# Understanding evidence in scientific disciplines: identifying and mapping ‘the thinking behind the doing’ and its importance in curriculum development

**Ros Roberts**

**School of Education, Durham University.**

**Rosalyn.Roberts@dur.ac.uk**

**0191 3348394**

## Abstract

Understanding research and the uncertainty of the resultant data and claims are widely held aims of HE science curricula. Expertise in scientific research involves, inter alia, an understanding of evidence which can be specified – the concepts of evidence – and represented as a conceptual network of ideas an understanding of which underpins decisions made during research and also forms the basis of evaluation of others’ research; the ‘thinking behind the doing’ of research expertise. Students’ responses to a module that explicitly teaches the concepts of evidence so that they can conduct open-ended investigations and evaluate others’ research are reported. The implications of this conceptualisation of expertise for curriculum specification, approaches to teaching and assessment and for Threshold Concept research are discussed.

## Key words

scientific research, expertise, network thinking, concepts of evidence, higher education, curriculum development, threshold concepts

## Introduction

In scientific disciplines, students’ understanding of research and the uncertainty of the resultant data and claims has been the focus of much recent work in different disciplines. Various potential Threshold Concepts (TCs: see for instance Meyer & Land, 2003; 2005; 2006) associated with this have been proposed including ‘measurement uncertainty’ (Wilson et al., 2010), the ‘testable hypothesis’ (Ross et al., 2010), ‘uncertainty in climate change’ (Hall, 2010), ‘quantitative numeracy’ (Frith & Lloyd, 2013) and ‘academic numeracy’ (Quinnell et al., 2013). All these different aspects are related, at their heart, to students' understanding of research where weight is given to data as evidence.

TC-inspired research has, in particular, emphasised the importance of teaching for understanding, 'disciplinary thinking' and making the implicit explicit. Expertise in a discipline includes understanding its research, yet this has been considered hard to articulate since research expertise has widely been viewed as encompassing tacit understanding (Polanyi, 1966; Sternberg, 1999; Kinchin & Cabot, 2010). How can we frame a curriculum based on tacit knowledge?

## Expertise

Kinchin (2008) has analysed expert practice, especially clinical practice, and has depicted 'chains of practice' that are seemingly manifest by experts from the underlying 'networks of understanding' which underpin and inform the practice (Figure 1). This paper considers expertise in scientific research from this perspective.

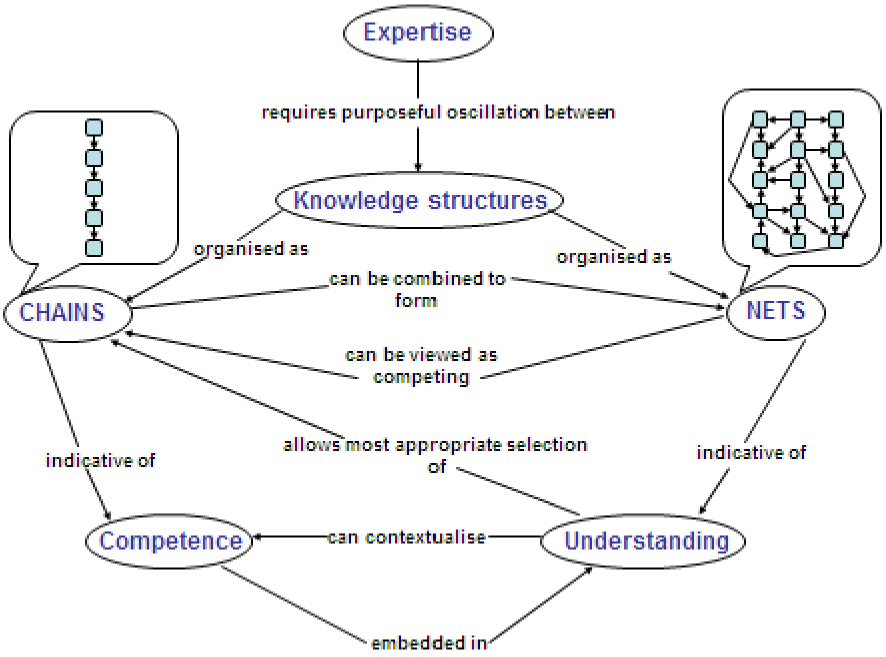

**Figure 1:** A dual-processing knowledge structures perspective on the nature of expertise (from Kinchin & Cabot (2010), p. 161)

### ***Expertise in scientific research***

Expertise in scientific research may, superficially have the appearance of a 'chain-like' activity but closer examination reveals some interesting details.

The seemingly chain-like expertise depicted on the left of Figure 1 involves making many decisions - 'thinking on your feet' - which draw on the seemingly tacit knowledge depicted on the right. However, close examination of expert scientific practice where researchers seek solutions to their research problems - the solution being to establish a valid pattern in their data that will be 'good enough' for the claims to be made – the resultant practice is seldom linear. Instead, the decisions made during practice are based on conducting trials and subsequent examination of the effects of all the decisions on the resultant data. The expert engages with the problem to see 'what works' and the way of working is more iterative than linear (Roberts, Gott & Glaesser, 2010) until the route to a solution is refined. The collaborative working and decision making familiar to researchers in science (see, for instance, Roth, 2013) also point to limitations of the linear model of practice. More 'chain-like' practice would be observed once all the 'design' decisions had been made although the expert researcher would always have an eye on the quality of the data as it was collected in relation to the ability to make a claim and this would influence ongoing decisions about, for instance, the amount of data collected and any further modifications required. Research practice is often iterative rather than linear practice. The 'quick thinking' of experts addressing these issues might give the appearance of chain-like practice; as do formal write ups of research which, by convention, report the end results of all the decisions and present the research as linear.

There are times, however, when expert practice *is* chain-like. For instance, 'thinking' is deliberately minimised during the enactment of specific procedures and protocols which ensure quick and standardised routes to a solution. In such 'routines' the 'thinking' to establish the procedure (using the network of underpinning ideas on the right of Figure 1) has already happened and few, if any, further decisions are required on the part of the practitioner – other than in the protocol's correct selection. Similarly, practice by experts in very familiar contexts would also appear more chain-like since nearly all the decisions requiring 'thinking' would have been made previously.

### *Teaching for scientific expertise*

The challenge for the lecturer is how to help develop these components of expertise in their students.

Traditionally, research depicted as a chain of practice has predominated in science education (in schools and in HE). Chains of practice can arguably be seen in some research methods courses, laboratory procedures, fieldwork protocols and study guides (Kinchin & Cabot, 2010). ‘Process skills’ such as hypothesising, planning, collecting, analysing and evaluating describe practice, yet arguably these provide little guidance of *what* to teach such that students can carry out these processes. As ‘skills’ they may be assumed to develop through practice, by imitation, and are often characterised by performance. Yet there are a limitless number of such ‘chains’ in science, thus presenting a challenge to lecturers. How, with the limitations of time and opportunity inherent in any undergraduate teaching, can students develop expertise if taught as such chains? That is not to say that students should not be taught the ‘chains’ of specifically selected protocols or techniques; after all, such standardised routines have been established to save ‘reinventing the wheel’ and address issues of standardisation and quality. Also, since ‘write ups’ of scientific research present practice as a linear chain, students need to be familiar with such conventions. Yet we know that students do not always develop a deep understanding of the net of ideas (on the right of Figure 1) if research is presented as practice (Kinchin et al., 2010; Roberts et al., 2010) and it is to that which we now turn.

So *what* is it that experts understand that forms the ‘network of understanding’, in Figure 1, about scientific research? In addition to the all-important substantive concepts, experts also understand evidence – the key over-arching concepts for understanding the quality of evidence being validity and reliability, underpinned by more detailed and inter-related concepts. The TC research (cited earlier) all identifies ideas like these as being important for students so that they can apply this understanding when making decisions ‘looking forward’ while solving research problems (in a variety of different ways, with different research designs appropriate to different subjects; e.g. in lab-based manipulations of variables, fieldwork, observations, RCTs) and when ‘looking back’, evaluating the quality of evidence in their own or others’ research (Gott & Duggan, 2003a).

Gott et al. (n.d.) have tentatively articulated the ‘concepts of evidence’, a phrase that emphasises the knowledge-base inherent in the ‘thinking behind the doing’; and have validated them against the work of experts (see for instance Gott, Duggan & Johnson, 1999). The concepts of evidence have informed the curriculum for 20 years of teaching and research into undergraduates’ understanding of evidence in Durham. For instance, we have found that an understanding of the ideas of evidence was a necessary condition for success in open-ended investigations (Glaesser et al., 2009) and that being taught about evidence enabled students to ask better questions about others’ research (Roberts & Gott, 2010).

Since ‘the thinking behind the doing’ is a knowledge base of concepts to be understood (rather than ‘processes’ to be mastered) such ‘networked understanding’ can be articulated (just as it can for the substantive ideas of science) and shown on a concept map (Roberts & Johnson, submitted).

### Networked understanding of the concepts of evidence

The concepts of evidence list (Gott et al., n.d.) turns out to be very useful when considering the science curriculum; the specification of these ideas has provided us with a domain specification for both teaching and assessment. Some may view the detailed list of ideas as implying a very reductionist approach. We do not intend this. By listing the ideas, which together can be constructed into a networked understanding – an understanding about evidence – we have attempted to specify the knowledge-base, reducing any possible ambiguity inherent in more general descriptions.

The extensive list of the concepts of evidence is unwieldy for some descriptive purposes (but critical for curriculum definition purposes). We have grouped them together into subsets, or nested layers (Figure 2, based on Roberts & Gott 2006). Essentially, understanding is about the reliability and validity in each layer and the networks between the layers. The inner three layers focus on the ideas associated with the conduct of an investigation while ideas in the other layers are important where the validity and reliability of the evidence may be affected by the broader context for the claim and its link to existing findings.

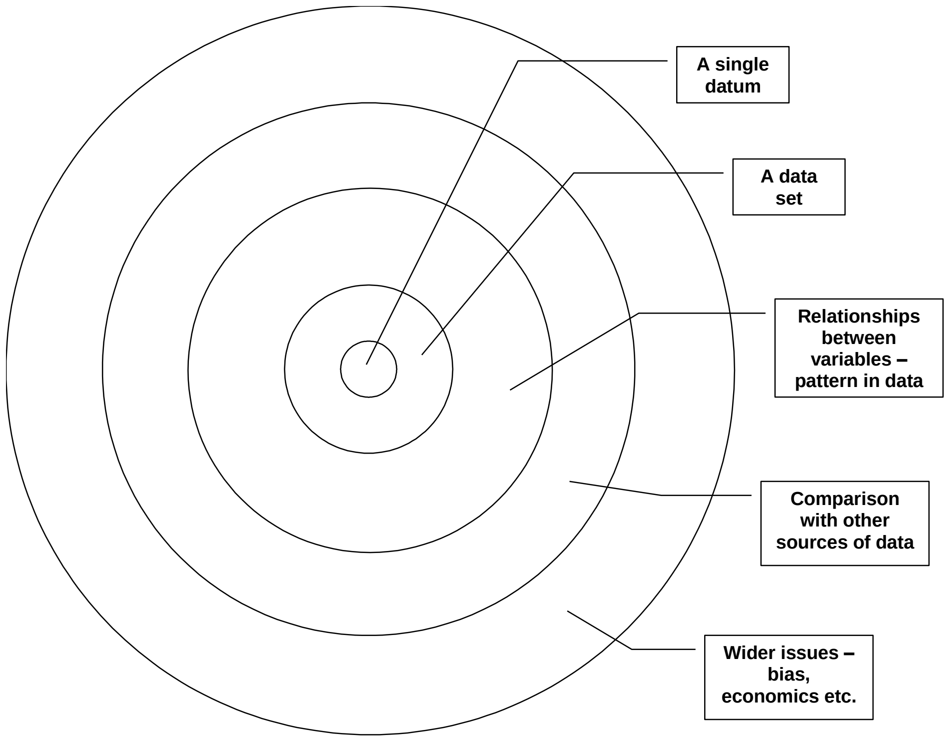

**Figure 2:** A framework for understanding the concepts of evidence (based on Roberts & Gott 2006).

### *A single datum*

In any empirical investigation something will need to be measured, either qualitatively or quantitatively. That may be the height of a tree, or the colour of a precipitate, or the speed of an object. This layer is to do with the making of a single measurement of that variable. It is at the heart of science and critical for the quality of evidence. The ideas in this layer must take into account, *inter alia*, the range and sensitivity of the instrument, the validity of the measurement and its accuracy. If the measurement is invalid or unreliable then the validity of the whole investigation and any claim made from it is called into question, which is why it is presented at the heart of the figure. But, of course, one measurement is not always enough and repeats may well be necessary.

### *A data set*

We are led, then, into the next layer which includes ideas to do with the validity and reliability of repeated measurements of the same variable under the same conditions: whether sufficient repeated readings have been taken to capture the variation – whether that is variation in the sample or variation inherent in the measurements themselves.

### ***Design and relationships***

Moving on, we arrive at the crux of the investigation; the establishing, or otherwise, of a relationship between one or more of the variables. When seeking relationships between variables – whether in controlled lab-based conditions or in surveys of naturally changing variables – the validity of the design must be considered as well as how the reliability of the data affects the interpretation of the relationship. For instance, decisions will need to be made about the range of the independent variable which will be needed to establish any potential relationship as well as the interval between such readings and also how potentially confounding variables are treated and their effect on the validity.

These three inner layers represent the core concepts of evidence in any investigation. These inner layers in turn may have been influenced by ideas subsumed in the other three layers: ideas from other similar research, which must itself be evaluated; the background of the investigators themselves and how this might affect the quality of the evidence being collected; and also any potential economic and social pressures influencing the design and conduct of the investigation. These ideas could potentially bias any stage of the data collection. An evaluation of the validity and reliability of the whole investigation should take account of all of these potential influences.

In an investigation, the ideas are put together (‘looking forward’) with a view to making a defensible claim. In a scientific literacy context, or in peer review of others’ work, the claim must be deconstructed by ‘looking back’. The layers do not imply any sequence. An investigator will make decisions using their understanding of each layer and its relationship to other layers.

During research, decisions are made about variables and their measurement such that the variation in any data in relation to the magnitude of any effect being examined enables a qualified claim to be made. The same ideas are important as a basis for decision making whether the research adopts a more ‘experimental approach’ (wherein often homogeneous variables can be isolated and manipulated and the biggest contributor to uncertainty in the data is often from the instrument and its use) or an ‘observational approach’ (usually involving less homogeneous and less easily isolated variables, whose values cannot usually be manipulated and with resultant variation in the data). Roberts and Johnson (submitted) illustrate how decisions based on an understanding of evidence, which can be represented on a concept map, underpin a diverse range of practices and that the map does not privilege any one approach.

Decisions when investigating – seen in experts' iterative ways of working - are based on nuanced application of the ideas, involving mental juggling as juxtapositions and contingencies are considered according to context. The creators of established 'chain-like' procedures and protocols have previously employed such an understanding. When evaluating research, the same 'joined up' thinking is used to determine whether the claims are supported by the evidence.

The intimate integration of substantive knowledge with evidence is essential for scientific practice. Neither stands alone; each is only as good as the other. While the substantive knowledge of each discipline will vary, the ideas of evidence are applicable across all scientific disciplines. Thus the map can be considered to make explicit previously 'underspecified' disciplinary knowledge.

## Reflections on teaching and learning from this networked thinking perspective

Our experience has been mainly working with undergraduate BA initial teacher training (pre-service) students, many of who have not studied science beyond GCSE (aged 16) and who tend to have low confidence in their scientific abilities. They may not be typical of all UG students. However there are some lessons from our work which may be of value to others.

In the following sections I will draw on our experience of 20 years of teaching about evidence in Durham and research that we have conducted. I include students' comments, gathered from their reflective journals and interviews, to give insights into the learners' perspective.

To understand science requires both a substantive understanding *and* an understanding of evidence. Viewing expertise as including a network of ideas about evidence to be understood has significant implications for teaching and learning which we have expanded on further elsewhere (see papers listed in Gott et al., n.d.).

### *The ideas can be specified and sequenced*

This enables the curriculum developer to plan systematically for coverage rather than relying on providing multiple experiences of diverse 'chains of practice' (wherein the understandings are implicit). We teach these ideas as a distinct element within a module (9 hour-long lectures, each followed by 90 minute workshops) and this model has been replicated successfully in Turkey (Roberts & Sahin-Pekmez, 2012).

In terms of sequence, we focus on different sections of Figure 2 in our teaching, starting with variables with little inherent variation in contexts where values can be manipulated and then considering how a valid design might be established in situations where they can't. Issues to do with the quality of measurement are then addressed so that students get an understanding of variation in the data before other sources of variation, such as in the sample, are introduced.

The focus on making the underpinning ideas explicit is different to many students' experience. With regard to exploring ideas associated with the validity of design a student commented in her reflective journal:

ID 72: *"Interesting to look at the variables to look what actually starts an investigation. I was used to having aims, objectives etc rather actually looking what is involved."*

For many, an instrument was just something they used, unquestioningly:

ID 12: *"Never realised it was an error about calibration during high school science experiments. It was always assumed that that was the right measurement"*.

Sometimes students have vague memories of having been taught about uncertainty associated with instruments, but when asked in an interview whether they'd ever been asked to apply this in their selection of an instrument in an investigation, this was a typical response:

ID 54: "*No. No, not at GCSE and neither at A level. I never… I may have done but it was just like learning the facts then just writing the facts on the exam paper. That was how I did it. Just… Not like making decisions on…. which is the most reliable and why I think it is the most reliable one to use".*

### *The ideas and their links can be articulated and shared*

Students then have a language to express their thinking and Figure 2, together with the list of the concepts of evidence (Gott et al., n.d ) and the concept map (Roberts & Johnson, submitted) can provide a structure, making the implicit links explicit. In our experience most students, after their school science education, are aware of many of the ideas but they are unable to express them or 'join them up'. For instance as these interview transcripts attest, prior to teaching, students are aware of the need to repeat readings but find it difficult to articulate why they are necessary and the links with other ideas they are aware of are not clear:

*RR: "Why do you think scientists may do repeated readings?"*
*ID2: "I think, err, they probably do repeated readings [pause] to, you know, make sure there's some sort of consistency. You can't just do one reading and say 'Oh that must be right', you need to, um, be able to, err, sort of validate it, check again and again, err …"*
*RR: "OK. What do you think they're checking for?"*
*ID 2: "To make sure, err, the variables remain [pause]. I don't really know how to put it in words. Err. They're checking for ….*"
*"ID 5: "Um. Just to make sure that the data's going to be more, um, equal. Um. Like, ensure that everything, the apparatus, aren't giving us false readings. Um, like, [laughs] make sure there's not other things, like, influences that are happening."*

Since much school work has been in lab-based 'fair testing' contexts with relatively straightforward measurements students, in our experience, find decisions requiring sampling of variables, matching of confounding variables and complex measurements much harder to 'juggle'. A recognition of the relationship between their decisions across the whole of Figure 2, their effect on the data and how this affects the strength of the claim is often lacking, but many students respond well to making the connections when the ideas are taught explicitly:

ID 33: "*I kind of liked being able to apply the things that I had learnt and the ways that I think about stuff and the science that I know. And apply the things that I know about errors and reliability of sources and… Yeah, it was kind of interesting."*
RR: "*So there was something about the complexity of it?"*
ID 33: "*Yeah, kind of. Yeah, I guess I liked having all these things that I had to think about. And I liked the idea of having to reach a conclusion that I thought was valid and kind of scientifically justifiable. That sort of thing."*
RR: "*Is that a learning activity that you're familiar with, or you did often?"*
ID 33: "*No. No, not at all. I've never done anything like that before. … It's very much been, before, kind of: here is the stuff, look at it, find out the answer, write the answer, tick! That sort of thing. I mean, I've had to kind of work it out myself but there's never been all these different variables and a real life situation. I don't think that I've had to kind of think through in such a way whilst remembering that it might affect me or other people or that sort of thing. That made me want to make sure I did it alright and take everything into account, I suppose."*

A student reflected in his journal after teaching:

ID 10: *"Now I see. When you take the questions apart then put them together it all makes sense. This actually is easy. I can see how this fits in!"*

The nuanced links between the ideas (representing higher order thinking) can open up new insights compared to the ritualized approach familiar to many from school, as this reflective journal entry shows:

ID 40: “*It is this idea of having a “fair test” however; it is not always simple to control certain things such as light intensity when you do an experiment outside, temperature, or humidity. … This idea of a fair test also ignores the fact that results are not always perfect and you don’t always get the desired result.”*

***Students require opportunities and time to develop this ‘network thinking’***

Since this approach represents a conceptual basis for understanding evidence, decisions can be made about the sort of activities best used to teach the ideas. As with the substantive ideas of science, these can involve using both practical and ‘non-practical’ opportunities.

Practical work is important in developing this understanding – during open-ended investigations students can make these decisions for themselves and see the effect of their decisions on the quality of the data and recognize this affects the strength of their claim. In genuinely ‘open-ended’ contexts where students are not focusing on getting ‘the right answer’ they can focus on the ‘trade-offs’ between sections of the map and their practice is characterized by trials and iterative working. For many this way of working is completely new to them and they find it unsettling – after all, they are familiar with ‘chains’ from their previous teaching and in accounts of research and not having ‘a right answer’ throws them. That this is ultimately satisfying is a point made by many:

ID 14: *“Because it’s a different view of what… I don’t know, A-levels and GCSEs just seem to be jumping through loopholes and memorising syllabuses and I don’t feel… like, here I feel like I’m being educated and I’m becoming, not more clever but just more integrated. All my knowledge is becoming more integrated,”*

We have found that practical contexts which generate ‘messy’ data highlight issues that can lead to fruitful discussions. The use of non-specialised, ‘everyday’ equipment, and instruments that students have had to create themselves, force students to confront the quality of their own evidence and better appreciate the decisions that must have gone into the design of instruments and protocols.

Asking students to write an account of their investigation as if ‘thinking aloud’, asking them to detail all the decisions that they’ve made and justifying them, has also been helpful.

We have also developed a database of real, messy, data (collected from students’ investigations over the years) which the students can use to simulate the effects of different decisions on the data, such as working with different designs, using different quality instruments and with different numbers of repeated readings (similar to that in Gott & Duggan, 2003b). This has the advantage of being much quicker than conducting many separate investigations and can be used by students to support out-of-class study (Gott, Duggan & Roberts, 1999). A student reflected in her journal:

> ID 20: "*Completed the two 1st activities on The Science ICT workshop. Really interesting program and easy to use - good demonstration in the lecture was helpful. Started the notebook for it too!. Feel like I'm starting to learn more about evidence- this was really helpful for thinking about different variables."*

We have also found that students learn lots from discussion involving their own and others' data. Acting formatively as 'peer reviewers' of others' investigations we have noticed, as did Nicol et al. (2014, p. 102) that 'producing feedback reviews engages students in multiple acts of evaluative judgement, both about the work of peers, and, through a reflective process, about their own work; that it involves them in both invoking and applying criteria to explain those judgements'. A reflective journal entry reflects the student's concerns about doing this:

> ID 16: *"It was useful to do the peer review sheets, although I was worried that because I wasn't totally confident about what I'd written, I didn't want to tell someone else what they'd done was right/wrong- although the review sheets did help. I've made my own version for my own revision including the key questions. I think I might understand this a little more, I just hope I can remember it in the exam! It was a useful activity and hopefully it will help me to write a well-structured claim in the exam".*

***Data generation seems to aid understanding of data handling and presentation***

Having generated their own data, students seem better able to understand the sources of variation in it. The importance and meaning of simple statistical tests and graphical forms of data presentation can then be appreciated more in our experience. Having grappled with 'messy data' and discussed how best to get meaning out of it students appear to be in a better position to understand the conventions employed in handling data and presentation.

> ID 5: "*This session has helped me to develop my knowledge further and learn to begin to break down more difficult concepts in order to see them as simply as this. I understood the frequency histogram completely and felt happy with the outcome of my ICT based graph."*

A student taught by us who had previously attended a traditional 'stats workshop' reported:

ID 22: *"I think this was a great way of bringing in that statistics stuff. I mean, you know that statistics workshop we had? It was awful. I thought it was absolutely horrendous. They presented you a load of formulas so no-one who hadn't done it before had any clue what they were doing. They just copied the formulas out and now don't know how to do it again."*

Journal entries record students' relief at finally understanding some of the stats:

ID 72: *"I found today quite interesting- it was useful to think about how and why standard error is used and how this helps support a claim. I actually feel like I'm understanding it a bit better ☺ I thought the bar graph with the different standard error confident levels was a great visual representation which supported the written aspect and helped me think about how the data can 'overlap' and how that may show if the population samples are the same (hopefully I'm thinking the right thing!)."*

ID 9: "*Both standard error and standard deviation has clicked!! Yey!! I feel quite proud of myself!"*

***A focus on evidence in simple contexts***

Although the substantive context of any research intimately informs the ideas of evidence, the same underpinning ideas are arguably relevant to all sciences and all scientific approaches. We have found that teaching these understandings, initially, in simple contexts (with very low substantive demand) enables students to focus on these ideas and develop their understanding of evidence before they then ought to be able to consider them in more substantively complex contexts. The focus of the sessions is on developing an understanding of evidence in contexts where students' potential limitations in their substantive understanding do not 'get in the way'.

ID 5: "*At the beginning I was really worried. I felt as though I'd have very little understanding of what was being taught and I felt my scientific knowledge could not meet the standard being addressed. After listening carefully to the 'real' life examples given I soon began to understand the topic and could develop a more thorough understanding of the concepts surrounding the lessons … During the workshop I felt I was 'keeping up' with the lesson and could carry the tasks out with confidence*."

***Ideas about evidence can be assessed***

Having specified the ideas and their inter-relationship, these understandings can be assessed. In terms of assessment tasks, critical evaluation of accounts of others' work, wherein the understandings in Figure 2 are applied, have provided particularly fruitful opportunities for both formative and summative tasks.

ID 20: *"Completed the formative task. I actually understood it, and attempted to complete it without looking at the book or my notes. Hopefully this will show me how much I know already and what I need to work on to improve my knowledge. It was really useful to complete as it helped me think about what*

> *we have been reading and doing in lectures and gave me a rough idea of what the exam will be like. … Will be intrigued to know how well I'm doing."*

We have also given students claims made by scientists and have asked them what they would like to ascertain about the scientists' research before deciding whether to accept the claim (Roberts & Gott, 2010). This requires the students to ask questions about the quality of the evidence which draws on the same ideas. For instance when provided with a claim that emissions from a chimney were 'safe', they asked more questions than before the teaching, such as:

> '*Did you collect enough data?*'; '*Was it consistent?*' and '*What was the sample?*'

Other questions were about what could be inferred from the data set:

> '*What is the uncertainty?*'; '*What is the significance level?*'; '*What is the probability of it being safe?*'

Both students' own investigations and their evaluations of others' accounts can be assessed in terms of their understanding and application of the concepts of evidence, rather than on some less-clearly defined notion of performance of 'process skills'. Students (and lecturers in their feedback) have found the clarification provided by the domain specification – i.e. the assessment of the application of their understanding of the concepts of evidence – particularly useful in terms of knowing what is required and how to improve.

## Conclusions

Kinchin and colleagues' model of expertise (Figure 1) has been developed in this paper to describe scientific research expertise. While some scientific practice does consist of linear chains these can be seen as the result of *prior* decision making; either on the part of the expert who developed the protocol, or by the expert when they had previously encountered the situation. An iterative process of determining 'what works', through trials and examination of the data, better reflects the work of expert researchers. The network of ideas about evidence (the key elements of which are shown in Figure 2) arguably underpin all this practice, regardless of disciplinary context, and the fact that it can be represented as a concept map (Roberts & Johnson, submitted) emphasises its conceptual basis.

There has been very little systematic research from this perspective into the teaching and learning of these ideas in HE and more is called for. The concepts of evidence have been validated against the work of experts in research and industry but how do

HE curriculum developers take to them? If the concepts of evidence are considered as an extension to the disciplinary knowledge-base for the curriculum, how might the curriculum best be organised to include it?

Since this networked thinking arguably underpins research in all scientific disciplines, despite their seemingly diverse research traditions and specialist substantive knowledge, interdisciplinary work would seem to be particularly important. What can we learn from each other to best develop research expertise in our students? Since understanding research is an important curriculum aim in HE, other disciplines that give weigh to data as evidence may also find this work of interest.

And finally, mapping out the conceptual basis of understanding evidence may also help to specify the understandings inherent in the TCs concerned with research expertise, since the TCs are all, at their heart, about the quality of evidence. The many conventions associated with disciplinary practice (whether that is about, for instance, the null hypothesis; accepted ways of quantifying uncertainty; or data presentation) may be better understood once the underpinning understanding about evidence, implicit within these conventions, is addressed. Disciplinary and inter-disciplinary research into the development and application of the map to address students' understanding of identified TCs associated with the disciplinary epistemes would help move the discourse from the identification of TCs to a focus on the networked understanding - which can then be explicitly taught - necessary for research expertise and practice.

[Word count = 5235]

## References

Frith, V., & Lloyd, P. (2013). *Quantitative Literacy and Epistemological Access at University: Reflections on Using the Threshold Concept Framework for Research*, Mathematics Education and Society 7th International Conference, 2-7 April 2013, Cape Town, South Africa.

Glaesser, J., Gott, R., Roberts, R., & Cooper, B. (2009): Underlying success in open-ended investigations in science: using qualitative comparative analysis to identify necessary and sufficient conditions. *Research in Science and Technological Education*, 27**(**1), 5-30.

Gott, R., & Duggan, S. (2003a). *Understanding and using scientific evidence*. London: Sage.

Gott, R., & Duggan, S. (2003b). *Science investigations in a virtual laboratory: Building Success in Sc1 Science* (CDRom), Dunstable: Folens Publishers.

Gott , R., Duggan, S., & Johnson, P. (1999). What do Practising Applied Scientists do and What are the Implications for Science Education? *Research in Science & Technological Education*, 17(1), 97-107.

Gott, R., Duggan, S., & Roberts, S. (1999). The science investigation workshop. *Education in Science.* Issue 183, 26-27.

Gott, R., Duggan, S., Roberts, R., & Hussain, A. (n.d.) *Research into understanding scientific evidence.* http://www.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm

Hall, B. M. (2010). *Teaching uncertainty: the case of climate change*. PhD Thesis. Gloucester: University of Gloucester.

Kinchin, I. M. (2008). *The qualitative analysis of concept maps: Some unforeseen consequences and emerging opportunities*. In A. J. Cañas, P. Reiska, M. Åhlberg & J. D. Novak, Eds. Concept Mapping: Connecting Educators. Proc. of the Third Int. Conference on Concept Mapping Tallinn, Estonia & Helsinki, Finland 2008.

Kinchin I. M., & Cabot, L. B. (2010). Reconsidering the dimensions of expertise; from linear stages to dual processing. *London Review of Education* 8(2), 153-166.

Kinchin, I. M., Cabot, L. B., & Hay, D. B. (2010). *Visualising expertise: revealing the nature of a threshold concept in the development of an authentic pedagogy for clinical education* (pp81-95). In Meyer J. H. F., Land, R., & Baillie, C. (Eds.). Threshold Concepts and transformational learning, Rotterdam: Sense Publishers.

Meyer, J., & Land, R. (2003). *Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practicing within the disciplines*. Enhancing teaching-learning environments in undergraduate courses: Occasional report 4. 1 – 12. Available online at: www.ed.ac.uk/etl/docs/ETLreport4.pdf

Meyer, J. H. F., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education*, 49: 373-388.

Meyer, J. H. F., & Land, R. (Eds.). (2006). *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge.* Oxford: Routledge.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment and Evaluation in Higher Education*, 39(1), 102-122.

Polanyi, M. (1969). *Knowing and Being.* Chicago: University of Chicago Press.

Quinnell, R., Thompson, R., & LeBard, R. J. (2013). It's not maths; it's science: exploring thinking dispositions, learning thresholds and mindfulness in science learning. *International Journal of Mathematical Education in Science and Technology*, 44(6), 808-816.

Roberts, R., & Gott, R. (2006). Assessment of performance in practical science and pupil attributes. *Assessment in Education,* 13(1), 45–67.

Roberts, R., & Gott, R. (2010). Questioning the evidence for a claim in a socio-scientific issue: an aspect of scientific literacy. *Research in Science & Technological Education*, 28**(**3), 203 – 226.

Roberts, R., Gott, R., & Glaesser, J. (2010). Students' approaches to open-ended science investigation: The importance of substantive and procedural understanding. *Research Papers in Education*, 25(4), 377-407.

Roberts, R., & Johnson, P. (submitted) Understanding the quality of data: a concept map for ‘the thinking behind the doing’ in scientific practice.

Roberts, R., & Sahin-Pekmez, E. (2012). Scientific evidence as content knowledge: A replication study with English and Turkish pre-service primary teachers. *European Journal of Teacher Education,* 35(1), 91-109.

Ross, P. M., Taylor, C. E., Hughes, C., Kofod, M., Whitaker, N., Lutze-Mann, L., & Tzioumis, V. (2010). *Threshold Concepts: challenging the way we think, teach and learn in biology* (pp. 165-178), in Meyer J. H.F., Land, R., & Baillie, C. (Eds.). Threshold Concepts and transformational learning. Rotterdam: Sense Publishers.

Roth, W-M. (2013). Data generation in the discovery sciences - Learning from the practices in an advanced research laboratory. *Research in Science Education,* 43(4), 1617–1644.

Sternberg, R. J. (1999). *What do we know about tacit knowledge? Making the tacit become explicit* (pp. 231-236). In Sternberg, R. J., & Horvath, J. A. (Eds.) Tacit knowledge in Professional Practice: researcher and practitioner perspectives. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wilson, A., Åkerlind, G., Francis, P., Kirkup, L., McKenzie, J., Pearce, D., & Sharma, M. D. (2010). *Measurement uncertainty as a threshold concept in physics*. Proceedings of the 16th UniServe Science Annual Conference, University of Sydney, Sept 29th to Oct 1st, 2010, 98-103.