

Children's science learning: a core skills approach

Andrew Tolmie*, Zayba Ghazali[†] & Suzanne Morris*

Department of Psychology and Human Development, UCL Institute of Education,

*University College London**

Moray House School of Education, University of Edinburgh[†]

Contact address:

Professor Andrew Tolmie

Department of Psychology and Human Development

UCL Institute of Education

University College London

25 Woburn Square

London WC1H 0AA

United Kingdom

e-mail: andrew.tolmie@ucl.ac.uk

tel +44 (0)20 7612 6224

Children's science learning: a core skills approach

Abstract

Background: Research has identified the core skills that predict success during primary school in reading and arithmetic, and this knowledge increasingly informs teaching.

However, there has been no comparable work which pinpoints the core skills that underlie success in science.

Aims and Method: The present paper attempts to redress this by examining candidate skills and considering what is known about the way in which they emerge, how they relate to each other and to other abilities, how they change with age, and how their growth may vary between topic areas.

Results: There is growing evidence that early-emerging tacit awareness of causal associations is initially separated from language-based causal knowledge, which is acquired in part from everyday conversation and shows inaccuracies not evident in tacit knowledge. Mapping of descriptive and explanatory language onto causal awareness appears therefore to be a key development, which promotes unified conceptual and procedural understanding.

Conclusions: This account suggests that the core components of initial science learning are 1) accurate observation, 2) the ability to extract and reason explicitly about causal connections, and 3) knowledge of mechanisms that explain these connections.

Observational ability is educationally inaccessible until integrated with verbal description and explanation, for instance via collaborative group work tasks that require explicit reasoning with respect to joint observations. Descriptive and explanatory ability are further promoted by managed exposure to scientific vocabulary and use of scientific language. Scientific reasoning and hypothesis testing are later acquisitions which depend on this integration of systems and improved executive control.

Background and aims

In recent years, research has made substantial progress in uncovering the core cognitive skills that predict success in primary reading and mathematics, using a strategy of systematic analysis of influences on individual variations in outcome. Specifically, evidence demonstrates support for the 'simple view of reading', which suggests that acquisition of reading skills rests on two distinct processes, learning of orthographic/sound relationships, and growth of the ability to extract meaning from assemblages of text (Hulme & Snowling, 2009). In mathematics, grasp of counting procedures and principles, recognition and use of symbols, translation between verbal and symbolic formats, and knowledge of number facts and calculation strategies have been identified as relatively independent key skills with an underpinning influence from working memory (Soltesz, Szucs & Szucs, 2010). Crucially, knowledge of core skills in both these subjects has begun to lead to changes in teaching practices with benefits for children's learning (Dowker & Sigley, 2010; Wyse & Goswami, 2008).

In contrast, there has yet to be any similar attempt to identify the underlying core competences necessary to engage with science. This absence is striking given the importance attached to science alongside literacy and numeracy by national governments, and its consequent inclusion in global measures of educational achievement such as the Programme for International Student Assessment (PISA). Vitality, scientific ability does not just concern disciplinary knowledge about how we and the world around us function, but also – like reading and mathematics – a crucial skill set, which includes the capacity to collect and assess evidence about causal relationships in unbiased fashion, and to draw from it appropriate and explicit conclusions that can be shared and debated with others. These quintessentially cognitive abilities are as fundamental to participation in modern societies as being able to read, write and calculate.

The current paper aims to consider how this gap in understanding might be filled, by examining existing evidence with regard to two essential questions:

1. What core skills might predict scientific understanding amongst primary school children?
2. How do these core skills relate to each other and to more general cognitive abilities?

What skills might underpin the growth of children's scientific understanding?

Construction of a skills-based account of science learning rests on how we define such learning. In terms of outcome, there is widespread consensus that *scientific reasoning* and *scientific knowledge* are the two main branches of scientific understanding since they underpin key aspects of professional science and are central to its focus on uncovering causation (Bybee, 1997; Koerber, Mayer, Osterhaus & Schwippert, 2015; Wellington, 1988). Scientific reasoning (or 'scientific thinking') refers to the *processes* of scientific enquiry that make it possible to isolate causal influences, particularly the key components of experimental procedure: generation of hypotheses, controlled testing, and use of evidence to evaluate hypotheses (Klahr, 2000, 2005; Zimmerman, 2007). Understanding of what is involved in each of these components has been argued to rest further on sub-skills such as the identification of variables and use of controlled tests that manipulate these one at a time, which make it possible to isolate causal influences. Scientific knowledge concerns the *concepts and theories* that capture and explain the mechanisms by which causal influences produce outcomes (Piaget, 1972) and the application of these to specific contexts. These two strands of ability are typically seen as the principal targets of science learning, and have long been treated as constituting a distinct and separable procedural vs conceptual dichotomy, with the former resting on domain-general skills (i.e., applicable in any context), and the latter on domain-specific knowledge (see e.g., Fugelsang & Mareschal, 2014).

However, as is outlined below, current evidence on the emergence and interrelationship between these two dimensions presents in fact a somewhat muddy picture.

Evidence on scientific reasoning

On the face of it, studies of children's procedural or investigative abilities indicate a straightforward willingness from 4 years to manipulate factors to assess their impact on outcome (i.e., in some sense to hypothesise and test their influence), especially when presented with contradictory or unclear prior evidence. For instance, Bonawitz, Ferranti, Saxe, Gopnik, Meltzoff, Woodward and Schulz (2010) found that from 47 months on – but not younger – the majority of children spontaneously intervened to manipulate a previously observed simple cause-effect relationship in which the movement of a block activated the operation of a toy plane. Drawing on the 'blicket detector' task devised by Gopnik, Sobel, Schulz and Glymour (2001), Legare (2012) presented preschool children with a series of trials in which only objects with certain features ('blickets') activated the light on a box when placed in contact with it. When these were followed by a further trial in which a 'blicket' failed to trigger the light, children produced causal function explanations when asked what had happened, and to the extent that they did so, engaged in greater amounts of exploratory manipulation of the objects, suggesting they were attempting to hypothesise and test expected effects.

Similarly, Cook, Goodman and Schulz (2011) examined children's activity when presented with ambiguous causal information. Preschool children were shown that 2 out of 4 beads made a toy play music when the beads were placed one at a time on top of it. They were then shown two pairs of further beads, one of which could be pulled apart into two individual beads, while the other pair was glued together. Both bead pairs activated the toy. The evidence about the bead pairs was ambiguous since it failed to distinguish which bead worked (or whether both did), and the baseline information suggested there was no inevitable relationship. Children were then allowed a brief period of free play with the bead

pairs, and assessed on how far they engaged in informative exploration by testing the effects of the separable beads individually. When given unambiguous evidence on comparison trials, children played indiscriminately, but when faced with ambiguous evidence, they both selected and designed informative interventions.

Given the apparent contingency of all these investigative behaviours on prior observation, they evidently depend on the ability to track and assess evidence, and to isolate factors that covary with outcomes. This ability appears to emerge even earlier. Using habituation paradigms, Leslie and Keeble (1987) found evidence that infants infer causal relationships from contact events in which one object collides with a second, apparently making it move, and carry expectations about these relationships forward into subsequent experience.

Gopnik et al. (2001) found that children as young as 2 years were able to reliably predict which objects would activate their 'blicket detector' on the basis of prior observation of the features associated with it being triggered. Schulz, Gopnik and Glymour (2007) found that 3 to 5 year olds were able to selectively identify diagrams which captured previously witnessed relationships between the movements of toy gears.

However, despite the consistent picture presented by these and related studies, the nature of the abilities that have been demonstrated – and their relation to scientific reasoning – is in fact unclear. There are three areas of doubt. First of all, the data obtained often fit poorly with the interpretation placed upon them. In Legare (2012), for instance, responses categorized as causal function explanations were more common among 2 to 4 year olds (c. 50% of responses) than 5 to 6 year olds (c. 30%), with the 5 year olds producing more non-causal explanations. Both points are developmentally surprising yet ignored by the author. Moreover, examples of causal function explanations included "it's broken" and simple descriptions of differences in object features, which carry little obvious causal intension or articulation of process, as the label 'explanation' ostensibly requires. More generally, both

here and elsewhere, the focus on modal patterns of response glosses over individual variations which might inform a richer story about the growth of abilities.

Second, as a result of concerns that existing concepts may bias the assessment of evidence (Fugelsang, Stein, Green & Dunbar, 2004), the paradigms involved in a great many of these studies (e.g., the 'blicket detector') are deliberately isolated from genuine scientific content, so as to capture 'pure', de-contextualised scientific thinking (cf. the procedural-conceptual dichotomy). The few studies that have utilised real contexts indicate that these do indeed influence performance in unpredictable fashion. For example, Ape, Flottmann and Leuchter (2015) used a simple multiple-choice paradigm with visual materials to assess 6 year olds' knowledge of effects in four areas (slopes, friction, object flotation and plant growth), along with their ability to identify a control of variables strategy (CVS) as an appropriate means of investigating effects. CVS responses varied across area from 10% to 42% of instances, but with no clear relationship to knowledge of effects. For instance, although the two least understood areas (slopes and object flotation) produced more CVS responses, which might suggest prior understanding can bias approaches to testing (cf. Penner & Klahr, 1996), the rate for slopes was twice that for flotation, which is harder to explain. Similarly, the two best understood areas (friction and plant growth) produced different response patterns: holding the focal variable constant (either providing water or sunlight) was dominant for plant growth, indicating a possible bias towards preserving an expected outcome, whereas no consistent choice of strategy was the dominant outcome for friction. Confusing the picture further, other studies have found that 4 year olds *can* identify conclusive experiments so long as the outcome is in line with their conceptual understanding, indicating knowledge may be a prerequisite, not a distractor (Piekny & Maehler, 2013; see also Croker & Buchanan, 2011; Koerber, Sodian, Thoermer & Nett, 2005).

These apparently unsystematic variations arguably justify blicket-style approaches to testing children's procedural skills, though they also undermine any claim that these skills operate in

domain-general fashion as that strategy presupposes (cf. Klahr, 2000; and see also Schauble, Klopfer & Raghavan, 1991, who argue that both straightforward domain generality and domain specificity are over-simplifications). The net result is that evidence from studies of this type tell us little that is certain about how children's abilities might relate to their capacity to investigate *genuine* phenomena – the ultimate point of developing such skills – or how that capacity improves with age. It is this which we most obviously need to understand from an educational perspective, and the apparently unpredictable effects identified by Ape et al. (2015) merely serve to underline why it is important that we do so.

The third area of doubt is that although the existence of a 'proto-investigative' propensity from early childhood appears undeniable, it plainly does not equate to genuine experimentation. Studies that have examined explicit hypothesis generation, testing and evaluation in the context of real phenomena have found that selection of test variables, controlled manipulation, and unbiased inference from observation are poor, regardless of topic area, and remain so into secondary school, albeit with some degree of individual variation. Howe, Tolmie and Sofroniou (1999), for example, examined the ability of 9 to 14 year olds to use physical materials to conduct systematic investigations of the effects of factors in four topic areas in order to establish whether or not their beliefs regarding these effects were correct. Very few children appreciated the need to manipulate variables in a controlled manner, and having introduced extraneous factors, most then struggled with predicting outcomes, observing effects and coordinating data into appropriate conclusions. Many also explicitly claimed that when the effect of a factor was 'known' (from their point of view), there was no point in testing it, and they therefore failed even to grasp the basic purpose of experimentation. Similar outcomes were reported by Schauble (1990), who found that invalid heuristics that preserved favoured theories persisted among 10 to 12 year olds even after eight weeks' work on problems.

Potentially explaining some of the difficulties, Kuhn, Iordanou, Pease and Wirkala (2008) found that 12 to 13 year olds exhibited high levels of inconsistency in the extent to which they implicated each of a set of established causal variables in predictions of multivariable outcomes, sometimes attributing causal influence to a given variable and sometimes not. This suggests that they have difficulty maintaining any stable focus on individual factors, perhaps because of the cognitive load involved, or even that they fail to appreciate that variables operate in consistent fashion. In line with the former, other research has reported that the number of variables to be controlled or manipulated impacts on performance, as does whether the task is to actually design experiments, as opposed to simply assessing the characteristics of presented designs (see e.g., Howe, 2014a, for a summary). In addition, though, Schauble et al. (1991) found that 10 to 12 year olds tended to favour an engineering model of experimentation, characterized by the goal of manipulating variables to produce a desired outcome, only shifting to a scientific model when they were supported to do so.

Piekny and Maehler (2013) also report that key aspects of scientific reasoning emerge at different times: evaluating perfectly covarying data develops early in primary school; understanding the difference between an experiment and obtaining a desired outcome emerges next; and the ability to generate hypotheses develops last. Moreover, the exact timing of this sequence may be unstable. Research by Shayer and Ginsburg (2009), comparing data on secondary school students' performance on control of variable tasks in 1976 and 2006/7, found a decline over that period of half a standard deviation or more. This may indicate the existence of unknown cultural influences which have perpetuated and increased biases towards demonstrating that one's existing beliefs are correct or obtaining a desired outcome, although other explanations may be possible (e.g. reduced familiarity with the questioning style or with the phenomena to be investigated due to changes in the curriculum).

At best then, all we are really able to conclude is that refined investigative understanding appears to be a relatively late development, dependent on earlier skills. It remains an open question as to whether early investigative propensities are among these, and if so, in what way they influence later ability. Efforts to theorise what the relationship might be are made harder by the loose definition in developmental research of what such propensities actually comprise and the accompanying tendency to focus on general patterns of performance rather than its precise details.

Evidence on scientific knowledge

Examination of past research on scientific knowledge indicates that any attempt to conceive of this as a coherent form of understanding faces similar kinds of difficulties. In terms of registration of evidence (as opposed to its generation via investigative activity), the early emergence of causal perception has already been noted. Importantly, Schulz et al. (2007) also found that young children's awareness of causal association across events corresponds to computer models of probabilities derived from the same information, suggesting it may be essentially driven by statistical sensitivity and direct registration of information about contingencies.

Further indications of such sensitivities are reported by Howe et al. (1999), who assessed 9 to 14 year olds' understanding of the four topic areas they considered (object flotation, water pressure, shadow size and motion down an incline) in preparation for the hypothesis testing phase of their research. They found that participants very rarely discounted as causally irrelevant any factor that actually affected outcome, though they were not always correct about the direction of effect, believing for instance that small things float and large sink, an oversimplification of the effect of size (a point we return to later). Such discounting happened on only 16 occasions in the course of over 1300 judgements (just over 1% of trials). This accuracy can only be attributed to children having previously tracked observed covariation over time, as these disparate topics were not areas covered in science lessons and the

study provided no opportunity for observation or feedback before beliefs were assessed. diSessa (1988) reports extremely similar findings.

Adult cognition in tasks analogous to the 'blicket detector' paradigm also indicates that their performance derives from estimations of relative probability. Indeed, even though adults are generally poor at deliberate reasoning about probability, their everyday estimations of outcome probabilities are sensitive to subtle distributional characteristics that are naturally captured by statistical models (Chater, Tenenbaum & Yuille, 2006). The evidence suggests then that statistically-based causal awareness emerges in infancy, and continues to operate into adulthood.

This accuracy in tacit awareness contrasts in striking fashion with inaccurate performance on tasks requiring explicit judgments that ostensibly ought to be able to utilise it, although as noted above accurate registration of contingency is not necessarily the same as being aware of the full characteristics of the effect produced. For example, Howe, Tavares and Devine (2012) compared 6 to 10 year olds' ability to either explicitly predict or recognise the trajectory likely to be followed by falling objects, half of which were in horizontal motion at the point of release. Prediction responses were characterised by an *increasing* tendency with age to portray a backwards trajectory – an outcome which was never correct – whereas backwards trajectories were selected in only 22% of recognition responses, and did not increase in frequency with age. Follow-up research (Howe, Taylor Tavares & Devine, 2014) focused solely on horizontal motion confirmed a consistent recognition-prediction gap of the same kind across two tasks addressing direction and speed respectively in the context of the motion of billiard balls.

There appear to be clear signs then that early-originating sensitivity to covariation between factors and outcomes is separated in some fashion from explicit knowledge, suggesting the latter is acquired in distinct and at least partly non-observational fashion. This disjunction

may possibly arise because tacit perceptual and explicit linguistic knowledge are initially held within separate or only partially overlapping neurocognitive systems (Fugelsang & Mareschal, 2014; Howe et al., 2014; Tolmie, 2012), with linguistic knowledge influenced by conversation. The work of Harris and colleagues on 'testimony' (see e.g., Harris & Koenig, 2006; Harris & Corriveau, 2014) certainly shows that young children readily attend to the knowledge implications of others' statements (especially their parents); are capable from 3 years of making sophisticated judgments about the reliability of different sources according to their track record and degree of consensus with others; and that these effects extend to learning about aspects of science.

In the specific context of scientific knowledge, an analysis of variations in primary age children's understanding of different types of physical phenomena by Howe (1998) found that better understood topics which exhibited theory-like organization corresponded with those which were commonly referred to in everyday conversation (e.g., heating and cooling). Research by Symons, Tolmie and Oaksford (2015) with 6 to 11 year olds in the context of motion down an incline found that although effects were more pronounced amongst older children, all age groups exhibited sensitivity to source reliability in being told that weight did not affect outcome, showing greater change in their own predictions when that information ostensibly came from a teacher as opposed to another child. Interestingly, a final shift to dismissing effects of weight only came when children subsequently tested outcomes themselves, suggesting that the combination of external information and personal observation has the greatest influence (see also Howe, 2014b).

Further evidence of both the hypothesised disjunction between tacit and explicit knowledge, and the influence of conversation is provided by a study of 5 to 10 year olds' understanding of physical state change i.e., melting, freezing, evaporation and condensation (Tolmie, 2014). Despite the essential equivalence of the physical processes involved in each, children's ability to describe and explain the different types of change differed significantly,

with performance best on melting and then freezing, and worst on evaporation and then condensation. This pattern appears to mirror again the extent to which a) children encounter each change type in everyday experience and b) it forms the topic of everyday conversation. Melting is commonly experienced and referred to, freezing is referred to but rarely directly witnessed, and evaporation and condensation are both harder to directly perceive (the gas-liquid transition is less tangible) and much less commonly mentioned. Moreover, explanation quality was only weakly related at best to description quality for melting and freezing, indicating that explanation and description were the product of at least partially distinct processes. However, the relationship between the two was substantially stronger for evaporation (performance on condensation was too poor to assess this relationship).

This research suggests conversation may create explicit conceptual grasp which is more advanced in terms of reportability, but potentially ill-grounded in observation if the source is perceived to be reliable but in fact is not (as may be the case for many parents, for example), and relevant events are either fleeting or even hidden. This would be a strong candidate mechanism for the origin of the naïve science concepts that in many instances children possess at entry to school (Driver, Guesne & Tiberghien, 1985). Bottom-up derivation of explanatory concepts via the description of observations appears to be more likely in the absence of conversational influence – possibly supported in some instances by more formally organised classroom encounters, as may be the case for evaporation – and may be more accurate. It may also be slower and have an uncertain trajectory. Whether it is assisted by skills of testing and manipulation is unknown.

A proposed set of core skills

The cumulative implication of these points is that despite its apparent appeal, the distinction between investigative and conceptual abilities has resulted in an artificial and unhelpful

distancing between them in past research, where systematic joint scrutiny would have been more informative. More crucially, we argue that this research has failed to provide a clearly articulated model of the skills on which performance actually turns, because it is in fact evident that neither investigative nor conceptual abilities are unitary in character and that the relationship between their different elements is complex.

The state change research suggests instead that coordinated scientific understanding may rest ultimately on a simpler and more fundamental set of core skills, about which, for the most part, we know surprisingly little in terms of their basic characteristics, let alone the relationships between them. The differential ability of children to describe melting and freezing indicates that opportunity to witness instances of phenomena is important, but the differences between melting and evaporation imply further that actively attending to available instances and noting the changes that occur is equally crucial. The differential relationship between explanation and description for melting or freezing and evaporation indicates that not only is explanatory ability inherently distinct from such observation, but that it can emerge in various ways. Finally, though the point is more implicit, any complete integration of description and explanation must require systematic tracking of effects across instances of phenomena and assessment of their consistency with proffered accounts. We therefore argue that the core skills underpinning science learning are as follows:

Accurate observation. Observation of covariation is a necessary first step in the identification of causal relations, but *accurate* scrutiny is the basis of unbiased observation, a key element of scientific procedure. Ability to *explicitly* describe specific observations makes this process fuller and more deliberate, and may also aid extension of causal perception to distal connections (i.e., causal effects over time) by making it possible to note and recall causal candidates among prior events. In terms of skill development, progress therefore appears likely to be along a dimension from tacit sensitivity – an ability which appears to be early-emerging and predominantly domain-general – to verbal encoding of relationships,

including identification of variables, which may be more domain-specific, initially at least. Despite generally good performance on measures of tacit awareness from 24 months, there is nevertheless clear evidence of individual variation in observational ability (Bonawitz et al., 2010; Sobel, Tenenbaum & Gopnik, 2004) which may potentially have long-term impact. It is evident from the state change research that descriptive ability varies to a greater extent, both individually and with age and context.

The ability to extract and reason about causal connections. Reasoning about causal connections (i.e., drawing causal inferences) relies on accurate observation, so should be related to it. However, it further requires the coordination of different observations (Kuhn et al., 2008); unbiased extraction of associations, discounting intrusive or even conflicting effects which arise from the noisy nature of many phenomena; recognition of gaps in evidence; and, ultimately, verbal summation of relationships (the description index used in Tolmie, 2014, distinguished between context-specific and generic responses, and found that the latter became more prevalent with age). It is this ability which may ultimately lead to an understanding of the importance of controlled tests, and to the explicit evaluation of evidence and scientific argument. The extent of individual variation is largely unknown, but the range of procedural ability found in secondary school students (see Kuhn et al.; Shayer & Ginsburg, 2009) suggests it is likely to be high.

Application of knowledge regarding mechanisms and processes that explain inferred connections. Explanation has crucial potential links to the data derived from observed causal connections since it provides the means of specifying the processes that result in those observations. At higher levels of performance, this relationship is likely to underpin hypothesis generation and testing. However, as noted above, explanatory concepts are often separated from accurate observation, at least initially, and evidence suggests individuals exhibit wide variation in explanatory ability between and even within topic areas (Ape et al., 2015; Howe, 1998; Tolmie, 2014), indicating a high degree of domain-specificity.

Whether learners who are more advanced in one area tend nevertheless to be more advanced in others is largely unknown, though, because few studies with primary age children have encompassed a number of areas, and those that have rarely report cross-topic correlations. Despite its importance to school curricula, it is also unclear whether there are predictable processes of conceptual extension and generalisation, with progress in one topic building on earlier gains in another. Some have argued that knowledge is represented in highly organised theory-like structures, with conceptual change occurring through theory-replacement processes – ‘knowledge as theory’ approaches (e.g., Carey 1985; Gelman, 2009; Vosniadou 2002, 2014; Vosniadou & Ioannides, 2006), suggesting there is a strong basis for extension and generalisation. However, others argue that concepts are more typically a collection of context-specific fragments, with development being more iterative in character – ‘knowledge in pieces’ approaches (e.g., diSessa, 1993).

Since any integrated performance demands all three of these skills, i.e. observation, extracting and reasoning about causal connections, and explanation, we hypothesise that it is in fact these – and not early proto-investigative abilities – which provide the foundation for fully-fledged reasoning and conceptual capacity, by establishing a sense of the relationship between theory and evidence. Interestingly, echoing this analysis, the National Curriculum for Science in England (Department for Education, 2013) now emphasises observation (including noting patterns and relationships), articulation of concepts and acquisition of scientific vocabulary, and at upper primary level, explanation and prediction, though it pays no attention to the coordination of these different skills, which we argue is crucial for children to progress. The designers provide no explicit account of why these skills matter, either, despite the fact that a clear rationale is far more likely to engage teachers than simple prescription.

Relationships of the core skills to each other and to more general cognitive abilities

The opening to this paper posed the question of how any identified core skills might relate to each other and to more general cognitive abilities. Returning to this question now, it is clear that there is no current basis for assuming any specific relationship between the abilities to observe, extract patterns of data and explain phenomena, and the nature of these relationships must therefore be determined by deliberate empirical investigation. The widespread strategy of examining scientific reasoning in contexts which do not involve real phenomena means that evidence on the relationship between observation and pattern extraction on the one hand and scientific concepts on the other is restricted – although the apparent emergence of reasoning abilities around 24 months (Schulz et al., 2007) when explanatory abilities are rudimentary at best suggests a disjunction between them.

Even where studies *have* looked at reasoning in real contexts, the data are inconsistent. Howe et al. (1999) found that late primary/early secondary participants were uniformly poor at conducting systematic tests of the effects of variables across four contexts, despite differing in their levels of understanding of these contexts, which also suggests there is no relationship between reasoning and explanatory abilities. Among pre-schoolers, the more recent multiple choice research by Ape et al. (2015) suggests if anything a negative relationship between understanding and reasoning. In contrast, using pictorial representations of a set of precursor conditions and outcomes, both covarying and non-covarying, Koerber et al. (2005) found that conceptual understanding was associated with better reasoning. The methods of assessment in these latter two studies differ, however, and in neither case is the direction of the reported effect particularly strong.

The relationship between observational and pattern extraction abilities is similarly unclear because we are not aware of any instance of them having been assessed separately: studies of reasoning uniformly concatenate them. Accurate observation does not inherently entail an ability to coordinate information across observations, though, and the state change

research (Tolmie, 2014) indicates that younger children in particular frequently fail to connect different instances of the same phenomenon (e.g. melting), suggesting that coordination does present specific challenges.

With regard to that coordination – and relationships to more general cognitive abilities – the above account suggests that *language*, expressive ability in particular, must significantly influence science learning, since it is language that allows children to build connections between observations and explicit explanation. Ghazali (2014) provides initial evidence of this influence. Using a lagged design, she found consistently strong relationships between measures of both receptive (i.e., recognition of vocabulary) and expressive language (production of definitions) and understanding of a range of biological concepts (biodiversity, ecology, inheritance and evolution) among 4 to 11 year olds. She also found that these relationships were not attributable to language having a general impact on performance within the test context, since they were almost entirely mediated by construct-specific language ability, with biodiversity related to ecology, and ecology related in turn to evolution and inheritance.

The implication is that external resourcing of science-related language skills via school input is likely to be an important contributory influence on understanding. This suggests effects of quality of classroom conversation in science, and also perhaps use of computer-based resources to support science activity, depending on their design and content. At a basic level, explanatory ability is likely to be influenced by exposure to explicit constructs, and therefore the quality of teacher or software input in terms of the use of vocabulary to convey science fact and to describe mechanisms and processes may be an important environmental variable. In line with this, Bonawitz et al. (2010) found that 24 month olds who had witnessed events described using causal language made attempts to manipulate causal effects; such manipulations were not otherwise seen until two years later. They conclude that language plays an important role in helping children extend causal representations. However, the

exact timing and nature of such input may be critical: Philips and Tolmie (2007) found that explicit explanations of balance scale problems by parental tutors promoted learning among 8 year olds only when they already had some grasp of the role of weight and distance; in the absence of this, such explanations appeared to create confusion.

The impact of language would appear at root, though, to be that it facilitates the coordination of different elements of conceptual grasp into a more coherent whole, and in this respect it is consistent with Karmiloff-Smith's (1992) account of conceptual change via the process of representational redescription (RR). According to this, conceptual development begins with the acquisition of fragmentary implicit representations of action-event relationships which are triggered when appropriate circumstances are encountered, but are not otherwise mentally manipulable. These become increasingly coordinated as the connections between them are made more explicit, initially via a heightened internal awareness, but eventually via encoding in language.

Crucially, in this account, at the initial stage of internal connection, much of the detail contained in implicit representations is lost and the resulting structures become ideas which are detached from data; but as the process of explication proceeds further, there is renewed attention to data and mapping this as part of more coordinated language-based structures. In other words, then, the process of explication of concepts via language could quite reasonably provide the basis of more systematic theory-building and scientific thinking, the sought-for bridge between early awareness of covariation and later reasoning skills. Before this happens, though, representations and observations are indeed separated, and might logically be more so if the former are externally resourced. Note too that within this account conceptual grasp *and* procedural knowledge tend to go hand in hand, albeit in complex fashion (cf. Schauble et al., 1991), and both must therefore be topic-specific, at least initially – as the evidence considered earlier suggests.

The emerging picture is a complex one, then, but basically:

- a) children and adults alike have accurate low level systems which cumulatively extract perceptual regularities, possibly to serve anticipatory mechanisms;
- b) some (but not all) aspects of experience get made more explicit via markers including language (cf. levels E1 to E3 in the RR model), and become coordinated into actual concepts;
- c) where this happens in bottom-up fashion, the result is slowly emerging, relatively accurate partially formulated explicit ideas (e.g., the effect of slope angle on motion), though these may sometimes be simplifications of the real state of affairs (e.g., beliefs about the effect of size on object flotation, when size is only relevant in combination with mass);
- d) where this process has top-down feeds from everyday conversation and accompanying experiences (perhaps including media footage), it may capture phenomena inaccurately and so lead to distorted perceptions of observable data and self-reinforcing misconceptions, which eventually become automated e.g., the belief in the backwards trajectories of falling objects (cf. Howe, McWilliam and Cross, 2005, who found that collaborative group work produces heightened sensitivity to the relationships that have been discussed when these are encountered in subsequent experience).

In this account, language is the driver for explicit concepts, and its effects are ubiquitous. However, it operates in variable fashion, produces different configurations of relationship between tacit and explicit understanding (cf. Howe et al., 2014), and is far from uniformly helpful. It needs to be systematically introduced alongside careful observation to be sure of being productive, which is where education comes in. Note that this position is more consistent with 'knowledge in pieces' (cf. diSessa, 1988, 1993) than 'knowledge as theory' approaches (cf. Gelman, 2009) to conceptual organisation. It also suggests an interesting potential interim point in development at which gesture might in some contexts capture direct readout from perceptual experience that may be in tension with explicit language (see Pine,

Lufkin & Messer, 2004, who found evidence of exactly this tension in a balance task, and also that children who exhibited it showed greater learning gains).

In addition to language, the ability to integrate different observations about the effect of a factor and link patterns of outcome to potential explanations seems likely to rest also on the ability to manipulate data (*executive function*), and on temporary storage capacity (*visual and verbal working memory*). There is evidence that this is the case with adults. Fugelsang and Dunbar (2005), for instance, found that processing of evidence which was inconsistent with a theory that was being assessed led to activation of the dorsolateral prefrontal cortex (DLPFC), which is associated with executive function – though consistent evidence was not apparently attended to in the same way, suggesting a tendency to simply accept this.

Nayfield, Fuccillo and Greenfield (2013) found that executive function is predictive of school readiness scores in science among 4 to 5 year olds. Other research with adults suggests *semantic inhibitory control* – the ability to suppress irrelevant associations – is key to supplanting misconceptions (Masson, Potvin, Riopel & Brault Foisy, 2014). It may therefore also influence the coordination of observed evidence and explanation by facilitating the separation of relevant from irrelevant information and the recognition of inconsistencies.

In general, empirical support for an influence of executive function and working memory on science learning is somewhat patchy. Rhodes, Booth, Campbell, Blythe, Delibegovic and Wheate (2012) and Gathercole, Pickering, Knight and Stegman (2004) report relationships between working memory and science achievement measures, but in both cases this was among secondary school students, and corresponding data were not available for primary age participants. Brookman (2015), again working with adolescents, found more specifically that inhibitory control measures were predictive of performance on a science and maths misconceptions task, in line with the hypothesis that suppression of incorrect ideas is required so that correct scientific or mathematical responses can be chosen. In contrast, the state change research with primary age children referred to above (Tolmie, 2014) included

measures of executive function, semantic inhibitory control and both verbal and visual working memory, and found little sign of general effects of any of these. Semantic inhibitory control was related to quality of descriptions for melting and freezing, however, consistent with the notion that received explanatory frames had to be suppressed for good descriptions of these to be produced.

The implication may be that executive function and working memory become more influential with age, which would be consistent with the relatively late maturation of the former (Zelazo, Carlson & Kesek, 2008). However, it may also be that we need more refined hypotheses of how executive function and working memory affect science learning, especially amongst primary age children. In particular, it may be unreasonable to expect their influence to manifest in relation to cumulative measures of understanding when their operation is much more likely to impact on the specific instances of processing from which cumulative understanding emerges over time. In other words, we should look for their influence on more moment-to-moment decisions, as in Brookman's (2015) task.

These are not the only broader functions that might be relevant. As already noted, sensitivity to observational data shows individual variation which might be attributable to differences in attentional control (cf. Bonawitz et al., 2010, on an apparent relationship between failure to look at the outcomes of actions and failure to engage in spontaneous manipulations). The balance between global and local processing (i.e., a focus on either broad patterns of information or detailed elements within this) may also be relevant to observational skill, and spatial ability may facilitate the building of representations that capture dynamic causal processes (cf. Piaget, 1972; Sanchez & Wiley, 2014). As with executive function and working memory, however, we need specific hypotheses about how these might impact on performance and learning before their potential influence can be properly assessed.

Some conclusions

At present there is no consensus on the nature of the skills involved in science learning, and little in depth analysis of their possible general cognitive precursors among school age children. There is a notable lack of research examining relationships between investigative skills at any level and grasp of scientific concepts, and almost no work looking at the extent to which conceptual grasp in one topic area predicts that in another. We need especially to establish how far skills are related between physics and biology topics, given the distinctions drawn between these domains (Carey, 1985), differences in the complexity of the physical systems involved, the proximal vs distal nature of effects, and contrasting manipulability of phenomena.

There is a clear need then for a systematic and broadly oriented programme of research on the key influences on science learning and the cognitive processes that these imply. The desire to understand these abilities, the processes they draw on and their relationship both to each other and to broader cognitive functions at different ages is more than academic. Their precise nature carries important implications for the design of effective pedagogical strategies in terms of what is promoted when, and how. In particular, we need to establish whether the developmental pattern is a shift from fragmented ability across different skills – and different topics – to gradual integration, or whether development in a specific skill area exerts an influence on growth in others. Uncovering this pattern has crucial implications: gradual integration would imply learning will be best promoted by developing each skill, possibly in a range of topic areas, and then building connections between them; a lead skill pattern would imply instead that effort should be focused on promoting that skill, at least initially. How far these skills and the relationships between them are facilitated or constrained by language and executive function is a further key consideration, because this too will impact on effective sequencing and timing. These are exactly the kinds of point on

which the National Curriculum for Science in England (Department for Education, 2013) is silent.

We believe that the core skills framework set out in this paper provides a clearer model for making progress than the procedural-conceptual dichotomy that has dominated most cognitively oriented research on science learning to date. One signal of its potential utility is its capacity to explain the established success of collaborative group work activities in science, in which learners engage in joint prediction, testing and interpretation of the effects of manipulating variables. Group work of this kind has been found to promote improved attention to disconfirming evidence and the development of more accurate concepts, especially where participants have differing initial conceptions (Howe, Tolmie, Duchak-Tanner & Rattray, 2000; Howe, Tolmie & Rodgers, 1992; Tolmie, Howe, Mackenzie & Greer, 1993). These outcomes have been found to apply not just under 'laboratory' conditions, but in standard classroom settings (Howe, Tolmie, Thurston, Topping, Christie, Livingston, Jessiman & Donaldson, 2007).

One plausible interpretation of these results is that collaborative work brings observations, extraction of patterns of data and explanatory language together under circumstances where there is both a constraint on bias and pressure to articulate ideas in relation to data because of the group context. The process of predicting, testing and explaining the effects of the manipulation of variables may create an informational load that primary age children find difficult to manage – hence their poor individual performance on hypothesis testing activities. However, the presence of others may serve to facilitate its management by effectively distributing it between group members, with each attending to different elements because of their differing initial ideas, and the shared nature of the task creating an incentive to pool these. In terms of the framework presented here, then, collaborative group work is successful because it creates exactly the conditions under which all three core skills need to be coordinated, whilst inherently providing support for the demands of doing so.

Even if this model is *not* correct, however, we still need to collect data across the primary school age range on these dimensions or something like them in order to establish 1) which skills are in fact core to scientific understanding, 2) how these relate to each other and to more general cognitive abilities, 3) how these skills and these relationships change with age, and 4) whether they vary with topic. These data might also allow us to examine why some children seem to have particular difficulties with science learning, and whether there may in fact be as yet unidentified deficits corresponding to dyslexia and dyscalculia in literacy and mathematics.

In the absence of a clear analysis of this kind, there is no certain basis for determining the pedagogical strategies that are likely to be effective or how the science curriculum might be sequentially organised to promote ordered growth in understanding – despite the clear message from work on reading and number that effective teaching rests on exactly this kind of detailed area-specific analysis. The skill set identified by the National Curriculum for Science in England represents a step in the right direction, but we currently have no real knowledge about how these pieces join together. The need to address this lack of understanding is urgent.

References

- Ape, M., Flottmann, J., & Leuchter, M. (2015). Coherence between science content and process knowledge in preschool age. EARLI Biennial Conference, Limassol.
- Bonawitz, E.B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A.N., Woodward, J., & Schulz, L.E. (2010). Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, *115*, 104-117.
- Brookman, A. (2015). The role of inhibitory control in adolescents' science and maths reasoning. Unpublished MSc dissertation, Birkbeck/UCL Institute of Education.
- Bybee, R. W. (1997). Toward an understanding of scientific literacy. In W. Gräber & C. Bolte (Eds.). *Scientific Literacy*. Kiel: IPN.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: conceptual foundations. *Trends in Cognitive Sciences*, *10*, 287-291.
- Cook, C., Goodman, N.D., & Schulz, L.E. (2011). Where science starts: spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*, 341-349.
- Croker, S., & Buchanan, H. (2011). Scientific reasoning in a real-world context: The effect of prior belief and outcome on children's hypothesis-testing strategies. *British Journal of Developmental Psychology*, *29*, 409–424.
- Department for Education (2013). Science programmes of study: Key Stages 1 and 2. National Curriculum in England.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239132/PRIMARY_national_curriculum_-_Science.pdf
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age*. Hillsdale, NJ: Lawrence Erlbaum.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*, 105-225.
- Dowker A., & Sigley, G. (2010). Targeted interventions for children with arithmetical difficulties. In R. Cowan, M. Saxton & A. Tolmie (Eds.), *British Journal of Educational*

Psychology Monograph Series II: Psychological Aspects of Education – Current Trends:
No. 7. Number Development and Difficulty. Leicester: BPS.

Driver, R., Guesne, E., & Tiberghien, A. (1985) *Children's ideas in science.* Buckingham: Open University Press.

Fugelsang, J.A., & Dunbar, K.N. (2005). Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, *43*, 1204-1213.

Fugelsang, J., & Mareschal, D. (2014). The development and application of scientific reasoning. In D. Mareschal, B. Butterworth & A. Tolmie (Eds.), *Educational Neuroscience.* Chichester: Wiley-Blackwell.

Fugelsang, J., Stein, C., Green, A., & Dunbar, K. (2004). Theory and data interactions of the scientific mind: Evidence from the molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology*, *58*, 132-141.

Gathercole, S.E., Pickering, S.J., Knight, C., & Stegman, Z. (2004). Working memory skills and educational attainment: evidence from National Curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*, 1–16.

Gelman, S. A. (2009). Learning from others: children's construction of concepts. *Annual Review of Psychology*, *60*, 115–40.

Ghazali, Z. (2014). New approaches to assessing the development of biological understanding in young children. BPS Developmental Section Annual Conference, Amsterdam.

Gopnik, A., Sobel, D.M., Schulz, L.E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three- and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*, 620-629.

Harris, P.L., & Corriveau, K.H. (2014). Learning from testimony about religion and science. In E.J. Robinson & s. Einav (Eds.), *Trust and skepticism: children's selective learning from testimony.* Hove: Psychology Press.

Harris, P. L., & Koenig, M. (2006). Trust in testimony: how children learn about science and religion. *Child Development*, *77*, 505-524.

- Howe, C. (1998). *Conceptual structure in childhood and adolescence*. London: Routledge.
- Howe, C. (2014a). The development of scientific reasoning. In A.J. Holliman (Ed.). *The Routledge International Companion to Educational Psychology*. London: New York.
- Howe, C. (2014b). Commentary II: 'If you've seen it before, then you know': physical evidence and children's trust in testimony. In E.J. Robinson & s. Einav (Eds.), *Trust and skepticism: children's selective learning from testimony*. Hove: Psychology Press.
- Howe, C.J., McWilliam, D., & Cross, G. (2005). Chance favours only the prepared mind: incubation and the delayed effects of peer collaboration. *British Journal of Psychology*, 96, 67–93.
- Howe, C., Taylor Tavares, J., & Devine, A. (2014). Children's conceptions of physical events: explicit and tacit understanding of horizontal motion. *British Journal of Developmental Psychology*, 32, 141-162.
- Howe, C., Tavares, J.T., & Devine, A. (2012). Everyday conceptions of object fall: explicit and tacit understanding during middle childhood. *Journal of Experimental Child Psychology*, 111, 351-366.
- Howe, C.J., Tolmie, A., Duchak-Tanner, V., & Rattray, C. (2000). Hypothesis testing in science: group consensus and the acquisition of conceptual and procedural knowledge. *Learning & Instruction*, 10, 361-391.
- Howe, C.J., Tolmie, A., & Rodgers, C. (1992). The acquisition of conceptual knowledge in science by primary school children: group interaction and the understanding of motion down an incline. *British Journal of Developmental Psychology*, 10, 113-130.
- Howe, C.J., Tolmie, A., & Sofroniou, N. (1999). Experimental appraisal of personal beliefs in science: constraints on performance in the 9 to 14 age group. *British Journal of Educational Psychology*, 69, 243-274.
- Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: towards organisational principles for supporting pupil learning. *Learning and Instruction*, 17, 549-563.

- Hulme, C., & Snowling, M.J. (2009). *Developmental disorders of language learning and cognition*. Chichester: Wiley-Blackwell.
- Karmiloff-Smith, A. (1992). *Beyond modularity: a developmental perspective on cognitive science*. Cambridge: MIT Press.
- Klahr, D. (2000). *Exploring science: the cognition and development of discovery processes*. Cambridge: MIT Press.
- Klahr, D. (2005). A framework for cognitive studies of science and technology. In M. Gorman, R. D. Tweney, D.C. Gooding, & A. P. Kincannon (Eds.), *Scientific and technological thinking*. Mahwah, NJ: Lawrence Erlbaum.
- Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology, 64*, 141-152.
- Koerber, S., Mayer, D., Osterhaus, C., & Schwippert, K. (2015). The development of scientific thinking in elementary school: a comprehensive inventory. *Child Development, 86*, 327-336.
- Kuhn, D., Iordanou, K., Pease, M., & Wirkala, C. (2008). Beyond control of variables: what needs to develop to achieve skilled scientific thinking? *Cognitive Development, 23*, 435-451.
- Legare, C.H. (2012). Exploring explanation: explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development, 83*, 173-185.
- Leslie, A. M. & Keeble, S. (1987). Do six-month old infants perceive causality? *Cognition, 25*, 265-288.
- Masson, S., Potvin, P., Riopel, M., & Brault Foisy, L-M. (2014). Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. *Mind, Brain and Education, 8*, 44-55.

- Nayfield, I., Fuccillo, J., & Greenfield, D.B. (2013). Executive functions in early learning: extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences, 26*, 81-88.
- Penner, D. E., & Klahr, D. (1996). The interaction of domain-specific knowledge and domain-general discovery strategies: a study with sinking objects. *Child Development, 67*, 2709–2727.
- Philips, S., & Tolmie, A. (2007). Children's performance on and understanding of the Balance Scale problem: the effects of parental support. *Infant and Child Development, 16*, 95-117.
- Piaget, J. (1972). *The principles of genetic epistemology*. New York: Basic Books.
- Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology, 31*, 153–179.
- Pine, K.J., Lufkin, N., & Messer, D. (2004). More gestures than answers: children learning about balance. *Developmental Psychology, 40*, 1059-1067.
- Rhodes, S., Booth, J., Campbell, E., Blythe, R., Delibegovic, M., & Wheate, N. (2012). The role of executive functions in science learning: chemistry and cognition. Paper presented at the British Psychological Society Developmental Section Annual Conference, Glasgow.
- Sanchez, C.A., & Wiley, J. (2014). The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction, 31*, 33-45.
- Schauble, L. (1990). Belief revision in children: the role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49*, 31-57.
- Schauble, L, Klopfer, L.E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching, 28*, 859-882.
- Schulz, L.E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science, 10*, 322-332.

- Shayer, M., & Ginsburg, D. (2009). Thirty years on – a large anti-Flynn effect? (II): 13- and 14-year-olds. Piagetian tests of formal operations norms 1976–2006/7. *British Journal of Educational Psychology*, 79, 409-418.
- Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, 28, 303–333.
- Soltesz, F., Szucs, D., & Szucs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: a developmental study. *Behavioral and Brain Functions*, 6:13, <http://www.behavioralandbrainfunctions.com/content/6/1/13>
- Symons, G., Tolmie, A., & Oaksford, M. (2015). Source reliability in the development of children's understanding of causal systems. EuroAsianPacific Joint Conference on Cognitive Science, Turin.
- Tolmie, A. (2012). Understanding core skills and influences in primary school science learning: taking a scientific approach. Institute of Education Professorial Lecture Series. London: Institute of Education Press.
- Tolmie, A. (2014). Observation, description and explanation in primary school science. BPS Developmental Section Annual Conference, Amsterdam.
- Tolmie, A., Howe, C.J., Mackenzie, M., & Greer, K. (1993). Task design as an influence on dialogue and learning: primary school group work with object flotation. *Social Development*, 2, 183-201.
- Vosniadou, S. (2002). Mental models in conceptual change research. In L. Magnani & N. J. Nersessian (Eds.). *Model-based reasoning: science, technology, values*. New York: Kluwer Academic/Plenum.
- Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology*, 11, 1–17.
- Vosniadou, S., & Ioannides, C. (1998). From conceptual development to science education: a psychological point of view. *International Journal of Science Education*, 20, 1213–1230.

- Wellington, J.J. (1988). The place of process in physics education. *Physics Education*, 23, 150-155.
- Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal*, 34, 691-710.
- Zelazo, P.D., Carlson, S.M., & Kesek, A. (2008). The development of executive function in childhood. In C.A. Nelson and M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience 2nd Edition*. Cambridge, MA: MIT Press.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27, 172-223.