is going to carry through to the particular domain with which you are working, whether this be verbal, spatial, mathematical or social. There are different domains of intelligence, only they are not as separate as Gardner would have us believe. There is no solid empirical evidence for multiple intelligences in the form that Gardner has described.

It’s not only empirical support for MIT that is lacking, but empirical support for its effectiveness in classrooms. It doesn’t work that well as pedagogy. It complicates the lesson, over-individualises it, and makes feedback really difficult because so many different things are happening at the same time. It obscures the content of the lesson by making the manner of presentation dominate the concept. Multiple intelligence pedagogies do work sometimes, but this is mostly due to gifted teachers working in optimal conditions; the point here being that these kinds of teachers can make anything work.

Why, if MIT is both wrong as a theory of intelligence and ineffective as a pedagogic strategy, does it still carry so much power and conviction for teachers and parents? One reason could be that it gives us hope when struggling with calculus, grammar, Shakespeare, a science equation or anything really difficult. Maybe we can’t get it because it’s just not suited to our particular intelligence
profile, that we have other intelligences. Everyone everywhere is intelligent in something. We just have to find what it is. We do have to find out what it is we do best, but this is a complex mix of intelligence, personality, environment, motivation and happenstance, not a simple finding of which intelligence profile we have from an ever-expanding list. The hard work is in squaring up to the task at hand with everything that you have and then persisting with it, not flipping around until you find what your own particular intelligence profile happens to be, after which everything will flow smoothly.

MIT gives us hope in the wrong place by attempting to improve student learning through trying to find just the right combination of factors for each student. It places the blame for poor performance on the lesson design being poorly adapted to the intelligence profile of the student; for if the design had been better adapted, then the student would have learnt, because the lesson would have inspirationally spoken to her own particular profile. Never mind that this is based on an over-simplified story of multiplicities of intelligences that results in over-complicated lesson design, over-stressed and fragmented teachers trying to do too many things, and learners who can blame their failure to learn on the inappropriateness of the lesson to their unique fingerprint. Never mind that MIT pedagogy results in a catastrophic increase in extraneous load. As parents and teachers we need to be far more careful about the fads pushing for admission into the precious sanctum of our classrooms.

That said, teachers across the world have resonated with MIT because it echoes something important in their own experiences. Different students learn differently and working out what these differences are has implications for teaching and learning. This is about learning styles, not multiple intelligences. Learning styles are different from multiple intelligences in that the former are closer to the reality of classroom life and the tangible practices of learning – how we concentrate, store, remember and make sense of knowledge. It looks at the practices of learning as a complex whole and explores the alternative ways different learners take through the process. For highly resourced schools with excellent and committed teachers, responding to a demanding parent body that insists their child gets individual attention, an explicit engagement with different learning styles is both rewarding and politically astute. But even here, with learning styles one has to be careful. Is it not better to induct learners into the rigours of the subject in its own terms, rather than bending the subject to the learning idiosyncrasies of a child? Will the child not learn more in the long run from engaging with what the subject demands rather than from what she demands of the subject? Humility is one of the key virtues of the educational imagination and at the heart of humility lies openness to the demands of the world, not what you demand of the world.

Let’s stand back from this chapter and formulate where it has taken us. We started by trying to find the smallest functional element inside our material being that is strongly related to education. It is very tricky because unlike the
materiality of a desk, the smallest functioning educational element inside us is part of a self-conscious and social animal. We started with a distinction between everyday experience and school knowledge and found a similar distinction within cognitive science between episodic and semantic knowledge. We then honed in on working memory as the key structure at a micro-level, tracking how knowledge is developed by getting through the astonishingly narrow and short-lived focus area of working memory into long-term memory. We showed how this basic functioning of our cognitive architecture has profound implications for teaching and learning; implications that have been best drawn out by cognitive load theory. We then broadened our focus to a more general model of the cognitive architecture of our developing minds and found it was vital to distinguish different specialised capacity spheres that operate as different streams or domains flowing at different speeds through developmental stages. We used this model to interrogate over-simplified models of development and the pedagogic dangers such over-simplifications hold, leading us finally into a critique of multiple intelligences. Just as chapter two went into the architecture of the school with its classrooms, desks and chairs, chapter three went into the architecture of our minds, looking for possible levels of movement up, down and across, always bearing in mind the instruction to stop when educational logics lose their purchase (figure 3.21).

**Figure 3.21 Shift from student brain to levels of interior development**

If we review the book so far, we see that chapter 1 provided the widest angle possible by expanding outwards to all the schools of the world. Chapter 2 then narrowed the focus and honed in on the interior architecture, shifting from the
school to the classroom, desks, chairs and writing equipment. We then jumped in chapter 3 from the desk to an individual learner sitting at the desk, but kept the narrow focus on the cognitive architecture of the learner’s developing mind and brain. In chapter 4 we shift from more material accounts of the mind into its interior depths, with the question: how high can the mind reach as it learns? Just as we expanded schooling outwards from a specific location to districts, provinces, countries, continents and the world, so too we are going to expand the mind to ever higher levels, reaching ever greater levels of development. The mind enlarges within itself, encompassing more and more within its interior functioning. With schools we travelled across the world; with the mind we travel into the interior depths of expansiveness.