detail about the likely scale of imperfection. Having an honest understanding of what can be achieved is important if we are to ensure that the demands placed upon an assessment system are realistic.

References

Benton, T. (2017, November). *Pooling the totality of our data resources to maintain standards in the face of changing cohorts.* Paper presented at the 18th annual AEA-Europe conference, Prague, Czech Republic.

Benton, T. (2018, November). *Exploring the relationship between optimal methods of item scoring and selection and predictive validity*. Paper presented at the 19th annual conference of the Association for Educational Assessment in Europe, Arnhem-Nijmegen, The Netherlands.

Benton, T., & Gallagher, T. (2018). Is comparative judgement just a quick form of multiple marking? *Research Matters: A Cambridge Assessment publication, 26*, 22–28.

Black, B., Suto, I., & Bramley, T. (2011). The interrelations of features of questions, mark schemes and examinee responses and their impact upon marker agreement, *Assessment in Education: Principles, Policy & Practice, 18*(3), 295–318.

Bramley, T. (2008, September). *Mark scheme features associated with different levels of marker agreement.* Paper presented at the annual conference of the British Educational Research Association, September 3–6, in Edinburgh, UK.

Clark, L. (2008, May 15). Pupils should not expect to receive 'accurate' results from exams, warns new watchdog. *Mail Online*. Retrieved 29th May, 2019, from http://www.dailymail.co.uk/news/article-1019938/Pupils-expect-receive-accurate-results exams-warms-watchdog.

Ofqual. (2016). *Marking consistency metrics*. Ofqual report, Ofqual/16/6121. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_da a/file/681625/Marking_consistency_metrics_-_November_2016.pdf.

Ofqual. (2018a). *Marking consistency metrics: An update*. Ofqual report, Ofqual/18/6449/2. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_da a/file/759207/Marking_consistency_metrics_-_an_update_-_FINAL64492.pdf.

Ofqual. (2018b). *A study of hard-to-mark responses: Why is there low mark agreement on some responses?* Ofqual report, Ofqual/18/6449/5. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_da a/file/759216/HardtoMark_-_FINAL64495.pdf.

Suto, I., Nádas, R., & Bell, J. (2011). Who should mark what? A study of factors affecting marking accuracy in a biology examination. *Research Papers in Education, 26*(1), 21–51.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S. Fourth Edition*. New York: Springer.

"Learning Progressions": A historical and theoretical discussion

Tom Gallacher GL Assessment[1] and **Martin Johnson** Research Division

Introduction

Learning Progressions (LPs) are a relatively recent approach that aim to support three aspects of education: teaching and learning, assessment, and curriculum design. According to Schmidt, Wang, and McKnight (2005) the effectiveness of these three aspects of education may be increased by better coherence, and the LP approach claims to improve coherence by providing frameworks of knowledge and skills called "LP models". These frameworks describe the progression that can be expected of learners through their education (Gotwals & Songer, 2013). LP approaches are popular and influential across the fields of education and curriculum development, with discussion being carried out across a number of international contexts (Australian Council for Educational Research, 2018; E. M. Kim, Haberstroh, Peters, Howell, & Nabors Oláh, 2017; H. Kim & Care, 2018). This suggests that the consideration of the approach is topical.

This article outlines the specific objectives of the LP approach, the mechanism by which LP models may attain these objectives, and finally, the likelihood of this attainment (based on previous evidence). LPs should only be expected to achieve their aims if the assumptions of the LP approach are correct; however, our view is that the evidence suggests that the assumptions embedded within the frameworks are overly simplistic. Education is complex and the implementation of the LP approach to teaching and learning, assessment, or curriculum design may have unintended consequences when implemented without consideration of other possible approaches.

Proponents of the LP approach display a minimal engagement with previous theories of learning, and their ideas have been criticised as being "the latest manifestation of a much older idea, that of regularity in the development of students as they learn a certain body of knowledge or professional practice" (Wilson, 2009, p.716). This suggests that LP proponents should also consider the similarities of their theory with previous work to derive an approach that is most likely to attain its desired objectives.

Objectives of LPs

In order for LPs to benefit teaching and learning, assessment, and curriculum design, the approach needs to have a theory of learning that satisfies the practical and theoretical demands of the professionals involved in all three areas (Black, Wilson, & Yao, 2011).

1. The work was carried out when the first author was a member of the Research Division.

Moreover, this central theory needs to be robust to the criticism of evidence so that it can satisfy its objectives of explaining important phenomena. As stated, the LP approach aims to improve teaching and learning, assessment, and curriculum design, by providing frameworks that model the process of change that learners go through when engaging in education. These frameworks cut across the three aspects, and are claimed to have benefits to each independently:

Firstly, in order to benefit teaching and learning, the LP framework aims to provide detailed instruction on the optimal order for presenting material within a subject. This structure can then support lesson planning, helping teachers to track student progress and identify actions that support the learners' learning (Alonzo & Gearhart, 2006, p.100).

Secondly, to support assessment, the LP approach aims to provide a framework for comparing different learners in order for the results of such comparisons to be useful for learners (Catley, Lehrer, & Reiser, 2005). This framework would also provide a validity argument for assessments (Gotwals & Songer, 2013). In addition, changing the emphasis of assessments so that they are demonstrations of problem solving that correspond to the way that an expert behaves (called "learning performances"), would provide rich and useful information on the abilities of learners (Coppola, 2006).

Thirdly, to support curriculum design, the LP approach aims to provide a method of refining the material presented to learners (Corcoran, Mosher, & Rogat, 2009). By empirical observation and research, a curriculum may be optimised to enable learners to derive the best possible education (Smith, Wiser, Anderson, & Krajcik, 2006).

The commonalities between these three areas are collectively referred to as the theory of learning that is prescribed by the LP approach (Black et al., 2011). It is this set of assumptions that can be tested against the evidence already accumulated within fields of teaching and learning, assessment, and curriculum. If the theory of learning is not contradicted by the previous findings, then we can presume that the theory works well enough to suggest that the implementation of an LP approach would be useful for achieving the stated aims. If the theory cannot account for previous findings, or worse, predicts the opposite, then we can conclude that the theory developers would benefit from more engagement with prior literature.

In our critique of the core issues around the LP approach, our intention is to contribute to the debate around conceptualisations of learning progressions and to suggest that it is also important to look at other areas of curriculum theory for insight.

An outline of the LP approach's theory of learning

As mentioned, the theory of learning that is prescribed by the LP approach underpins the three aims of the approach since it is this theory that allows the coherence between the three areas. The LP approach can broadly be summarised into four points:

- LP models are domain-content specific. Subjects like Science or Mathematics have distinct ways of thinking and distinct bodies of material that need to be taught to be understood (e.g., Smith et al., 2006). This means that there are central concepts and principles of a discipline, which ties the area to the notion of coherent "big ideas" (e.g., see Harlen, 2010).
- LP models incorporate knowledge and practice. Learners learn (and demonstrate through assessment) the "what" and the "how" of a subject domain-content (e.g., Lehrer & Schauble, 2015; Wilson, 2009).
- LP models are successive and progressive. If learners fail to master a particular first thing, they are not able to do a specific second or third thing (e.g., National Research Council, 2007).
- LP models are based on research about what learners can do at different ages and stages of progression (e.g., Duncan & Hmelo-Silver, 2009).

At the simplest level, therefore, LP theory analogises learning a subject to be like climbing a ladder. Climbing each rung is dependent on climbing the previous rung, and it allows the climber (learner) a better view of the subject. Therefore, the main priority for LP developers is to design the ladder appropriately.

For illustrative purposes, a simple LP model of two ideas ("matter" and "colour") is conceptualised (by the authors) and presented in Figure 1. "Naïve understanding" is taken as the starting point, so these learning progressions start with learners who have no prior empirical or "scientific" understanding (in a conventional sense) of matter or colour, but might have a variety of views about what matter and colour are.

In this article we argue that the LP theory of learning is made up of four simplifications, which we can unpick to begin evaluating whether the adoption of the LP approach at the expense of other approaches will meet the aims above. A lot of the arguments presented have not been addressed since the advent of previous theories of hierarchy development (Phillips & Kelly, 1975).

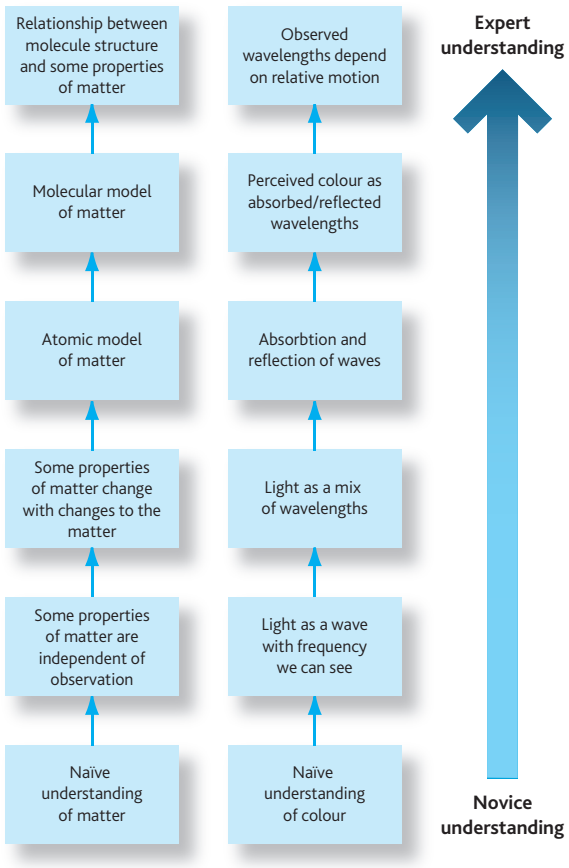

Figure 1: Illustrative Learning Progressions model of matter and colour (authors' own example)[2]

2. Please note that that the progressions presented here are open to discussion, for example, see Taber (2000) and his commentary on the "pedagogic pitfalls of the atomic ontology".

Simplification 1: LPs include a mix of cognitive processes

The first simplification inherent in the theory of learning proposed by the LP approach stems from the second point above: that knowledge and practice should be combined or bundled into a unit of "concept". Moreover, a learner's grasp of this concept is a part of "the developmental pathway in which students' understanding…become[s] more sophisticated over a long period of time" (Paik, Song, Kim, & Ha, 2017, p.4965). In the analogy of a ladder, these bundled concepts comprise the rungs of the ladder, since they are on the same level. Despite the legitimate concern that models can lack the sophistication to describe complex realities (e.g., Goldstein, 1998), some theorists have employed the metaphor of a ladder to exemplify the learning process (e.g., Hess, 2008; Masters & Forster, 1996; Vorst, 2018). Our concern is that the conflation of knowledge and practice into a concept may lead to an insecure inference about what a learner "knows" based on their performance.

Catley et al. (2005) are very explicit about how concepts implicate the bundling of different activities together, stating that "we represent this blend of knowledge, skills and forms of activity that support the development of knowledge and skill as learning performances" (2005, p.5). Other authors are less explicit, but make some reference to differentiating cognitive processes, since all parts of understanding are "enacted" by a "learning performance" with the material (Smith et al., 2006).

Downplaying the differences in cognitive processes into one single unit allows a potentially problematic assumption about what learners can and cannot do. Any successful performance with learning materials can be taken to indicate, according to the LP ladder analogy, that the learners can demonstrate successful performances when the material is presented in different ways. This is because the learner is assumed to have gained understanding (have climbed that rung of the ladder). This assumption, as will be seen, is not always true or useful for educators in practice.

Some LP authors seem to have a preference for some cognitive processes over others, such that declarative memory recall is negatively contrasted by Smith et al. (2006, p.93) with "important aspects…of understanding and reasoning." Other authors are less explicit, except that by emphasising the development of problem-solving skills, there is little mention of developing the knowledge required by the beginner levels (Messick, 1984, p.216) where "in beginning or low-level achievement a major issue is the acquisition of a critical mass of information on the subject," with more advanced levels reflecting more complex cognitive processes.

Simplification 1 Evidence

Firstly, the claim that different processes can be effectively bundled together is considered a simplification given evidence from how the development of different processes happen at different points, and are likely to happen in cycles (Fischer, 2008). The specific cognitive development of learners might enable them to perform some tasks with the material, but not all, while failure at a task might be due to the failure of several different cognitive processes. This makes it difficult for teachers to identify how to help different learners who have a range of successes at different tasks. If learners are not consistently able to demonstrate a "concept" then the ladder may be a dangerous analogy on which to base decisions.

Within alternative theories of learning, memory is typically specified as crucial to "higher" application of knowledge, and so in Bloom's Taxonomy of Educational Objectives (Bloom, 1956) the cognitive processes are shown as a hierarchy or network with memory at the bottom. More recently revised versions of the taxonomy, such as by Anderson and Krathwohl (2001) and Webb (1997, 1999), retain a distinction between memory and problem solving. While recognising the foundational nature of memory for "higher" processes, years of teacher observation data suggests that memory, although insufficient of itself, is a necessary requirement for higher cognitive functions.

Studies of experts and non-experts show crucial differences in how memory – rather than understanding – is changed by learning (e.g., National Research Council, 2000, 2001), and a theory of learning that downplays these changes will not be able to account for such evidence. For these reasons, if LPs were to be implemented as a system of learning, we would expect insufficient consideration of the different cognitive processes that support learning, and therefore that the implementation might provide ineffective education. When "Assessment without levels" was introduced in England, the system "encouraged undue pace and progression onto more difficult work while pupils still had gaps in their knowledge or understanding" (Department for Education, 2015, p.17).

This simplification has focused on the rungs of the ladder, and is essential to understanding the next two simplifications, as will be seen.

Simplification 2: Hierarchies of concepts

The second simplification inherent in the theory stems from the third "successive and progressive" point, such that within a subject domain there is a hierarchy of "understandings" that proceed over the course of learning a subject. In the analogy of the ladder, some rungs are higher than others, with each successive rung being higher than the same set of previous rungs, plus one more. Within each LP, some concepts are more advanced, and are therefore closer to the concepts of experts within that field.

The purpose of education within the LP approach is to bring the understandings of non-experts closer to that of experts (Duschl, 2006; Lehrer & Schauble, 2015). In addition, the role of teachers is to mediate the material and to scaffold the learning so that learners are brought closer to the end goal of the LP (Duncan & Hmelo-Silver, 2009).

This simplification is an attractive one, since it implies the simple progression derived from learning (Fensham, 1994). Once a leaner has progressed beyond a stage of learning, their understanding of a subject is closer to that of an expert, and therefore they are able to solve more problems than before and are ready for the next stage. It has been noted that high performing educational jurisdictions incorporate such a process (Valverde & Schmidt, 1998), which might support a conclusion that such a process is useful for learning. However, this model assumes that subject experts have a monolithic set of concepts to be worked towards, that might not be applicable to all subject areas (Bernstein, 1999), while teaching such a view might damage the process of later learning (Efland, 1995).

Simplification 2 Evidence

The evidence for hierarchies of knowledge and skills is mixed. Gagné (1968) reviews the evidence to support the idea of "Learning Sets", which supports a theory of learning that claims that the optimal ordering of material can be found empirically. Like LPs, the theory is implicitly Vygotskian, in that the main determinant for whether a material can be learnt is the prior learning and knowledge, rather than any formal stage of cognitive development. For illustrative purposes, a Learning Sets curriculum of matter and colour that builds toward one idea is presented in Figure 2. Notice that Learning Sets allows connections between parallel curricula, allowing the possibility of a network analogy, rather than the ladder prescribed by Learning Progressions.

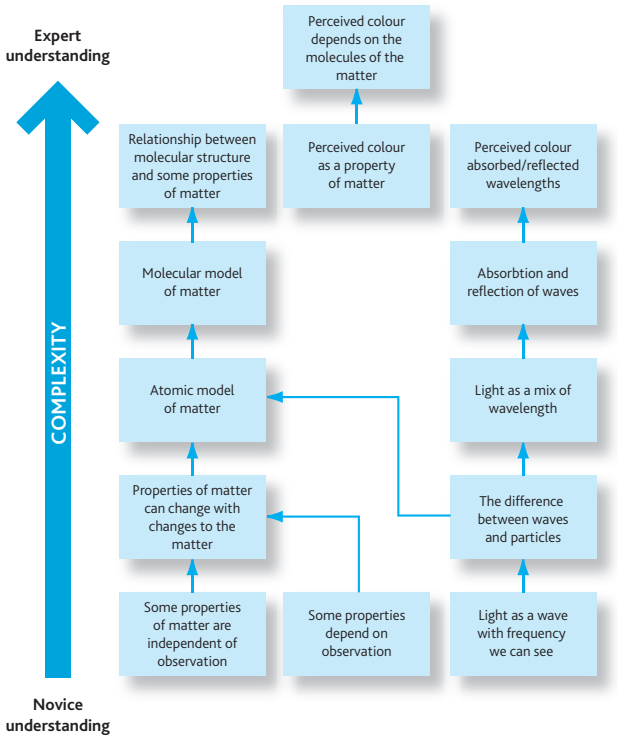

Figure 2: Illustrative Learning Sets model of matter and colour (authors' own example)

Gagné & Bassler (1963) found that the forgetting of subordinate learning sets may occur independently of, and without effect upon, retention of the total task which has been achieved through learning. This forgetting has implications for assessment if discrete task performance at a particular time is taken as a signifier of learning or ability. This undermining of the theory was also confirmed by Kolb (1967), particularly for declarative knowledge. This suggests that although an optimal method of presenting material could theoretically exist, the order will not determine the retention of material by the learner, undermining the analogy of a ladder since rungs are not retained in the order they were climbed. More recent evidence goes one step further to show that intermediate learners are dependent on the context and presentation of problems to guide how they apply the skills that they have learnt (Bao, Hogg, & Zollman, 2002). This inconsistency of applying a skill or knowledge generalises across different cognitive activities from declarative memory recall to problem solving beyond secondary education. Future skill development has been argued to *require* inconsistency as part of the process of consolidating learning (Fischer, 2008). This suggests that the trajectory of learning is less like a ladder, but more "three steps forward, one step backwards", and suggests that despite its use a s a model for learning (e.g., Hess, 2008; Masters & Forster, 1996; Vorst, 2018), the ladder analogy is inadequate to describe the complexity of the learning process.

Applications of the LP approach that adhere to the ladder-progression analogy are likely to be problematic if they do not consider the contexts in which different learners can or cannot demonstrate a technique appropriately. Such a problem has been observed in England where "teachers planned lessons which would allow pupils to learn or demonstrate the requirements for specific levels. This encouraged teachers to design and use only classroom assessments that would report a level outcome. As a result, formative classroom assessment was not always being used as an integral part of effective teaching" (Department for Education, 2015, p.13). It may be that the conclusion from Valverde and Schmidt (1998) failed to identify other differences between jurisdictions that contributed to the observed high performances in those jurisdictions.

Another problematic issue that may pertain to the ladder-progression analogy relates to learner equity. Lehrer and Schauble (2015) note that conforming to generalised learning models may restrict the landscape of possibilities and deprive students the opportunity of (a) encountering concepts that have traditionally been considered too difficult to learn, but which can be made accessible through appropriate teaching (White & Frederiksen, 1998); or (b) lead teachers to fail to consider that some ideas that are presumed to be self-evident may turn out to be more challenging when encountered from a student's perspective (Sandoval & Millwood, 2005).

Simplification 3: Assessment of progression

The third simplification rests on the previous two simplifications, and states that the stage of a learner's progression towards a goal can be assessed reliably, and reported to learners and teachers in order to support the overall education process (Steedle & Shavelson, 2009). Results from assessments are taken to be a reflection of the concepts grasped by a learner (how far up the ladder they have climbed), and their degree of expertise. Learners are positioned not relative to each other, but are located on the ladder of progression (Corcoran et al., 2009).

Simplification 3 Evidence

The idea that progression through levels can be reliably assessed is a simplification since learners can inconsistently demonstrate a range of abilities that do not support a hierarchy based solely on conceptual difficulty. Hart (1981), as cited in Simons and Porter (2015), shows examples of students who can demonstrate an ability when asked one way, but not another. For example, in the case of the conceptual difficulty of "knowledge of fractions", 90 per cent of students can respond that 5/7 is greater than 3/7, but only 15 per cent can respond that 5/7 is greater than 5/9. This difference exists despite the assumed conceptual commonality of denominator and numerator knowledge that underpins fraction knowledge. Differences have also been found between students' performances on the same skills depending on whether they were assessed by a class test or by an individual interview (Denvir & Brown, 1987, p.106).

The same type of conclusion was reached by proponents of LP assessments. Graf and van Rijn (2016) report that the likelihood of a learner successfully completing a task related to three things: 1) degree of progression through the learning pathway; 2) non-progression related complexity, such as computation (i.e., systematic sources of difficulty that covary with the levels of the progression but which are not specified by the learning pathway through conceptual complexity); and 3) sources of difficulty which are not related to the levels of the progression, such as reading demands. Alonzo and Steedle (2009) observed that students' responses were only 60 per cent consistent within one level of a learning progression, with some of the rest of the variance being explained by features of the items.

This evidence also goes some way to discredit the first simplification: that learners routinely solve "easy questions" on "advanced concepts" and fail "hard questions" on "basic concepts". This means that not all parts of a concept are a single unit, and that teaching and assessment need multiple dimensions to understand variance in performance rather than the unidimensional ladder analogy. Feedback which reflects a learner's inconsistencies would arguably be more useful for planning future educational activities required by educators.

Attempts to make grades from summative assessments in England reflect the trajectories of progression came to the same conclusion in the 1980s: "The larger obstacle appeared in the distribution of performances of students. Each set of criteria of this type had to assume a model of students' progress…and students' performances did not fit these models" (Black & Wiliam, 2002, p.25). Similar concerns have also been voiced more recently over the model of progression implied by "Single Level Tests" (Whetton, 2008) and National Tests (Oates, 2011).

Simplification 4: Big Ideas

A fourth simplification within the LP theory of learning is the generalisation of learning from specific "big ideas" to the breadth of a subject area. The claim here is that learners who gain knowledge and skills from one area of a subject are able to apply these to untaught areas, if the original area is fundamental enough to that subject.

This claim is most relevant for curriculum designers, who may wish learners to achieve a wide range of objectives in a fixed time. The LP approach argues that learners who master the concepts of a specific big idea are better placed to answer new questions from an unrelated sub-area than are learners who master a breadth of areas to a lesser depth. What is and what is not a "big idea" is not obvious *a priori*, but criteria may include that the teaching of the idea should facilitate understanding of current issues, be satisfying to learn, and have cultural significance (Harlen, 2010, p.19). Although the idea of a "big idea" is found elsewhere (Bruner, 1960, p.18), it is something that is hard to falsify, since evidence of "no transfer" could be taken as evidence that the taught idea was not "*big*" enough.

Bruner (1960) advocated a curriculum where topics are revisited at intervals, with different ways of presenting the topic. The theory takes an explicitly Piagetian view that learners go through stages of representation of ideas, from enactive to iconic to symbolic. The challenge of education, according to the Piagetian principles embodied in such "spiral curricula", is to present ideas in ways that correspond to the developmental stage of the learners (Bruner, 1960, p.39; Efland, 1995). When learners are ready for the next stage of representation, educators need to present the same topic correspondingly, which thereby allows a more complex understanding of the topic for the learner. This idea echoes the thinking of LPs, but with an explicit theoretical understanding of knowledge development. For illustrative purposes, the LPs for matter and colour that were used in previous figures are produced within a spiral curriculum framework (Figure 3). Here, the connection between the nodes is unimportant because each node is *qualitatively* distinct from those coming before and after.

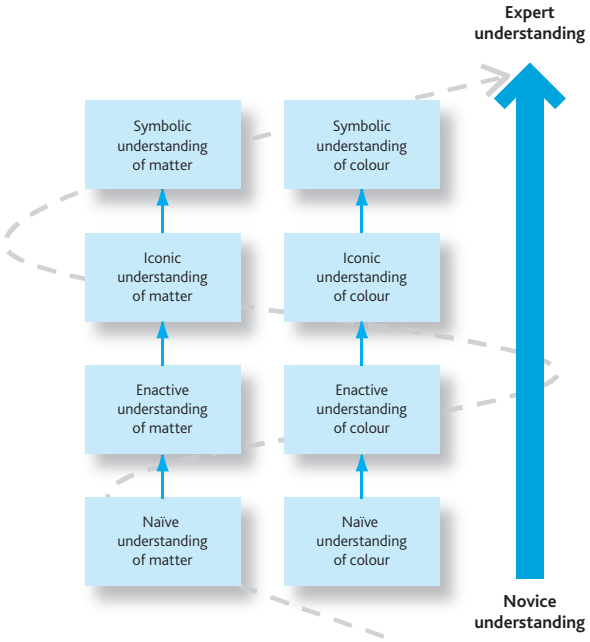

Figure 3: Illustrative Spiral Curricula model of matter and colour (authors' own example)

The concept of "big ideas" in the modern sense arose from studies of experts' thinking (National Research Council, 2007, p.37), where it describes the ways that experts group problems that they have seen and how they identify new problems in relation with these problem groups (Chi, Feltovich, & Glaser, 1981). This does not, however, logically support the idea that reducing the breadth of a subject to increase depth of study improves the transfer of knowledge beyond the types of problems presented. The experts studied to support the existence of "big ideas" had themselves come through a broad curriculum before the study, where the groupings might only emerge with sufficient exposure to a breadth of problems encountered, rather than being directly teachable.

The design of LP curricula is taken to be evidence based, so studying the abilities and concepts of experts, learners, and novices is the method to suggest useful changes to the order of an LP curriculum.

Simplification 4 Evidence

Given the inconsistency of learners' application of knowledge within a subdomain of content, it is unlikely that concepts (skills, knowledge, etc.) should be applied consistently across a subject. There is little evidence in the literature relating to transfer of learning across domains to support a claim that reducing the breadth of a subject will improve transfer (Salomon & Perkins, 1989). Secondly, the role of evidence in building effective curricula predates the debates that have been spurred

by the LP approach (e.g., Black & Simon, 1992; Bruner, 1960, p.19; Department of Education and Science/Welsh Office, 1987, p.37).

Conclusion

The findings we have presented suggest that the theory of learning that is inherent to the LP approach is unhelpfully simplistic. This is because the theory does not reflect the inconsistencies and complexities of the actual process of change that learners go through, or how inconsistently they can demonstrate their learning. Therefore, implementation of a theory based on the ladder analogy, replacing other approaches and models of learning, is likely to be counterproductive for learning since learners are never on one rung of a ladder at a time. This is not to say that no learner makes progression, or that simplifications cannot be useful in some contexts (such as for creating a scheme of work from a curriculum), but that the theory described by "Learning Progressions" authors, if implemented with no extra consideration of curriculum and learning, would not lead to positive educational outcomes.

From the perspective of a subject expert who has made the learning journey through a subject, it may seem that the journey was smooth in retrospect, but this is unlikely to have been the case. Learners develop cognitively as well as neurologically, with performance on solving problems being at best inconsistent during intermediate phases of the journey. The highest demonstrated level of performance from a learner might not be maintained across different contexts, and should not be expected to indicate mastery of activities learned previously, since forgetting can occur independently of the order of presentation.

Our consideration of the simplifications in the underlying learning theory of the LP approach suggests that the three aims of the approach (i.e., to support teaching and learning, assessment, and curriculum design) are unlikely to be met.

Firstly, tracking students' progress and identifying actions to support learning (Alonzo & Gearhart, 2006, p.100) is unlikely to support learning if success or failure at a task is taken as being unproblematic evidence of similar success and failure at related tasks. The breadth of possible learning across contexts needs support, and this will not be provided by an LP approach that considers these activities to be on the same rung of the ladder and therefore unimportant. Learners' progress is not ladder-like, but complex, and so learning activities that only consider conceptual difficulty are unlikely to diagnose students' *particular* weaknesses (Briggs, Alonzo, Schwab, & Wilson, 2006). This point may help to explain why it has been observed that the use of LPs to inform teachers' formative assessment practices has not lived up to expectations (Hammer & Sikorski, 2015).

Secondly, the utility of the framework of levels is unlikely to provide a valid way of comparing learner abilities since demonstrations of ability can be inconsistent (Leahy & Wiliam, 2011, p.5). Assuming that success or failure at one task will mean success or failure at related tasks fails to reflect the complexities of abilities that learners have, and assessments that report level-based outcomes will not be any more useful than currently reported outcomes such as grades reflecting atheoretical levels of attainment (Department for Education, 2015).

Generally speaking, the LP theory of learning, although not spelled out in any consistent way by any author, is similar to those that have come before, such as the Piagetian Spiral Curriculum (Bruner, 1960), and Gagné's Learner Sets (e.g., Gagné, 1968). Many of the criticisms laid out here are modified versions of observations of these earlier theories (Phillips & Kelly, 1975), and LP theory would benefit with engaging more with such work.

References

Alonzo, A. C., & Gearhart, M. (2006). Considering Learning Progressions From a Classroom Assessment Perspective. *Measurement: Interdisciplinary Research and Perspectives, 4*(1–2), 99–126.

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education, 93*(3), 389–421.

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. New York: Longman.

Australian Council for Educational Research. (2018). The role of learning progressions in global scales. Retrieved June 18, 2019, from [rd] ACER website: https://rd.acer.org//article/the-role-of-learning-progressions-in-global-scales

Bao, L., Hogg, K., & Zollman, D. (2002). Model analysis of fine structures of student models: An example with Newton's third law. *American Journal of Physics, 70*(7), 766–778.

Bernstein, B. (1999). Vertical and Horizontal Discourse: An essay. *British Journal of Sociology of Education, 20*(2), 157–173.

Black, P., & Simon, S. (1992). Progression in learning science. *Research in Science Education, 22*(1), 45–54.

Black, P., & Wiliam, D. (2002). *Standards in Public Examinations*. London: King's College London.

Black, P., Wilson, M., & Yao, S.-Y. (2011). Road Maps for Learning: A Guide to the Navigation of Learning Progressions. *Measurement: Interdisciplinary Research and Perspectives, 9*(2–3), 71–123.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives Book 1: Cognitive Domain* (2nd edition). New York: Longmans, Green.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic Assessment With Ordered Multiple-Choice Items. *Educational Assessment, 11*(1), 33–63.

Bruner, J. S. (1960). *The Process of Education*. Cambridge, Mass: Harvard University Press.

Catley, K., Lehrer, R., & Reiser, B. (2005). *Tracing a Prospective Learning Progression for Developing Understanding of Evolution*. Washington, DC: National Academy of Sciences.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science, 5*(2), 121–152.

Coppola, B. P. (2006). Learning progressions provide a clear map for designing Standards-based assessment [Peer Commentary on: "Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic Molecular Theory." by Smith, C. A.; Wiser, M.; Anderson, C. A.; Krajcik, J.] *Measurement: Interdisciplinary Research and Perspectives, 14*(1/2), 111–115.

Corcoran, T., Mosher, F., & Rogat, A. (2009). Learning Progressions in Science: An Evidence-Based Approach to Reform. *CPRE Research Reports*.

Denvir, B., & Brown, M. (1987). The feasibility of class administered diagnostic assessment in primary mathematics. *Educational Research, 29*(2), 95–107.

Department for Education. (2015). *Commission on Assessment Without Levels: final report*. Retrieved from DfE website: https://www.gov.uk/government/publications/commission-on-assessment-without-levels-final-report

Department of Education and Science/Welsh Office. (1987). The National Curriculum 5–16: A consultation document. London: Department of Education and Science/Welsh Office.

Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching, 46*(6), 606–609.

Duschl, R. (2006). Learning Progressions as Babushkas. *Measurement: Interdisciplinary Research and Perspectives, 4*(1–2), 116–119.

Efland, A. D. (1995). The Spiral and the Lattice: Changes in Cognitive Learning Theory with Implications for Art Education. *Studies in Art Education, 36*(3), 134–153.

Fensham, P. (1994). Progression in school science curriculum: A rational prospect or a chimera? *Research in Science Education, 24*(1), 76–82.

Fischer, K. W. (2008). Dynamic cycles of cognitive and brain development: Measuring growth in mind, brain, and education. In A.M. Barrio, K.W. Fischer, & P.J. Léna (Eds.), *The educated brain: Essays in Neuroeducation* (pp.127–150).

Gagné, R. M. (1968). Presidential address of division 15 learning hierarchies. *Educational Psychologist, 6*(1), 1–9.

Gagné, R. M., & Bassler, O. C. (1963). Study of retention of some topics of elementary nonmetric geometry. *Educational Psychology, 54*(3), 123–131.

Goldstein, H. (1998). *Models for reality: new approaches to the understanding of educational processes*. Professorial Lecture presented at the Institute of Education University of London. Institute of Education, University of London.

Gotwals, A. W., & Songer, N. B. (2013). Validity Evidence for Learning Progression-Based Assessment Items That Fuse Core Disciplinary Ideas and Science Practices. *Journal of Research in Science Teaching, 50*(5), 597–626.

Graf, E. A., & van Rijn, P. W. (2016). Learning Progressions as a Guide for Design: Recommendations Based on Observations from a Mathematics Assessment. In S. Lane, M. Raymond, & T. M. Haladyna (Eds.), *Handbook of Test Development* (pp.165–189). Abingdon, Oxon: Routledge.

Hammer, D., & Sikorski, T.-R. (2015). Implications of Complexity for Research on Learning Progressions. *Science Education, 99(*3), 424–431.

Harlen, W. (Ed.). (2010). *Principles and big ideas of science education*. Retrieved from https://www.ase.org.uk/documents/principles-and-big-ideas-of-science-education/

Hart, K. M. & Chelsea College. CSMS Mathematics Team (1981). *Children's understanding of mathematics: 11–16*. (reprinted 1986). London: John Murray.

Hess, K. K. (2008). *Developing and Using Learning Progressions as a Schema for Measuring Progress*. Dover, New Hampshire: National Center for the Improvement of Educational Assessment.

Kim, E. M., Haberstroh, J., Peters, S., Howell, H., & Nabors Oláh, L. (2017). A Learning Progression for Geometrical Measurement in One, Two, and Three Dimensions. *ETS Research Report Series, 2017*(1), 1–26.

Kim, H., & Care, E. (2018, March 27). Learning progressions: Pathways for 21st century teaching and learning. Retrieved June 18, 2019, from https://www.brookings.edu/blog/education-plus-development/2018/03/27/learning-progressions-pathways-for-21st-century-teaching-and-learning/

Kolb, J. R. (1967). Effects of relating mathematics to science instruction on the acquisition of quantitative science behaviors. *Journal of Research in Science Teaching, 5*(2), 174–182.

Leahy, S., & Wiliam, D. (2011, April). *Devising learning progressions*. Presented at the Symposium on How to Build Learning Progressions; American Educational Research Association Annual Conference, New Orleans.

Lehrer, R., & Schauble, L. (2015). Learning Progressions: The Whole World is NOT a Stage. *Science Education, 99*(3), 432–437.

Masters, G., & Forster, M. (1996). *Progress Maps*. Melbourne, Australia: ACER.

Messick, S. (1984). The Psychology of Educational Measurement. *Journal of Educational Measurement, 21*(3), 215–237.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: The National Academies Press.

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: The National Academies Press.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K– 8*. Washington, DC: The National Academies Press.

Oates, T. (2011). Could do better: using international comparisons to refine the National Curriculum in England. *The Curriculum Journal, 22*(2), 121–150.

Paik, S., Song, G., Kim, S., & Ha, M. (2017). Developing a Four-level Learning Progression and Assessment for the Concept of Buoyancy. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(8), 4965–4986.

Phillips, D., & Kelly, M. (1975). Hierarchical Theories of Development in Education and Psychology. *Harvard Educational Review, 45*(3), 351–375.

Salomon, G., & Perkins, D. N. (1989). Rocky Roads to Transfer: Rethinking Mechanism of a Neglected Phenomenon. *Educational Psychologist, 24*(2), 113–142.

Sandoval, W. A., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction, 23*(1), 23–55.

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: an examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37*(5), 525–559.

Simons, J., & Porter, N. (2015). *Knowledge and the Curriculum: A collection of essays to accompany E. D. Hirsch's lecture at Policy Exchange | Policy Exchange*. Retrieved from Policy Exchange website: https://policyexchange.org.uk/publication/knowledge-and-the-curriculum-a-collection-of-essays-to-accompany-e-d-hirschs-lecture-at-policy-exchange/

Smith, C. L., & Wiser, M. (2015). On the Importance of Epistemology–Disciplinary Core Concept Interactions in LPs. *Science Education, 99*(3), 417–423.

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). FOCUS ARTICLE: Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic-Molecular Theory. *Measurement: Interdisciplinary Research and Perspectives, 4*(1–2), 1–98.

Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching, 46*(6), 699–715.

Taber, K. S. (2000). *Teaching chemistry without (too much emphasis on) atoms?* [A Royal Society of Chemistry Teacher Fellowship discussion paper]. Retrieved from British Education Index website: http://www.leeds.ac.uk/educol/documents/00001559.htm

Valverde, G. A., & Schmidt, W. H. (1998). Refocusing U.S. Math and Science Education. *Issues in Science and Technology, 14*(2), 60–66.

Vorst, H. van. (2018). Structuring learning processes by ladders of learning: results from an implementation study. *Chemistry Education Research and Practice, 19*(4), 1081–1095.

Webb, N. L. (1997). *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*. (No. National Institute for Science Education Research Monograph 6). Washington, DC: Council of Chief State School Officers.

Webb, N. L. (1999). *Alignment of Science and Mathematics Standards and Assessments in Four States*. (No. National Institute for Science Education Research Monograph 18). Washington, DC: Council of Chief State School Officers.

Whetton, C. (2008, September). *The Big New Idea: Single Level Tests*. Presented at the International Association for Educational Assessment Conference, Cambridge.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction, 16*(1), 3–118.

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching, 46*(6), 716–730.