

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/347522873>

# Auditing summative assessments: the need to increase creative reasoning in mathematics and science at lower secondary in Ireland

Article in *The School science review* · December 2020

CITATIONS

0

READS

111

4 authors, including:



**Majella Dempsey**

Maynooth University

27 PUBLICATIONS 97 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Covid-19 rapid response in schools [View project](#)



Task Evaluation and Development [View project](#)

# Auditing summative assessments: the need to increase creative reasoning in mathematics and science at lower secondary in Ireland

Ailbhe Garry, Aishling Reilly, Majella Dempsey and Ann O'Shea

**Abstract** This article is concerned with the level of reasoning needed to complete state examinations in mathematics and science at lower secondary level in Ireland. The authors used Lithner's (2008) Creative Reasoning Framework to classify tasks in three years of mathematics examination papers. They adapted this framework to do the same for science. It was found that most examination tasks require imitative reasoning rather than creative reasoning, and that small modifications can alter the type of reasoning required in tasks. The authors recommend that attention to task design needs to be at the forefront of teaching, learning and assessment.

The types of tasks that students engage with have been shown to influence their development of skills such as reasoning and problem-solving. Smith and Stein (1998) asserted that the highest learning gains in mathematics were related to how tasks were implemented in teaching and they highlighted the importance of students being engaged in high levels of cognitive thinking and reasoning. It is commonly accepted that what is assessed in the curriculum is what gets valued and taught in the classroom. Assessment has impact not only on teaching and learning but also on student motivation (Harlen, 2006). Therefore, it is imperative that we have tools to help us audit tasks we use in summative assessment. In this article we outline how using frameworks to analyse the level of reasoning in tasks set as summative assessment at lower secondary in Ireland indicates an over emphasis on examining imitative reasoning. This process is easy to carry out and would be of benefit to all teachers when setting assessments.

## Literature review

Lithner (2008) defined reasoning as '*the line of thought adopted to produce assertions and reach conclusions in task solving*' (p. 257). This understanding of reasoning led him to develop his framework (Table 1), which defines

**Table 1** An overview of mathematical reasoning (Bergqvist, 2007; Lithner, 2008)

Imitative reasoning	Memorised reasoning	Algorithmic reasoning		
		Familiar	Delimiting	Guided
Creative reasoning	Local creative reasoning			
	Global creative reasoning			

the different levels of reasoning required for completion of mathematics tasks. He divides mathematical reasoning into two broad categories: imitative and creative. Within imitative reasoning (IR) there is memorised reasoning (MR) and algorithmic reasoning (AR). MR and AR are both low levels of reasoning, with MR '*founded on recalling a complete answer*' (p. 258) and AR requiring the solver to '*recall a solution algorithm*' with '*no need to create a new solution*' (p. 259). In relation to AR, Lithner identifies three different categories of reasoning that can be employed by the reasoner when approaching a task. Each type only requires the reasoner to establish surface connections with the underlying mathematical concepts involved. Familiar AR involves utilising a known algorithm to solve a task that is already familiar to the reasoner. Delimiting AR takes place when the reasoner chooses an algorithm from a pool by identifying features of the algorithm that seem superficially related to the task. If a plausible solution is not reached, the reasoner does not engage in critical reflection on why their chosen method led to this result and may go back to their original pool of algorithms to choose another. Guided AR tasks provide the reasoner with prompts for solving, resulting in a correct solution but requiring little to no understanding of why the strategies suggested by the prompts are suitable and no verification of their suitability (Lithner, 2008).

Creative reasoning (CR) is characterised by novelty (to the reasoner), plausibility and mathematical foundations (Lithner, 2008). Bergqvist further subdivided CR into local creative reasoning (LCR) and global creative reasoning (GCR). The two differ in that LCR tasks only require CR to modify an existing algorithm already known to the reasoner, while GCR tasks require the creation of a new solution algorithm by the reasoner (Bergqvist, 2007). Further to this, when classifying tasks as LCR or

GCR one must take into consideration the amount of modification needed by the reasoner to complete the task. Mac an Bhaird *et al.* (2017) classify tasks as LCR if only one sub-procedure needs to be adapted and as GCR if more than one sub-procedure is new.

Scientific reasoning has been the subject of much discussion. Tekkumru-Kisa, Stein and Schunn (2015) emphasised the importance of student engagement with high-level reasoning tasks that *'demand interpretation, flexibility and the construction of meaning; they must learn how to persevere when the "right" answer or preferred method is not obvious'* (p. 660). Their *'Task Analysis Guide in Science'* (TAGS) framework outlines different levels of cognitive demand required by a student to complete a task. It is divided into *'science content'*, *'scientific practices'* and the *'integration of content and practices'*. The authors emphasise that tasks that require a high level of cognitive demand must involve the integration of both science content and scientific practices.

In the National Curriculum GCSE mathematics specification for higher tier, three assessment objectives are outlined with the following weightings:

- AO1: Use and apply standard techniques – 40%
- AO2: Reason, interpret and communicate mathematically – 30%
- AO3: Solve problems within mathematics and in other contexts – 30% (Department for Education, 2013: 13).

The first of these requires imitative reasoning (both memorised and algorithmic). The second and third assessment objectives involve creative reasoning in that students are required to model phenomena, create arguments, draw conclusions and interpret results. However, Jäder, Lithner and Sidenvall (2020) looked at textbooks in 12 countries (including the UK and Ireland) and found the majority of tasks in all textbooks analysed required only algorithmic reasoning, where solution templates were available to students, and very few required students to use creative reasoning or problem-solving skills. They also found that the percentage of different types of tasks was relatively similar in the textbooks from all 12 countries. They recommend the need for careful selection of tasks from textbooks (by teachers and students) to create opportunities to develop reasoning and problem-solving competency. We suggest that the method used in our research could enable teachers to make more informed choices on tasks they use in teaching and assessment.

## Context for the research

The mathematics curriculum at lower secondary level in Ireland has an emphasis on problem-solving and makes multiple explicit references to reasoning in relation

to problem-solving. *'A task must engage learners and present them with a challenge that requires exploration. Problem-solving tasks activate creative mathematical thinking processes as opposed to imitative thinking processes activated by routine tasks'* (National Council for Curriculum and Assessment, n.d., p. 11). The mention of *'creative mathematical thinking'* and *'imitative thinking'* echo Lithner's understanding of reasoning and it is clear that the written curriculum places a primary focus on CR tasks rather than those involving purely IR.

The science curriculum shares this goal and addresses the concept of reasoning with its emphasis on scientific inquiry. A key component of the curriculum is that of *'scientific literacy'* and it defines a scientifically literate person as one who is capable of designing and assessing methods of scientific inquiry (National Council for Curriculum and Assessment, 2015). It aims to develop students' reasoning skills using tasks involving innovative and inquisitive thinking, encouraging students to critically evaluate and validate methods and data. The importance being placed on scientific inquiry is new to the science curriculum; the previous curriculum contained no references to inquiry.

## Methodology

The Junior Certificate is a set of examinations undertaken by third-year students (typically 14- to 15-year-olds) in Irish schools. The course of study is 3 years long and all examinations are set and marked by the State Examinations Commission (SEC). Examination papers can be found at [www.examinations.ie](http://www.examinations.ie). While a low-stakes examination, it is the first set of state examinations encountered by students and typically determines what level students study particular subjects in upper secondary school. We chose to analyse the three most recent years of the Junior Certificate Higher Level (HL) mathematics and science papers, reasoning that these best aligned with the current curriculum specifications.

As part of a summer student research project, the first two authors extended Lithner's framework to classify tasks in science examination (Table 2). The key alternative framework, which might be used instead, is the TAGS framework; this works on a gradient of cognitive demand rather than sorting tasks clearly into their types of reasoning. While the TAGS framework is useful for identifying different types of tasks, Lithner's framework is more useful for thinking about the type of reasoning required for the task.

We believe that CR cannot take place using only content or only practice. Content and practice must be intertwined to succeed in authentic scientific enquiry; therefore, we chose to differentiate between them in this framework. When extending Lithner's framework, we further developed his ideas of reasoning to form the

**Table 2** Framework for analysis of science tasks – adapted from Lithner's Creative Reasoning Framework (2008)

<p><b>Imitative reasoning</b> Memorised</p> <p>Recalling:</p> <ul style="list-style-type: none"> <li>• Scientific definitions</li> <li>• Chemical formulae</li> <li>• Experimental procedures</li> <li>• Diagrams and labels for apparatus</li> <li>• Functions of components</li> </ul>	<p><b>Imitative reasoning</b> Algorithmic</p> <p><i>Familiar:</i></p> <ul style="list-style-type: none"> <li>• Applying formulae</li> <li>• Practices already seen, e.g. graphing, following a solution outlined previously</li> </ul> <p><i>Delimiting:</i></p> <ul style="list-style-type: none"> <li>• Correct practice to follow is not explicitly suggested. However, script for solving can be obtained by student thinking through learned practices and selecting one that, at its surface, appears connected to task, e.g. describing an experiment from a chapter containing multiple experiments on one topic.</li> </ul> <p><i>Guided:</i></p> <ul style="list-style-type: none"> <li>• Prompts are given in text/by questioner with no connection to underlying concept</li> <li>• All problematic strategic choices are made for student and explicitly suggested in prompts (Lithner, 2008)</li> </ul>
<p><b>Creative reasoning</b> Local creative reasoning</p> <ul style="list-style-type: none"> <li>• Modifying a previously seen practice based on conceptual knowledge, e.g. graphing with unfamiliar variables, experiments with a change in parameters</li> <li>• Some level of flexibility in decision making, i.e. what modification can be made to practice in order to suit task at hand</li> <li>• Student provides justification for modifications based on requirements of task</li> </ul>	<p><b>Creative reasoning</b> Global creative reasoning</p> <ul style="list-style-type: none"> <li>• Novel, e.g. designing an experiment/investigation based on a novel/different scenario</li> <li>• Linking topics and the use of multiple representations</li> <li>• Higher level of flexibility as solution must be original (to student), and student must take into account various factors of task in order to produce a fully rounded solution</li> <li>• Plausibility: Rigorous justification of approach based on deeper scientific understanding</li> </ul>

conjecture that one cannot achieve CR without first having mastered the memorisation and algorithmic components of the concept in question. Anderson and Schunn (2000) explored this idea: they wrote about the ACT-R theory of cognition, which states that in order to learn a complex competence, each component of that competence must be mastered. This provides a clear link between AR and CR, advising that if CR is to be developed on a particular topic, the practices involved within that topic must be learned correctly and repeated in order to gain full competence over them and be able to move on to understanding them.

For the mathematics examination tasks, we chose to use Lithner's Creative Reasoning Framework (2008), in conjunction with Mac an Bhaird *et al.*'s interpretation of the framework (2017) to classify the tasks. As

Lithner's framework and our adapted framework are broadly similar and use the same categories of reasoning, we were able to compare the prevalence of each type of reasoning across the mathematics and science tasks in the relevant examination papers.

## Findings

In this section we present our findings regarding the number of different task types in each examination paper before going on to discuss examples of each task type identified (Table 3). We note that although the majority of tasks were IR level in both science and mathematics examinations, the science tasks mostly required MR while AR was the most frequent category in the mathematics papers.

**Table 3** Task types as percentages of total tasks per paper

Reasoning	Science 2017	Science 2016	Science 2015	Mathematics 2017	Mathematics 2016	Mathematics 2015
MR	86.54% (90)	85.98% (92)	83.5% (86)	13.10% (11)	2.22% (2)	8.54% (7)
AR	10.58% (11)	14.02% (15)	15.53% (16)	73.81% (62)	80.00% (72)	84.15% (69)
LCR	1.92% (2)	0.00%	0.97% (1)	13.10% (11)	16.67% (15)	6.10% (5)
GCR	0.96% (1)	0.00%	0.00%	0.00%	1.11% (1)	1.22% (1)
Total count	104	107	103	84	90	82

(The counts of different reasoning types are presented in brackets.)



**Question 3**

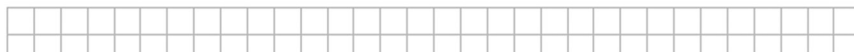
(Suggested maximum time: 25 minutes)

**Table 1** on the right shows the percentage of female members of parliament in each of the current 28 EU countries in 2005 and 2015. The figures are given in increasing order for each year.

Table 1
% of female members of parliament

- (f) Display the data **graphically** in a way that allows you to compare the data for the two years. Label your graph(s) clearly. Show any calculations that you make.

You may use the data from **Table 1** or **Table 2**. The tables are reprinted on the next page.



**Figure 4** 2016 Junior Certificate HL mathematics paper 2, question 3(f)

With the aid of a labelled diagram, describe an experiment to investigate the conversion of the chemical energy in food into heat energy.

**Figure 5** 2017 Junior Certificate HL science paper, question 3(a)

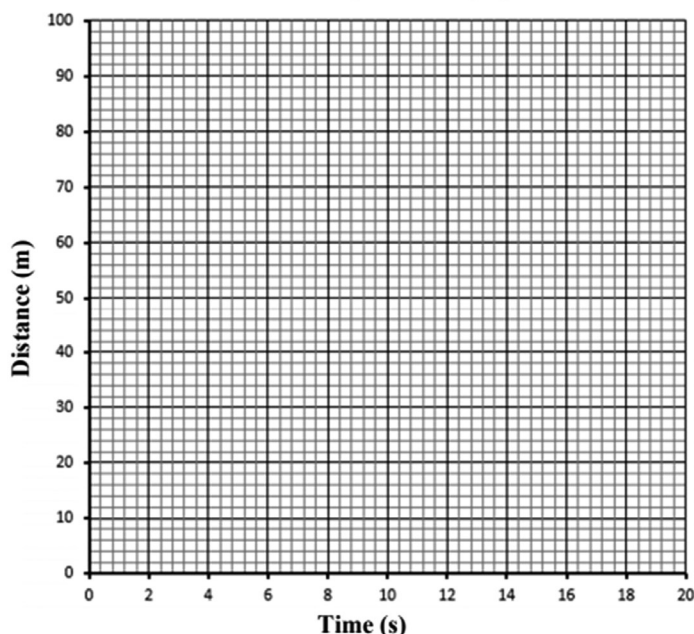
- (c) In a ‘soapbox’ competition a driver raced against the clock in a straight line down a track in a vehicle with no power source. (18)



The table below gives the distances (from the start) travelled by the driver at various times during the run down the track.

<b>Time (s)</b>	<b>0</b>	<b>4</b>	<b>8</b>	<b>12</b>	<b>16</b>	<b>20</b>
<b>Distance (m)</b>	<b>0</b>	<b>6</b>	<b>14</b>	<b>24</b>	<b>44</b>	<b>100</b>

- (i) Use this table to draw a distance against time graph.



- (ii) Find the time taken for the vehicle to travel 80 m. \_\_\_\_\_
- (iii) Calculate the average speed of the vehicle during the last four seconds of the run.

**Figure 6** 2016 Junior Certificate HL science paper, question 9(c) (i), (ii) and (iii)

**Science tasks**

Figure 5 is an example of a type of MR task that commonly appears in the Junior Certificate science examination paper and involves recalling a rote-learned experimental procedure and diagram.

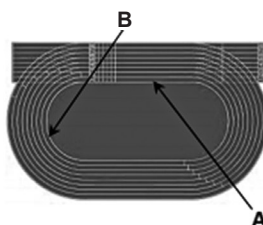
Figure 6 is an example of an AR task taken from the physics section of the science examination paper for

2016. This question involves guided AR. There is no ambiguity associated with what is being asked. Part (i) involves reproducing an algorithmic method of graphing which is already familiar to students due to its prominence as an exercise in both science and mathematics textbooks as well as in past examination papers. Part (ii) involves following a learned algorithmic method while part (iii) involves using a learned practice rehearsed by students prior to examinations.

One of the components of LCR present in Figure 7 is a justified answer. The main concept involved is the distinction between speed and velocity. This requires the student to identify that, due to the shape of the running track, the runner is changing direction at different times, which leads to a change in velocity. Velocity itself is a learned concept, but applying it to this situation may prove difficult for some students, as this is a step above a simple definition of velocity. Students must engage in the scientific practice of constructing an explanation using their conceptual understanding of velocity.

In Figure 8, similar to the previous task, students are required to construct an explanation. However, this question differs in that this concept is novel to the reasoner. They may have used this piece of equipment regularly, but may not have considered the link between the properties of a metal such as iron and the metal’s suitability for use in a piece of equipment as shown. The solution requires a level of flexibility, as students must consider all mitigating factors, such as what the equipment is used for and the conditions it must be able to withstand. Therefore, we have classified this task as GCR.

- (b) Micheál completes one lap of a running track at a constant speed of  $6 \text{ m s}^{-1}$ .
- (i) Does he have the same velocity at A and B? \_\_\_\_\_
- (ii) Explain your answer. \_\_\_\_\_



**Figure 7** 2017 Junior Certificate HL science paper, question 7(b) (i) and (ii)

- (d) (i) Why is the tripod shown in the diagram usually made of iron?
- \_\_\_\_\_
- \_\_\_\_\_



**Figure 8** 2017 Junior Certificate HL science paper, question 4(d) (i)

### Discussion

As can be seen from Table 3, IR tasks are featured much more commonly than CR tasks in Junior Certificate HL mathematics and science examination papers, despite the professed interest in CR in the respective curricula. The prevalence of IR tasks in examinations indicates a clear disconnect between curriculum specifications and assessment material.

We must then ask ourselves why this is the case. The first issue arises within the definition of ‘doing mathematics’ by Smith and Stein (1998), which highlights that high levels of CR may cause feelings of uncertainty and nervousness in the reasoner. The anxiety associated with CR and its perception as being ‘difficult’ is largely due to the unfamiliarity of the thought processes involved. Our analysis of the examination papers demonstrated that students are not familiar with CR tasks in assessments. Students may feel anxious when presented with a CR task, since they are used to regurgitating definitions and formulae. If we wish to tackle this image of CR as being complex and unattainable, we must increase students’ exposure to it by building their reasoning skills through activities that scaffold them through the different levels of reasoning; as per Table 2, each type of reasoning is dependent on the students’ experience with the level of reasoning below it.

Time constraints could be a contributing factor to the prominence of IR questions in both science and mathematics examinations. The Junior Certificate mathematics examination is split into two papers that are both 150 minutes long, while the science examination consists of one 150-minute paper. The examinations cover a wide range of topics, resulting in questions that test a surface knowledge

of each topic rather than focusing on in-depth understanding of a smaller number of topics. This is particularly evident in the extremely high percentages of MR questions in the science examination papers. Examinations formatted in this manner are not conducive to CR, as there is not sufficient time to engage in the complex cognitive processes that characterise CR. Harlen (2006) emphasises the importance of allowing students time to think about a question, so students confronted with a CR question during an examination could struggle to complete it

due to time-pressure.

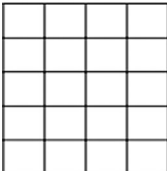
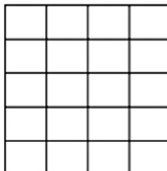
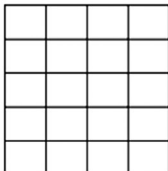
### Recommendations

A task involving CR does not have to be a complex task requiring a lot of resources. Many of the tasks in the examination papers that we analysed could easily be modified to increase their level of required reasoning. To demonstrate this, we have included modified versions of the MR and AR science and mathematics tasks.

The original version of the question modified in Figure 9 tested students’ ability to remember the definitions of rational and irrational numbers. In this modification, students must realise that 0.25 and 3.2 can be represented by shading the diagrams as they can both be written as fractions.  $\sqrt{10}$  cannot be converted into a fraction; therefore, it cannot be represented by shading a grid. This question links decimals to fractions, which in turn link to rational and irrational numbers. Students will already be familiar with representing decimals using diagrams, but must now extend that conceptual knowledge to irrational numbers in a way they may not have previously encountered. As such, we have classified this task as GCR.

The original task in (Figure 5) relied solely on reproducing a rote-learned experimental procedure. The aim

Can you represent the numbers below by shading the grids appropriately? If not, why not? Indicate how you are using each grid – i.e. if you are taking the entire grid to be 1 unit or each row to be 1 unit etc.

0.25	3.2	√10
		

**Figure 9** Modified version of 2016 Junior Certificate HL mathematics paper 1, question 6(b)

A long-distance runner is training for an event and needs to increase their energy intake. They have a choice between two types of energy bars.

Bar A contains 30g carbohydrates and 15g protein. Bar B contains 50g carbohydrates and 10g protein.

(a) Which bar would you recommend for the runner?  
Explain your reasoning.

The bar the runner chooses is sold by two different brands. The runner wants to see which bar will provide them with the chemical energy for the longest amount of time.

(b) Design an experiment to compare how much energy is released by the two bars. Describe fully all the reasoning behind your experiment design.

(c) How would you make sure the experiment you outlined in part (b) is a fair test?

**Figure 10** Modified version of 2017 Junior Certificate HL science paper, question 3(a)

of our adaptation (Figure 10) was to include a greater link between energy and food while challenging the reasoner to engage with CR. For part (a) the reasoner must be able construct a plausible answer by taking into account the various factors of the task, for example the energy requirements of a long-distance runner and the rate of energy release of different biomolecules in the body. Part (b) involves the reasoner adapting a learned experiment to suit the novel scenario presented in the task. The experiment, which students will have carried

out, involves burning a food sample, resulting in a flame, which demonstrates the energy conversion from its chemical form to heat. However, the reasoner must use their prior scientific knowledge to work out a plausible way to compare the amount of heat energy produced by each sample. Part (c) involves the reasoner reflecting on the experiment they have outlined and potentially adjusting its design to produce an accurate experimental procedure. As a result of the modifications made to the learned experiment, this task could now be classed as a high LCR task. However, it would take longer to answer this task in an examination.

As mentioned above, there is scope for CR in revised forms of assessment. How we assess learning is one of the key motivators for learning (Harlen, 2006). Attention to task design needs to be at the forefront of teaching and assessment.

To successfully foster CR in the classroom, supports must be in place for teachers. To strengthen students' reasoning abilities, teachers must lay the groundwork using carefully chosen tasks that scaffold students to the high level of CR and provide students with additional time in class and in examinations. Using the framework outlined above to audit the current level of reasoning required by tasks will enhance the validity of assessments and align curriculum, teaching, learning and assessment.

## References

- Anderson, J. R. and Schunn, C. D. (2000) Implications of the ACT-R learning theory: no magic bullets. In *Advances in Instructional Psychology: Educational Design and Cognitive Science*, ed. Glaser, R. Vol. 5, pp. 1–34. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bergqvist, E. (2007) Types of reasoning in university exams in mathematics. *Journal of Mathematical Behavior*, **26**, 348–370.
- Department for Education (2013) *Mathematics GCSE Subject Content and Assessment Objectives*. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/254441/GCSE\\_mathematics\\_subject\\_content\\_and\\_assessment\\_objectives.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/254441/GCSE_mathematics_subject_content_and_assessment_objectives.pdf).
- Harlen, W. (2006) The role of assessment in developing motivation for learning. In *Assessment and Learning*, ed. Gardner, J. London: Sage.
- Jäder, J., Lithner, J. and Sidenvall, J. (2020) Mathematical problem solving in textbooks from twelve countries. *International Journal of Mathematical Education in Science and Technology*, **51**(7), 1120–1136. Available at: [www.tandfonline.com/doi/full/10.1080/0020739X.2019.1656826](http://www.tandfonline.com/doi/full/10.1080/0020739X.2019.1656826).
- Lithner, J. (2008) A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, **67**, 255–276.
- Mac an Bhaird, C., Nolan, B., O'Shea, A. and Pfeiffer, K. (2017) A study of creative reasoning opportunities in assessments in undergraduate calculus courses. *Research in Mathematics Education*, **19**(2), 147–162.
- National Council for Curriculum and Assessment (n.d.) *Junior Certificate Mathematics Syllabus*. Available at: [www.curriculumonline.ie/getmedia/4f6c8a68-ac41-485c-85a0-32ae6c3559a7/JCSEC18\\_Maths\\_Examination-in-2016.pdf](http://www.curriculumonline.ie/getmedia/4f6c8a68-ac41-485c-85a0-32ae6c3559a7/JCSEC18_Maths_Examination-in-2016.pdf).
- National Council for Curriculum and Assessment (2015) *Junior Cycle Science Curriculum Specification*. Available at: [www.curriculumonline.ie/Junior-cycle/Junior-Cycle-Subjects/Science/](http://www.curriculumonline.ie/Junior-cycle/Junior-Cycle-Subjects/Science/).
- Smith, M. S. and Stein, M. K. (1998) Selecting and creating mathematical tasks: from research to practice. *Mathematics Teaching in the Middle School*, **3**(5), 344–350.
- Tekkumru-Kisa, M., Stein, M. and Schunn, C. (2015) A framework for analyzing cognitive demand and content-practices integration: task analysis guide in science. *Journal of Research in Science Teaching*, **52**(5), 659–685.

**Ailbhe Garry** teaches mathematics and science in Coláiste de hÍde, Ireland. Email: [ailbhe@colaistedehide.ie](mailto:ailbhe@colaistedehide.ie). **Aishling Reilly** teaches chemistry and mathematics at Ardee Community School, Ireland. Email: [aishling.reilly@ardeecs.ie](mailto:aishling.reilly@ardeecs.ie). **Majella Dempsey** lectures in STEM education at Maynooth University Department of Education, Ireland. Email: [majella.dempsey@mu.ie](mailto:majella.dempsey@mu.ie). **Ann O'Shea** lectures in mathematics at Maynooth University Department of Mathematics and Statistics, Ireland. Email: [ann.oshea@mu.ie](mailto:ann.oshea@mu.ie)