This is the author's manuscript version.

The version of record is
Taber, K. S. (2021). Making sense of a pedagogic text. Foundations of Chemistry. doi:10.1007/s10698-021-09405-8
https://link.springer.com/article/10.1007/s10698-021-09405-8

Making sense of a pedagogic text

Review of Reid, N., & Ali, A. A. (2020). ***Making Sense of Learning: A research based approach. Evidence to guide policy and practice, with an emphasis on secondary stages***. Cham, Switzerland: Springer Nature Switzerland AG.

Softcover, ISBN 978-3-030-53676-3, £74.99. 1st ed. 2020, XXI, 496 p.

Introduction

'*Making Sense of Learning: A research based approach. Evidence to guide policy and practice, with an emphasis on secondary stages*' (henceforth, '*Making Sense of Learning*') is a book that aims "to make available, hopefully in straightforward language, what is *directly relevant* to the actual practice of teaching along with the pressing issues that face education today" (p.vii, emphasis in original). It is a substantial volume of 19 chapters spread across about 500 pages. The book includes a large number of figures, many being simple illustrations to highlight and reiterate key points in the text, and an index. This review offers a brief introduction to the scope of the volume, and discusses some key features of the book in relation to its nature as a pedagogic text. Page references to quotations below refer to '*Making Sense of Learning*', except where otherwise indicated.

Reviewing a textbook

'Making Sense of Learning' is a textbook, which is a form of text somewhat different to a scholarly review of literature published in a journal. Such a review would be written for (and usually subject to pre-publication evaluation by) peers in a scholarly community. Those peers would mostly not know as much about the *specific* focus of the review as the author(s), but would generally know enough about the wider field of study in which it is located to have a strong familiarity with the nexus of theory, research, methodological tradition, and so forth, current within that particular field. For the purpose of this review I take the general area of research and scholarship relating to classroom teaching and learning as the field in which *'Making Sense of Learning'* is located.

There is always a pedagogic aspect to *any* academic writing (the person setting themselves up as an expert in some small area wishes to teach others about their insights, speculations, or specific research findings), but in a textbook the assumption is that intended readers are relative novices in the field, and they need an expert to *mediate* for them (Vygotsky, 1978) between the current state of the field (the extensive and diverse technical literature) and their relatively limited background knowledge.

This influences how the author should approach writing. One methodological consequence is that the textbook author should seek to think about the subject matter as it would appear from the readers' starting points - so, for example, being *selective* to help highlight what is of most significance, and looking for optimal *simplifications* that communicate the gist of complex ideas without stripping away their essence. A key consideration in evaluating a textbook, then, is how disciplinary knowledge has been *represented* in a form likely to be accessible to the intended readership.

Arguably, another consequence is that whereas when the author of a journal article offers an idiosyncratic take on a topic her peers are well-placed to recognise this and evaluate whether it is a useful (even if not a currently canonical) contribution; by contrast, the author of an introductory textbook has a responsibly to seek to offer a broad and balanced view of the field for readers who are not well-placed to gauge the authority of the text. However, sometimes (as here), a strong motivation to write the textbook may be that the author feels that the current state of the field invites some restructuring, refocussing, and even perhaps a somewhat heretical reframing (i.e., given "the pressing issues that face education today"), in which case the author is faced with a fascinating dilemma in 'balancing' two contrary ethical imperatives: to reflect, and yet at the same time to reconfigure, the field. One focus of this review is how Reid and Ali have responded to this 'essential tension' (T. S. Kuhn, 1959/1977). [1]

A focus on general educational issues rather than subject pedagogy

The scope of *'Making Sense of Learning'* is impressive. Its focus is learning in formal educational settings: especially secondary school contexts. It seeks to be relevant to teachers, student teachers, "teacher trainers" (i.e., teacher educators) and educational managers (p.viii) - across different educational contexts, and regardless of curriculum area. Responding to this broad scope would be a challenging task even for an edited volume, but *'Making Sense of Learning'* is co-written by just two authors - which makes for a somewhat gargantuan task. One advantage of such a small author team is a general sense of coherence and uniform approach across the volume, which can be especially useful in a textbook, although this is at the cost of not involving specialists from each of the different topics discussed.

This breadth may be considered especially 'brave' given that one point that is emphasised at various places in the text (and this is a book that utilises repetition extensively, as discussed below) is that secondary school teaching needs specialist teachers and that it is therefore unreasonable and inappropriate for teachers to be asked to teach out of specialism. In that regard, it is worth noting that the authors are associated with the strong tradition at the University of Glasgow of research into chemistry and physics education, where Reid was for some years the head of a centre for science education (Johnstone, 2006). Indeed, the book is dedicated to the memory of one of the most influential chemistry educators, Alex Johnstone, who also taught and undertook research there for many years,

> In his long career of research - much in chemistry education - Johnstone generated major insights into the world of conceptual learning. His work still has a huge impact at university levels throughout the world but, sadly, it has not influenced much at school level, partly because the decisions about what is taught, how it is taught and assessed as well as the conditions for teaching and learning are largely decided OUTSIDE the schools in most countries. (p.29, emphasis in original)

Between them, Reid and Johnstone must be responsible for training a very large number of current academics and researchers in Universities spread across the world. A careful reader may notice that a disproportionate number of the studies cited in this volume derive from post-graduate projects at the Glasgow Centre, in keeping with the volume's dedication.

This raises the issue of how the authors of *'Making Sense of Learning'* deal with the challenge of writing a text to support teachers of mathematics, biology, geography, history, music, religious education, and so on. In effect they do this by largely avoiding the issue, by nearly always keeping the text generic, and offering an account intended to be applicable to the learning of skills and the development of understanding in general terms. Sometimes the curriculum context in which particular research results were obtained is noted, and sometimes there are vague references to how much of the research in which they are basing an argument tends to come from one curriculum area (presumably chemistry education) - but, mostly, the book presents accounts and claims that are offered as being *generally* applicable to learning (and so teaching) in all formal educational contexts.

What has been achieved is considerable. The book includes material about a wide range of issues, including for example, cognitive development, learning styles (which are, sensibly, largely debunked), curriculum, assessment, thinking skills, attitudes, intelligence, and so forth. The book also considers aspects of research itself - "to tell something of the story of research" (p.viii). There seems little doubt that any teacher or teacher-in-preparation who has previously had limited opportunity to engage with such a range of issues will find a wealth of useful background here. Whether such a learner would be advised to select this book rather than perhaps a *range* of more modest titles from different authors each examining a specialist topic in depth, and - in particular - *contextualising key ideas within the context of their own teaching subject* (and so offering practical examples from the particular subject material that readers will be asked to teach), is a moot point.

One major area of research in teaching in recent years has concerned pedagogic content knowledge (Kind, 2009) - the specialist knowledge a teacher develops that relates to the teaching and learning of specific topics (what the learners find difficult, accessible examples, useful resources, and so forth) and it has been mooted that through planning, executing and evaluating teaching, and assessing learning, a teacher comes to also develop a form of specialist subject knowledge (Taber, 2020a). Texts to support teacher development are therefore often aligned with specific curriculum areas to exemplify principles in ways that teachers can see are relevant to their classroom practice.

In a more generic treatment, responsibility is placed on the reader to work out how to apply the ideas and principles in their own professional work. A consequence of seeking to write for all secondary level teachers is that a greater onus is placed on the reader to interpret the text in the context of their own classroom work. Indeed, in considering the textbook as an instrument of pedagogy (that is, any textbook author's purpose is pedagogical, as a kind of teacher working with learners indirectly through a fixed text) it makes sense to look to link abstract theoretical perspectives or research-derived principles with specific examples in the experience of learners (i.e., readers - here, mainly teachers and those preparing for teaching). This is a key principle of what is sometimes termed 'constructivist' pedagogy, that 'meaningful' learning of abstract ideas requires them to be relatable to the mental resources a learner already has available to make sense of teaching (Ausubel, 2000).

So, in a book aimed at teachers and teachers-in-preparation in *a single school subject area such as chemistry*, it is possible to introduce and/or exemplify the theory with examples relating to what the reader will be engaged with in their own lesson preparation, classroom teaching, and student assessment. This becomes very difficult in a generic volume meant to be equally applicable to the history teacher, the ethics teacher and the physics teacher. Whilst this is a severe limitation, the authors seem to be genuinely concerned to address a serious lacuna in certain teacher preparation / teacher development contexts where they have worked and where they consider that other relevant resources are not available to educational professionals (p.viii). In such contexts *'Making Sense of Learning'* can make a very valuable contribution. (At least, if this worthy aim has not been seriously undermined by the price point the publisher sells the book at.)

Moving beyond that general evaluation, two features of the text seem to invite more detailed, and perhaps more critical, examination. One of these is the lack of balance in the treatment of some topics, and the other is what struck this reader as an intriguing juxtaposition between the explicit model of learning that is the core message - relentlessly championed in the text - and some key implicit assumptions about learning which seem to have informed the pedagogic style and structure of the book.

Taking a polemic approach

As suggested above, the authors of a book such as this face an interesting choice. If this book is intended as a comprehensive resource which may be the main source of information for teachers and student teachers, then it might be suggested that the authors have a responsibility to offer a balanced account of any topics that might be considered controversial. However, where scholars have a strong commitment to a position that current policies and practices are ill-informed and damaging, they would also seem to have a responsibility to do their best to make a case for the changes they feel are indicated. Reid and Ali clearly do have some very strong views on some important educational issues, and appear to have decided that they should not compromise on those views (through presenting alternative views or counter arguments), perhaps seeking to avoid any potential risk of diluting the message they feel should be communicated.

One of these areas has already been referenced above. Readers are informed:

> It is incredibly difficult to keep up to date, even with one part of one science discipline. To try to keep up to date in all three sciences is asking the impossible for any teacher. It is, therefore, very foolish to ask one person to teach all three sciences: no one can know enough, understand enough or keep up to date……[sic]! (p.154)

I have much sympathy with this view, and indeed was very critical of a shift that took place in England some decades ago to make 'science' a school curriculum subject, and then train graduates from across the natural and behavioural sciences as 'science teachers' expected to teach biology, chemistry *and* physics.[2] I consequently have some reservations about the current international trend towards 'STEM' (i.e., science, technology, engineering, and mathematics) education (Chesky & Wolfmeyer, 2015) when this involves replacing specialist teaching in the component areas.

That said, opportunities for secondary learners to engage in cross-disciplinary activities and projects can be valuable. Moreover, there could be an argument for continuity of teaching in the *early* secondary years (where children have often moved from being with one teacher for all subjects in primary school to suddenly having a fragmented curriculum experience) when it may be most important to offer a consistent introduction to the natural sciences and to develop a strong base in key principles, common scientific values, and the nature of science, as a foundation for subsequent more specialist teaching. Different science teachers have very different disciplinary backgrounds, different levels of interest and confidence in different areas of science; and different schools have very different resources, local contexts, and cohorts of children. I would agree that *generally* the pupils in the *upper* secondary grades should *normally* be taught by different disciplinary specialists, but I am much less convinced this is *always* in the best interest of *all* pupils across *all* secondary grades.

This is just one place where Reid and Ali take rather definitive and absolute positions that would not be seen as reflecting a general consensus within the field, as there are different views among experts. Some of their other claims include:

• "There is no evidence that teaching critical thinking as a separate course has been shown to be successful" (p.240).

• "far too much teacher education is based on personal experience, coupled to the educational folk-lore and practices in any educational culture, with limited reference to the very clear findings from worldwide research" (p.457).

• "It is a sad fact of life that the curricula and assessment systems developed in most countries *almost never* take into account the findings from educational research" (p.2, emphasis in original).

• "Conceptual ideas simply are inaccessible up to age 11 or 12 and, even then, the growth of the ability to handle concepts comes slowly…any science taught up to about age 14 has to recognise that conceptual ideas are largely inaccessible at that stage" (p.19).

• "In the late nineteenth century, education was seen entirely as the transfer of knowledge from the head of the teacher to the heads of learners. In most countries of the world, this still largely applies" (p.116).

• "Science subjects were not an established part of the school curriculum at any level in any country before the twentieth century" (p.151).

• "It is essential to resist the pressures for national or international testing regimes. The evidence shows that such testing does immeasurable harm…" (p.224).

• "Most learning in schools and universities is passive. The learner absorbs (or maybe does not absorb!) all that '*washes by*' during a lecture or a lesson. It is rare for the learner to interact mentally with the new material" (p.293, emphasis in original).

• "examination marks all tend to follow the normal distribution…If we have a big enough sample, we shall always obtain a marks distribution that follows the pattern…" (p.403, p.396).

Again, I have considerable sympathy with some of the positions presented here. Yet, I am less comfortable with what seem (albeit informed) opinions being presented as factual statements. I do not have strong expertise to challenge the claim about critical thinking courses, but am aware there are studies that come to a different view (Halpern, 2000). I suspect what is claimed about curricula and assessment systems developed in most countries *may* be *generally* true, as may, *possibly*, what is claimed about 'much' teacher education (though I worked in teacher education in a context where it was certainly not the case). Yet, I wonder if *these* statements are based on a careful review of all the research evidence.

I tend to agree with the gist of the view about testing regimes, but surely the issue is more nuanced than is suggested here? Too much learning in schools and universities may well be passive (as surely, any is too much?), but I wonder what evidence-base justifies the claim that 'most' is? I am not sure what to read into "The learner absorbs (or maybe does not absorb!) all that *'washes by'* during a lecture or a lesson" as this is certainly counter to my experience as a teacher (or scholar), is contrary to widely accepted educational thinking, and seems antithetical to Reid and Ali's core message about the critical role of working memory (WM, discussed below).

The idea that "conceptual ideas simply are inaccessible up to age 11 or 12" would seem to be a gloss on Piaget's (1970/1972) extensive research. However, it is not clear what Reid and Ali mean by 'conceptual ideas' as Piaget himself extensively discussed the conceptions of much younger children (Piaget, 1929/1973). Moreover, there has been extensive work challenging a simplistic interpretation of Piaget's findings. Such a general statement as this stands in stark contrast to Jerome Bruner's (1960) equally bold claim that the basis of any subject can be taught to a child at any level in an intellectually honest way. Neither of these extreme claims stand scrutiny without particular interpretations of their terms - what counts as a 'conceptual idea' ('plant', 'melt', 'float'?) and 'inaccessible' or indeed as a 'subject' (science, quantum gravity?) and as 'intellectually honest' - as well as considerable caveats. The useful 'truth' is likely found where such elaborations allow these seemingly contrary claims to actually overlap. There is work, for example, that shows quite complex scientific conceptual ideas can be effectively taught in primary school when the level of treatment (simplification) and teaching approach are designed to match the learners' starting points (Browning & Hohenstein, 2015) - that is, when a so-called 'constructivist' approach is followed.

I was surprised to read the suggestion that examination marks always tend to follow a normal distribution, as I had understood that mark distribution would depend on test design features (for example, it should not be difficult to intentionally design a test paper that would lead to a bimodal distribution). It is common for examinations to be deliberately *designed* with a range of items having a wide spread of difficulty levels (facility indices) when measured for the intended population, which when successfully achieved will tend to give a bell-curve shaped distribution of scores - but this is by no means inevitable. Although in one study an analysis of 400 undergraduate examinations found the the data was not well described by a normal distribution (Arthurs, Stenhaug, Karayev, & Piech, 2019), the authors acknowledged that "historically, grades have been assumed to be normally distributed, and…this is a good assumption for [those] tests comprised of equally-weighted dichotomous items" (p.1). So, although Reid and Ali's claim is an over-simplification, it could be considered a reasonable 'first approximation' for introducing the topic.

A gloss on the history of education

Two other claims listed are, strictly, contrary to the available evidence. There are plenty of examples of thinkers prior to the twentieth century who recognised the poverty of the 'transfer' metaphor for learning (Rousseau's 'Emile', for example, was published in 1762). Fröbel (1827/2012), actually writing in the nineteenth century, argued that the purpose of education was to support the development of a self-determined, independent learner who would be guided by his or her own conscience - a perspective reflected today in the notion of education as 'Bildung' (Hansen & Olson, 1996). Sadly, even today, the 'transfer of knowledge' notion still operates as an influential folk model of learning (Taber, 2013b), and it certainly was once *prevalent*: but claiming this was 'entirely' the case seems an unnecessary distortion. This raises the question of whether the introductory nature of this textbook justifies this degree of ahistorical simplification.

A related over-generalisation would be that "formal education, involving schools and colleges, was in its infancy in the nineteenth century…People at that time thought that the only difference between children and adults was that the children were like empty pots to be filled with endless knowledge given to them by adults" (p.15). Whilst free education for all young people is certainly a relatively 'modern' institution, formal education in schools and colleges, for a minority at least, goes back many centuries. This may have long been limited to males, and may have been only systematically accessible to the wealthy, but as long ago as medieval times there was a tradition of identifying and educating boys from modest backgrounds with scholarly potential (albeit this usually depended upon patronage of local nobility or the church) - and some of those sponsored went on to influential positions (such as bishop, university chancellor, or renown scholar) (Hannam, 2010). Social mobility was far from the norm at that time, but it is not an exclusively modern phenomenon.

Similarly, the rather absolute claim that "science subjects were not an established part of the school curriculum at *any level in any country* before the twentieth century" [emphasis added] simply does not stand up to scrutiny. A UK Parliamentary report of 1868, having found existing science provision in British schools uneven, recommended that although "the evidence of schoolmasters goes to show that a great majority of those who were examined have accepted natural science as a part of the school work…more space must be given to the discussion of the place which ought to be assigned, according to the evidence and according to our judgment on it, to natural science" (Schools Inquiry Commission, 1868). The report pointed out how science was already being taken more seriously in schools elsewhere, citing France and Germany. Reid and Ali might be interested to know this report especially recommended experimental physics and chemistry as "the subjects best adapted for teaching…[and which] have generally been preferred in schools, and particularly in those in which natural science teaching has proved successful".

In Greece, a Royal Decree of 1837 required Hellenic (lower secondary) schools to teach three hours per week of natural history in the first two grades, and three hours of physics and geometry in the third grade, and Gymnasia (upper secondary schools) to teach two hours of natural history and physics in each of the first three grades and three hours of natural history, physics and chemistry in the fourth grade; although it has been suggested that it was only "in the period after 1874, that science began to be extensively taught" (Tampakis, 2013, p. 797). In Belgium, "science courses were developed in the curriculum of secondary schools from the 1880s onwards" (Onghena, 2013). Science education in schools in the nineteenth century could be reasonably characterised as patchy and underdeveloped, but it was part of the formal school curriculum in a number of countries. So, again, exaggeration is used rhetorically to make a point, but at the cost of being (strictly) inaccurate.

Textbook authors *do* need to focus their readers on key points, to strip away superfluous detail, to simplify and generalise. That is their pedagogic task. Reid and Ali certainly do this, but in places perhaps err on the side of over-simplification and over-generalisation.

Researching professional practice

Reid and Ali describe one of their target reader groups as "teacher trainers". Those involved in the preparation of school teachers often prefer the term 'teacher *education*' to 'teacher *training*', arguing that teaching is not just a craft that one can be trained-up for, but a professional activity where theoretical perspectives can provide insight into classroom experiences, and where such experiences provide the ground for interpreting and testing theory. That is, the fully professional teacher engages in a critical, questioning, practice through a kind of dialectic that reflects how science itself proceeds through the interplay of theory both guiding, and being informed by, empirical investigations. (Or, if one prefers, empirical work being both motivated by, and motivating, theory.)

Teaching is then not just a craft, but needs to be a professional activity, because learning is always contextualised, and generalisation has to be undertaken very carefully. Anyone who has taught the 'same' course to different classes, especially at school level, will appreciate how teaching can only be planned to a certain extent in advance, as every class is unique, every learner is different, and the lesson or activity that worked so well last year (or even last week) may not automatically go so well with the 'equivalent' class today. This is even more extreme when we look to generalise across different classrooms, different teachers, different schools, different curriculum contexts, and different cultural settings. This is a major challenge for educational research. Even the most carefully carried out study may have little to tell us about what would have been found on another day in another teacher's class. Reid and Ali claim that "it is relatively easy to show that, for every study which shows [a particular strategy] brings improvements, it is possible to find another study where it makes things worse!" (p.9). While it may not be that *easy* to demonstrate this (after all, there are a lot of different strategies that have been studied), I suspect in essence they may be largely right. However, it is easy to over-emphasise this point: there are areas of research which have produced *largely consistent* findings across diverse classrooms, age ranges, topics, and countries. Indeed, contrary to Reid and Ali's reading of the literature, there has been so much reiteration of the effectiveness of many of the teaching techniques designed according to a constructivist perspective on learning that the succession of studies assigning 'control' groups to contrasting 'traditional' teaching conditions, fully expecting to find them inferior to some 'constructivist' approach, has been considered ethically dubious (Taber, 2019).

Notions of replication that may be assumed, even if sometimes over-simplistically (Alger, 2020), in the natural sciences, simply do not apply unproblematically in education (Taber, 2020b). It is sometimes suggested that randomised-controlled trials address this issue, but educational research is seldom resourced sufficiently to allow the sample sizes to make such work viable. In any case, a large-scale experimental study only has the potential to offer guidance on what is most likely on average across a diversity of contexts sampled in the research, which may mean little when, as can happen, the spread of results *within* any condition is wider than any statistically significant difference identified *between* conditions (thus explaining Reid and Ali's complaint). There is not usually the information base available to identify truly representative samples (of teachers, of classrooms) to support such work - and random sampling from target populations of interest (e.g., high school teachers in Western Australia) is seldom feasible. Moreover, there are inherent features of educational studies which produce confounding variables which cannot, in principle, be controlled by research design when double-blind protocols are not possible (i.e., whenever teachers and students are aware they are in an experimental or control condition - which is, surely, usually the case). All in all, the challenges to undertaking experimental research in teaching and learning capable of leading to generalisable results are extensive, and perhaps often even insurmountable (Taber, 2019).

Meaningless comparisons

Reid and Ali cite a study by a researcher, Hussain, who sought to investigate if attitudes to learning chemistry might be improved by developing new teaching materials - a well motivated study in a particular context.[3] A base-line was established,

> with a random sample of 115 (60 boys and 55 girls) aged about 15-16…He then used this major modification with another random group of 400 students (200 boys, 200 girls) of the same age. The group was identical in every other way… (p.283)

This example intrigued me in part because past experience when a journal editor had taught me not to trust claims of 'random selection' unless both the sampled population was clearly defined and some method of randomisation was reported. When I accessed the original thesis cited as the source of this study, I could not find any details of a technique used to select random samples, nor indeed any mention of the samples being random at all (an electronic search of the thesis only found one instance of 'random' as part of the name of a publishing house in the reference list). In the original account (Hussein, 2006) reported:

> To assess the impact of the booklets in action, a total sample of 800 students was selected. There were 400 students (200 boys and 200 girls) from year 10 and 400 students (200 boys and 200 girls) from year 11. The students were aged between 16-17 years and came from a [sic] large typical secondary schools in the United Arab Emirates. (p.103)
>
> With the previous year group, 115 year 10 and 110 year 11 pupils completed the questionnaire. The following year, after using the new teaching materials, 400 year 10 pupils (aged between 15 - 16 years) and 400 for year 11 (aged between 16 - 17 years) were involved. The two successive year groups were similar in other ways, being typical pupils in Emirates schools. (p.141)

This illustrates one of the challenges of making comparisons in educational research: students from 'typical secondary schools' in successive years were compared adopting the (reasonable, but far from certain) assumption that any differences found would relate to the intervention being evaluated. So, Reid and Ali's claim that the intervention group "was identical in every other way" to the base-line group seems to actually mean the two groups were recruited from schools that were subjectively judged typical enough to be considered equivalent. Often, this is the best a researcher can hope for, but this does not *assure* comparability. Experimental method requires randomisation at the level of the unit of analysis (often individual learners) but this is seldom possible working with schools, and comparisons between supposedly equivalent classes, schools, or year groups are often all that can be practically undertaken. Reid and Ali themselves argue that "it is important to realise that test and examination scores can vary quite markedly from year to year in a school for many reasons unrelated to quality of education…" (p.353), and that:

> Learners can vary enormously, even in schools which appear to have similar catchments. There can be high variations between successive year groups in any one school, simply on the grounds of randomness: the laws of probability cannot be broken! Thus, comparing student performances between schools or between successive years in one school are totally meaningless - we cannot avoid the normal randomness of nature. (p.364)

Perhaps Hussein's comparisons were neither "totally meaningless" nor between "identical" groups, but rather comprised the best, imperfect, comparison that was practicable. So, such comparisons need to be read with appropriate caveats, and ideally complemented by the analysis of other kinds of data (such as from in-depth interviews and close observations in the intervention conditions) to see if what the statistics suggest can be triangulated by qualitative data offering rich descriptions of educational processes. Educational phenomena are often messy, and so drawing conclusions from educational research is seldom straight-forward.

Criticisms of educational research

Reid and Ali are very critical of research practices in the field,

> Much educational research
> • Is conducted by those who do not have extensive *'hands-on'* experience of teaching…
> • Is simply descriptive…
> • Relies on 'one-off' projects…
> • Is clothed in unnecessarily complex language…
> • Spends so much time arguing over the trivia of methodologies that it loses focus on the need to gain findings that relate to the real-life of teaching and learning in schools and universities.
> • Is published in journals which are neither accessible (or often meaningful) to practising teachers.
> •Uses questionnaires, focus groups, and interviews, failing to recognise that these only reveal what people THINK: collations of opinion and assertion.
> • Starts with the demand for a 'project proposal' which, by its very nature, more or less pre-determines the outcomes.
> • Fails to measure and, in the end, quantitative evidence is needed to reveal insights into the complex processes that underpin all all [sic] educational provision. (p.6, emphasis in original)

It is certainly true that there is a vast body of published educational research, of variable quality. As with most areas of scholarship today, there are a great many journals competing for submissions, and it is not difficult to get poor work published without a high level of peer review scrutiny, if one is prepared to pay for publication and not too choosy about the outlet. Yet, the more prestigious journals do have rigorous standards of review. Of course, education is rather different from chemistry in the degree of diversity of acceptable perspectives, methodologies, and data collection and analysis methods. In Kuhn's (1959/1977) terms, perhaps educational studies may not be considered a 'mature' field, but it also seems that the complexity of social phenomena mean that particular educational foci can often justifiably and advantageously be studied from diverse perspectives (Taber, 2014). There are scholarly debates about the extent to which pluralism is admissible or valuable in a natural science such as chemistry (Chang, 2012): but in educational studies this is largely taken for granted.

Researching complexity

Teaching and learning is a complex system, and any particular classroom is always embedded in its specific institutional and cultural context (from which it cannot be excised with integrity for laboratory study). So, for example, if a school class were found to largely fail to learn a canonical account of some area of chemistry there could be a range of factors at work. It might be down to an inappropriate curriculum specification of the science for the age range, poor teacher subject knowledge, or the use of inappropriate pedagogy (for example, teaching younger children through long periods of direct lecturing). Or it might relate to common alternative conceptions that students brought to class leading to them misinterpreting teaching.[4] Or it may largely be down to issues such as classroom disruption, or lack of student self-efficacy and low levels of motivation that could primarily derive from issues beyond the classroom (such as social deprivation and inequalities). The 'same' problem could arise from very different causes in different cases, and may commonly reflect a profile of interacting factors. Clearly a range of different perspectives and approaches might be relevant to exploring any one such issue.

One might suggest a contrast here with a Bachelardian model of how concepts that have developed historically are found to be understood in chemistry. Where plurality in the natural sciences can sometimes be considered as the unhelpful existence of fossilised thinking due to the strata of philosophical perspectives through which current canonical understanding has evolved (Bachelard, 1940/1968), plurality in education can be understood as the way teaching and learning are necessarily embedded within layers of context (e.g., classroom, curriculum, institution, language, culture, societal, natural) each of which can be relevant to the same episode.

Reid and Ali rightly raise some good reasons for their readership to be critical readers of research. Yet some of their arguments lack depth. For example, they are right to question the tendency for so much research to be undertaken with questionnaires (which are difficult to design such that they are valid and reliable, and may not transfer unproblematically between research contexts), and there is a tendency to accept self-reports when it is easier to use questionnaires than observe activities. A questionnaire that asks a student how many hours homework they do each week does not offer a direct measure of the *actual* time spent on homework, and asking a teacher whether they tend to teach in a learner-centred or teacher-centred way may elicit an honest response (in that the teacher reports their genuine belief) without being a valid way to characterise what the teacher *actually does* in the classroom. That said, there are many research questions where "reveal[ing] what people THINK" is just what is called for. If you want to find out whether a student understands the redox concept, then a well-designed interview (i.e., to investigate what the student thinks) is likely to be the most sensible choice of instrument (Gilbert, Watts, & Osborne, 1985).

The idea of measurement being more useful than description reflects an ideal of educational enquiry modelled on the physical sciences. Against Reid and Ali, I would suggest there are probably *too many* small-scale studies looking to measure such things as student motivation - based on ratings scales reporting how the class average was perhaps 2.45 on this sub-scale and perhaps 3.12 on that sub-scale - which offer very little useful insight. Such findings may perhaps be compared with published results from some years earlier arising from a sample drawn from a very different population. Or they may be used to make a 'before and after' comparison to gauge the effect of some innovation (creating the problem of knowing how to account for both noise due to limited instrument reliability, and the effect of other simultaneous changes that learners may be experiencing - even such banal matters as the changing seasons outside the classroom). [5]

A good case study that offers a detailed description of a complex phenomenon can sometimes be vastly more informative. I am thinking of work like Petri and Neidderer's (1998) study of a student's shifting thinking about models of the atom, or Duit and colleagues' (Duit, Roth, Komorek, & Withers, 1998) account of a group of students engaging in talk to learn about chaos theory using a magnetic pendulum. One cannot generalise directly from such cases, but *in education* that also tends to be true of most quantitative studies, even when we think we know how to interpret the numerical measures reported. I am certainly not suggesting, contra Reid and Ali, that we should avoid quantitative studies - but I am suggesting, contra Reid and Ali, that different types of study have different strengths and weaknesses, and therefore, in education, we ultimately learn more by using complementary approaches - and qualitative research is often more suited to address particular research questions.

Reid and Ali reject (as "complete nonsense") this argument from complexity as the basis for treating education differently from the natural sciences, and support this evaluation with a comparison of teaching with medicine,

> There are many who argue that all aspects of learning are so complex that they cannot be reduced to numbers. This argument is complete nonsense. The human body is perhaps the most complex system of all and yet medical advances have often depended on the careful use of statistical evidence in making its greatest advances. Medical research depends heavily on statistical analysis to gain insights. Similarly research related to learning must also depend heavily on statistical analyses to gain its insights. (p.387)

As a general point, this is reasonable as far as it goes - even if a careful examination of the modality reveals the paucity of logic here: absolute statements ('all aspects', 'complete nonsense') are justified by conditional arguments ('perhaps', 'often') and what is necessary ('depended upon') but may not be sufficient. The practice of medicine is indeed much like the practice of teaching in that clinical practice can be usefully informed by studies of what is commonly the case.

The clinician is informed by the statistical evidence, but is aware that there is variation between cases, and that a particular diagnosis does not lead to the same prognosis or preferred treatment each time (Alzu'bi, Zhou, & Watzlaf, 2019). This is why medical research contains so many detailed case descriptions as well as statistical studies. As Reid and Ali suggest, "medical research depends heavily on statistical analysis to gain insights", but it *also* depends heavily on careful description in case reports, that is "descriptive studies that are prepared for illustrating novel, unusual, or atypical features identified in patients in medical practice" (Sayre, Toklu, Ye, Mazza, & Yale, 2017). It has been argued that

> Case reports and case series have their own role in the progress of medical science. They permit discovery of new diseases and unexpected effects (adverse or beneficial) as well as the study of mechanisms…have a high sensitivity for detecting novelty and therefore remain one of the cornerstones of medical progress; they provide many new ideas in medicine. (Vandenbroucke, 2001, p. 330)

Research in education, as in medicine, depends upon a toolkit of different instruments and techniques.

The trivia of methodologies

This leads to the comments on the "trivia of methodologies" and "the demand for a 'project proposal' which, by its very nature, more or less pre-determines the outcomes". Reid and Ali's readers are told: "The key things to avoid [include]:

• *Research questions* These close off options and destroy the real nature of research enquiry

• *Research proposals* These reduce research to finding out what is mostly already known…" (p.479, emphasis in original)

The suggestion that details of research design and procedures are trivial seems hard to sustain, and is certainly at odds with expectations in the natural sciences where logical experimental designs, careful techniques and well-calibrated instruments, are generally welcomed as indicators of quality research. I therefore suspect that Reid and Ali's criticism is meant to be aimed at *the amount of space given over to discussing such issues* in educational research papers. Here there is often a stark contrast with what is typical in many natural science fields. However, there is a rationale for this difference. For one thing, where new scientific results are often published in specialist journals, many educational studies are placed in journals with much broader remits. Perhaps even more significantly, as suggested above, educational phenomena are often studied from diverse perspectives and using myriad approaches, unlike in the natural sciences where there is usually a strong programmatic flavour (Lakatos, 1970) to work in any particular field at a given time. So, in the natural sciences, a paper on 'this' topic, placed in 'that' journal, will usually adopt the current canonical conceptualisation of 'the' topic, and will apply the currently accepted procedures, instrumentation, and analytical tools. So, much that is important about the methodological context of the study can then be inferred by those working in the field based on very short descriptions and limited citations to core literature. Authors can take much less for granted about the more promiscuous readerships of educational studies.

Reid and Ali's call for educational research that is *not supported by a careful conceptualisation of existing studies to inform research questions*, and which does *not have a plan for how the research will proceed*, is not convincing. In simplistic terms, there are two basic types of research question and design.[6] One approach has a tightly defined research question along the lines of a hypothesis. Such research does need to be designed in some detail so before research begins is it clear who will be sampled, what measurements will be made (for here, quantification is usually appropriate), how much data will be collected, and how it will be analysed to see if the hypothesis is supported or not. Whilst the hypothesis needs to be motivated, in genuine research it is not known in advance what will be found. This type of research is sometimes called 'confirmatory', but in spirit it might be considered 'confirmatory/disconfirmatory'.

In the other main type of research there is a phenomenon of interest, perhaps in the form of an issue or problem to be addressed, but the researcher is open-minded about how to best frame the enquiry focus given the existing state of knowledge. A research question will then be more exploratory, and different methods will be needed (this is where description is likely to be more useful than measurement as it will not yet be clear what it would be sensible to measure), and the research design will be more fluid, so that it can be refined as the research proceeds.

In my experience, where new researchers are guided to identify a research question (when appropriate, an open-ended one) and to plan how they will seek to answer it, they are better prepared to make progress. Indeed, it is usually expected that the proposal should set out what is potentially novel about the proposed study which makes it much less likely that research following a proposal to address a specified research question will only "find…out what is mostly already known". Reid and Ali are right to recognise that over-restrictive parameters for research are likely to constrain creativity and mitigate against highly novel findings, but that is an argument for ensuring that, in exploratory or 'discovery' research, designs are seen as starting points rather than straight-jackets, and that research questions always offer the potential for finding something unexpected and are not simply routine. Popper's (1953/1989) recommendation to seek bold conjectures is useful advice here.

Why are there so many studies claiming constructivist pedagogy is superior?

Certainly there are 'rhetorical' studies that find what they look for - but this actually tends to reflect poor conceptualisation of the problem or insufficient attention to the design of the study. This can be the case with qualitative studies where supportive evidence is interpreted as indicating a preferred narrative because the research design did not include appropriate safeguards to ensure rigorous sampling and analysis. In these cases interrogation of the methodology would suggest the *implicit* research question that is *in effect* being asked (regardless of any explicit question cited) is merely of the form "can I find some data which can be interpreted in terms of hypothesis/theory/model/perspective X?" which is clearly a very weak question (Taber, 2013a). Anyone observing classrooms can almost certainly find some data that can be construed as evidence for the theories of Piaget, Vygotsky, Bruner, Ausubel, Skinner, Dewey, Fröbel, Rousseau - or just about anyone else who has written about education. Whilst that might make for a useful undergraduate exercise in applying educational ideas, it does not make for quality research.

There are also 'rhetorical' quantitative studies. For example there are a great many studies demonstrating the superiority of what is sometimes called constructivist pedagogy. Reid and Ali suggest that "for every study which shows [e.g., constructivism] brings improvements, it is possible to find another study where it makes things worse!" (p.9). Perhaps they have checked this claim carefully in the educational literature across curriculum areas more widely, but certainly within science education there seems to be a vast excess of studies which find 'constructivist' approaches bring about greater learning when compared with 'traditional' approaches.

This pattern of published findings could simply reflect the superiority of constructivist approaches. [7] The author of this review (wishing to acknowledge his own possible bias) does think that, when understood and applied correctly, constructivist pedagogy is generally effective. However, it is widely considered that there tends to be a publication bias where studies with 'positive' findings are more likely to end up in print (Franco, Malhotra, & Simonovits, 2014) - and that is likely to be important here given that many of the published studies (and so therefore, presumably, many unpublished studies) are set up to 'test' the hypothesis that constructivist teaching will outperform other approaches. This also means that the well-established phenomenon of expectancy effects (Rosenthal & Rubin, 1978), where investigator bias inadvertently influences outcomes, is likely to often be operating. It is the importance of such effects that justifies the additional cost and complexity involved in organising double-blind studies (as used in drug trials) when these are feasible.

Given these factors, we might not be surprised that there are so many studies that show various constructivist teaching techniques are more effective than so-called traditional teaching. However, quite a number of studies reporting the superiority of constructive pedagogy are 'rhetorical experiments' (Taber, 2019). In these studies the investigators actively constrain the comparison condition (e.g., no technology, no discussion, no group-work, no practical work…just a teacher presenting to a class and answering any student questions) to such an extent that it is difficult to consider the 'control' authentically reflects typical classrooms in schools in most countries today. In effect, many of these studies are not actually set up to *test*, but rather to *demonstrate* the hypothesis that constructivist teaching will outperform other approaches.

Simplifying human learning

Reid and Ali's tendency to seek to reduce complex matters to very simple mottoes and claims extends to the way cognition is treated in the book. There is much discussion of WM, but the treatment of long term memory (LTM) is very limited. Knowledge and thinking are not examined critically, but rather are treated as though they can be taken-for-granted.

The critical role of WM

The educational psychologist David Ausubel (1968, p. vi) famously prefaced his textbook on educational psychology by suggesting "if I had to reduce all of educational psychology to just one principle, I would say this: the most important single factor influencing learning is what the learner already knows. Ascertain this and teach him [or her] accordingly". It seems safe to assume that if Reid and Ali wished to reduce "all of" what has been discovered about learning to just one principle - it would be that "the limited capacity of the WM is the controlling factor in all thinking, understanding and problem solving" (p.90).

It is *almost* impossible to over-emphasise the critical importance of WM in explicit learning, and so its significance for teaching. Almost impossible, but Reid and Ali make so many repeated claims about this, that their book gives the impression that WM capacity is not *a* critical issue in learning but *the* only critical issue. The following examples of the many references to WM spread throughout the book gives a flavour of the treatment,

- "The actual process of dissonance occurs in the working memory…" (p.47)
- "Fundamental Principle: Extent of understanding is *controlled* by the limited capacity of working memory. This explains why the teaching strategy is not the key issue." (p.69, emphasis in original)
- "Working memory capacity and the way it controls performance is nothing to do with abilities or intelligence… There is nothing a learner can do to alter his/her genetically-determined working memory capacity." (p.77)
- "Working memory capacity correlates with performance in intelligence tests in just the same way as it correlates with tests in school and university subjects. That does not prove that working memory is simply a measure of intelligence for that deduction assumes that so-called intelligence tests measure intelligence (whatever that is!)" (pp. 79-80)
- "The importance of field dependency cannot be underestimated. It has been found to be a key learner characteristic in enabling success in formal tests and examinations. What has been found consistently across many studies is that those who are more field independent *always* perform much better in the kinds of tests and examinations we use

today…It is helpful to see this learner characteristic in relation to working memory capacity." (p.98, emphasis in original)

- "It is well established that working memory can often be a rate-controlling feature in the way information is processed, understood and accessed. The student with a high working memory capacity will *always* have an advantage when faced with situations when understanding, thinking, and searching long-term memory are involved." (p.107, emphasis in original)
- "There are very good reasons why too many categories [such as response options on a questionnaire scale] will not work. We are back to the limitations of working memory." (p.255)
- "The central importance of limited working memory capacity has been shown repeatedly to be critical in all assessment." (p.467)
- "There is a statement attributed to Confucius (sixth century BC) that says: I hear and I forget. I see and I remember. I do and I understand. This is very misleading and very often NOT true. The evidence shows very clearly that the key to all learning (seen as related to understanding) lies in understanding the limitations of working memory capacity." (p.465, emphasis in original)

Reid and Ali recognise the very important role WM plays in learning, as a limit on the amount of material a person can mentally juggle, 'keep in mind', at any one time. Some of these quotes may sound fatalistic, but Read and Ali make the important point that although the learner cannot do anything to increase their working memory, teaching can be designed to work within the working memory capacity of the learner. This is a very valuable message, but it is accompanied by an unjustified corollary along the lines that because WM is critical, other things may be less important:

> The working memory is the central key for ALL understanding. This is the same for everyone. In essence, we all learn the same fundamental way. The strategy the teacher adopts is NOT the key at all when it comes to understanding although different strategies can achieve other important goals. (p.460, emphasis in original)

The constructivist approach to pedagogy is designed around the idea that (as Ausubel had implied) the starting point for new learning has to be what the learner currently knows / thinks /understands, and that teaching is about making the unfamiliar familiar by finding ways to build on, and sometimes seek to modify, learners' existing conceptual 'resources'. It is not usually feasible to introduce a whole new area of learning all at once, or expect to totally overcome a tenacious existing alternative conception through presenting one argument, demonstration or counter-claim. Teaching for conceptual learning needs to be undertaken iteratively, through modest learning quanta. That is, working within the learner's WM capacity.

One useful strategy that is a key part of constructivist teaching is to use diagnostic assessment at the start of a new topic or concept area to check (a) that learners have the expected prerequisite learning and (b) whether they have acquired unhelpful common alternative conceptions (Driver & Oldham, 1986). This provides the teacher with a guide to the kinds of gains that will be feasible given a learner's WM capacity. Reid and Ali discuss assessment in some depth, but do not have a section on diagnostic assessment. They do discuss the related idea of formative assessment, which has been described as an "approach to improving student learning involving activities undertaken by teachers and students to assess themselves in order to provide feedback for modifying the activities in which they are engaged" (Park, Liu, & Waight, 2017, p. 273). This is usually understood to be undertaken *during* the teaching of a concept area or topic, so that feedback can be acted on during the learning process. (Reid and Ali suggest instead that formative assessment comprises "assessment tasks set at the end of a unit of work", p.175).

Another common constructivist approach is to use such devices as narratives, analogy, metaphor and simile to show how what is being taught is *somewhat like* something already familiar. In discussing Ausubel's distinction between meaningful and rote learning Reid and Ali suggest that "in rote learning, the learner does not possess the relevant prior knowledge in his cognitive structure to link with new information" (p.34), but that is not necessarily the case (Ausubel, 2000). Ausubel pointed out that there needed to be *potentially* relatable material in a learner's cognitive structure to which teaching could be anchored, but also that there had to be a suitable 'learning set'. Meaningful learning does not occur if the learner does not *recognise* what is being presented as related to something they already know about. So, a learner who knows about atoms having 'outer shells of electrons' may not be able to assimilate new related learning if a teacher refers instead to 'the valence shell' unless the learner recognises that this is something she already knows about. Although many of the comparisons used in teaching are quite shallow (the nucleus is like the brain of the cell, energy levels are like the rungs of a ladder, the mechanism of an S_N2 reaction is like an umbrella turning inside out, a transition state is like when a ball rolls over the top of a hill), they often do useful work in making abstract ideas seem familiar, and so making them meaningful to learners when first introduced.

What is thinking?

In chemistry, core terms are often 'proprietary' and are operationalised through scientific education (T. S. Kuhn, 1977) in a way that excludes the uninitiated from professional discourse: oxidation, hybridisation, adduct, nucleophilic substitution, metalloid, hyperconjugation, anti-aromaticity… However, a common issue that arises in educational texts is that professional discourse encompasses a good many terms which are used *as if* technical terms, but which may not have been operationalised as they derive from the 'mental register' of informal everyday concepts that we all take for granted (Taber, 2013b). That is, everyone uses terms such as 'thinking', 'learning', 'remembering', 'forgetting', 'understanding', 'knowledge', and so forth.

One example is 'thinking'. I think (sic) I know what thinking is. When I read Reid and Ali, I suspect they mean something rather different:

> Thinking takes place in the working memory and the capacity of the working memory is very limited. Thus, the extent of thinking is controlled by the capacity of the working memory…Much has been written about thinking but most are [sic] our attempts to analyse thinking into some kinds of categories. However, the working memory just thinks and the categories may be quite artificial. (pp. 227-8)

Leaving aside the question of whether the working memory itself is best said to think rather than to facilitate thinking in the person, or to be the location or source of the thinking process, I do not think thinking (as I use the term) is *restricted* to the working memory. But Reid and Ali clearly do: "The central component is the working memory. This is the part of the brain that *works*: thinks, understands, solves problems. It is the ONLY part of the brain to do this" (p. 459, emphasis in original).

However, this is likely just a difference in the boundary conditions of thinking. If thinking is restricted to that process that we are immediately aware of, that 'stream of consciousness', then perhaps this does always involve the WM. However, if (as would tend to be my own sense of the term) thinking has a broader meaning, encompassing much preconscious processing that takes place out of our awareness, then this is much more widely distributed across brain structures. From this perspective it is not reasonable to suggest that working memory "controls all thinking, understanding, and problem-solving" - especially given that there are many examples where scientists have highlighted the value of incubation of ideas and the roles of the 'unconscious' in leading to problem solutions (Brock, 2015). Whether suggesting that thinking "occurs in the working memory which is limited by its capacity" (p.240) might mislead readers then likely depends on how those readers themselves understand 'thinking'.

For Reid and Ali, however, not only does thinking occur in working memory, but the type of thinking that will be possible depends on working memory capacity,

> scientific thinking is not really accessible until about age 15-16. This fits with the insights from the research of Jean Piaget…By contrast, critical thinking can be developed at a younger age…Systems thinking is complex but may have considerable benefits for university students.
>
> It is interesting that all these findings can be linked back to limited working memory capacity. Without a higher working memory capacity (this has developed by about age 16), scientific thinking is unlikely to be possible. Similarly, systems thinking is probably impossible for those with developing working memories. By contrast, critical thinking does not place such great demands on limited working memory capacity and is possible at younger ages. (p.247)

If teachers were to base their teaching on these suggestions too closely, it could have major implications on the kinds of tasks set for students at different ages - why would we try to teach 'scientific thinking' before age 16 if we believe that the cognitive apparatus needed to carry out this type of thinking will not have matured? Research suggests that scientific thinking starts to develop much earlier than Reid and Ali claim (D. Kuhn & Pearsall, 2000). Whilst there is evidence that working memory capacity does slightly change with age, it also varies from person to person. This is the basis for Reid and Ali's very sensible arguments about how suitably structuring learning and assessment tasks can ensure that otherwise intelligent learners do not fail because they have slightly less WM capacity than their classmates. By the same token, the lesson here should be about structuring classroom activities intended to develop scientific thinking to be within the learners' capacity, rather than to imply that most students under the age of 16 will not be able to engage in it.

A knowledge paradox

Knowledge, like thinking, is one of those terms that is so much a part of everyday discourse that it is easy to assume we agree on what it refers to. A traditional notion along the lines of true, justified, belief (Bhaskar, 1981; Goldman, 1995) would exclude much that is commonly meant when people talk of 'knowing' and 'knowledge' (Aaron, 1971) including so-called implicit or tacit knowledge (Collins, 2010; Polanyi, 1962). Reid and Ali recognise "considerable confusion" in how the word knowledge is understood (p.415). They themselves make a distinction between knowledge and understanding, suggesting "we *cannot* construct knowledge (this is external to us) – we can construct understanding" (p.415, emphasis in original). As has been very extensively documented (in science classrooms at least) what students understand is not always what their teachers intended, so "…the student will probably develop their own idiosyncratic understandings of what is provided. In this, the learner is NOT constructing knowledge" (p.414, emphasis in original). This is why ongoing formative assessment is so valuable to teachers.

This treatment seems to suggest that understanding is a mental phenomena (in the sense of being associated with, found in, minds) whereas knowledge, which "we *cannot* construct" is not a mental phenomena as it "is external to us". This could be understood as linked to a kind of Platonic model where knowledge refers to something like Plato's forms, and so whereas understanding would be located in Popper's (1979, p. 106) second world of "states of consciousness, or of mental states", knowledge itself is one of the 'inmates' of the third world of "objective contents of thought, especially of scientific and poetic thoughts and of works of art" (p.106).

Yet Reid and Ali also inform their readers that "research shows us that knowledge is stored in long term memory in compartments. This is inevitable in that we do all of our understanding in the working memory and this is of very small capacity" (p.53), suggesting that knowledge is the kind of thing that can be located in the brain in the first world "of physical objects or of physical states" (Popper, 1979, p. 106). Reid and Ali are right to suggest that research shows learning often tends to be compartmentalised (for example, according to curriculum subjects) and this has long been recognised as one reason why students successful in formal assessments of classroom science learning may nonetheless often continue to use alternative everyday ideas in informal contexts (Solomon, 1983). However, there is something of a non sequitur here: there is no reason why the "very small capacity" of WM necessarily requires LTM to be compartmentalised. Indeed, the very nature of meaningful learning is that new learning is understood in terms of, and linked to, what has already been represented in LTM. Human beings develop extensive conceptual networks related to topics that interest and engage them, and experts may have highly integrated knowledge structures in their domains of expertise.

There is also something of a mystery here, as if "we *cannot* construct knowledge", then how can it become "stored in long term memory"? One potential response might be that learners *construct* understandings, sometimes "idiosyncratic understandings", but that knowledge is not constructed but rather *transferred* intact. This would seem to be supported by the suggestion that often "teachers, understandably, focus their efforts in transferring as much knowledge as they can to the learners…" (p.469). So, this would mean that knowledge (something that we cannot construct) is sometimes successfully transferred through teaching, but sometimes becomes distorted when the learner misinterprets teaching and constructs their own understandings. Yet that interpretation would be inconsistent with much in *Making Sense of Learning* where effective teaching is seen as supporting learners in constructing understanding, and where overpacked curriculum specifications encourage teachers to seek to directly transfer information instead: something that often results in WM overload. Further, if knowledge is something external to us, and which "we *cannot* construct", it is not all clear how any of us (whether school pupil, teacher, scientist, or textbook author) came to knowledge in the first place. There seems to be an infinite regress here: even if 'transferred', knowledge has to be 'transferred' from someone who already has acquired it.

Long Term Memory is not a store

Whilst Reid and Ali focus extensively on the nature and importance of WM, LTM is given much less attention, and is seen as somewhere where learning can be stored ("the long-term memory is a permanent memory store which seems to be unlimited in capacity", p.53), and from which it can be readily retrieved,

> The actual linking together of ideas happens in working memory which then enables the information to be stored as a chunk…The working memory is always looking [sic] for ways to chunk information. The chunked information can be stored as chunks in long term memory and when brought back into working memory, the information is seen as one chunk. (p.78)

This raises the question of the level of sophistication of understanding of memory processes that is useful for a teacher or other educational professional. It *is* common to consider memory as some kind of store, but this reflects a folk model which is not consistent with current scientific accounts. Memory is a dynamic, and plastic, faculty. A simple model would have WM accessing perceptual data from the senses and accessing memories from some kind of storage. Yet this misrepresents the role of memory in both processes (Taber, 2013b). Past experience influences what we perceive by modifying the structures that process - interpret and select - sensory data, so that the perceptions accessed by WM are already heavily influenced by memory. That is, a modern understanding of memory does not see it as a distinct function of some specific brain structure but distributed across the cortex in both functional and structural terms (Fuster, 1995). Moreover, what are consciously experienced as memories are not simply intact structures pulled out of storage, but *re*constructions based upon accessing what has been represented in LTM. Representations that are activated in this way are modified by the process, both being more readily activated thereafter, and also often actually changed through the process of the recalled material being reconstructed. [8] It is a common experience of school teachers that what seems to be have been successfully understood and learnt by students often later proves to be forgotten or misunderstood. Indeed, research shows how the unconscious processes that take place after material is initially represented ('stored') in memory (Walker & Stickgold, 2004) can completely change what is later remembered to better cohere with preconceptions or other learnt material (Gauld, 1989).

This matters because teachers who are aware of this can adopt teaching approaches and strategies accordingly. Formative assessment can tell the teacher what learners are taking away from today's lesson, but unless new learning is reinforced to support memory consolidation the concepts and principles that seem to be understood today are often likely to become inaccessible or distorted some weeks or months later. So 'drip feed' teaching is indicated, where the teacher keeps returning to taught material, to ask students to revisit, apply, and link it to new examples and material in different contexts at every opportunity. "Understandings developed in the working memory can then be stored in the long-term memory for later use" (p.458) offers only part of the story if we actually want learners to be able to access and apply those understandings later.

I doubt Reid and Ali would actually disagree with that, because as textbook authors they employ the equivalent technique. This is one reason why another 'expert' cannot review a textbook such as this from the perspective of the intended readership. To someone familiar with much of the material in this volume, the book is repetitious in the extreme. Sometimes the same point is made several times with slight variation in wording in the space of a page or so. Key ideas are previewed, presented, reviewed, and then brought out again for revision at every opportunity. This level of repetition may seem excessive, but it is exactly how a school teacher might operate. That is, this may be exactly what a learner new to the ideas being presented needs. Reid and Ali implicitly model this process in their writing, but unfortunately the apparent imperative to set up the WM as the critical factor in all aspects of learning seems to exclude a more explicit treatment. In a similar way, Reid and Ali implicitly employ what is known as dual code theory which suggests that learning may be more effective when the learner is presented with both verbal/textual and imagistic representations of the same material (Cheng & Gilbert, 2013). Reid and Ali supplement their text in many places with simple diagrams that reinforce key messages.

A teaching model of teaching

Teachers of science at school level seldom teach completely authentic canonical science. Curriculum statements are simplifications, and so are textbook treatments, as are the models teachers use to get ideas across to novices. Authors of a textbook about teaching and learning are subject to the same considerations. In order to emphasise and communicate key ideas it is necessary to be selective about content, to simplify complexity, to avoid subtle nuance and confusing detail, to use models that can be readily grasped - to present materials in modest learning quanta that do not exceed the learners' WM capacities. Read and Ali's book is an excellent exemplification of this principle, and of the tensions it creates. In being selective, in keeping things simple, in limiting unfamiliar language, in repeating key points (and then repeating them again, and then later repeating them once more), decisions must be made about the degree of selection and the level of simplification. Their approach to a textbook offers a model for educators of the process of designing teaching that will support learning by keeping things simple and constantly reviewing core principles.

This reviewer found quite a few points where these decisions may have been misjudged on the side of oversimplification and overgeneralisation. Yet this book was not written for someone with extensive experience of classroom teaching, educational research, and teacher development, but rather for those who are approaching much of this material for the first time. I suspect that at times Reid and Ali *have* underestimated their professional readership. However, it is very easy for those already familiar with material to overlook the challenges it may present when still unfamiliar, and it is very difficult for the relative expert to return to the novice learner's perspective. So, the real test of this book is how useful the *intended* readership find it. Whatever its faults, *'Making Sense of Learning'* is an impressive undertaking and should prove incredibly valuable to those for whom it was prepared.

Keith S. Taber

University of Cambridge, kst24@cam.ac.uk

References

Aaron, R. I. (1971). *Knowing and the Function of Reason*. Oxford: Oxford University Press.

Alger, B. E. (2020). Opinion: Is Reproducibility a Crisis for Science? *HPS&ST Newsletter*(February), 9-18.

Alzu'bi, A. A., Zhou, L., & Watzlaf, V. J. M. (2019). Genetic Variations and Precision Medicine. *Perspectives in health information management, 16*(Spring), 1a-1a.

Arthurs, N., Stenhaug, B., Karayev, S., & Piech, C. (2019). *Grades Are Not Normal: Improving Exam Score Models Using the Logit-Normal Distribution*. Paper presented at the International Conference on Educational Data Mining, Montreal, Canada.

Ausubel, D. P. (1968). Educational Psychology: A cognitive view. New York: Holt, Rinehart & Winston.

Ausubel, D. P. (2000). The Acquisition and Retention of Knowledge: a cognitive view. Dordrecht: Kluwer Academic Publishers.

Bachelard, G. (1940/1968). The Philosophy of No: a philosophy of the scientific mind. New York: Orion Press.

Bhaskar, R. (1981). Epistemology. In W. F. Bynum, E. J. Browne, & R. Porter (Eds.), *Macmillan Dictionary of the History of Science* (pp. 128). London: The Macmillan Press.

Brock, R. (2015). Intuition and insight: two concepts that illuminate the tacit in science education. *Studies in Science Education, 51*(2), 127-167. doi:10.1080/03057267.2015.1049843

Browning, E., & Hohenstein, J. (2015). The use of narrative to promote primary school children's understanding of evolution. *Education 3-13, 43*(5), 530-547. doi:10.1080/03004279.2013.837943

Bruner, J. S. (1960). *The Process of Education*. New York: Vintage Books.

Chang, H. (2012). Is Water H2O? Evidence, Realism and Pluralism. Dordrecht: Springer.

Cheng, M. M. W., & Gilbert, J. K. (2013). Students' Visualization of Metallic Bonding and the Malleability of Metals. *International Journal of Science Education*, 1-35. doi:10.1080/09500693.2013.867089

Chesky, N. Z., & Wolfmeyer, M. R. (2015). Philosophy of STEM Education: A critical investigation. New York: Palgrave Macmillan.

Collins, H. (2010). *Tacit and Explicit Knowledge*. Chicago: The University of Chicago Press.

Driver, R., & Oldham, V. (1986). A constructivist approach to curriculum development in science. *Studies in Science Education, 13*, 105-122.

Duit, R., Roth, W.-M., Komorek, M., & Withers, J. (1998). Conceptual change cum discourse analysis to understand cognition in a unit on chaotic systems: towards an integrative perspective on learning in science. *International Journal of Science Education, 20*(9), 1059-1073.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345*(6203), 1502-1505. doi:10.1126/science.1255484

Fröbel, F. (1827/2012). *The Education of Man*. Mamphis: General Books.

Fuster, J. M. (1995). Memory in the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate. Cambridge, Massachusetts: The MIT Press.

Gauld, C. (1989). A study of pupils' responses to empirical evidence. In R. Millar (Ed.), *Doing Science: images of science in science education* (pp. 62-82). London: The Falmer Press.

Gilbert, J. K., Watts, D. M., & Osborne, R. J. (1985). Eliciting student views using an interview-about-instances technique. In L. H. T. West & A. L. Pines (Eds.), *Cognitive Structure and Conceptual Change* (pp. 11-27). London: Academic Press.

Goldman, A. (1995). Knowledge. In T. Honderich (Ed.), *The Oxford Companion to Philosophy* (pp. 447-448). Oxford: Oxford University Press.

Halpern, D. F. (2000). Thinking critically about critical thinking: lessons from cognitive psychology. In S. L. Riedel, R. A. Morath, & P. T. McGonigle (Eds.), *Training Critical Thinking Skills for Battle Command: ARI Workshop Proceedings* (pp. 22-32). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Hannam, J. (2010). God's Philosophers. How the medieval world laid the foundations of modern science. London: Icon Books.

Hansen, K.-H., & Olson, J. (1996). How teachers construe curriculum integration: the Science, Technology, Society (sts) movement as Bildung. *Journal of Curriculum Studies, 28*(6), 669-682. doi:10.1080/0022027980280603

Hussein, F. K. A. (2006). Exploring attitudes and difficulties in school chemistry in the Emirates. (Ph.D. Ph.D.), University of Glasgow, Glasgow.

Johnstone, A. H. (2006). Chemical education research in Glasgow in perspective. *Chemistry Education Research and Practice, 7*(2), 49-63. doi:10.1039/B5RP90021B

Kind, V. (2009). Pedagogical content knowledge in science education: perspectives and potential for progress. *Studies in Science Education, 45*(2), 169-204. doi:10.1080/03057260903142285

Kuhn, D., & Pearsall, S. (2000). Developmental Origins of Scientific Thinking. *Journal of Cognition and Development, 1*(1), 113-129. doi: 10.1207/S15327647JCD0101N_11

Kuhn, T. S. (1959/1977). The essential tension: tradition and innovation in scientific research. In T. S. Kuhn (Ed.), *The Essential Tension: Selected studies in scientific tradition and change* (pp. 225-239). Chicago: University of Chicago Press.

Kuhn, T. S. (Ed.) (1977). The Essential Tension: selected studies in scientific tradition and change. Chicago: University of Chicago Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrove (Eds.), *Criticism and the Growth of Knowledge* (pp. 91-196). Cambridge: Cambridge University Press.

Onghena, S. (2013). A Blend of Romanism and Germanism: Experimental Science Instruction in Belgian State Secondary Education, 1880–1914. *Science & Education, 22*(4), 807-825. doi:10.1007/s11191-012-9512-2

Park, M., Liu, X., & Waight, N. (2017). Development of the Connected Chemistry as Formative Assessment Pedagogy for High School Chemistry Teaching. *Journal of Chemical Education, 94*(3), 273-281. doi:10.1021/acs.jchemed.6b00299

Petri, J., & Niedderer, H. (1998). A learning pathway in high-school level quantum atomic physics. *International Journal of Science Education, 20*(9), 1075-1088.

Piaget, J. (1929/1973). *The Child's Conception of The World* (J. Tomlinson & A. Tomlinson, Trans.). St. Albans: Granada.

Piaget, J. (1970/1972). *The Principles of Genetic Epistemology* (W. Mays, Trans.). London: Routledge & Kegan Paul.

Polanyi, M. (1962). *Personal Knowledge: Towards a post-critical philosophy* (Corrected version ed.). Chicago: University of Chicago Press.

Popper, K. R. (1953/1989). Science: Conjectures and refutations *Conjectures and Refutations. The Growth of Scientific Knowledge* (pp. 33-65). London: Routledge.

Popper, K. R. (1979). *Objective Knowledge: an evolutionary approach* (Revised ed.). Oxford: Oxford University Press.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: the first 345 studies. *Behavioral and Brain Sciences, 1*, 377-386. doi:10.1017/S0140525X00075506

Sayre, J. W., Toklu, H. Z., Ye, F., Mazza, J., & Yale, S. (2017). Case Reports, Case Series - From Clinical Practice to Evidence-Based Medicine in Graduate Medical Education. *Cureus, 9*(8), e1546-e1546. doi:10.7759/cureus.1546

Schools Inquiry Commission. (1868). Report of the Commissioners [a.k.a. The Taunton Report]. London: H. M. Stationary Office.

Solomon, J. (1983). Learning about energy: how pupils think in two domains. *European Journal of Science Education, 5*(1), 49-59. doi: 10.1080/0140528830050105

Taber, K. S. (2013a). Classroom-based Research and Evidence-based Practice: An introduction (2nd ed.). London: Sage.

Taber, K. S. (2013b). Modelling Learners and Learning in Science Education: Developing representations of concepts, conceptual structure and conceptual change to inform teaching and research. Dordrecht: Springer.

Taber, K. S. (2014). Methodological issues in science education research: a perspective from the philosophy of science. In M. R. Matthews (Ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (Vol. 3, pp. 1839-1893). Dordrecht: Springer Netherlands.

Taber, K. S. (2019). Experimental research into teaching innovations: responding to methodological and ethical challenges. *Studies in Science Education, 55*(1), 69-119. doi:10.1080/03057267.2019.1658058

Taber, K. S. (2020a). Foundations for Teaching Chemistry: Chemical knowledge for teaching. Abingdon, Oxon.: Routledge.

Taber, K. S. (2020b). Is reproducibility a realistic norm for scientific research into teaching? *HPS&ST Newsletter*(April), 13-23.

Tampakis, K. (2013). Science Education and the Emergence of the Specialized Scientist in Nineteenth Century Greece. *Science & Education, 22*(4), 789-805. doi:10.1007/s11191-012-9538-5

Vandenbroucke, J. P. (2001). In Defense of Case Reports and Case Series. *Annals of Internal Medicine, 134*(4), 330-334. doi: 10.7326/0003-4819-134-4-200102200-00017 %m 11182844

Vygotsky, L. S. (1978). Mind in Society: The development of higher psychological processes. Cambridge, Massachusetts: Harvard University Press.

Walker, M. P., & Stickgold, R. (2004). Sleep-Dependent Learning and Memory Consolidation. *Neuron, 44*(1), 121-133. doi:10.1016/j.neuron.2004.08.031

Notes:

[1] Kuhn's original essential tension was between following tradition and introducing novelty in the practice of science itself.

[2] This was one aspect of the introduction of the English National Curriculum at the start of the 1990s. One consequence was that engineering and physics graduates who might have wished to prepare to teach physics and mathematics in schools found that they had to choose to train only as mathematics teachers or alternatively only as science teachers and be prepared to teach some chemistry and biology.

[3] It should be pointed out that this was only one aspect of a more substantial study. Student performance with the new teaching materials was found to be significantly greater than with the standard materials: this was tested by dividing the 400 students in each year into two groups, each of which studied a different topic with the new materials so that any differences between the groups would cancel out across the study as a whole.

[4] There is an extensive literature reporting many of the common alternative conceptions that tend to be found among students of different age ranges in various science topics, or as Reid and Ali report "endless studies, each of which sought to identify some topic or theme where learners show misconceptions or alternative conceptions" (p.9), and key findings have been published in articles in teacher professional journals (not just academic journals) and a number of popular books written for teachers. This material helps alert teachers to the kinds of alternative ways of thinking that may be found among their pupils (Key Stage 3 National Strategy, 2002) and is often included in teacher preparation and development courses, and curriculum support materials. This area of research is also informing the work into 'learning progressions' (Alonzo & Gotwals, 2012) as part of curriculum development.

[5] At an anecdotal level, my experience working with science graduates undertaking classroom research as part of a teacher preparation course left me with the strong impression that many tend to, like Reid and Ali, assume that obtaining some kind of quantitive measurement inherently offers a superior approach. This seemed to derive from previously working with natural phenomena, often under laboratory conditions, when using well-established and carefully calibrated, precise and reliable, instrumentation (https://science-education-research.com/publications/miscellaneous/why-do-natural-scientists-tend-to-make-poor-social-scientists/).

[6] This is a gross simplification, but nonetheless has considerable heuristic value. This distinction is commonly made in educational research texts, where it can act as a useful starting point for conceptualising research philosophies and strategies, as long as it is presented as a pedagogic model intended to *introduce* the topic rather than comprehensively cover the complexity and nuances of all cases (Taber, 2013a). This might be understood as similar to introducing elements as being metals or non-metals, or compounds as having ionic or covalent bonding - which can be useful as long as such a dichotomy is not presented as an absolute distinction, but rather as a general model which is often approximately correct - a kind of first order approximation that offers a useful jumping-off point for later development.

[7] A difficulty here is that different authors use different terms and make different judgments about what are included as examples of classes of pedagogic approaches. For example, Kirschner and colleagues (Kirschner, Sweller, & Clark, 2006) published a much cited paper arguing that approaches such as constructivism were inferior to what they called 'direct' teaching. However, they identified constructivist teaching with minimal guidance from teachers, which is completely inconsistent with constructivist pedagogy as understood by those espousing constructivist teaching (Taber, 2011), and their notion of direct instruction is a dialogic form of teaching that many would consider a constructivist approach. Whilst there is a very important debate here, it is necessary to examine the nature of the different forms of pedagogy being evaluated (where there is quite wide agreement on what works) rather than just the labels used.

[8] If it is assumed (as currently believed) that memories are somehow facilitated by 'traces' in the physical structure of the brain, then there is clearly a major ontological difference between (a) initial experience that is later recalled, and the experience of remembering (i.e., subjective experiences); and (b) the physical traces that enable the later remembering (Taber, 2013b). Therefore this cannot be a matter of simply moving or copying something from consciousness to memory and back again.