When there isn't a right answer: interpretation and reasoning, key skills for 21st century geoscience.

Journal:	<i>International Journal of Science Education</i>
Manuscript ID:	TSED-2009-0233.R2
Manuscript Type:	Research Paper
Keywords:	reasoning, earth science education, university, conceptual development, argumentation
Keywords (user):	uncertainty, vocational education

When there isn't a right answer: interpretation and reasoning, key skills for 21st century geoscience.

Abstract

A key challenge in university geoscience teaching is to give students the skills to cope with uncertainty. Professional geoscientists can rarely be certain of the 'right answer' to problems posed by most geological datasets, and reasoning through this uncertainty, being intelligently flexible in interpreting data which are limited in resolution and spatial distribution, is an important skill for students to learn. Understanding how interpretative and reasoning skills are, or might better be, developed alongside foundational geological concepts and techniques is crucial for the effective training of our future geoscientists. In the study presented here, a seismic interpretation exercise was undertaken by 36 geoscientists ranging in experience from novice to expert. During the exercise observations were made of the approaches taken by participants, what we term the interpretational process, and an evaluation was made of what they actually accomplished, what we term the final interpretational outcomes. From a comparison of the different experience cohorts, we show that even as a novice it is important to develop geological reasoning skills to enhance the interpretation of datasets where there is indeed no 'right answer'.

Introduction

Frodeman (1995) contrasted geoscience with the more classical sciences of chemistry and physics, describing the nature of these more traditional sciences as based on performing a set of critical tests that isolate individual variables to determine if a given hypothesis is correct. In geoscience, hypothesis-testing is an important element

of the research process, but precisely isolating the variables required for testing hypotheses is rarely easy. In general, evidence will be gathered to test a hypothesis, but in reality it will need to be assessed against more than one hypothesis. For example, the data collected by a geologist undertaking geological mapping will be dependent on the availability of rock exposures, which usually leaves significant uncertainty in the position of geological boundaries (e.g. Bond et al., 2007a).

In his conclusion, Frodeman argued that the skills of intelligent but flexible reasoning and interpretation will be much needed for dealing with the types of questions that the 21st century may pose to geoscience, while Martin and Jucker (2005) underline the imperative of fostering such skills in the process of educating more 'earth-literate leaders'. The next generation of geoscientists will play key roles in the policy and practice development for many 21st century issues, such as climate change, natural hazard mitigation, energy exploration, development of resources and waste management. Addressing questions like how to deal with climate change or how best to predict flooding patterns (AGI, 2008) clearly necessitates the development of multiple (i.e. best and worst case) scenarios, and the calculation of risks associated with those scenarios. These geoscience questions do not necessarily have a straightforward or singular 'right answer', and distinguishing realistic or likely scenarios requires judgement, forward modelling skills and imagination as well as classic hypothesis-testing. We can draw an analogy with medical practice: when a patient visits a doctor, they present a limited number of signs and symptoms for diagnosis. In many cases, similar symptoms can indicate different conditions or diseases, and in some cases a final diagnosis is not available until a postmortem examination (analogous to the information retrieved by drilling an unsuccessful oil

well). Geological reasoning is also often the product of decisions made by an individual or team about what evidence to analyse, the manner of its interpretation, the inductions made and even the possible role of the geoscientist's 'intuition'. An increased understanding of how both geoscience students and professionals deal with datasets and questions possessing an inherent uncertainty therefore has the potential to improve training at all levels, but with particular importance for 'apprentice' geoscientists in universities.

Here we present the results of a study that investigates how geoscientists deal with uncertainty and test hypotheses for datasets and problems that do not have a 'right answer'. The study involved running a geological data interpretation exercise with differing cohorts of geoscientists, from undergraduates (relative novices) to professionals (relative experts). The exercise required participants to interpret a geological dataset, answer a questionnaire and participate in discussion groups. We aimed to contrast both the outcomes of the exercise and the processes integral to how the cohorts approached the interpretation exercise. To capture the processes, we employed observational techniques from the social and psychological sciences to observe the approaches taken by participants (e.g. Englebrecht et al., 2005; Knight, 2004). Specifically, we compared our differing cohorts of geoscientists from observations of: i) how experience levels influence the conduct (i.e. techniques used, time taken, use of materials, etc.) of a geoscientist's data interpretation; ii) the outcome of their interpretation, in part with respect to how realistic or likely it might be relative to the original dataset; and iii) 'feelings' regarding confidence (or not!) in their interpretations. More broadly, we aim to reflect upon: i) how do geoscientists

1
2
3 deal with the uncertainty of there not being a 'right answer'?; and ii) what can we
4
5 learn from the observations to enhance geoscience education.
6
7
8
9

10 **Methodology**

11 *Interpreting Geological Data*

12
13 Our research used two-dimensional (2D) seismic images as the geological dataset for
14
15 the interpretation exercise. Seismic-images show slices through sub-surface geology
16
17 in a similar way to a CAT scan of the human body. To create a seismic image, an
18
19 acoustic source on the Earth's surface radiates energy in the form of seismic waves
20
21 and the waves are reflected back from sub-surface layers. The reflected waves are
22
23 picked up by a network of receivers at the surface, and their amplitudes and relative
24
25 time delays are processed into images of the reflecting layers (e.g. Brown, 1986).
26
27 Although seismic interpretation is now commonly computer-based, using specifically
28
29 designed software, the act of printing out and annotating paper copies of seismic
30
31 images remains commonplace for ease of visualisation and trying ideas for discussion.
32
33 Specific techniques that geologists use to create a 'picture' of the rock layers from a
34
35 seismic image include: identifying consistent seismic reflectors across the image
36
37 which define the tops of sedimentary packages; identifying faults that cut the
38
39 stratigraphy; identifying folds where layers have been deformed; and identifying flat
40
41 (younger) layers that are undeformed. This information is often combined with data
42
43 from boreholes to build up a picture of the type and sequence of the rocks. The
44
45 complete interpretation is then used to build a hypothesis of the geological processes
46
47 that resulted in the formation of the deformed and undeformed layers..
48
49
50
51
52
53
54
55
56
57
58
59
60

Seismic images are commonly used in many commercial geoscience contexts to identify subsurface geological features, such as structures that could trap oil. Information from the seismic data has an inherently limited resolution, and spatial uncertainty, in a typical commercial oil field setting features with a wavelength less than 10-30m, will be not be resolved in a seismic image. Given the large volume of rock involved in commercial oil exploration (10s -100s of km³) and the relatively small volume of any direct investigations (e.g. 10cm diameter borehole cores), there is a substantial mass of rock for which the ‘right answer’ (regarding exact form and genesis) will never be ascertained. Indeed, both Rankey and Mitchell (2003) and Bond et al. (2007b) have used seismic interpretation exercises to show that many interpretations can plausibly be made of the same seismic image, some implying radically different geological processes for their genesis,.

Interpretation Exercises

Rather than using ‘real’ data, our exercises were based on synthetic seismic images. Synthetic seismic images are computer-generated from a known geological model and are often used as a cross-check for real seismic data in industrial applications. The authors created three different 2D geological models from which synthetic seismic images were generated, which means that, unlike with real seismic data, the authors do know the ‘right answer’¹. For more information on the creation of synthetic seismic images from geological models, see Bond et al. (2007b) and the associated data respository. The three synthetic seismic sections used for the experiment were based on geological models of: i) a compressional tectonic setting; ii) an extensional setting; and iii) an inversion setting (Figures 1a, b, c). The Figure’s are annotated with

¹ This sense of there being a ‘right answer’ here is purely an artifice of our methodology for producing seismic images to deploy in our exercises; it is not meant as a closet smuggling in of ‘rightness’ through the back door.

the types of features that a geoscientist interpreting the seismic images may have identified.

Experience Cohorts

The interpretation exercise was run with 36 geoscientists from four different experience cohorts, targeting geoscientists at different levels of their training and professional career. Three of the cohorts were student groups and the fourth was comprised of professional geoscientists working for a geological consultancy company. The two undergraduate cohorts were recruited from students studying for an Earth Science degree at the same UK University; the MSc cohort was recruited from students studying for an MSc in geoscience at another UK university; while the professional cohort was recruited from a geoscience consultancy based in the UK. All participants were volunteers,² and no effort was made to control for variables such as gender, age, ethnicity or social class. That said, the undergraduate cohorts comprised mainly of UK nationals, whereas the MSc and professional cohorts comprised of mixed nationalities.

The first and second cohorts were penultimate and final year undergraduate students. The penultimate year students had no experience in interpreting seismic data, but had been taught the theoretical side of seismic acquisition, how seismic data is collected and processed into images. The second cohort were final year students who had some limited experience in the interpretation of seismic images. The third cohort was Masters students studying for an MSc that has a significant amount of seismic

---

² The penultimate and final year undergraduate students gave up their own time to participate in the exercise. This self-selection may mean that the diversity of skills within a student year group is not fully represented in the cohort, but since we are not comparing the results directly with data from other student groups this is not an issue. In general, the study reveals most about the contrasts between how the different experience cohorts approached the overall exercise.

interpretation as part of the course. This cohort contained a mixture of both UK and international students who had been educated to Bachelor degree level, and the students here were beginning to acquire extensive experience in seismic interpretation and structural geology from different tectonic settings, as well as geophysical techniques. In the fourth cohort, the professional geoscientists, all the participants had a higher degree in geoscience while the professional experience in the group ranged from newly qualified PhD and MSc students to professionals who had worked in the petroleum geoscience industry for 10+ years, and had considerable experience of working with seismic data.

Exercises

One interpretation exercise was run for each of the experience cohorts following the workflow outlined in Figure 2. Each interpretation exercise had three phases: i) the seismic interpretation; ii) ranking of statements about the interpretation exercise and writing down thoughts; and finally iii) a group discussion. The interpretation exercises started with a brief introduction to the project and an outline of what would be required from participants during the interpretation exercise. Printed copies of the instructions, the seismic section to interpret and the statements to rank, were given to the participants at the relevant stages of the interpretation exercise. Additional background information, three printed A4 sheets of paper containing undergraduate-level material, was given to the penultimate and final year undergraduate cohorts, but not to the other two cohorts, at the start of the interpretation exercise. This printed copy offered basic information on seismic interpretation which the students were told that they could use for reference during the interpretation exercise. The penultimate year and final year students were also given longer, 30 minutes rather than 20

minutes, to complete the interpretation part of the interpretation exercise. In all other aspects the workshops had the same format for each cohort.

In the instructions the participants were asked to: “Interpret the seismic section, by highlighting the main horizons and structural features, so you can identify an oil prospect, or place on the seismic section to bury hazardous waste”. This was followed by the following statements: “Oil/Hazardous Waste is buoyant and will move upwards unless it is trapped” and “At the end of the 20 (or 30) minutes we need a fair copy of your interpreted section, a confidence value for your interpretation and a one line summary of your interpretation written in the box below”. For a geoscientist, the meaning of ‘interpret’ in this context would be to draw on and annotate the seismic image, highlighting and interpreting the key features (and examples of key features that might have been identified accompany Figure 1). The materials available to all cohorts during the interpretation exercise included: pencils, erasers, rulers, pencil sharpeners, coloured pencils and felt-tip pens. These items were laid out on a table for the participants to use, but were not given to the participants directly. A summary of the interpretation exercise formats, the seismic section used for each cohort and the number of participants, is given in Table 1.

After the seismic interpretation interpretation exercise, the participants were given a printed A4 sheet of paper containing 14 statements, for which they had to rate their ‘feelings’ on a 1-10 scale. The statements included:

‘I found the interpretation exercise easy’,

‘I was concerned about getting the structure correct’

‘I felt awkward during the interpretation’

The full list of statements and the mean results for each experience cohort are shown in Table 2. On the reverse of the printed copy, participants were asked to “write any words to describe your thoughts and feeling here”. The participants were given two minutes to complete the ratings of their feelings to the 14 statements and to write down any words, the idea being to achieve a ‘gut feeling’ by forcing a quick response to the interpretation exercise. At the end of the exercise, the participants came together to discuss their thoughts and to record any further words that they felt represented their feelings about the interpretation exercise. Participants were given five minutes to complete this final part of the exercise, and they were also asked what geological factors had been taken into account when completing the interpretation exercise.

Data Collection

The data collected fell into two categories: close observation of the interpretation exercises in terms of exactly how the task was undertaken by members of each cohort, and the material output in terms of the seismic sections, confidence ratings, statement ratings and the words and sentences written by participants on the printed copy provided. To monitor the interpretational behaviour of the participants, there were between one and three observers for each interpretation exercise, and for three of the four interpretation exercises digital video recorders were used to supplement direct observations at the time. The exercise observers included a geoscience lecturer, a postdoctoral researcher in geosciences with petroleum industry experience and a PhD student in human geography, while the video recordings were watched by the same

team plus a professor in human geography. This mix of experience and expertise was used to maximise insights from the observations, and included scholars versed in the treatment of qualitative (participatory) research evidence. The interpretation exercise observers had a printed proforma on which they wrote down the time at which various events took place such as: reading; collecting pencils and pens for use; discussion with peers; time taken before the first mark was made on the seismic section; erasing lines; the techniques used, etc. The observers also made notes during the discussion session. The material produced by participants, including their writing and interpretation of the seismic image, was collected both to assess interpretational behaviour (process), including techniques used, and the final interpretational results (outcome). The statement ratings and the words used by participants to describe their thoughts were collated and used to assess their thoughts and feelings during the interpretation exercise.

Findings

The data collected during the interpretation exercise chiefly illustrates the differing behaviours and results of the different experience cohorts. That said, there were variations in process and outcome *within* each experience cohort, and the conclusions drawn should be considered as initial reflections, given the uncertainties resulting from small sample groups.

Process –Interpretational Behaviour

Our study reveals a number of things about the process that the individuals and groups went through in making their interpretations of the 2D seismic images. In terms of basic contrasts between the cohorts, it is unsurprising that we witnessed a decrease in

time taken, from the penultimate year undergraduate student cohort through to the professional geoscientist cohort, to start the interpretation exercise (Figure 3). It was clear that the cohort with the least experience hesitated for longest in starting the exercise, taking longer to put pen to paper and in moving from pencil marks that could be erased to the bolder strokes of a felt pen or coloured pencil. These actions should be distinguished from what we call the ‘ghost tracing’ displayed by individuals in three of the four cohorts, occurring when individuals would trace lines in the air, just above the paper, effectively rehearsing what might be a plausible interpretation. From our observations, it was often the more ‘capable’ individuals who did the latter, reflecting a careful and quite sophisticated intellectual operation entailing prior visualisation – and, we might suggest, conceptualisation – that was then being tested against the seismic data before them. Interestingly, none of the MSc students used this technique, although one MSc student was observed bending the paper to match seismic reflectors across the section, while another looked along the seismic lines, a common visualisation technique in paper-based seismic interpretation. One final year undergraduate student was seen using his/her hands to ‘think’ about the 3D structures involved and how they might have evolved to produce something like the 2D image.

Observations of the cohorts during the interpretation suggest that in the two least experienced (i.e. the undergraduate) cohorts, there was evidence of what might be termed ‘herding’ (Hamilton, 1971). This was notable with respect to collecting materials, pens and pencils to start their interpretations, in that participants in both cohorts waited for one of their colleagues to get up for pens and pencils to undertake the interpretation before immediately following this behaviour. Members of the penultimate year student cohort were also observed to be watching others in their

cohort, apparently to see when they had stopped reading and started the interpretation, as a cue to what they themselves should be doing next. Evidence of such ‘herding’ was not obvious in the more experienced cohorts.

The differing conducts of the cohorts were evidently related to their respective backgrounds. At no point were the cohorts told that they must work either alone or in groups, but the students tended to work by themselves or in pairs (i.e. they had whispered conversations with the person sat next to them), rather than indulging in more inclusive group discussion. This may be due to the more individualistic character of study, and perhaps more importantly assessment, in university (which we discuss further below). Conversely, the professional cohort quickly began to discuss the tasks together as a large group. The observers of the professional geoscientist cohort remarked on: *‘much more chat than the students’* and *‘completely different to student groups, much more banter and more geo-based discussion’*. This fact would seem to reflect the usual working practices of the individuals involved, who, as employees of a company that divides staff with complementary skills and specialisms into teams to tackle projects, would be familiar with groupwork in seeking solutions to problems. One member of the professional cohort wrote *‘not often individuals have done their own interpretation’*, which we take to mean that in their normal working practices an individual is unlikely to be the sole interpreter. It might be conjectured that working cooperatively, allowing time for discussion and if necessary argument, results from an increase in confidence and ability. Training students to be open to ideas, to work openly with others, rather than in competition, and to be able to test and deal with multiple possible solutions, in a situation demanding the arguing

through and indeed defence of one's reasoning, is a key teaching challenge in preparing student geoscientists for professional environments.

The final year undergraduate and MSc students seemed in a hurry to get past the initial reading to work on the seismic images. As one observer of the final year undergraduate students noted, '*all read through very quickly and turned over to look at the section*', while observers for the MSc cohort noted, '*not much reading of background notes, people just got straight on and did it*' and '*hardly anyone reads through the background despite telling them they must do this*'. Perhaps this is a reflection of the classic time pressure faced by students in examinations in which they are so focused on making sure they finish the interpretation exercise that they do not actually read what they are suppose to be doing. Ironically, both of the latter two cohorts still complained about the '*lack of information*' in both their written personal comments and the group discussion.

There was a marked step change in the style, detail and use of available resources between the penultimate and final year undergraduate cohorts, most noticeable being the use of colour by the final year students. None of the ten penultimate year undergraduates used multiple colours on their interpretation, whereas four of the six final year undergraduate students did. The MSc cohort also used colour with five of the 12 participants using colours to highlight sedimentary horizons (e.g. Figure 4e). Their seismic interpretation exercise was less complicated structurally and the horizons were only marginally off-set by faults, so the need to define them so carefully by colour could be considered unnecessary. Use of colour was prolific

amongst the professional geoscientists, with five of the eight participants using colour effectively in their interpretations.

Outcome - Interpretational Results

The interpretations themselves show an interesting evolution between the experience cohorts. The interpretations produced by the penultimate year student cohort were restricted in detail and geological insight, as might be expected from their limited seismic interpretation experience. The physical marking of their interpretations was limited, in that not much pen was put to paper during the interpretation in terms of both lines drawn as a geological model and annotations, and the geological concepts applied by the cohort were relatively simple. The penultimate year student who produced the interpretation in Figure 4a had used the concept of syncline and anticline fold pairs, common nomenclature to describe folds, but the use of the annotation '*almost a syncline*' is hinting at their uncertainty of its applicability in this case. The two examples in Figures 4a and b show the interpretations with the least and most physical marking within the penultimate year undergraduate cohort, while examples of individual interpretations from the other cohorts are shown in Figures 4c-h. These examples have been chosen to be representative of the interpretations produced by each experience cohort.

The penultimate year undergraduate students did not use colour effectively in their interpretations, in contrast to the final year students who used colour to define sediment packages. The 'better' final year students picked tectonically distinct sedimentary packages: i.e. packages of sediments that had been affected by distinct tectonic events and can be used to determine the timing of sediment deposition with

respect to movement on a fault or faults (Figure 4c). Despite four out of the six final year students successfully using colour to define distinct sedimentary packages, the majority of them failed to make the link between the sedimentary package and fault placement. Consider the example in Figure 4c, in which the student has carefully identified tectonically distinct packages, but has drawn a fault cutting through one of them. There appears to be a mismatch between observational ability and the ability to link observations to underlying geological concepts. In colouring the sediment packages, the students have observed and annotated the discontinuities between them, but have not appeared to appreciate that these discontinuities must therefore be defined by a concept that explains the discontinuity: i.e. faults, unconformities, etc. It should be noted that the colouring of the sedimentary packages did not connect directly to a colour coding of stratigraphy, with students often using the same colour for more than one layer. This may have been the result of a limited availability of colours, or that the student could not be bothered, or felt too self-conscious, to get up to collect more pencils. In the final year undergraduate cohort discussion, one student did indicate that the interpretation exercise would have been easier if they had been given a stratigraphy. Overall, though, the interpretations by the final year undergraduate student cohort show a higher order of skill in terms of observations and concepts than the penultimate year cohort, particularly evidenced by one interpreter whose annotations gave geological reasons for choosing potential locations for oil or gas reserves (Figure 4d). Those MSc students who deployed colour were consistent in using different colours for distinct horizons or sedimentary packages. The clearest change here from the less experienced cohorts, however, was the greater use of annotations by the participants to explain their interpretation geologically (Figure 4f),

and there is a strong trend of increasing use of geological annotation with experience (Figure 3).

The apparent confidence of the professional geoscientist cohort seemed to be reflected in the style of interpretations produced by them. The simplest interpretation highlighted the main faults and had the annotation '*drill anywhere you like to get some information*', while others showed careful correlation of rock layers across the seismic section, with colours and numbering/lettering: e.g. Figure 4g. There was also evidence of interpretation through the use of sketching scenarios and explaining geological structures through annotation: e.g. Figure 4h. The interpretational ability of those in each sequential experience cohort improves, as we would have expected. This said, we would argue that it is not simply their ability in terms of known concepts or methodologies to apply to the interpretation, but the mindset, or mental approach, of the interpreter that changes. Indeed, one might go as far as to say that *this* is where we locate the clear jump to being a professional geoscientist, as we will now elaborate.

Confidence –Thoughts, feelings, statement ratings and group discussion

After the interpretation part of the exercise, the participants ranked a set of statements (Table 3) and wrote down how they had felt and their thoughts during the interpretation exercise individually and as a group. It is clear from all of this evidence that the ease by which participants in each cohort interpreted the seismic images increased considerably from the more junior to the more senior cohorts. While the penultimate year undergraduate students tended to use words or phrases like '*unconfident*', '*confused*', '*lost*', '*puzzled*', '*not a clue*' and '*not enough guidance*' to

1
2
3 describe their feelings about the interpretation exercise individually and in the group
4
5 discussion, the equivalents for the final year undergraduates and MSc students were
6
7 more along the lines of '*quietly confident*' and '*happy!*' or being more challenging of
8
9 the exercise (i.e. indicating that any problems lay more with the nature of the
10
11 interpretation exercise than with their own capabilities: see Table 3). The word
12
13 '*confident*' did nonetheless appear once amongst the words written down by the
14
15 penultimate year undergraduate cohort, indicating that a single individual was happier
16
17 with their interpretation and the exercise than his or her fellow cohort participants.
18
19 Both the final year and the MSc student cohorts included geological terminology that
20
21 related to the problem facing them such as: '*obvious structure*', '*tight seal*' and
22
23 '*onlap*'.
24
25
26
27
28
29
30
31

32 The words used by the professional geoscientists to describe their thoughts and
33
34 feelings focused on: geological reasoning, client discussion and potential outcomes of
35
36 drilling. Their comments were cast more in terms of assessing what could or could not
37
38 realistically be expected from a geologist given the materials before them (e.g. '*you*
39
40 '*would not drill on the basis of the information provided*' and '*insufficient data to*
41
42 '*make a sound interpretation*'). As such, 'confidence' here became more a matter of
43
44 expressing 'confidence limits', in a loosely statistical sense, around interpretations
45
46 made on the basis of the quality and quantity of information available, rather than a
47
48 matter of confidence levels in one's own abilities *per se*. One professional
49
50 geoscientist expressly stated that he/her '*couldn't justify more than a 20 percent*
51
52 '*confidence level*' in their interpretation, but mischievously added that '*such a figure is*
53
54 '*still higher than exploration wells success rates!*'
55
56
57
58
59
60

With increasing experience, there was an increase in confidence of the participants about the interpretation exercise, as indicated by the mean answers of each cohort to question 11 (Table 2) and from the differences between each cohort's group words (Table 3). This qualitative evidence was not reflected in the numerical confidence values given by individuals to their own interpretation, or the mean for each cohort, which range from c.49-63% (Table 3). This perhaps highlights the difficulties for geoscientists at all levels in expressing confidence about their 'geological reasoning' in terms of numerical values, and precisely this difficulty in assigning numerical probabilities to judgements has been shown in several studies (e.g. Kahneman et al., 1982; Anderson, 1998).

The answers given by the participants in each cohort to the questions in Table 2 do not show highly distinct trends, or reflect what was observed either directly or in the videos. For example, question 2 asked the participants to rank the following statement from a lot to a little: *'I was reluctant to put pen to paper'*. The final year undergraduate student cohort had the lowest mean result here, suggesting they were least reluctant, whereas in practice they were the second slowest cohort to start the interpretation. We might suggest that the penultimate year cohort answered more 'truthfully' than the final year undergraduate or MSc student cohorts, who perhaps chose what might be considered the right answer or what they thought that they should put (i.e. Table 2, question 3: *'I went about it in a logical manner, based on what I had learnt in class'*).

It is obvious that the three student cohorts found it hard to escape from what was usually expected of them in a teaching, learning and assessment context or, at least,

what they *assumed* to be expected of them. In the main, they appeared to suppose that there was still a ‘right answer’, and that it was their task to discern the correct response in order for their efforts to be judged a success (by their instructors). Despite clearly outlining at the beginning of the session that the work was anonymous and would not be used in anyway for assessment, one student remarked ‘*I felt like I was in an exam without having been able to study for it*’, although another admitted ‘*I wasn’t scared because I wasn’t getting marked on it*’. For the students, therefore, the sort of consequence that most mattered was academic success or failure, with some seemingly believing – notwithstanding how the exercise had been introduced to them – that they really *were* to be assessed in some way. The penultimate year and final year students may have been influenced by the fact that the interpretation exercise took place in their normal teaching laboratory, although this was also the case for the MSc students and the exercise for the professional geoscientists took place in their own work environment. The penultimate and final year undergraduate students may also have been influenced in their approach by the fact that one of the observers was one of their lecturers, but the exercise was not run by the teaching member, and a post-doctoral researcher and a PhD student (from a different subject area) were also observing the interpretation exercise. To mitigate against the feeling that the interpretation exercise might form part of the students’ university work, coffee and pastries were supplied in the laboratory. The participants in all three student cohorts still harboured some hesitation, even in some cases resentment, about what it was they were being asked to do, complaining at the end that they did not have enough information or that they felt ‘*suspicious*’ or even ‘*tricked*’ (Table 3). It is worth noting that this was also the case for the MSc students even though they had no professional or personal links to the observers. The professional cohort offered what might be

1
2
3 deemed a more 'mature' reaction in this regard, and were aware that on occasions a
4
5 geologist will have relatively fulsome data on which to base decisions, including the
6
7 likes of well core and information from adjacent oil fields, but that there will
8
9 frequently be occasions where the data is much more patchy and uncertain. For the
10
11 professional cohort, there was also much more awareness that one's interpretations
12
13 really could have a consequence beyond merely the academic or technical assessment
14
15 of one's work, making comments like *'would recommend client to choose a better*
16
17 *site'*.
18
19
20
21
22
23
24

25 **Educational Learning Outcomes**

26
27 Having discussed the interpretational processes and outcomes of the seismic
28
29 interpretation exercise, there is a question of whether we have simply observed the
30
31 difference between the increasing 'knowledge' and deepening skills base of the
32
33 cohorts from novice to expert, or if there is something to be learned from the
34
35 professional cohort that can be taught to future students. Clearly it is not possible to
36
37 teach experience gained through practice over time, but maybe there are elements of
38
39 *how* the professionals apply their experience that lend themselves to be taught, or at
40
41 least be better taught, within formal education systems.
42
43
44
45
46
47

48
49 There has been much research into comparisons between experts and novices (e.g.
50
51 Chi et al., 1981; Chi et al., 1982; Kozma, 2003; Patrick et al., 2005). Many of these
52
53 studies have attempted to define the differences in approach and outcome between
54
55 novices and experts when problem-solving in a scientific context. For example,
56
57 Patrick et al. (2005) showed that novice learners identified differences in DNA colour,
58
59 but were unable to interpret meaning. Similarly, Kozma (2003) suggested that,
60

because novices' mental models do not go beyond a perceptual processing level, they are unable to coordinate observations to understand underlying concepts. We observe this in the use of colour by a final year undergraduate student to define a tectonic package (observation), but which they then cross with a fault (i.e. producing a conflict between observation and concept: Figure 4c). Chi et al. (1981) and Schnotz et al. (1993) argued that experts' knowledge enables them to focus on the information in a graphic image that is most relevant for constructing a mental model, implying a continual feedback loop between observation, reasoning and interpretation. These studies and others imply that novices rely mainly on observed phenomena to develop understanding and do not use inference or reason-based understanding to develop observations into concepts. Bransford et al. (2000) summarised the key differences between experts and novices, proposing that experts' knowledge cannot be reduced to key facts because their knowledge reflects context and applicability. In the field of engineering design, Ho (2001) demonstrated that experts used explicit problem decomposition strategies by deploying constraining assumptions which non-experts did not employ. Lloyd and Scott (1994) found that experienced design engineers made conjectures to define the problem and solution space, whereas those with less experience used deductive reasoning. Given these key differences in approach to problem-solving with experience, it is unsurprising that Cook (2006) called for more research to explore the continuum of the transition from novice to expert, most specifically to look at the particular effects of educational intervention.

The main knowledge-based differences observed in this study were between the penultimate and final year undergraduate cohorts, with the penultimate year students having no experience of taught courses in seismic interpretation. These differences

1
2
3 seem fairly self-evident in the physical outcomes and observed behaviours of the
4
5 penultimate year cohort interpretations as compared to the other cohorts. In the more
6
7 experienced student cohorts, the final year and MSc students, the main basis of
8
9 seismic interpretation and geological knowledge was clearly there, but the students
10
11 seemed to lack confidence in applying that knowledge. This lack of confidence
12
13 appeared to be primarily focused in their ability to apply techniques and concepts
14
15 rather than in their observational skills or actual conceptual knowledge.
16
17
18
19
20
21

22 Perhaps the clearest difference between the student and professional cohorts, though,
23
24 was simply in attitude and approach, some of which can be equated to differences in
25
26 maturity, but why is there such a difference in attitude and approach when the
27
28 interpretation exercises were run with the same format for all cohorts? We argue that,
29
30 despite the contextual scenarios provided to participants, it was only the professional
31
32 cohort who fully appreciated the potential 'real world' context in which such an
33
34 interpretation might be made, and this lack of context hampered our student cohorts
35
36 significantly. The professionals appreciated that in the real world there is no 'right
37
38 answer' and, crucially, were comfortable with this state of affairs.
39
40
41
42
43
44

45 The student cohorts were so focused on getting the 'right answer' that they seemed
46
47 almost oblivious to what they were being asked to do, and more importantly were
48
49 unsure how to apply their skills of reasoning and hypothesis-testing. In many ways
50
51 our student groups behaved like children rather than adults, displaying attributes of
52
53 childhood behaviour that are well acknowledged within childhood psychology (e.g.
54
55 Piaget 1932; Kohlberg, 1958, 1963). The answer for many of the student cohort
56
57 participants was what would benefit them best, get top marks, and make the 'teacher'
58
59
60

happy. This idea of not only getting the 'right answer' but also choosing the right approach was particularly evident in the answers given by the final year undergraduate and MSc student cohorts to question 3 in Table 2 (which asked the participants if they approached the interpretation exercise in a logical manner). Both of these cohorts had a mean score of 8/10 for their answer to this question, the highest mean score for any of the 14 questions.

An immediate response to these findings is to set off an alarm-bell: are there perhaps problems in how we are training geosciences students if they are so focused on getting a 'right answer', especially when there is not necessarily one to be found. How do we allow students to develop reasoning skills and show development of interpretations, hypothesis-testing and use of multiple scenarios (e.g. Chamberlin, 1890)? Many geoscience teachers would argue that this is a clear element of the training that they deliver, from recording hypothesis development and testing in field notebooks through to requiring descriptions of observations made of minerals in a thin section, before naming the minerals and the rock type (e.g. Reynolds and Peacock, 1998). Yet, it would seem that only the best students truly pick up on this teaching, and the majority are still focused on the 'right answer', be it the name of the rock or the 'correct' placement of a fault line on a field map. Given traditional educational assessment schemes, we do not think this is particularly surprising.

In contrast, rather than being focused on getting the 'right answer', our professional cohort participants were more pragmatic, realising that there is not a 'right answer' to the data or, to put it another way, that more knowledge was required to answer the question and having the confidence to acknowledge this fact. This reasoning and

confidence may be seen to be emerging within the MSc cohort: see Table 3, and group words such as '*unanswerable as too little variables are known*'. It is almost as if the professionals, and those with greater experience, have learnt that there are other questions and/or answers that may be more pertinent. In essence, they seem to recognise that it is often better to apply what is known in Physics as the Fermi rule "better to be approximately right, rather than precisely wrong". In some ways they have learnt that in the professional world a definitive answer is not necessarily the best solution.

Discussion with the less experienced student cohorts suggests that they are interested in the application of geological interpretations in society, with one student suggesting that if the scenario had been really real it may have affected his input and interpretation. This expressed view suggests a significant problem in making taught exercises feel 'real' in a classroom environment, feelings perhaps echoed by the views of the MSc cohort who expressed the sense of being '*tricked*'. Why? Because it was not within their range of expectations; it was not like a normal class where they would be marked and assessed? Of course in the real world, the professionals are effectively marked and assessed by their clients and/or boss, albeit in a less direct but probably even more consequential way.

The obvious proposal from our observations and interpretations is to suggest an increasingly applied approach and vocational training aspect to teaching in geoscience, *but* at the same time nurturing a more sophisticated sensibility about the limitations on what (geo)science can ever accomplish. It is to demand a shift from a 'teach me' mentality on the part of students to a 'help me to learn' mentality

(Chappell, 2006). These suggestions are supported by the work of several other studies in different fields. Shin et al. (2003), for instance, showed that domain-specific knowledge is important to solve ill- and well-defined problems in astronomy, but that ill-defined problems require a type of reasoning that students largely do not have; and they proposed that this lack is due to the limitations of their educational experience. The use of reasoning and assumptions to help solve ill-defined problems was advocated by Reitman (1964), a view also supported more recently by Fortus et al. (2009), who used well-defined physics problems, akin to those given in a class physics assessment, alongside more ‘real-world’ or ill-defined problems. They concluded from the experiment that the process of making constraining assumptions to define the ‘real world’ problem was the most difficult part of the process for all students, and that only those who had previous experience of using constraining assumptions were able to make the assumptions needed to solve the problem. They added that the findings suggest a need to support physics students in developing this important skill through restructuring science teaching around such ‘real world’ problems. Others calling for such changes include: AAAS (1990), Bartel et al. (1992), Blumenfield et al. (1997), Duschl et al. (2007), Heller and Hollabauh (1992), Lipman (1991), National Research Council (2000) and Bryce and MacMillan, (2009). On the merits of problem-based learning, or PBL, specifically in the environmental sciences, see, for instance, Beringer (2007) and Chappell (2006).

As Fortus (2004) claimed with respect to product design, it is impossible to know if the product designed is the best for the job, only that it fulfils the overall requirements. It might not be, and probably is not, a unique solution or definitive ‘right answer’ to the problem. In much of geoscience the lack of a unique solution for

the data available has similar implications, such that the skills of reasoning through realistic or likely solutions are critical for success. These skills are particularly important for those approaching the time when they will be looking for employment and leaving full-time education, but we suggest that intelligently flexible scientific reasoning should be developed through time alongside the geological knowledge and techniques that students built up during a geoscience student's whole university career.

A greater focus on the actual practices of 'real' science-in-action has also led Lehrer and Schauble (2006) to propose that we must not only focus on scientific products (outcomes, conclusions, papers, etc.) but also on the scientific process, including and perhaps most particularly ways of talking about phenomena. Scientific debate through argument and persuasion has been shown to comprise a key skill in developing reasoning skills in science, and Duschl et al. (2007) suggested that engaging in scientific debate makes students take critical stances, increases their willingness to ask questions and helps them to develop a sense of both trust and scepticism in scientific evidence. The importance of scientific 'advocacy' is echoed by Jimenez-Alexandre et al. (2000), Driver et al., (2000) and Sandoval and Millwood (2005), all highlighting how the construction and defence of scientific explanations to persuade others is a key component of science. Berland and Reiser (2008) tested three scientific skill sets of a student group – sense making, articulating and persuading – and they found that the students were least effective at persuading, suggesting that students lack this key skill because the relevant *social* interactions are constrained by normal classroom practices.

To develop this final point and to underscore our main conclusion, we suggest a revised pedagogic approach, one demanding a more subtle grasp of (geo)science and of the trust that we can place in such (geo)science. This approach should include an explicit acknowledgement that there is indeed often no unique ‘right answer’ in the interpretation of geological data, connecting with other studies that have shown the importance of exposing students to the effects of different data-sampling strategies and interpretations on model outcome (e.g. Priestnall, 2009; Sturkell et al., 2008). Students should instead be taught how to formulate strategies for being intelligently flexible when reasoning through what may (or may not) be plausible interpretations of uncertain geological data, and – also as a result of importing vocational and PBL elements into our teaching – we might further urge more collaborative group work, complete with discussion and debate occurring on a less individualistic, more group-based footing than is usually the case (also Knight, 2004).

Conclusions

In this paper we have concentrated on the differences in the interpretational processes used and the interpretational outcomes of four different experience geoscience cohorts invited to work with equivocal seismic datasets. The cohorts ranged from penultimate year undergraduate students to professional geoscientists. We have used the differences observed to show the development of interpretation and reasoning skills with training and experience, and have used these findings to reflect on possible ways to enhance such skills in an higher-educational environment. Our observations and reflections suggest that there are aspects of the interpretational process and outcome that are not simply a factor of increasing knowledge, either concepts or techniques. Rather, what must be acknowledged are the distinct differences between the

professional cohort and the less experienced cohorts manifested in the attitude, and hence approach, of the former cohort to the interpretation exercise. Unlike the professional cohort, the characteristics of the responses from the student cohorts can be broken down into three distinct, but inter-related components:

- 1) Students are focused on getting the right answer.
- 2) Students, despite having the relevant geological knowledge, show a reluctance/lack of confidence in applying interpretational techniques.
- 3) Students lack 'real world', more problem-based contexts for their interpretational reasoning, and have a limited sense of how to construct arguments and thereby advocate 'their' science.

We suggest that, in order better to prepare geoscience students for the challenges that will face them in the 21st century, they should be encouraged – for some if not all components of their academic training – to foster their interpretative reasoning skills in an applied context, one that itself demands a more nuanced understanding of what (geo)science can ultimately achieve. In doing so, they should learn the broader interpretative reasoning skills of the geoscientist, including an acknowledgement that, after all, the 'right answer' is not straightforwardly 'out there' awaiting its revelation.

The project was relatively small-scale, and hence its potential for affording generalisations must be acknowledged, but a fairly clear pattern emerged in the distinctions between the different experience cohorts with respect to both

interpretational process and outcomes. Moreover, we are confident – given our wider experiences as university teachers, and given parallel findings elsewhere in the literature – that the educational inferences drawn above have considerable validity. Nonetheless, our project was in part configured as a ‘pilot’, the anticipation being that further research could be undertaken in this field. More specifically, we have since received funding for a PhD study that will greatly extend the numbers of geoscientists at all experience levels participating in seismic image data response exercises akin to those described above. The aim is to inquire more systematically into the complex mix of factors that influence the interpretative reasoning skills of geoscientists, from the most personal through domains of education, social status and cultural attributes, with the potential for generating both statistical findings and qualitative conclusions. We can also envisage ways of building more directly upon aspects of the specific project reported here; perhaps to explore further the extent to which it is simply differences in a geoscientist’s ‘knowledge base’ that is the issue, rather than their level of experience in actually working *with* that geological base; perhaps to explore further the influence of the location of and personnel present in the exercises, notably asking whether students taken out of an environment pre-associated with learning, assessment and hierarchical relations with staff might perform differently; and perhaps comparing exercises run with a more obvious vocational scenario (complete with known budgets, stakeholders and policy implications) and ones run with a more purely academic scenario (concerned with ‘proving’ given geological concepts). These are just suggestions, ones that the authors are considering, but there may be many other possible angles and questions that press upon us once we accept that educating/training geoscientists is about far more than hunting for the mythical ‘right answer’.

Acknowledgements

The participants, universities and companies involved are thanked for their input into this research. The research was funded by an RGS-IBG small grant to CEB.

GXTechnology are thanked for creating synthetic seismic images for use. Jo Norcup is thanked for help with running and observing the interpretation exercises. This paper was significantly improved by the suggestions of two anonymous reviewers.

References

American Geological Institute (2008). Critical Needs for the Twenty First Century: The Role of the Geosciences. AGI. p19.

American Association for the Advancement of Science. (1990). Science for all Americans. New York, Oxford University Press.

Anderson, J. L. 1998. Embracing uncertainty: the interface of Bayesian statistics and cognitive psychology. *Conservation Ecology*, 2, 2. [Online] World Wide Web Address: <http://www.consecol.org/vol2/iss1/art2>.

Bartel, A. P., Lichtenberg, F. R., & Vaughan, R. J. (1992). Technological change, trade, and the need for educated employees: Implications for economic policy (NCEE brief). Washington, DC: National Centre on Education and Employment.

Beringer, J., 2007. Application of problem-based learning through research investigation. *Journal of Geography in Higher Education*, 31, 445-457.

Berland, L.K., and Reiser, B.J., 2008. Making Sense of Argumentation and Explanation. *International Journal of Science Education*, 93, 1, 26-55.

Blumenfeld, P. C., Marx, R. W., Patrick, H., Krajcik, J. S., & Soloway, E. (1997). Teaching for understanding. In B. J. Biddle, T. L. Good, & I. F. Goodson (Eds.), *International handbook of teachers and teaching* (pp. 819 – 878). Dordrecht, The Netherlands: Kluwer.

Bond C.E., Shipton Z.K., Butler R.W.H., Jones R.J. and Gibbs A.D. 2007a. Knowledge transfer in a digital world: uncertainty, visualisation and data management. *Geosphere* 3, 568-576 doi: 10.1130/GES00094.1

Bond C.E., Gibbs, A.D., Shipton, Z.K., and Jones, S., 2007b. What do you think this is? “Conceptual uncertainty” in geoscience interpretation. *GSAToday*, 17, 11, pp.4-10. DOI:10.1130/GSAR01711A.

Bransford, J. D., Brown, A. L., Cocking, R. R., Donovan, M. S., & Pellegrino, J. W. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC, National Academy Press.

Brierly, G., 2009. Communicating geomorphology. *Journal of Geography in Higher Education*, 33, 3-17.

Brown, A.R., 1986. Interpretation of 3D seismic data, 6th ed: AAPG Memoir 42, 541p.

Bryce, T.G.K. and MacMillan, K. (2009) Momentum and Kinetic Energy: Confusable concepts in secondary school physics. *Journal of Research in Science Teaching*. 46.7, 739 – 761. Also published on-line, January 2009. DOI 10.1002/tea.20274.

Chamberlin, T.C., 1890. The method of multiple working hypotheses. *Science* (Old series) 15. Reprinted in 1965 in *Science* 148, 754_759].

Chappell, A., 2006. Using the 'grieving process' and leading journals to evaluate students' responses to problem-based learning in an undergraduate geography curriculum. *Journal of Geography in Higher Education*, 30, 15-31.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121 – 152.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, NJ, Erlbaum.

Cook, M. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *International Journal of Science Education*, 90, 6, 1073 –1091.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84, 3, 287 – 312.

Duschl, R. A., and Osborne, J., (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education*, 38, 1, 39 – 72.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. E. (Eds.). (2007). Taking science to school: Learning and teaching science in grades K-8. Washington, DC, National Academies Press.

Engelbrecht, A.C., Mintzes, J.J., Brown, L.M., and Kelso, P.R., 2005. Probing Understanding in Physical Geology Using Concept Maps Clinical Interviews. *Journal of Geoscience Education*, 55, 3, p.263-270.

Fortus, D., Dershimer, R. C., Krajcik, J. S., Marx, R. W., & Mamlok-Naaman, R. (2004). Design-based science (DBS) and student learning. *Journal of Research in Science Teaching*, 41(10), 1081 – 1110.

Fortus, D., 2009. The importance of Learning to Make Assumptions. *International Journal of Science Education*, 93, 1, 86-108.

Frodeman, R., 1995, Geological reasoning: Geology as an interpretive and historical science. *Geological Society of America Bulletin*, 107, 8, 960-968.

Hamilton, W. D., 1971, Geometry for the Selfish Herd. *Journal of Theoretical Biology*, 31, 2, pp. 295-311.

Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping: Part 2—Designing problems and structuring groups. *American Journal of Physics*, 60, 7, 637 – 644.

Ho, C. H. (2001). Some phenomena of problem decomposition strategy for design thinking: Differences between novices and experts. *Design Studies*, 22, 27 – 45.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction*, 17, 379 – 432.

Jimenez-Aleixandre, M. P., Bugallo Rodriguez, A., & Duschl, R. A. (2000). “Doing the lesson” or “doing science”: Argument in high school genetics. *Science Education*, 84, 6, 757-792.

Kahneman, D., Slovic, P., and Tversky, A. (eds), 1982. Judgement under Uncertainty: Heuristics and Biases. Cambridge, Cambridge University Press.

Kohlberg, L., 1958. "The Development of Modes of Thinking and Choices in Years 10 to 16". *Ph. D. dissertation, University of Chicago*.

Kohlberg, L., 1963. The Development of Children's Orientations Toward a Moral Order. I. Sequence in Development of Moral Thought. *VITA HUMANA*, 6, 11-33.

Knight, J., 2004. Comparison of student perception and performance in individual and group assessment in practical classes. *Journal of Geography in Higher Education*, 28, 63-81.

Lehrer, R., & Schauble, L., 2006. Scientific thinking and science literacy: Supporting development in learning in contexts. In W. Damon, R. Lerner, K. A. Renninger, & E. Sigel (Eds.), *Handbook of child psychology* (6th ed., Vol. 4, pp. 153 – 196). Hoboken, NJ, Wiley.

Lipman, M. (1991). *Thinking in education*. New York, Cambridge University Press.

Lloyd, P., & Scott, P. (1994). Discovering the design problem. *Design Studies*, 15, 125 – 140.

Martin, S. & Jucker, R., 2005. Educating earth-literate leaders. *Journal of Geography in Higher Education*, 29, 19-29.

National Research Council. (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. Washington DC, National Academy Press.

Patrick, M. D., Carter, G., & Wiebe, E. N. (2005). Visual representations of DNA replication: Middle grades students' perceptions and interpretations. *Journal of Science Education and Technology*, 14, 3, 353 – 365.

Piaget, J. 1932. The Moral Judgment of the Child. London: Kegan Paul, Trench, Trubner and Co. [ISBN 0-02-925240-7](#). DOI 2004-02-23 00:00:00

Priestnall, G., 2009. Landscape visualisation in fieldwork. *Journal of Geography in Higher Education*, 33, S104-S112.

Rankey, E.C., and Mitchell, J.C. (2003). That's why it's called interpretation: Impact of horizon uncertainty on seismic attribute analysis: *The Leading Edge*, 22, 820–828. DOI:10.1190/1.1614152.

Reitman, W. (1964). Heuristic decision procedures, open constraints, and the structure of ill-defined problems. In M. W. Shelley & G. L. Bryan (Eds.), *Human judgments and optimality* (pp. 282 – 315). New York, Wiley.

Reynolds, S.J., and Peacock, S.M., 1998. Slide Observations – Promoting Active Learning, Landscape Appreciation, and Critical Thinking in Introductory Geology Courses. *Journal of Geoscience Education*, 46, 421-426.

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

Schnotz, W., Picard, E., & Hron, A. (1993). How do successful and unsuccessful learners use texts and graphics? *Learning and Instruction*, 3, 181– 199.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Shin, N., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching*, 40, 1, 6 – 33.

Sturkell, E., Jakobsson, M., and Gyllencreutz, R., 2008. How true are Geological Maps? An Exercise in Geological Mapping. *Journal of Geoscience Education*, 56, 4.

Figure 1. Examples of features in seismic images that can be used to interpret the geology. The three images shown are the ones used in the experiment (see table 1). In all three images the parallel sub-horizontal lines are reflections from the layers of sedimentary rock. The order and thickness of these layers is known as the stratigraphy and can be calibrated with well logs and cores. In image 1a) Faults cut through these layers, offsetting the rocks on either side. Movement on the faults has created a fold in the centre of the image. Figure 1b) Similar faults are seen in this figure note that younger sedimentary layers deposited on top of the fold are flat (undeformed). Figure 1c) Four steeply dipping faults cut the thin layers near the bottom of the image. The upper most horizontal layer is undeformed. In figure 1b the layers have been moved up and over one another by the fault, accommodating compression. In the lower layers of figure 1c the right handside of each fault has been moved down and along accommodating extension. In figure 1a the layers have been extended and then compressed. Fault A is accommodating pure extension, while fault B has accommodated extension followed by compression (producing a fold similar to that seen in figure 1b).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2. A flow chart showing the format and timing of the interpretation exercises.

Figure 3. Observed interpretational behaviours (processes) and interpretational results (outcomes) of the four cohorts. The x-axis shows the different experience cohorts, the blocks indicate the time taken for the first participant in each cohort to: start the exercise, to employ a ghost tracing technique and to use coloured pencils. The line graph indicates the percentage of participants in each cohort to use geological annotations in their interpretation of the seismic data.

Figure 4. The seismic images given to the different experience cohorts. Interpretations A) and B) are from the penultimate year undergraduate students, the participant in A) produced the most detailed interpretation from this cohort including annotating some of the geological features. B) This interpretation shows that hardly any physical interpretation has been made on the seismic section. Both participants have indicated that they have a 40% confidence in their interpretation. Interpretations C) and D) are from the final year undergraduate student cohort. C) shows the use of colour by the participant to highlight sedimentary packages, whilst participant D) has not used colour, but has annotated the different layers and geological features. Participant C) has indicated a 70% confidence in their interpretation, participant D) 65% confidence. Examples E) and F) are from the MSc cohort; E) has used different colours for horizons, whilst F) has used a combination of colour and annotation to highlight horizons, geological features and processes. Neither of the interpreters assigned confidence values to their interpretations. Finally examples G) and H) are from the professional cohort, interpreter G) has used different colours for different horizons and annotations of layers, as well as highlighting faults and other geological features. H) The participant has used colour to annotate horizons and sketched potential geological features. The interpreters have assigned 20% and 50% confidence values, respectively for their interpretations.

Table 1. Summary of the workshop formats for the different experience cohorts. A single seismic image was used in each workshop, the seismic images used, without the annotations, are shown in Figure 1.

Experience group	Time for interpretation	Seismic image	Background information	Number of participants
Penultimate year undergraduate students	30 mins	inversion (Fig 1a)	Yes	10
Final Year undergraduate students	30 mins	compressional (Fig 1b)	Yes	6
MSc students	20 mins	extensional (Fig 1c)	No	12
Professional	20 mins	extensional (Fig 1c)	No	8

Table 2. The participants were asked to score the statements in column A) by circling a number 1-10 where 1 was described as a little and 10 a lot. The numbers in columns B -E) are the mean resultant score from each experience cohort (rounded to an integer).

Questions	Mean results for each experience cohort			
A) All cohorts	B) Penultimate year students	C) Final year students	D) MSc students	E) Professionals
1) I found the interpretation exercise easy	4	6	5	6
2) I was reluctant to put pen to paper	6	3	5	4
3) I went about it in a logical manner based on what I had learnt in class	6	8	8	7
4) I was concerned about leakages and focused my interpretation on this	5	7	6	4
5) I was concerned about traps and focused my interpretation on this	8	7	6	5
6) I was concerned about seals and focused my interpretation on this	5	7	6	4
7) I was concerned about getting the structure correct	6	6	8	8
8) I felt unconfident in my interpretation	7	4	5	4
9) I felt awkward during the interpretation	6	4	4	4
10) I felt unsure during the interpretation	6	5	6	5
11) I felt confident during the interpretation	4	5	5	6
12) I felt positive during the interpretation	5	7	6	7
13) I felt hesitant during the interpretation	6	5	5	5
14) I was concerned about getting the interpretation wrong	6	7	6	5

Table 3. The confidence values and group words of the four experience cohorts. The confidence values (%) that each participant assigned to their own interpretation and the average for each cohort are shown as well as the descriptive words written down by participants at the end of the interpretation exercise, note that these are direct quotes of the group words.

1) Penultimate year students													
Confidence %	40	0	65	30	65	75	70	25	75	40	Average 48.5%		
Group words	Hard, didn't know where to start looking, lack of information, Additional Information sheets-helped a little although much was irrelevant, lack of knowledge from 3 rd year students, initial shock, better if you could talk to someone, not a clue, confused, lost no key to picture, alright, good fun, ok, confident, puzzled, not enough guidance, lack of info, interesting												
2) Final year students													
Confidence %	10	70	65	70	90	70	Average 62.5%						
Group words	(Picture of a flower), Headache, Hungry (pastries?). Simplified (looked like it was laid out so we could interpret it easily), Obvious structure, Would have been easier to have general lithologies, Varying confidence levels looked for offlaps and onlaps, Happy!, Unsure!, Folds, Faults (thrust), Quietly confident!, 3D visualisation, Traps, Understand section then apply to area, Looked for tight seal (problem first), Arrows!, Onlap!, 2 phases of deposition												
3) MSc students													
Confidence %	60	N	N	50	60	50	60	75	20	30	60	20	Average 48.5%
Group words	Suspicion, unanswerable as too little variables are known, various models proposed, just went through work flow with little emotion More confident in lower (below 2 sec) area of section, Better/stronger images, More fuzzy towards top, Undecisive, uncertain, cap rock?, Reflective – on experience learnt in class and past experience.												
4) Professionals													
Confidence %	90	50	80	50	50	70	70	20	Average 60%				
Group words	You would not drill on the basis of the information provided. Discuss interpretation first. Not often individuals have done their own interpretation. Compared one line summaries, Didn't like the error statistics angle. Would choose to recommend client to choose a better site. Insufficient data to make sound interpretation.												