

# A scaffold, not a cage: progression and progression models in history

The need to understand ways of defining progression in history becomes ever more pressing in the face of a target-setting, assessment-driven regime which requires us to measure progress at every turn. We must defend our professional expertise in terms of measurable outcomes. Did we add value? Have our end of Key Stage levels improved? Have we met our targets at GCSE? Superficially, being able to measure 'progress' in these ways seems sensible. How else are we to monitor the performance of teachers, departments and schools against national benchmarks? But of course, as we have noted in previous editions, being able to *produce* simple, seemingly persuasive data is not, by itself, evidence that true progress in students' historical understanding has actually *occurred*. Reaching Level 5 does not automatically define a student's understanding of history as a discipline and it certainly does not capture the complexities that historical understanding entails. Project Chata (Concepts of History and Teaching Approaches) has, over a number of years, attempted to map changes in students' ideas about history between the ages of seven and fourteen. In this article, the first in a series commissioned by *Teaching History*, Peter Lee and Denis Shemilt discuss the fundamental flaws in the current system of National Curriculum assessment and argue that other, research-based models of progression can be more helpful in our day-to-day planning and teaching. They do not claim any model to be perfect, nor indeed that what is most valuable in history – 'the wisdom, perspective and understanding' – can be captured in this way. They do, however, explore ways in which models based on empirical research – rather than on educational expediency – can be genuinely useful in understanding how to move students forward in their historical understanding by identifying and clarifying the misconceptions which hold them back.

## What is progression?

'Progress' has a wide range of meanings. No usage is compulsory, but however we decide to use words, we must ensure we know what job we want them to do. One important distinction to keep open is between 'progress' in general and 'progression' in particular.

Pupils can make progress in any area of history, whether it is in keeping better notes, writing better essays, or giving better presentations to the class. But what counts as 'better' is likely to differ considerably for different activities. Better notes will probably mean a higher degree of organization and quicker retrieval. Better essays will involve (among other things) clearer structure and greater sensitivity to the question. Better presentations will mean stronger engagement with audiences and more acute judgement of what must be explained.

The kind of achievement at stake impacts on how we measure progress. Progress in the amount of information students are able to remember is often measured quantitatively. Thus, in public examinations in history in Year 11, ticks used to be given for items recalled, subject only to fairly weak criteria of relevance: the more ticks the better. We may be unhappy with the simplicities involved in this, but we can hardly dispute that it involved a kind of progress.

The relatively recent appearance of another word – 'progression' – hints that we sometimes want to talk about something rather different from a general notion of progress. 'Progression' was juxtaposed with 'aggregation' to emphasize that progress in history could be more than an increase in the amount of information pupils could recall: learning history was not just learning 'one damn thing after another'. Research suggested that children's ideas about history and about the past changed as they grew older and that it was possible to view these changes in terms of *development*. (Note that even moving systematically from information on one period to information on another is not enough to count as progression in this sense, even though OFSTED usage sometimes appears to assume that it is.) The research itself grew out of teaching and was 'use-inspired'; it arose from dissatisfaction with the aggregationist assumptions that seemed to be implicit in examinations, widespread among curriculum managers in schools and enshrined in classroom practice. So 'progression' came to focus on the way in which pupils' ideas – about history and about the past – develop.

If a plausible case for this kind of progression is to be made, it has to be able to show some structure in the way children's ideas change. Early research took a Piagetian approach to children's thinking, placing emphasis on the formal characteristics of historical

Peter Lee is Senior Lecturer in Education at the University of London Institute of Education. He previously taught history in secondary modern, grammar and comprehensive schools. Denis Shemilt has worked at the University of Leeds for more than 25 years, many of them as head of the School of Education at Trinity and All Saints. He was extensively involved in the Schools History Project.

**Peter Lee and Denis Shemilt**

Figure 1: Second-order and substantive concepts

Substantive concepts are part of the substance or 'content' of history. 'Entrepreneurs', 'politicians' and 'plagues' all explicitly figure in historical accounts of the past. But the second-order concepts that give shape to the discipline of history, like 'evidence' or 'explanation', are not the focus of those accounts. They are higher order organizing concepts that guide historians' work on the substance of history, hence 'second-order'.

Confusion about this distinction can arise with concepts like 'change', where we want to talk about both substantive and second-order ideas. Students' pictures of the changes that have taken place in periods they have studied (like the move from rural to urban life, or the increasing use of energy) can be thought of as substantive, and sometimes the ambiguous phrase 'the nature of change' is used for such pictures. But we may ask questions about students' ideas of change at another level, and notice that some of them conceive it as event-like, whereas others understand that change is often impossible to locate at any particular moment. These are second-order ideas about 'the nature of change'.

reasoning.<sup>1</sup> It tended to concentrate on the substantive concepts that historians employ in dealing with the past, like 'peasant', 'bishop', 'parliament' and 'queen', as opposed to ideas that shape the discipline of history, like 'evidence', 'historical account', 'empathy' and 'change' (see Figure 1). It is too simple to say that research on substantive concepts failed to find patterns of change in students' ideas, but it ran into problems about whether the concepts were in any clear sense 'historical', why some should be taught rather than others, and how they related to one another.<sup>2</sup> Children seemed to approach them in so many different ways that patterns were hard to establish, and while a Piagetian framework appeared to offer the prospect of a very general pattern, it carried a theoretical burden that was arguably inappropriate for history.<sup>3</sup>

Hence there was already a disposition among teachers and researchers to look beyond substantive ideas when school history began to move towards teaching about the discipline of history as well as about passages of the past. Work on pupils' second-order ideas began to provide evidence that it was possible to treat history as progressive in a somewhat analogous way to physics: pupils did not simply add to their information about the past, but acquired understandings that changed in patterned ways as they learnt about history. The concept of progression began to mean something more specific than progress. It meant that history was not just about aggregation.

### Progression as acquiring more powerful ideas

It is sometimes said that history is a common-sense subject, but research in the UK and elsewhere suggests that this may be a mistake.<sup>4</sup> Students' ideas are learned in everyday life, and work well in that context, but may create difficulties in history. Think how children learn what it is to tell the truth. A window is broken or clothes are torn, so mum wants to know what happened. The question for the child (and mum too) is simply whether or not she tells it like it was. From the child's point of view the past is known: it is given and fixed. Because mother and child are working with shared assumptions about what matters in the past, the past can become a touchstone for telling the truth; once it has happened, it cannot be changed, and there can only be one true account of it.

But we have history precisely because the past is neither given nor fixed. Notoriously, people do not share assumptions about the significance of events, so the past cannot be treated as given. Moreover, it can never be fixed, because as new events occur they change the significance of what has already happened. What could truthfully be said about the Treaty of Versailles changed dramatically in 1933, and what may be said about 9/11 is changing almost day by day. Because history's key ideas can run against common sense assumptions, it makes sense

**Learning  
history is  
not just  
learning  
'one damn  
thing after  
another'.**

to speak of history as counter-intuitive. Indeed, some common-sense ideas make history impossible.

The common sense idea that we can only really know what we witness ourselves makes history a dubious proposition for many children. Once young children start to wonder how we know about the past (rather than taking it as given) they often assume that we can't know anything 'because no-one was there in them days'.<sup>5</sup> For pupils who believe this, history is a non-starter. Of course many students soon recognize that people who did witness events may have left reports of what they saw, and so history becomes possible once more. Unfortunately, testimony has its own problems: children know that we don't always tell the truth and adolescents are only too aware that we can slant our stories of what happened for ulterior motives. History is once again revealed as a highly suspect activity. But as they begin to grasp that we can ask questions of sources that they were not intended to answer, pupils come to see how 'bias' is not the disaster they thought it was and that historians can operate successfully without being dependent on reports. The later ideas in this series are more powerful than the earlier ones in that whereas the earlier ones bring history to a grinding halt, the later ideas allow it to go on. In this way progression is the acquisition of more powerful ideas.

### **What are progression models?**

If key historical concepts are counter-intuitive, it is clearly important to understand students' preconceptions.<sup>6</sup> In mapping the 'ideas' students are likely to hold about

history as a discipline, a progression model is, of course, uncovering students' prior conceptions. Understanding such prior conceptions is essential if our teaching is to correct misconceptions or to build on students' ideas. Ignorance of preconceptions risks the assimilation of what we fondly think we are teaching to sets of ideas the children already have.

Models of progression are much misunderstood, so as well as trying to say what they are, we should perhaps emphasize what they are not. Such models derive from research employing inductive categories to pick out broad divisions of ideas in children's responses to tasks, but they also owe much to the early days of SHP analysis of examination responses, which added considerably to our knowledge of children's ideas.<sup>7</sup>

Progression models grounded in research do not, like the National Curriculum attainment target (NCAT), simply combine complex ideas into a single target, and then cut it into an arbitrary number of convenient slices. Nor can they fall back on the conceptual crudity of generic and imprecise language like 'simple', 'begin to' and 'show some independence' as a substitute for identification of important shifts in understanding. Moreover, because there is some research evidence that students' ideas are decoupled – so that, for example, a student's ideas about evidence may remain the same while his or her ideas about historical accounts are changing quite rapidly – it is a mistake to try to bundle progression in different concepts together (See Figure 2).<sup>8</sup> We must therefore construct separate models for key concepts like 'change', 'evidence', 'accounts', 'cause' and 'empathy'.

**We have  
history  
precisely  
because the  
past is  
neither  
given nor  
fixed.**

**Figure 2:** Decoupled ideas

Chata research suggests that children's ideas do not develop at the same rate. For instance, in a sample of 92 students in years seven and nine, followed over two terms, ideas about 'cause' and 'empathy' did not appear to develop in parallel. The same kind of decoupling was apparent in a longitudinal sample of 20 students (followed from year three to year five inclusive). Four students showed no progression at all, one showed similar progression in understanding of both cause and empathy, and fifteen showed different degrees of progression in the two concepts.

We have to be careful here: we cannot say in quantitative terms what the gaps are between categories, let alone assume that the gaps in one concept area are equivalent to those in the other. However, the difference for ten students in the longitudinal sample was at least two categories, and for six was three or more categories. Given that some students showed no category change at all, it seems likely that differences of two or more categories represent substantial decoupling.

Complaints about 'atomism' are misplaced here. Until we understand how students' prior conceptions relate to one another we cannot indulge in simplistic syncretism and lump everything together. It will be enormously valuable to be able to see ideas in their proper relationships to one another, but such 'holism' has to be earned. Pious phrases about the indivisibility of history are not a substitute for research.

**Progression models grounded in research do not simply combine complex ideas into a single target and then cut it into slices.**

There is nothing privileged about any particular list of concepts that researchers have tended to work with and models of progression can be constructed for other concepts. (See, for example, Lis Cercadillo's recent exploration of students' ideas about historical significance, so far the only major research on the concept.<sup>9</sup>) When we explore children's ideas about history, we necessarily translate them into a form that will make sense to adults. We talk about an 'evidence' model, but students may be working with pre-evidential concepts like 'information' or 'testimony': we have an idea of what a sophisticated 'higher level' understanding entails, and we group their ideas for our convenience because we are interested in the development of their concepts. Progression models are therefore hierarchical and normative: they carry an implication that students are working from less to more powerful ideas. This should not be surprising given that the context is history education.

This normative feature of progression models (together with a misplaced tendency to assimilate them to Piagetian 'stage' theories) leads to some misconceptions. The 'levels' in a progression model are not a sequence of ladder-like rungs that every student must step on as he or she climbs.<sup>10</sup> Indeed, a model of the development of students' ideas does not set out a learning path for individuals at all. Assuming it is well founded, it is valid for groups, not for individuals. That is, it sets out the ideas likely to be found in any reasonably large group of children, the likely distribution of those ideas among students of different ages, and the pattern of developing ideas we might expect.

A progression model can therefore help us to predict the range of ideas we are likely to encounter, and the kind of changes we are likely to see as students' ideas develop. But research suggests that in history (as in science) there is a seven-year gap.<sup>11</sup> The ideas that some seven year-olds have about 'cause', for example, will be the same as those found among most fourteen year-olds. And some fourteen year-olds will be working with the same ideas that most seven year-olds employ, while a few will already be at the level we would expect our A2 students to reach. A model gives no warrant for the kind of 'stagism' that says that if a pupil is in year nine, he or she must be at level 'n', and even less to the assumption that the next move the pupil must make is to level 'n+1'.

A progression model can show us how most students of a given age are likely to be thinking, given teaching as it is and the ideas current in our society. It does not tell us what students must necessarily do. The changes in students' ideas displayed by a model are like the paths made by sheep as they move across a mountainside. The sheep move through the terrain in more or less regular ways, but a sheepdog can change their route. Nor is there any guarantee about a model's shelf life; changes in either teaching or social mores may compel its reconfiguration. However, such models may not be as parochial or transient as might have been expected. While progression research outside the UK is still scarce, it looks as if the models

developed in the UK may successfully predict the range of ideas with which students operate, even if the age distribution in other cultures is rather different.<sup>12</sup>

## **Progression and the National Curriculum attainment target**

The closest thing that we have to an 'authorised' model of progression is the National Curriculum history attainment target (NCAT). Whatever virtues of holism and inoffensiveness this displays, it may be criticised on several counts. For example, the liberal use of such weasel phrases as 'simple observations', 'beginning to identify', 'a few reasons' and so on, serves to blur the boundaries between level descriptors.

More serious problems follow from the use of language, which, if taken literally, could reinforce common conceptual confusions. For example, the phrase 'sources of information' is used throughout the NCAT. There is no doubt that pupils must learn to transform information from one form into others (e.g. burial register data may be transformed into histograms), to collate and to generalize from data. But since we are also concerned to establish what can and cannot be said about the past it is important for students to use *sources of evidence* rather differently than they use *bodies of information*. This distinction is especially sharp when 'relic' sources are used. As conventionally understood, a seventeenth century musket ball contains no information – in this respect it is akin to a blank page or an unpainted canvas – but, when subject to interrogation as to the incidence and distribution of similar finds, it may contribute to evidence that serves to augment or even to modify the established body of information about the Battle of Naseby.

The distinction between evidence and information should not be dismissed as sterile pedantry. It is, in fact, one of the most critical conceptual steps that pupils ever take since, until the distinction is understood, they are condemned to approach source-based tasks and, in particular, those involving textual materials in one or more of three ways:

1. Comprehension tasks akin to those presented in English lessons.
2. Text-based jigsaw puzzles in which chopped-up bits of information must be assembled into more-or-less coherent paragraphs.
3. 'Winnowing' and 'red-herring' games in which true and relevant statements are selected from texts contaminated with distracters. Pupils may construe such tasks as somewhat tiresome multiple-choice questions.

More serious than the occasional infelicities of phrasing are problems with the *nature* of the learning progressed within the NCAT. In particular, some strands of the NCAT represent models of progression that are tidy and logical but fail to connect with the ways in which pupils construe what they are taught. Worse still, the step changes from

Figure 3: The NCAT model of progression for historical interpretation

It is at the Level 4 – Level 5 interface that this model of progression ceases to be defensible. The progression steps hereon form a neat logical hierarchy of cognitive operations when what is required is a developmental model of history-specific understanding.

- Know that some events, people and changes have been interpreted in different ways and suggest possible reasons for this. (Level 5)

The key operation here is that of conjecture, the ability to suggest possible reasons as to why non-fictional accounts sometimes conflict. The next step involves the ability to move from possible to probable reasons:

- Describe, and begin to analyse, why there are different historical interpretations of events, people and changes. (Level 6)

Analysis of texts is necessary before pupils are able to move beyond conjecture but this depends upon progression in learning no more than does the filling of a kettle with water before plugging it in and switching it on! The priority afforded to one operation over the other is *logical* not *developmental*.

level to level frequently reference generic skills and operations rather than historical understanding. For example, for the historical interpretation strand, higher levels present a hierarchy of cognitive operations that moves from 'knowledge' to 'evaluation and judgement' via the steps of 'conjecture', 'analysis' and 'explanation'. (See Figure 3.)

The NCAT model of progression is flawed in two respects. First, it treats such operations as evaluation as all-or-nothing accomplishments accessible to no more than a small minority of pupils who are beyond the normal Level 3 to Level 7 range for Key Stage 3. In reality, many young or low attaining pupils can evaluate conflicting accounts to some degree, even though many such evaluations are egocentric, e.g.

- 'A is better because shorter and easier to read'
- 'B is better because it's more exciting and has pictures'

Evaluations and judgements may also be impersonal and based upon some degree of textual analysis:

- 'A tells you more about King John; it goes into more detail'
- 'B is better because it backs up what it says about King John by giving the opinions of other people'
- 'A is best because it uses words like "probably" and "perhaps"'

It follows that operations like evaluation and judgement are more accessible than the NCAT appears to suggest and can be taught before a battery of more elementary operations is in place.

The NCAT model of progression is flawed in a second and more critical respect, namely its failure to consider how pupils make sense of what they learn and are asked to do. Evaluations of contrary interpretations of the past depend not only upon certain operational competences of a generic nature but also upon the relationship that pupils conceive such accounts as having to the real past. For example, if they think that accounts should present true pictures of the past and believe it possible for the past to have its photograph taken, pupils will equate differences in interpretation with differences in accuracy of representation. In consequence, they will hunt for signs of bias (seen as doctoring of photographs of the past) and intimations of accuracy in the accounts analysed, often finding plausible signs in the backgrounds of authors and in the texts themselves. Pupils' generic skills of analysis and explanation, evaluation and judgement may be exceptional but, if they construe accounts in this way, what is written will remain very limited.

Models of progression used to inform teaching and assessment should look nothing like the NCAT. In contrast to this 'authorized version', they should:

- be developmental rather than, or as well as, formally hierarchical;
- relate to pupils' ideas and, in particular, to the ways in which they make sense of what they are taught as well as to what they can do;
- progress beyond the commonsense of the street and the media to embrace that which is counter-intuitive and particular to the discipline.

**The sheep  
move  
through the  
terrain in  
more or  
less regular  
ways, but a  
sheepdog  
can change  
their route.**

## Research based models of progression

The NCAT is not the only kid on the block. Alternative models of progression may be derived from the now considerable body of constructivist research conducted in the U.K. and overseas over the past 25 to 30 years. It must be admitted, however, that research-based models of progression are neither as comprehensive nor as utilitarian as the NCAT.

It is only fair to acknowledge the revolution in assessment signified by the NCAT and associated procedures for statutory teacher assessment at the

end of Key Stage 3. For decades, teachers and examiners have struggled to ensure the reliability and validity of assessments. One revolutionary masterstroke was to dispense with rules of evidence for end of key stage assessments. The authors have seen work copied from the board and written by 'peer tutors' being used for this purpose and, in the absence of any rules of evidence, such material is (absurdly) as admissible as formal test and examination papers. Since any material produced under any conditions may be 'best-fit' matched against the NCAT nothing is deemed to be unreliable and assessments become 'error free'.

Figure 4: NCAT versus research based models of progression

Robust research-based models of progression may be produced for such second-order concepts as evidence, accounts, change, cause and empathy, and, more tentatively, of the big picture of the past. But, unlike the NCAT, research-based models cannot encompass the totality of knowledge, skills and understanding required for Key Stage 3. They cannot accommodate elements like 'communication of knowledge and understanding'.

Research suggests that pupils' understanding of substantive history may be enhanced or limited by their grasp of second-order concepts. It is not possible, however, to offer a single research-based model that applies to the subject as a whole because ideas about second-order concepts do not develop in parallel. In contrast, the NCAT is holistic and, in this respect at least, goes far beyond anything currently known.

Research-based models can be levelled in ways that correspond with current practice, but do not naturally break down into nine or any other convenient number of levels. Indeed, for research-based models there is an ineluctable trade-off between complexity and robustness: the greater the number of levels, the less robust the model becomes. The integrity of the NCAT, on the other hand, cannot be degraded by anything whatever. The nine-level version is as valid as any other – how many levels would you like?

A research-based model sets levels in an order. But we cannot say whether the 'gaps' between levels are equal, or how wide they are. The scale offered by the NCAT presumes that the intervals are equal, another illegitimate pretension to sophistication. For the average pupil, each gap between each pair of NCAT levels takes exactly two years to cross, neither more nor less. If the NCAT scale really were allied with error-free assessments, it would enable precise targets to be set and value-added calculations to be made. Needless to say, research-based models cannot be used for such purposes.

Even more revolutionary are the ways in which issues of validity are resolved. The NCAT goes some way towards eliminating this problem by specifying multi-strand descriptors with indeterminate boundaries. This allows for interminable debate about which of many possible ‘fits’ is in fact the ‘best’. Better still, the whole concept of ‘best fit’ actually enables assessment to take place whether or not the data actually ‘fit’ the performance criteria. Imagine a darts match in which three darts miss the board but hit the ceiling, the barmaid and the dog in the corner. With the aid of a tape-measure each dart can be ‘best-fitted’ to a particular cell in the board; the dart in the ceiling, for example, might ‘best-fit’ to double-twenty! In like manner, it is possible for assessment data to be ‘best-fitted’ to a level descriptor that they fail to match on the grounds that the mismatch with other levels is even greater. Thus it is that issues of validity are sidestepped.

Despite the fact that the NCAT is an unrivalled tool for assessment at the end of Key Stage 3, research-based models of progression may claim to be superior in two respects. First, they connect more closely with pupils’ ideas, errors and misconceptions. Second, they are developmental, highlighting the watersheds (or crisis points) in learning that teachers must plan to negotiate and which assessments should seek to register. These advantages are evident in the six-level model of progression in ideas about evidence given in Figure 5.

### What have models of progression ever done for me?

Ten years ago Paul Black wrote, ‘Anyone planning teaching has to have some way to decide in what order pupils’ thinking should be encouraged to develop – it is inconceivable that a subject’s teaching be planned without some model of progression as a basis.’<sup>13</sup> We are now more disposed to conceive the inconceivable because, like many other bright ideas for filling the waking hours of serving teachers, implementation of progression models can become an obituary-enhancing activity with no perceptible impact on what pupils actually learn. This does not mean that Paul Black is wrong, but rather that we need to think carefully about what might count as appropriate models and how they may be used. Progression models must be capable of being applied in the ways and to the extent that individual teachers find useful. Such uses will vary greatly from classroom to classroom but might extend to some or all of the following:

- Planning within and across key stages.
- The identification, evaluation and remediation of common misconceptions.
- Recording and analysis of attainment and progress over the long term.

Options with respect to these categories of use will be examined in turn.

## Models of progression and curriculum planning

Research-based models of progression may be used to calibrate the outcomes of National Curriculum History with GCSE specifications and assessment objectives. Models that are applicable to the whole of the secondary age-phase may serve to mitigate the dislocations attaching to shifts from mandatory to elective courses of study, and from compulsory to post-compulsory age-phases.

Models may also be used to facilitate the planning of progression both within a scheme of work and across units of work. For instance, historical enquiry and source work figure in several of the units in the QCA model scheme of work, but the manner in which historical enquiry is revisited appears to be *ad hoc* and to betray little sense of direction or purpose. Historical enquiry seems to have been mapped onto a set of units rather than to progress through them in any systematic fashion. This is not to suggest that historical enquiry and evidence concepts be addressed in all or most units, but these concepts should be regularly revisited and relevant unit objectives should progress from each other in a rational and calculated fashion. Research-based models of progression may provide empirically grounded compass bearings and sets of way markers for this purpose.

Although such models can be enormously valuable in curriculum planning, we must avoid falling into a dangerous trap. We cannot treat models of progression as if they set out a list of ideas that must be taught one after the other. The ‘evidence’ model given in Figure 5 suggests ways in which we might expect students’ ideas to develop. It enables us to recognize real achievements, even when the new understanding achieved remains a misconception. It helps prepare us to recognize how students may assimilate what we think we are teaching to their existing prior conceptions. But it does not mean that we should set out to teach students an inadequate idea simply because it seems to be the next one on the list.

The development of pupils’ ideas about the reliability of sources of evidence may serve as an example of some important conceptual shifts that should inform planning. (‘Reliability’ is used as an example because most teachers address it at KS3, not because it is the most important idea at stake.)

(a) The notion of reliability means little to pupils until they distinguish between information about the past (Level 2 in Figure 5) and testimony from the past (Level 3). Pupils who are still at Level 2 assimilate the distinction ‘reliable v unreliable’ to that of ‘correct v incorrect’ as applied to information. This assimilation of new jargon to old knowledge reinforces rather than challenges the latter. It is only when pupils treat sources as testimony about the past from putative eyewitnesses to the events reported that they can

**Models of progression should go beyond the commonsense of the street and the media to embrace the counter-intuitive.**

**Progression  
models  
must be  
capable of  
being  
applied in  
the ways  
and to the  
extent that  
individual  
teachers  
find useful.**

understand the notion of reliability as part of a methodology for testing testimony.

(b) Some ways of evaluating the reliability of source-based data are likely to be understood long before other ways will make sense to pupils. Pupils thinking in Level 3 terms (see Figure 5) are usually able to understand that (i) some witnesses are more credible than others, e.g. they are disposed to trust clerics more readily than lawyers (the *ad hominem* test); and (ii) that first-hand reports should be accepted before hearsay (the witness test). Although such ideas are seldom adequate for their purpose, they represent real achievements for pupils for whom the question, ‘How do we know?’ had hitherto been meaningless or unanswerable.

(c) A shift to Level 4 ways of thinking involves the coordination of several new ideas (see Figure 5). Most significant is the realization that the several statements within a source (made by an eyewitness) can be more or less reliable and that the all-or-nothing approach to testimony characteristic of Level 3 is not the optimum means of constructing a true (sic) report about the past. Because it is hard to apply the *ad hominem* and witness tests to fractions of testimony, they tend to reinforce the ‘all-or-nothing’ approach. If students thinking like this encounter new reliability tests, such as ‘access to information’ and ‘author intent’, they will have tools to move them on. But, as experience has shown only too well over the past decades of ‘source work’, students already have an everyday concept of ‘bias’ and may take it to mean nothing more than the reason why someone was lying. In consequence, the new tool may be assimilated to prior conceptions and become just a new variant of the *ad hominem* test. If ‘bias’ is taught as a complex phenomenon that extends beyond partisanship to include ‘interest’ in its widest sense, and if identification of ‘bias’ is also used to demonstrate the reliability of particular statements as well as the converse, students may avoid this kind of assimilation. (E.g. statements about the treatment of wounded Zulus after the Battle of Rourke’s Drift gain credibility because they were made by people with an interest in silence, evasion and self-exculpation.) Above all, the move from considering the reliability of testimony on an all-or-nothing to a statement-by-statement basis requires teaching pupils to cross reference statements contained in independent and unrelated sources.

(d) The major shift at Level 5 is from the common-sense idea that our knowledge about the past depends on witnesses’ reports, to an understanding that we can make valid statements without accepting the literal truth of reports in whole or in part. The key ideas for students here are first, that the reliability of sources or their component statements is not absolute but depends on the questions that we wish to answer; and second, that it makes no sense to question the reliability of some sources. Some sources are completely mute until they are interrogated. A midden heap in a Neolithic settlement, for example, is a relic from the past that

reports nothing. The evidence it yields is conditional upon the questions posed, and will differ against such questions as ‘How nutritious was the diet of these people?’ ‘How large was the settlement?’ and ‘Was inhabitation of the site continuous or seasonal?’ It makes sense to query the authenticity of the source and the validity of the conclusions derived from evidence, but it does not make sense to challenge its reliability. This may seem to be an easy thing to teach but it is not. Pupils assimilate midden heaps, musket balls and castle walls into the ranks of eyewitnesses that are thought to tell, or show, it as it was. It is only when they understand that evidence taken from both relics and reports varies in both kind and quality according to the questions posed that pupils realise that relics ‘say’ nothing in and of themselves. Reports are more difficult to handle precisely because they do say something about what happened in the past. But students asked about the reliability of school reports in answering questions about their school work, their relations with teachers, and their teachers’ ideas about standards of work and behaviour, may be persuaded that what we can say on the basis of any particular record of the past will vary with the questions we pose. (Indeed the very distinction between relic and record sources depends in part on the questions posed: the Bayeux Tapestry may serve as a record of the Battle of Hastings for some purposes and as a relic artefact for others).

(e) At Level 6 a key idea is that we need to reconstruct the meanings that attach to sources in the contexts in which they were produced and used. This can be a very sophisticated process. For example near contemporary accounts of the Battle of Hastings are replete with echoes of classical epics and Dark Age romances. It is tempting to dismiss such allusions as literary artifice but there is good reason to believe that noble protagonists consciously mimicked episodes from tales about Aeneas, Roland, Oliver and so on. Indeed, at the onset of battle, Taillefer rode before the Norman lines singing lines from the *chanson de geste* that invited comparison between the coming fight and the Battle of Roncesvalles. The extent to which art embellished life and life imitated art may be difficult to determine, but it is not always so tricky. When taken in conjunction with other evidence, the fact that blinding was an established punishment for oath-breaking may lead us to question near contemporary reports that, at a critical point in the battle, Harold was struck in the eye by a Norman arrow.

It should be apparent that the model of progression for evidence must not be applied to the history curriculum in any rigid or mechanistic sense. There is no suggestion that levels are fixed by age or even that teaching should follow an invariant sequence. What is argued is, first, that we should plan to progress pupils’ understanding of such second-order concepts as evidence across units of work and key stages; second, that it is important to take account of how pupils make sense of what they are taught rather than to plan for progression solely in terms of some notional

Figure 5: Progression in ideas about evidence

### **Pictures of the past**

The past is viewed as though it were the present, and students treat potential evidence as if it offers direct access to the past. Questions about the basis of statements about the past do not arise. Stories are just stories.

### **Information**

The past is treated as fixed and known by some authority; students treat potential evidence as information. Given statements to test against evidence, students match information or count sources to solve the problem. Questions arise about whether the information offered is correct or incorrect, but no methodology is attributed to history for answering such questions beyond an appeal to books, diaries or what is dug up. These, although sometimes seen as being connected with the past, provide transparent information that is either correct or incorrect.

### **Testimony**

The past is reported to us by people living at the time. Like eyewitnesses today, they do this either well or badly. Questions as to how we know about the past are regarded as sensible: students begin to understand that history has a methodology for testing statements about the past. Conflicts in potential evidence are thought appropriately settled by deciding which report is best. Notions of bias, exaggeration and loss of information in transmission supplement the simple dichotomy between truth-telling and lies. Reports are often treated as if the authors are more or less direct eyewitnesses: the more direct, the better.

### **Scissors and Paste**

The past can be probed even when no individual reporter has told us truthfully or accurately what happened. We can put together a version by picking out the true statements from different reports and putting them together. In one student's words: 'You take the true bits out of this one, and the best bits out of that one, and when you've got it up, you've got a picture.' Notions of bias or lies are supplemented by questions about whether the reporter is in a position to know.

### **Evidence in isolation**

Statements about the past can be inferred from sources of evidence. We can ask questions of sources that they were not designed to answer, so that evidence will bear questions for which it could not be testimony. Many things may serve as evidence that do not report anything. (Nineteenth Century rail timetables were not constructed for the benefit of historians.) This means that historians may 'work out' historical facts even if no testimony survives. Evidence may be defective without questions of bias or lies. Reliability is not a fixed property of a source, and the weight we can rest on any piece of evidence depends on what questions we ask of it.

### **Evidence in context**

A source only yields evidence when it is understood in its historical context: we must know what a source meant to those by and for whom it was produced. This involves the suspension of certain lines of questioning and a provisional acceptance of much historical work as established fact (a known context). We cannot question everything at once. Contexts vary across time and place and thus a sense of period is important.

hierarchy of cognitive operations (as with the NCAT); and third, that research-based models of progression can prove useful in this connection.

## Models of progression and day-to-day monitoring

The previous discussion has already emphasized the importance of addressing students' prior conceptions. Many of the misconceptions that pupils exhibit derive from ways of thinking described in research-based models of progression. For example, a pupil who justifies her failure to refer to certain sources on the grounds that 'they only tell you things that you already know' because the information they contain is 'about the same thing' as that in another source clearly fails to see information about the past as in any way problematic (Level 1: Pictures of the Past). Attempts to teach her why cross-referencing is important are unlikely to be fruitful because she has yet to ask the question, 'How do we know about the past?' or to accept that some information about it can be false. For her, it would be an advance were she to understand the importance of 'counting sources' to resolve discrepancies in bodies of information.

Reference to research-based models of progression can thus be useful in enabling us to see where some common misconceptions might come from and in suggesting the level at which attempts at remediation may be pitched. More important though is the fact that such models represent progression in understanding as the gradual transformation of misconceptions rather than as the gradual accumulation of correct ideas. Only on rare occasions can we remedy fundamental misconceptions; most of the time success is limited to the slow transmutation of less into more useful and sophisticated misconceptions.

## Models of progression and the assessment of learning

This is territory claimed by the NCAT for statutory teacher assessment and reporting at the end of Key Stage 3, but a surprising number of schools use the NCAT for non-statutory purposes including:

- ➔ The routine marking of internal assessments and, in a few extreme cases, of homework.
- ➔ Reports to parents in years 7 and 8.
- ➔ Target-setting for pupils on an annual, term or half-term basis.
- ➔ Departmental evaluation and target setting

We should not let the dubious claims of the NCAT model seduce us into trying to use research-based progression models in these ways. But progression models can legitimately be used in assessment tasks. Space limits us to two examples.

1. We can provide notes in pupil records that Boris seems to have made a secure transition to, say, Level 4 understanding of evidence but his big picture overviews of the past remain at Level 2, his understanding of cause does not seem to fit the normative model and so on. These have the status of descriptive notes, not scores or measures of whatever kind.

2. We can calculate cohort progress against one or more models over a year or a key stage. Progress may be validly represented as the difference between two histograms, medians or modes (but because we do not know that the intervals between levels are equal, it may only be represented in terms of mean gain scores – the difference between two average levels – if the deputy head-teacher suffers from methodological thought disorder). Under no circumstances is it valid to report levels to parents as 'measures' of individual attainment or progress, to set levels as targets for individual pupils or colleagues, or to use levels as a basis for grade predictions or value-added calculations.

## Research-based progression models: the dark side

Models of progression are vulnerable to abuse of many kinds and it is important that their limitations be understood. The first limitation pertains to their range of convenience. Most of history, and most of what is of most value in the study of the past, cannot be captured by models of progression. Indeed, it is impossible to legislate for the wisdom, perspective and understanding that can come from historical study since this is particular to the knowledge, creativity and charisma of individual teachers. This is not to devalue the significance of the concepts and understanding that can be systematically captured in models over the medium and long term. Wisdom and insight need not preclude intellectual rigour. Models of progression can supply a scaffold for the teaching and learning of history, but, if ill-used, this scaffold can become a cage.

The second limitation derives from the fact that research-based models of progression are *normative* not *universal*. There are many roads to Rome and some pupils prefer to head for Geneva. Normative models work well for groups of pupils but the progress of some pupils may be best described in terms of the ways in which they depart from a standard model.

The third and final limitation relates to the fact that research-based models have low resolution and do no more than pick out the main features of progression over the long term. For units of work we need high-resolution models that sequence achievable objectives over a number of lessons. We cannot simply deduce high-resolution models from research-based models, but the former should be informed by and compatible with the latter.

**There are many roads to Rome and some pupils prefer to head for Geneva.**

Current research-based progression models are necessarily provisional, but they are grounded in empirical investigation, not conjured from the air. A growing body of research, increasingly international, is becoming available to test them. They offer a picture of the development of students' prior conceptions, and a basis for thinking about the conceptual tools students need to make sense of the past. Most important of all, they offer teachers a starting point for exploring their own students' ideas. Marking need never be the same again.

#### REFERENCES

1. Examples include Charlton, K (1952), *Comprehension of Historical Terms*, unpublished B.Ed. thesis, University of Glasgow; Coltham, J. (1960) *Junior School Children's understanding of Historical Terms*, unpublished PhD. thesis, University of Manchester.
2. Note that outside the UK (especially in Italy) valuable work – influenced by Piaget – has been done over two decades on children's ideas about political and economic life. See, for example, Berti, A.E. and Bombi, A.S. (1988) *The Child's Construction of Economics*, Cambridge: Cambridge University Press; Berti, A.E. (1994) 'Children's understanding of the concept of the state'; in *Cognitive and Instructional Processes in History and the Social Sciences*, M. Carretero and J.F. Voss (eds), New Jersey: Erlbaum. Neglect of children's preconceptions appears to be a serious problem in citizenship education.
3. We do not subscribe to the apparatus of Piagetian stages, and would want to emphasize distinctive characteristics of historical thinking, but we also think that simplistic criticism of Piaget (for example that he omitted social aspects of development) has persuaded too many commentators in history education to ignore more interesting elements of his theories, like his concepts of 'centration' or 'field'. The issue is more complex than 'rejection', let alone 'disproof', of Piaget.
4. This has not been explicitly discussed in published papers from Project *Chata* (Concepts of History and Teaching Approaches at Key Stages 2 and 3), but see Lee, P.J. and Ashby, R. (2000) 'Progression in historical understanding among students ages 7-14'; in *Knowing, Teaching and Learning History*, P.N. Stearns, P. Seixas and S. Wineburg (eds), New York: New York University Press; and Lee, P.J. and Ashby, R. (2001) 'Empathy, perspective taking and rational understanding'; in *Historical Empathy and Perspective Taking in the Social Studies*, O.L. Davis Jr., S. Foster and E. Yaeger (eds), Boulder: Rowman and Littlefield. Keith Barton's work in the US seems to point in the same direction – Barton, K. (1996) 'Narrative simplifications in elementary students' historical thinking'; in *Advances in Research on Teaching Vol. 6: Teaching and Learning History*, J. Brophy (ed), Greenwich: JAI Press.
5. Project *Chata* response.
6. Understanding preconceptions is the first of the three key principles ¾ derived from 50 years of research on human learning ¾ picked out by the US National Research Council's report *How People Learn: Brain, Mind, Experience and School*, J.D Bransford, A.L. Brown and R.R. Cocking (eds) (1999) Washington DC: National Academy Press.
7. Henry Macintosh and John Hamer produced CSE examiners' reports from the Southern Regional Board giving teachers feedback on a scale that shames current reports.
8. Lee, P.J. and Ashby, R. (2000) *op. cit.*
9. Cercadillo, L., (2001) 'Significance in history, students' ideas in England and Spain'; in A.K. Dickinson, P. Gordon and P.J. Lee (eds), *Raising Standards in History Education, International Review of History Education Volume 3*, London: Woburn Press.
10. Talk of 'linear progression' is also misleading if this implies that analysis of students' ideas about a particular concept commits researchers to the notion that learning in that concept follows a linear path. But it is seldom clear what the idea of 'linear progression' actually means to those who use it. See, for example, E. Vermeulen (2000) 'What is progress in history', *Teaching History 98, Defining Progression Edition*.
11. Lee, P.J. and Ashby, R. (2000) *op. cit.*
12. The work of Isabel Barca and Marilia Gago in Portugal, Lis Cercadillo in Spain, and Project CHIN (Children's Ideas about Narrative: Understanding Historical Accounts) in Taiwan suggest that students in those countries employ surprisingly similar ideas to those of UK students. Explorations of students' ideas by Keith Barton and Alan McCully in the USA and Northern Ireland also reveal some marked similarities. Caution is obviously required, but such unexpected similarities merit further exploration.
13. Black, P.J. (1993) 'The shifting scenery of the National Curriculum', in O'Hear and White, J.(eds) *Assessing the National Curriculum*, London: Paul Chapman.

**Models of progression can supply a scaffold ... but, if ill-used, this scaffold can become a cage.**