# “It’s all just a bit too much!”

# Exploring the use of taught interventions to develop students’ word problem solving strategies in post-16 chemistry.

Stuart Phillips

MSc Learning and Teaching, 2024

# DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

| | |
|---|---|
| Signed [an electronic signature is sufficient]: | |
| Date: | 2nd December 2024 |

# Abstract

Mathematical problems in post-16 chemistry courses such as stoichiometry problems can be challenging for a number of students. By viewing these chemistry problems as mathematical word problems, this study explored the extent to which taught interventions developed students' chemistry word problem solving strategies. Employing a practitioner research approach with a mixed methods design, the study consisted of three main phases: the pre-intervention, the text editing intervention, and the visual representations intervention. Following the intervention phases, analysis of responses to the problems showed an improvement in the structure and coherence of students' problem-solving strategies. This was also accompanied by improvements in performance in relation to the proportion of problem steps attempted and completed correctly. Additionally, perceived difficulty had decreased slightly, and perceived confidence had increased slightly. Improvements in students' problem-solving strategies and performance are proposed to be due to the interventions enabling the development of problem-solving schema, which resulted in a reduction in cognitive load. Finally, four profiles are suggested in relation to chemistry word problem solving: innovative expertise, localised proficiency, aspirational beginner and opportunistic novice.

# Table of Contents

# List of Tables

# List of Figures

# List of abbreviations used

| Abbreviation | Full Term |
|---|---|
| CLT | Cognitive load theory |
| ECL | Extraneous cognitive load |
| GCL | Germane cognitive load |
| ICL | Intrinsic cognitive load |
| LTM | Long-term memory |
| TE | Text editing |
| VR | Visual representation |
| WM | Working memory |
| WP | Word problem |

# 1 Introduction

## 1.1 Is everything OK?

“Is everything OK?” I asked Julia as she sat with her forehead on the desk.

“It’s all just a bit too much!” she replied, half-joking, as she sat up and pointed down to the paper in front of her.

Julia was attempting a particularly challenging calculation in my Year 12 chemistry class. She was having difficulty understanding what the problem on the paper in front of her was asking her to do. The volume of information on the page had left her feeling confused and overwhelmed, unsure of how to best approach the problem in front of her.

A number of authors have investigated students’ challenges in relation to mathematics in chemistry in secondary education (Ardura *et al.*, 2021; Scott, 2012) and in undergraduate courses (Clark *et al.*, 2022; Hoban *et al.*, 2013). Some studies have focussed upon specific concepts such as stoichiometric problems (Clark *et al.*, 2022; Gulacar *et al.*, 2013; Scott, 2012), whilst others have sought to view these challenges in relation to students’ problem-solving skills (Johnstone, 1993; Johnstone & El-Bana, 1984; Overton & Potter, 2011). In a similar way to Julia, many students seem to find the comprehension of chemistry problems challenging. Whilst this has been acknowledged in some of the chemistry education literature (e.g., Ngu *et al*, 2002; Ngu & Yeung, 2013), it does not seem to have been explored in-depth. However, studies into mathematical word problems have noticed these challenges and sought to understand them further (e.g., Daroczy *et al*., 2015; Heggarty *et al*, 1995; Leiss *et al*., 2019).

## 1.2 School context

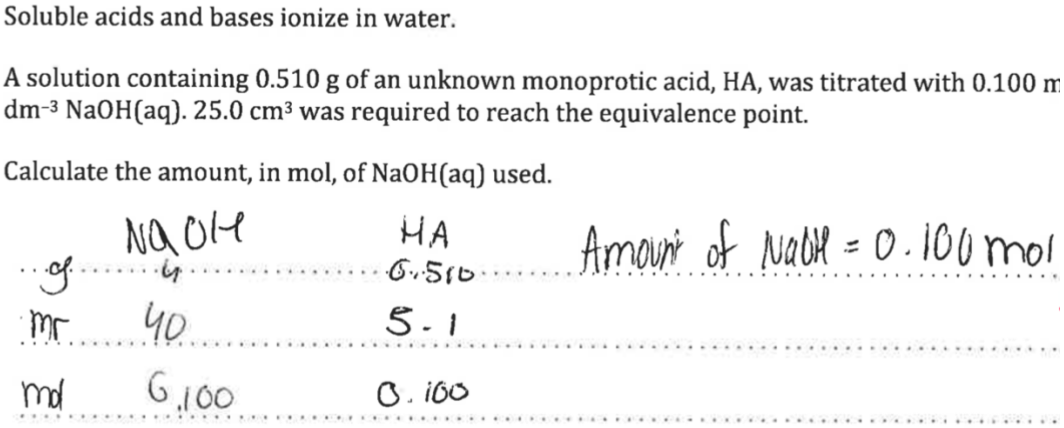
The school is a fee-paying British curriculum international school located in Central Europe, providing early years through to post-16 education.

In relation to post-16 sciences (biology, chemistry and physics), the school offers the International Baccalaureate (IB) Diploma Programme (DP) at Higher level (HL) and Standard level (SL), and the General Certificate of Education Advanced Level (A-level).

The high performing nature of the school means that many students perform very well in public examinations for the chemistry courses that are offered. However, in my experience, there are still a number of students that seem to find the mathematical elements of chemistry challenging.

During lessons, where support is available, many students complete mathematical chemistry problems well. However, during assessments, this tends to be an area where students perform less well. Figure 1 shows a student's answer to a stoichiometry problem from a Year 12 test.

Figure 1: Year 12 student response to a stoichiometry problem



Soluble acids and bases ionize in water.

A solution containing 0.510 g of an unknown monoprotic acid, HA, was titrated with 0.100 m dm$^{-3}$ NaOH(aq). 25.0 cm$^3$ was required to reach the equivalence point.

Calculate the amount, in mol, of NaOH(aq) used.

|  | NaOH | HA |
|---|---|---|
| g | 4 | 0.510 |
| mr | 40 | 5.1 |
| mol | 0.100 | 0.100 |

Amount of NaOH = 0.100 mol

The student's answer shows they have tried to use the data in the question to calculate an answer. Unfortunately, they have misunderstood the focus of the question, and applied an incorrect formula, calculating moles in relation to mass and relative mass, rather than using concentration and volume.

Therefore, the purpose of this study is to understand the reasons as to why students find these problems challenging. Furthermore, the use of taught interventions that may assist with developing their problem-solving skills in relation to these problems will be explored.

# 2 Literature review

## 2.1 Considering the chemistry-maths problem

A number of the challenges posed by the mathematical aspects of chemistry have been studied in detail, especially regarding students' understanding of the mole concept, molar calculations, and stoichiometric relationships (Furió *et al.*, 2000; Gulacar *et al.*, 2013; Raviolo *et al.*, 2021; Strömdahl *et al.*, 1994). Some studies have focussed upon seeking to understand students' performance in chemistry calculations alongside similar areas in mathematics (Scott, 2012; Hoban *et al.*, 2013). Within their methodologies, these studies assume that the mathematical problems are analogous, assessing the same mathematical skills but within a chemistry context. The conclusions of such studies, advocate a deficit view. Arguing the main reason for students' difficulties in these areas of chemistry is that they lack the necessary mathematical skills, or fail to transfer mathematical skills to a chemistry context (e.g., Britton *et al.*, 2005; Hoban *et al.*, 2013; Scott, 2012).

Redish and Kuo (2015), and Wong (2023) have challenged this deficit view. Instead, they assert that mathematics in science is used in different ways, and for different things due to the meanings that are applied to numbers (e.g., units of measurement).

Therefore, considering this argument, mathematics and the mathematical areas of science may not be able to be considered to be analogous with one another.

Within my own professional experience, I have indeed encountered students that do not possess the required mathematical skills to be able to access these areas of chemistry curricula. However, these students have needed significant support and intervention in both chemistry *and* mathematics. Nevertheless, the deficit view does not explain why students viewed as strong mathematicians also find the mathematical areas of chemistry challenging. Therefore, it would seem wise to examine these areas of chemistry from different perspectives.

## 2.2 Chemistry calculations as exercises in problem solving

Whilst the extent to which mathematics and the mathematical areas of chemistry may or may not relate to one another is a matter of debate, the idea that chemistry calculations are exercises in problem solving seems to be agreed upon. One area of chemistry calculations is stoichiometric problems. These are numerical problems that require students to use information about the different substances in chemical equations, to calculate an unknown property of a specific substance. Whilst not an exhaustive list, these properties may include: the mass, amount (moles), or volumes and concentrations of solutions (volumetric analysis).

This type of problem solving in chemistry education literature is often referred to as a closed problem, whereby there is a single, definite answer, often solvable by algorithmic methods (Johnstone, 1993; Chandrasegaran *et al.*, 2009; Tsaparlis, 2021a). In direct contrast, open problems are defined as having several reasonable answers and require students to use unfamiliar methods to identify a solution pathway (Johnstone, 1993; Cooper & Stowe, 2018; Randles & Overton, 2015). Some argue that most stoichiometry problems that students experience within secondary education are of a closed nature (Reid & Yang, 2002). In many instances this may well be true. Assessment within courses such as A-level and the International Baccalaureate (IB) is heavily, if not entirely examination-based, and therefore require students to solve problems from given data, with a definite outcome. However, the fact remains that stoichiometry problems can be problematic.

Literature from chemistry education research tends to focus on three main factors that contribute to problem solving difficulty in relation to stoichiometry problems: their multiple-step nature; understanding of chemistry concepts; and problem format.

A stoichiometry problem such as finding the mass of a product from a known mass of reactants, may require between five and seven separate calculation steps. This may include writing chemical equations and the manipulation of mathematical formulae (Gulacar *et al.*, 2021). The multiple-step nature of stoichiometry problems is a potential source of difficulty for many students. The high number of independent, yet interconnected steps that must be taken increase the chances for error. A mistake in an individual step of the problem causes the final answer to be incorrect (Gulacar *et al*., 2021). Another contributing factor is that some steps may be more difficult than others (Tsaparlis, 2021b).

Students' conceptual understanding of the problem content has also been shown to hinder students' problem-solving ability. For example, misconceptions of stoichiometric ratios and how they relate to a chemical equation may lead students to incorrectly calculate the amount of other substances in a reaction (e.g., Clark *et al*., 2022). Gulacar *et al*. (2013) studied undergraduate students' difficulties with stoichiometry problems in relation to their understanding of chemistry concepts. They found that writing chemical equations was associated with low attempt and successful completion rates (45% and 43% respectively). However, balancing a provided chemical equation was associated with much higher attempt (90%) and successful completion (89%) rates. This may suggest that students experience difficulty in writing chemical equations but are able to apply learned methods to successfully balance provided equations. Additionally, the study found that the concepts of limiting reagent, stoichiometric ratio, and the mole concept were characterised by high attempt rates but lower successful completion rates (Gulacar *et al*., 2013). This indicates that students are aware of the steps required to solve the problem but are not able to complete these steps correctly. Indeed, the study concludes that "students in the study lacked the ability to properly plan how to solve a multi-step problem like a stoichiometry problem” (Gulacar *et al*., 2013, p. 514).

Problem format has also been found to affect problem solving success. A study conducted by Lazonby *et al*. (1985) investigated the impact of posing the same stoichiometric problem in different formats. Firstly, only the problem itself was given without any guidance, as shown by Figure 1 below:

Figure 2: Stoichiometry problem example (Lazonby *et al*., 1985, p. 60)

> Silver chloride (AgCl) is formed in the following reaction: $AgNO_3 + HCl \rightarrow AgCl + HNO_3$
>
> Calculate the maximum yield of solid silver chloride which can obtained from reacting 25 $cm^3$ of 2.0M hydrochloric acid with excess silver nitrate (AgCl = 143.5).

The subsequent formats showed the question in separate parts, guiding students through the problem-solving process. The findings revealed that when problems were broken down into separate parts, students were able to demonstrate their understanding more easily (Lazonby *et al*.,1985). In this respect, Figure 1 represents an unstructured problem, whereas separating it into a set of smaller subproblems leading to the same answer represents a structured problem.

Considering stoichiometry problems from a problem-solving viewpoint gives useful insights into why students may experience difficulties with these types of problems. Their multiple-step nature, reliance on subject-specific knowledge and varying formats provides many potential sources of error during the problem-solving process. However, even with the above points in mind, it is likely that there are other factors in relation to the structure of these problems and the problem-solving process that warrant further investigation.

## 2.3 Mathematical word problems

A small number of studies have considered stoichiometry problems as mathematical word problems (e.g., Ngu *et al*., 2002; Ngu & Yeung, 2013; Tang *et al*., 2014). However, it would seem that the influence of the specific elements of word problems has not been examined explicitly within a chemistry-specific context.

Word problems (WPs) can be defined as mathematical tasks whereby task information is usually communicated in written form, rather than in a mathematical format such as an equation (Daroczy *et al*., 2015; Strohmaier *et al*., 2022; Verschaffel *et al*., 2020). Whilst a staple in mathematics curricula globally, WPs are often considered to be one of the more difficult types of problem that students face (Jaffe & Bolger, 2023; Verschaffel *et al*., 2020).

### 2.3.1 The role of selected linguistic and numerical factors

Stoichiometry problems may be considered as WPs due to the fact that they are mathematical tasks communicated to students through written means (see Figure 1). However, research into WPs shows the existence of complex interactions between the linguistic and numerical factors of the problem. Further analysis of these factors of WPs may offer valuable insights into why stoichiometric problems are often challenging for students. Therefore, a selection of the linguistic and numerical factors that are most relevant to chemistry WPs will be discussed.

Solving WPs involves developing a situational model, which is most simply understood as the process of understanding the content of a task. Development of a suitable situational model dependent upon students' general and domain-specific reading and comprehension skills (Kintsch & Greeno, 1985; Leiss *et al*. 2019). This is further supported by the work of Vilenius-Tuohimaa *et al*. (2008) who found that WP solving success was strongly related to reading comprehension. Additionally, technical reading abilities further improved WP solving. However, when technical reading level was controlled for, performance in WPs was still related to reading comprehension. Therefore, it was suggested that fluent technical reading skills "may work as a 'capacity enhancer', helping students to further leverage their prior knowledge related to the problem" (Vilenius-Tuohimaa *et al.*, 2008, p. 422). The role of comprehension has been found to be fundamental in successfully solving WPs, with this step taking approximately 41% of the total solution time. Although this is strongly affected by the linguistic complexity of the task (Leiss *et al.*, 2019).

Jerman & Rees (1972) examined the effect of the word count of the problem text, finding that it contributed to the difficulty of WPs. Conversely, Lepik (1990) found that word count did not significantly predict WP solving but was strongly linked to the time taken to solve WPs. More recently, Walkington *et al.* (2018) found that WP length and WP solving accuracy were negatively correlated. However, in studies where word count was changed without changing the situational or conceptual context of the WP, no significant differences in performance were found (Vondrová *et al.*, 2018; Walkington *et al.*, 2019).

The presence of irrelevant linguistic and numerical information within the WP statement has been found to result in lower WP solving performance (Muth, 1992). However, this is disputed by Englert *et al.* (1987) who found that irrelevant numerical information has been found to negatively affect WP solving performance, whilst irrelevant linguistic information did not. Furthermore, the presence of implicit information within the WP statement has been found to affect performance. Unsuccessful problem solvers may often use methods such as direct translation (looking for numbers and keywords), meaning that important implicit information may be missed when employing this strategy (Hegarty *et al.* 1995).

Problems become more complex with a larger number of calculation steps (Terao *et al*., 2004). Quintero (1983) studied students' WP performance in relation to ratio, finding that two-step problems were more difficult than single-step problems. Additionally, when all other factors are controlled, Muth (1992) observed that two-step problems experienced a larger number of errors than one-step problems. However, it should be noted that Daroczy *et al*. (2015) warn that it would be difficult to categorically conclude that the reason two-step problems are more difficult is solely due to mathematical complexity because, "in two-step problems, the WP has also become more difficult linguistically as it usually contains more phrases and semantic distractors" (p. 348).

### 2.3.2 Word problem complexity

Based on the linguistic and numerical factors contained within the problem, WPs can be considered to be simple or complex. Strohmaier (2020) contrasts these WPs and concludes that simple problems follow a simple syntax, within a superficial context that requires a single mathematical operation. Conversely, the syntax of complex problems does not mirror the solution process. Furthermore, complex problems contain irrelevant information, with multiple representations, within a context that is central to the problem solution (Strohmaier, 2020).

Complex WPs contain a mathematical concept within a real-life situation whereby both must be understood to successfully solve the problem, requiring connections between the mathematical and extra-mathematical (real-world) contexts (Leiss *et al.*, 2019). Therefore, complex WPs require considerable comprehension processes and strong engagement with of the problem context (Leiss *et al.*, 2019). Furthermore, the central mathematical aspect of the WP is not able to be solved by fact retrieval, therefore requiring calculations (Strohmaier *et al.*, 2022).

## 2.4 Positioning stoichiometry problems as complex word problems

As previously mentioned, chemistry education literature in relation to the challenges of stoichiometry problems tends to focus on the multiple-step nature of problems (Gulacar *et al.*, 2021), students' conceptual understanding of chemistry topics (Chandrasegaran *et al.*, 2009; Gulacar *et al.*, 2013), and problem type and structure (Cooper & Stowe, 2018; Lazonby *et al.*,1985; Tsaparlis, 2021a). Although some studies do refer to stoichiometry problems as WPs (Ngu *et al.*, 2002; Ngu & Yeung, 2013; Tang *et al.* 2014), there is little evaluation of the complex relationships between the linguistic and numerical factors of these problems, and how it may influence students' performance.

On the other hand, mathematics education literature has studied these factors extensively, especially in relation to students' comprehension of problems and the role of multiple solution steps (Daroczy *et al*., 2015; Leiss *et al*., 2010, 2019; Muth, 1992; Terao *et al*., 2004). It would therefore seem sensible to examine these factors in relation stoichiometry problems, and how they may affect students' ability to solve these problems.

Figure 2 (Section 2.2) and 3 (below) show two different stoichiometric problems in chemistry.

Figure 3: A-level titration problem (Edexcel, 2016)

Table 1 below gives a summary of the linguistic and numerical factors for each problem, along with other relevant factors.

Table 1: Comparison of factors for two stoichiometry problems

| | **Problem** | **Figure 2** | **Figure 3** |
|---|---|---|---|
| Linguistic Factors | Word count | 42 | 179 |
| | Technical vocabulary | Yes | Yes |
| | Syntax follows the problem solution | Partially | No |
| | Irrelevant/superfluous information | No | Yes |
| | Real-world situation | No | Yes |
| Numerical factors | Irrelevant/superfluous numbers | No | Yes |
| | Number of steps required to solve | 5 | 13 |
| Other relevant factors | Structured or unstructured | Unstructured | Unstructured |
| | Different units of measurement | 2 | 3 |
| | Unit conversions needed | Yes (1) | Yes (2) |
| | Number of different formulae needed | 2 | 4 |

Both problems contain many elements of linguistic and numerical complexity. In relation to linguistic complexity, both contain significant chemistry-specific vocabulary. Without an understanding of this vocabulary, it would be impossible to solve the problem. However, Figure 3 contains more elements of linguistic complexity than Figure 2, including a real-world context that is central to understanding and solving the problem. The numerical complexity of each problem is high, requiring multiple solution steps. When examining the level of numerical complexity, Figure 3 could be considered to be significantly more complex than Figure 2. When considering some other relevant factors that may add to problem complexity, both are unstructured problems, and require the use of more than one formula, and unit conversions to reach a solution.

Considering all of the above, there may be a strong argument for viewing stoichiometry problems as complex WPs. By doing so, it may offer valuable insights into why students find them challenging and enable teachers to better support students.

Therefore, for the purposes of this study, the term ‘chemistry word problem’ will be used to refer to problems of this nature.

## 2.5 Working memory, cognitive load theory and chemistry word problems

With respect to WPs in mathematics, the role of working memory (WM) has been strongly linked to WP solving ability (Fuchs *et al.*, 2020). Within the literature from mathematics education, the role of WM tends to focus on the interactions between the linguistic and numerical factors of a WP, and how these may place increased demands on WM. For example, Fuchs *et al*. (2020) argue that “WM resources involved in solving word problems differ from and are more numerous than those underlying number knowledge or calculation skill” (p.11). Peng *et al*. (2016) further acknowledge the multiple-step nature of solving WPs, stating that they “draw upon significant WM resources” (p. 457). However, in relation to stoichiometric problem solving in chemistry education, the main consideration in terms of WM seems to be related to the number of problem steps and centres around the Johnstone-El-Bana Model (Johnstone & El-Bana, 1984, 1986).

The Johnstone-El-Bana Model proposes a bottleneck in information at around five to six problem steps resulting in WM overload and significantly decreased performance (Tsaparlis, 2021b). However, a study by Opdenacker *et al*. (1990) found that as the number of problem steps increased, performance decreased smoothly, with no sharp decrease that may indicate sudden WM overload. Furthermore, Tsaparlis (2021b) argues that the Johnstone-El-Bana Model may not be sufficient, as it does not account for unsuccessful attempts when problems may be regarded as less difficult. In explanation, four possibilities are offered: incomplete conceptual understanding (including forgetting); problem steps not being of equivalent difficulty; being able to work out the part, single step subproblems but not the composite problem; and the presence of irrelevant information (Tsaparlis, 2021b). Tsaparlis (2021b) also offers two explanations for when the model may be violated (successful solution by lower achieving students): familiarity with the problem type, and successful 'chunking' of the problem, decreasing WM demand.

WM is often referred to as one's ability to retain and process small amounts of information for short time periods (Baddeley, 2012; Cowan; 2017). Miller (1956) proposed that the WM is limited in capacity, being able to hold seven (± 2) 'bits' such as individual numbers in a sequence. The limited capacity of WM seems to effectively apply only to novel information (Sweller *et al*, 2019). Consequently, this may have direct impacts on solving complex WPs, as they centre around situations that are not experienced by learners in their daily lessons (Verschaffel *et al.*, 2020).

However, this limited capacity for bits of information can be overcome by the process of chunking (Miller 1956). Through chunking, individual bits of information are organised and recoded into single units or 'chunks'. Therefore, using only chunks of information reduces the storage load on WM, freeing up its capacity to maintain other information concurrently (Thalman *et al*., 2019). The WM is believed to be able to hold around four chunks of information at a given time (Cowan, 2001), with chunk size being found to have no influence on how much WM capacity it requires (Cowan, 2001; Thalman *et al*., 2019). Furthermore, chunks of new information held in WM can be compared to information held in the long-term memory (LTM). If matching information is found, it is recalled from the LTM and encoded into the WM. This further reduces WM load, facilitating the encoding of subsequent information into the WM (Thalman *et al*., 2019; Waterman & Miller, 2022).

During the process of comparing chunks of information in the WM to information stored in the LTM, an existing schema may be activated. Schema or schemas (plural) are mental frameworks that allow information to be organised and processed (Anderson, 1977). Schema can be used to transform a complex problem into smaller, more familiar problem-solving sequences (Overton & Potter, 2011). Additionally, according to Piaget (1999) new sources of information may be able to be combined with existing schema through assimilation.

Cognitive load theory (CLT) (Sweller, 1988; Sweller *et al*., 1998) posits that to enable the successful completion of problem-solving tasks, the cognitive load needed to complete the task (or as a result of the task) must not exceed the capacity of the WM (Sweller, 1994; Sweller et al., 2019). Chandler & Sweller (1991) outline three types of cognitive load: intrinsic cognitive load (ICL), extraneous cognitive load (ECL), and germane cognitive load (GCL). ICL depends upon task complexity-the number of interacting information elements in the task. These elements must be processed simultaneously in order to successfully carry out the task (van Gog *et al.*, 2011). Due to the limited capacity of the WM, tasks that have a higher number of interacting elements impose a greater ICL on the WM (van Gog *et al.*, 2011). ECL is caused by the way in which tasks are designed. It can arise from cognitive processes such as processing redundant or irrelevant information (Kalyuga, 2011), or by engaging in less effective problem-solving strategies, for example, means-ends-analysis (trial-and-error) (Sweller *et al.*, 2019; van Merriënboer & Sweller, 2005). GCL is a product of processes caused by the design of the task that contribute to learning, such as the use of schema from the LTM (van Merriënboer & Sweller, 2005). Therefore, in order for problem solving (and learning) to be effective and efficient, ECL should be minimised, and GCL optimised in order to maximise available WM resources (Sweller *et al.*, 1998, 2019).

The Environmental Organising and Linking Principle of CLT (Sweller *et al.*, 2019) relates to the linking or combining of novel information with schemas held in the LTM, allowing it to be more easily recalled. Therefore, the use of schema causes “the limitations of working memory disappear, and the information can be transferred back to working memory using the Environmental Organising and Linking Principle to generate appropriate action” (Sweller *et al.*, 2019, p. 274). However, on occasions when information stored in LTM is not available it will need to be generated (Sweller *et al*, 2019). In this case, new information is generated using a “random generate and test procedure” (Sweller *et al*, 2019, p. 273) during problem solving. This is known as the Randomness as Genesis Principle, whereby, only effective strategies are retained (Sweller *et al*, 2019).

These two principles from CLT may go some way to explaining why WPs and more specifically to this study, chemistry WPs are so challenging. Becoming proficient in solving chemistry WPs places high demands on WM. Additionally, the development of new and adaptation of existing schemas may require additional investments in relation to time and cognitive resources. Understanding this may be especially important in relation to examinations, as the ICL and ECL of problems may be high and out of the control of students and teachers. Therefore, enabling students to understand and independently use strategies that are able to optimise GCL is especially important.

## 2.6 Expert vs. novice problem solving

High levels of subject knowledge and its application is often regarded as a prerequisite for expertise (Gulacar *et al*., 2019; Larkin *et al*., 1980), with experts tending to view their subjects more holistically than novices (Chi *et al*., 1981; Silver, 1979). Silver (1979) found that in relation to mathematical problems, experts categorised them based upon underlying mathematical structures or shared problem structures. In physics, Chi *et al*. (1981) observed that experts categorised physics problems in relation to scientific principles, whereas novices tended to focus upon features such as the specific topic mentioned in the problem. Being able to take a broad view of their subject could enable experts to use their familiarity with the problem content and structure to more readily formulate a solution plan than novices.

The notion of experts being able to make explicit use of their experience and familiarity is exemplified by Gulacar *et al*. (2019) who observed how experts' familiarity with stoichiometry problems enabled them to complete unfamiliar problems. They found many incidences of participants recognising specific elements of problems and relating them to methods they had used previously (Gulacar *et al*., 2019). Related to experts' familiarity is their effective use of chunking. More effective chunking occurs with increased familiarity with the information (Overton & Potter, 2011). In relation to chess, Gobet and Clarkson (2004) found that the size of the chunks used by Masters were larger than those of novices, but there was little difference in the number of chunks used by both. These findings may relate directly to the solving of chemistry WPs. If experts are able to combine new information from WPs into larger chunks, more WM resources may be available for other processes related to the problem.

Another element of expertise in relation to WM is the possession of larger amounts of relevant, domain-specific knowledge in the form of diverse and complex schemas (Tricot & Sweller, 2014). These schemas may significantly reduce WM load as even highly complex schema can be dealt with as one element in the WM (van Merriënboer &

Sweller, 2005). In this respect, it could be argued that the WM of experts is more efficient than that of novices. This is likely due to experts being able to draw on their extensive library of complex schemas through The Environmental Organising and Linking Principle (Sweller *et al.*, 2019) and optimising GCL. This may also mean that when confronted with unfamiliar problems, experts are likely to spend less time and exert fewer WM resources in relation to the Randomness as Genesis Principle. Whereas, for novices, the opposite may be true, as without effective schemas or chunking, the WM may become overloaded and cognitive overload occurs (Overton & Potter, 2011).

Linked to the WM characteristics of experts and novices but worth discussing separately, is the fact that experts (and those developing expertise) have been observed to use fewer solution steps when completing multiple-step problems (Blessing & Anderson, 1996; Koedinger & Anderson, 1990; Ngu & Yeung, 2013). Ngu and Phan (2022) propose that this step-skipping performance of experts is likely due to the presence of a schema linking conceptual knowledge and procedural knowledge, and that “the schema allows the experts to use conceptual knowledge to refine the solutions steps, resulting in fewer solution steps” (p. 2).

Studies have also shown that expert and novice problem solvers exhibit different attributes and behaviours. Overton *et al*. (2013) studied the problem-solving approaches of chemists from different backgrounds (undergraduates, academics and industry) in relation to open-ended problems. They suggested that expert problem solvers were characterised by adopting approaches such as: understanding the problem and employing a structured method, being able to cope with a lack of data, and evaluate their solutions. Novices' approaches were defined as: lacking a structured strategy, being unable to define the problem, very limited evaluation, being unable to detach themselves from the context of the problem and seeking an arithmetic approach, and no clear problem-solving approach (Overton *et al.*, 2013). Despite these categories being related to open-ended problems, they are also applicable to stoichiometry problems. However, references to arithmetic approaches are not considered as stoichiometric problems are arithmetic by nature.

Being able to understand problems suggests the importance of problem comprehension, as the successful solving of complex problems requires domain-specific knowledge in addition to domain-specific reading and comprehension skills (Kintsch & Greeno, 1985; Leiss *et al.*, 2019). Therefore, in addition to their strong subject knowledge, experts may have increased levels of comprehension skills, or at least higher levels of technical reading skill that act as a 'capacity enhancer' (Vilenius-Tuohimaa *et al.*, 2008). Consequently, experts may be able to construct more effective situational models to assist in problem solving than novices.

The role of evaluation of answers and coherence of strategy in both categories may indicate differences in the metacognition and metacognitive skills between experts and novices. Experts may be able to employ evaluative skills more effectively due to their more developed metacognitive skills, which apply throughout the problem-solving process (Depaepe *et al*., 2010). Thus, experts may engage in evaluative behaviour more readily than novices. The ability to formulate a more logical and structured approach could be a function of experts' metacognitive knowledge. Metacognitive knowledge is an important element of metacognition (Flavell, 1979), and can been viewed as knowledge of "one's own cognitive processing, experiences during learning and the outcomes of their activities" (Efklides, 2011, p. 8). Experts' increased metacognitive knowledge may assist them in choosing effective problem-solving strategies, as metacognitive knowledge also includes strategic knowledge about the advantages and disadvantages of applying a particular strategy in a given situation (Jia *et al*., 2019).

Due to their relative success, experts may experience higher levels of self-efficacy. Self-efficacy can be defined as a judgement as to one's ability or capacity to perform domain-specific tasks (Bandura, 1997). Moreover, self-efficacy can be thought to act and vary across three dimensions: facet-specificity (general or specific tasks), level of difficulty (perceived as easy or hard), and strength (confidence in one's capability to perform the task) (Bandura, 1997; Street *et al*., 2017; Street *et al*., 2022). In this respect, experts could be viewed as having increased self-efficacy as a product of higher levels of confidence across a wide range of tasks (general and specific), which they perceive as less difficult. More positive experiences across the three dimensions of self-efficacy may be due to experts' familiarity with the subject; ability to free up WM resources by more effectively chunking information and the use of schemas; and possessing more developed metacognitive skills and knowledge.

## 2.7 Interventions to support the development of chemistry word problem solving strategies

There are a number of different interventions that may develop students' chemistry WP solving strategies. The following section contains a brief discussion of some interventions that may be the most appropriate for chemistry WPs.

### 2.7.1 Text editing

Originally developed by Low and Over (1990), text editing (TE) was used as a tool to measure schema acquisition in mathematics education and was consequently developed as an instructional technique in relation to algebraic WPs (Low *et al.*, 1994). Later, Ngu *et al.* (2002) used TE in relation to chemistry WPs, due to algebraic and chemistry WPs sharing similar characteristics (Ngu *et al.* 2002).

The premise of TE is relatively straightforward, requiring students to categorise WPs as having sufficient information to be solved; missing information; or containing irrelevant information. Ngu *et al*. (2002) investigated the use of this strategy in comparison to conventional problem solving for molarity problems and stoichiometric problems. They found that for molarity problems the TE group performed better than the conventional problem-solving group. However, when completing stoichiometric problems, the performance of the two groups was reversed. This difference was attributed to the fact that "text editing did not give students practise in the skills required to solve stoichiometry problems" (Ngu *et al*., 2002, p.395). However, Ngu and Yeung (2013) did find that students using TE made few computational errors, possibly due to them being able to identify useful numerical values in the problem statement more readily. Therefore, TE may have relevance in relation to assisting students with making sense of the information contained in the WP statement. Due to the limitations identified above, TE could potentially be more useful if combined with another problem-solving approach. This could be viewed as especially important with respect to stoichiometry problems where there are many different numerical values for a range of substances that must be evaluated before solving the problem.

### 2.7.2 Visual representations

The use of visual representations (VRs) is another strategy that has been noted to be useful in WP solving. The use of VRs in WP solving has been well studied (Boonen *et al*., 2014; Jitendra *et al*., 2007; Stylianou & Silver, 2004) and there is evidence to show that students who are able to construct an accurate VR of a problem may be more successful (Boonen et al., 2014; Hegarty & Kozhevnikov, 1999). However, when students construct their own VRs, they may fail to construct an appropriate VR or may omit it altogether (Ayabe *et al*., 2022). Moreover, the accuracy of students' VRs influence WP solving performance (e.g., Krawitz & Schukajlow, 2020). This may be especially important when considering what Ayabe *et al*. (2022) call the problem-appropriateness of a VR. Studies such as Hurley and Novick (2010) found that performance was lower in cases where the VR was not matched to the problem. Further highlighting the importance of the correct type of VR in relation to the problem.

Some studies have examined the effect of VR in relation perceived cognitive load in addition to WP solving performance. Beitzel and Staley (2015) found that the overall, the use of VRs in undergraduate probability problems reduced students perceived cognitive load and increased solution accuracy. Related to this, it has been proposed that some VRs such as tables may act as, or in a similar way to schema. Ayabe *et al*. (2022) and Zahner and Corter (2010) argue that empty cells in a table draw attention, enabling solvers to identify missing information. This in turn allows the problem to be broken down into smaller, more manageable steps, and may allow connections to relevant formulae (Zahner & Corter, 2010). This may be especially important as the use of a table to structure stoichiometry problems is a common teaching method, especially at GCSE-level.

### 2.7.3 Deliberate practice

One facet of expertise is the development of complex cognitive schemas that allow the recognition and retrieval of large chunks of appropriate information from the LTM. A commonly agreed requirement for the development of expertise within a domain is extensive practice (Pachman *et al*., 2013). The process of deliberate practice (Ericsson *et al*., 1993) may be best thought of as the completion of well-defined tasks that are challenging but achievable (Ericsson, 1996, 2018). Furthermore, these tasks are specifically tailored to a learner's weaker areas, leading to the development of specific improvement goals (Ericsson, 2006). However, the empirical studies into the effects of deliberate practice in a classroom setting appear to be limited.

Pachman *et al*. (2013) studied the effect of deliberate practice on 8th Grade students' performance in geometry. They found that students in the deliberate practice condition performed better than students in the free-choice practice condition, as the latter tended to select practise activities that were less challenging. This choice was found to effect higher achieving students the most, as they did not challenge themselves enough to reach higher levels of performance. The study also found that deliberate practice was more effective when it was focussed on fewer problem areas. The suggested reason for this was that concentrating on too many problem areas at a time may overload WM and negatively affect schema development, especially in relation to novice learners. According to Pachman *et al*. (2013) further research is required to better understand how deliberate practice could be operationalised in cognitively rich domains. However, from a teaching and learning perspective, the use of deliberate practice requires careful consideration of students' areas for development, selection of appropriate practise materials, monitoring of students' progress and ensuring that areas of competence are not neglected.

### 2.7.4 Worked Examples

Worked examples provide students with a step-by-step demonstration of a task that makes the process of completing the task and how to achieve the answer clear (EEF, 2021). When faced with new problems to solve, novice learners will generally engage in less effective problem-solving strategies such as direct translation (Hegarty *et al*., 1995) and means-ends-analysis. It is therefore argued that the use of worked examples as a teaching and learning strategy reduces ECL caused by less effective problem-solving methods, and focusses learner's attention on problem structures and useful solution steps (Leppink *et al*., 2014; van Merriënboer & Sweller, 2005). The reduction of cognitive load therefore allows novice learners to assign more of their WM resources to dealing with ICL and to eventually develop schema in relation to the problem (Leppink *et al*., 2014).

Worked examples have been shown to improving learning, particularly for students with lower prior knowledge (Cooper & Sweller, 1987; Sweller & Cooper, 1985). By contrast, more expert learners, are more likely to benefit from autonomous problem solving as they have already acquired the schema to solve the problem, which assists their problem-solving approach (Leppink *et al*., 2014). Therefore, worked example formats designed for novice learners can actually result in negative effects as learners gain expertise, this is known as the 'expertise reversal effect' (Kalyuga *et al*., 2003). Additionally, it has been observed that worked examples are less effective when examples contain redundant information (e.g., Tarmizi & Sweller, 1988). Therefore, from a teaching and learning perspective, it would be prudent to take account of the nature of the problems being solved, in addition to learners' prior knowledge and how to monitor their progress during the use us worked examples (EEF, 2021).

## 2.8 Summary

In order to fully understand students' difficulties with chemistry WPs such as stoichiometry problems, it may be necessary to examine them from a number of different perspectives. The complex relationship between elements of chemistry education, mathematics education, and CLT must be considered in order to effectively support students within this challenging area of chemistry.

Viewing stoichiometry problems from a chemistry problem solving perspective enables understanding of the important interactions between students' conceptual understanding of chemistry and the multiple-step nature of chemistry WPs (Gulacar, 2013; Tsaparlis, 2021a). However, supplementing this understanding with the findings from studies into mathematical WPs may shed light on new perspectives as to why students find these areas of chemistry challenging. For example, WP research tends to have a strong focus on the underlying linguistic structure of problems (Leiss *et al.*, 2019), and how this interacts with the numerical factors of problems (Daroczy *et al.*, 2015). Furthermore, WP research in mathematics also offers valuable insights into how considering stoichiometry problems as complex WPs may further contribute to knowledge in this area. In relation to the latter point, this may be especially important as much of the literature in chemistry education shows a tendency to classify stoichiometry problems as 'routine exercises', categorising them as closed problems that are often solvable by algorithmic means (Cooper & Stowe, 2018; Reid & Yang, 2002; Tsaparlis, 2021a). However, mathematics WP research broadly tends to take the view that a true problem is dependent upon familiarity with the problem, and proficiency in the types of required knowledge and skills (Verschaffel *et al.*, 2014).

Ideas from CLT and WM research contribute further to our wider understanding of chemistry WPs. This is especially important given the fact that WPs place significant demands upon WM resources (Peng *et al.*, 2016). Understanding the role of WM and CLT in relation to the structure of chemistry WPs may help with enabling students to make use of their existing schemas and familiarity with related problems. Additionally, it may further enable students to better cope with the presence of irrelevant or superfluous information, and the number of steps required to solve the problem.

Although the number of problem steps has been explored in relation to chemistry problem solving through the Johnstone-El-Bana model, this model has been subjected to criticism.

Enabling the successful solution of chemistry WPs by students of all abilities is the eventual goal, facilitating the transition from novice to expert. Therefore, by considering how the attitudes and behaviours of expert learners can be operationalised may facilitate the design of effective interventions that lead to improved chemistry WP performance.

Although there is a wealth of research and literature detailing students challenges with stoichiometry problems, there only seems to be a handful that view mathematical chemistry problems, such as stoichiometry problems as WPs. Additionally, it appears to be difficult to find studies that seek to understand the effect of interventions from the perspectives of students. This seems to be especially true in relation to qualitative or mixed methods studies into students' perceived difficulty and confidence in relation to these problems. Therefore, this study seeks to contribute to this area by positioning stoichiometry problems as complex WPs. The study will seek to understand how taught interventions influence student performance, perceived confidence and difficulty. Additionally, how students use interventions to develop their chemistry WP solving strategies will be explored.

## 2.9 Research questions

This study seeks to contribute to the field of problem solving in chemistry by exploring the following research questions:

1. Which types of chemistry word problem do students find the most challenging?
2. How do taught interventions impact students' chemistry word problem performance, perceived difficulty and confidence?
3. How do students use taught interventions to solve chemistry word problems?

# 3 Methodology

This study used a practitioner research approach, with elements inspired by design-based research (DBR). Although small-scale applications of DBR in relation to graduate studies have been discussed (e.g. McKenney & Reeves, 2012; Pool & Laubscher, 2016), it was decided that the time scale of this study was too short to allow for the important, iterative nature of a DBR approach.

Dinkelman (2003) defines practitioner research as “the intentional and systematic inquiry into one’s own practice” (p. 8), focusing on questions that emerge from teachers’ observations, everyday experiences, and own practice (Campbell, 2013; Menter *et al.*, 2016; Westbroek *et al.*, 2022). In this respect, teachers seek to find possible solutions to specific practical problems and questions, with the aim of offering evidence to illustrate what may work in practice and why (Cochran-Smith & Lytle, 2009; Ponte, 2005).

Practitioner research is participatory in nature, whereby educators are integral members of the process (Mertler, 2017), being involved in the practices themselves and understanding them from an ‘insider position’ (Anderson & Herr 1999). In relation to this particular study, two positionalities were employed based on the categorisations provided by Herr and Anderson (2005, p. 31). At the beginning of the study, the “insider” positionality was employed as I began researching my own practice and identifying a problem to be investigated. However, as the study progressed, my positionality evolved to that of “insider in collaboration with other insiders”, with the study participants acting as the ‘other insiders’. This allowed me to gain a deeper insight into their perspectives and an increased awareness of their needs and motivations (Levin & Rock, 2003).

The absence of a control group could be viewed as potentially problematic. However, in keeping with the DBR-inspired approach used in this study, a traditional control group was not deemed necessary as DBR compares iterations of the intervention and that the agenda of the designer(s) is not seen as a threat to validity (Hoadley & Campos, 2022). Additionally, a DBR approach seeks to focus on changing practice, rather than a ‘frozen’ input–output model of an intervention that can be seen in much experimental and educational research” (Baumgartner *et al*., 2003, p. 7).

Practitioner research often uses all of the data that is available from both qualitative and quantitative sources (Ivankova, 2015). Therefore, mixed methods designs may be considered extremely useful when conducting practitioner research. This study employed a mixed methods convergent design, whereby both quantitative and qualitative data are collected and analysed for the purpose of comparing or combining the results (Creswell & Plano Clark, 2017). This design was chosen in order to combine quantitative and qualitative findings for a more complete understanding of my practice-based problem.

# 4 Design of the intervention and data collection methods

## 4.1 Participants

The study was undertaken with a mixed ability, Year 12 International Baccalaureate (IB) chemistry class, comprised of eight students. Seven students studying the higher level (HL) course, and one student following the standard level (SL) course. This class were chosen for three main reasons:

1. Having previously been introduced to stoichiometry topic at the beginning of Year 12, and showing a good understanding in lessons, this class performed less well than expected in the end-of-topic assessment. The marks for this assessment ranged from 9-23 out of a maximum of 44, with a mean mark of 15.5 (SD=4.44).
2. There is no difference in the HL and SL content of the stoichiometry topic in the IB Chemistry Diploma Program, and so the same interventions were able to be used with all students in the class.
3. The findings of my Year 1 assignment that investigated students' perceptions of the mathematical elements of IGCSE and post-16 chemistry courses in the school, showed that Year 12 students found these areas of their chemistry course more challenging than any of the other year groups.

## 4.2 Intervention summary

Table 2: Summary of study phases

| Study Phase | Description | Purpose |
|---|---|---|
| 1 | Pre-intervention assessment and interviews. | – Identify WPs students find the most challenging and why.<br>– Understand how students approach these problems. |
| 2 | Analyse data from pre-intervention assessment | – Measure student performance in different types of WP as a baseline for comparison.<br>– Identify any current WP solving strategies.<br>– Understand students' perceptions and experiences of WPs in chemistry.<br>– Identify a focus for the study: type(s) of WP. |
| 3 | Investigate possible interventions | – Review of the literature to inform the intervention strategies based upon the pre-intervention findings.<br>– Design intervention(s). |
| 4 | Introduce the text editing intervention | – Introduce students to the intervention and model its use.<br>– Allow students to practise using the intervention strategy. |
| 5 | Post-text editing assessment | – Assess the impact of the intervention.<br>– Assess students' use of the intervention strategy.<br>– Assess students' perceptions of the intervention. |
| 6 | Post-text editing data analysis | – Identify trends in student performance related to the intervention.<br>– Analyse students' strategy usage.<br>– Understand students' perceptions of the intervention and identify themes.<br>– Identify possible next steps for the study. |
| 7 | Introduce visual representations intervention | – Introduce students to the intervention and model its use.<br>– Allow students to practise using the intervention strategy. |
| 8 | Post-visual representations assessment | – Assess the impact of the intervention.<br>– Assess students' use of the intervention strategy.<br>– Assess students' perceptions of the intervention. |
| 9 | Post-visual representations data analysis | – Identify trends in student performance related to the intervention.<br>– Analyse students' strategy usage.<br>– Understand students' perceptions of the intervention and identify themes. |
| 10 | Identify implications for practice | – How will the findings of this study impact my teaching practice and that of my colleagues? |

## 4.3 Description of the main study phases

The study took place over three main phases: pre-intervention; the text editing intervention; and the visual representations intervention.

It should be noted that for each of the interventions, students were introduced to the strategy during a lesson with the strategy modelled by myself.  Students were provided with acquisition problems with answers and encouraged to use them to practise and check their progress.  The answers provided took the form of problem-example pairs. This was decided as students face challenges, and make errors, they may be more inclined to study examples closely and pay close attention to the steps that couldn't be solved (van Gog *et al*., 2011).

To increase exposure to each strategy, six lessons over a three-week period were used. Time was given to practising and discussing the strategy, and for students to seek assistance where necessary.  This decision was based upon the meta-analysis of WP solving interventions by Myers *et al*. (2022), who found that interventions provided at least three times per week had a much greater effect size than those provided once or twice per week.  As the number of lessons per week could not be changed, the length of the intervention was increased.  Myers *et al*. (2022) also found that interventions that offered students additional opportunities to receive explicit modelling, develop their skills, and receive feedback on their performance were more effective.

Each intervention strategy was trialled with six Year 13 students and was amended based upon their feedback before use with the study participants.

### 4.3.1 The pre-intervention phase (Pre)

The purpose of this phase in the study was twofold. Firstly, to understand the types of WP that students found the most challenging and why. Secondly, to provide a baseline of student performance in different types of WP and to establish a focus for the interventions.

### 4.3.2 First intervention phase: text editing (TE)

Following analysis of data from the pre-intervention, it was evident that most students possessed the required chemistry-specific knowledge in relation to stoichiometry problems. The main challenges students faced were finding and organising the relevant information in complex titration problems. Due to this, the intention of this intervention phase was assisting students with finding information, helping to develop more a structured approach to solving chemistry WPs. Therefore, TE was chosen as the basis for the first intervention to assist with finding information as it required students to actively engage with the content of the question.

For the purposes of this intervention, the TE strategy outlined by Ngu *et al*. (2002) was adapted to provide students with a framework within which to work, enabling them to move beyond simply categorising information as sufficient, missing, or irrelevant. The final version of the strategy along with feedback from the trail can be found in Appendix 1.

### 4.3.3 Second intervention phase: visual representations (VR)

The main purpose of this intervention phase was to assist students with organising information so that it could be used to begin solving the problem. The intervention was based upon the "Draw-it" problem solving cycle developed by van Garderen and Scheuermann (2015). Originally, developed as a teaching tool to help students with learning difficulties to develop WP solving skills in mathematics, it was believed that it provided a sound structure within which students could work. The final version of the strategy developed for this intervention along with feedback from trials can be found in Appendix 2, along with a copy of the original "Draw-it" strategy.

In addition to being introduced to the VR strategy, students were also taught the role of VR in WP solving. For example, what constitutes a VR, and what different representations can be used for (diagrams, tables, graphs, etc), and given guidance on selecting appropriate VRs in relation to the problem, as per the findings of Ayabe *et al*. (2022).

## 4.4 Data collection methods

### 4.4.1 Pre-intervention and post-intervention assessments

The pre-intervention assessment contained four WPs that focussed on different types of stoichiometry problem in post-16 chemistry. Complex titration questions were identified as the focus for the study. Post-intervention assessments for the TE and VR phases contained three titration questions with an increasing word count and number of solution steps. Post-intervention assessments for each phase were closely matched in terms of the number of words and solutions steps, so that performance in each of the problems would be comparable.

To account for the individual needs and differences of each student, and reduce test anxiety, a single 55-minute lesson was allocated. The reason for the removal of a set time limit (other than that of the lesson itself) to complete the problems was based upon the suggestions of Gernsbacher *et al*. (2020). They argue that time-limited tests are less valid, less reliable, less inclusive, and less equitable than untimed tests

(Gernsbacher *et al.*, 2020). All students completed the problems within the allocated 55-minutes.

Problems for the assessments were either written by myself, in collaboration with the class co-teacher, adapted from Edexcel GCE A-level examination papers, or adapted from Clarke (2005) *Calculations in AS/A Level Chemistry* (7th Ed).

Appendices 13-15 shows the WPs used in each of the study phases.

### 4.4.2 Questionnaires

Whilst completing problems in each of the post-intervention assessments, students were asked to complete a short paper-based questionnaire following each problem. The questionnaire contained of a mixture of two Likert response format questions, and three open-ended questions. The questions sought students' perceptions of their confidence and difficulty of the problem, what they felt was most and least challenging about the problem, and what approach or strategy they used whilst completing the problem.

Likert format questions had a 6-point scale with labelled scale points. The lowest numerical value was assigned to the most negative option (extremely difficult, and extremely unconfident) following the suggestions of Fink (2002). An even numbered scale was chosen to reduce the potential problems of survey satisficing, and self-protection (Champagne, 2014). Additionally, when the questionnaire was trialled with Year 13 students, they thought that a ‘neutral’ option should not be included, with one student stating, “you’re either confident, or you’re not” and many agreed. This sentiment was also shared by departmental colleagues, reinforcing the decision to use an even-numbered scale and the removal of a mid-point. Although it is argued that a 7-to-11-point Likert scale may generate more reliable data (Krosnick & Presser, 2010), a 6-point scale is within the acceptable 5-to-11-point range (Friedman & Amoo, 1999). Moreover, it enables effective differentiation between response categories, without providing too many response options, reducing the task difficulty of providing a response (Dillman *et al*., 2014; Krosnick & Presser, 2010).

Open-ended questions were included as they are able to provide authenticity, richness, and depth of response (Cohen et al., 2018). These benefits far outweigh the associated problems with additional cognitive and time demands placed on respondents, data handling, data volume and possible difficulties in making comparisons between respondents (Cohen *et al*., 2018).

Although think-aloud data collection methods have been successfully used in problem solving research in chemistry and mathematics education (e.g., Csíkos *et al*., 2011; Overton *et* al., 2013), a questionnaire-based approach was used for this study. This was due to the associated increase in participants mental workload during think-aloud methods, in comparison to completing tasks in silence (Hertzum *et al*., 2009). Additionally, participants have been found to take more time to complete tasks during think-aloud methods (Ericsson & Simon, 1993; Hertzum *et al*., 2009). A copy of the questionnaire can be found in Appendix 3.

### 4.4.3 Student interviews

Semi-structured interviews were used to gain a deeper insight into the practice-based problem, as they enable greater depth of data collection (Cohen et al., 2018). The semi-structured interview is flexible, yet allows the specific topics related to the study to be addressed whilst enabling participants to bring new meaning and insight to the study focus (Galletta, 2012).

Small group interviews ranging from 20-30 minutes in length took place in the school conference room during the school day. Appendix 4 shows examples of interview guides.

## 4.5 Data analysis

### 4.5.1 Quantitative data analysis

#### 4.5.1.1 Pre-and-post-intervention assessments

WPs were assessed to determine the word count and number of required solution steps, similar to the study conducted by Pongsakdi *et al*. (2020). The process of deciding each of the steps required to solve individual problems was done in collaboration with the class co-teacher until an agreement was reached.

Following a similar, method to that used by Gulacar *et al*. (2013, 2019) students' responses were analysed by the proportion of solution steps that were attempted (correctly or incorrectly) and completed correctly. A success ratio was calculated as a measurement of the proportion of how many of the attempted problem steps were completed correctly.

$$\text{success ratio} = \frac{\text{number of correct problem steps}}{\text{number of attempted problem steps}}$$

As all of the above measurements are reported as proportions, the minimum value is 0.00 and the maximum value is 1.00.

Based upon students' responses, the difficulty and discrimination of the problems was assessed. Difficulty rating ranges from 0.00-1.00. Higher difficulty values represent an easier question, whereby more students achieved the correct answer (Towns, 2014). Discrimination values also range from 0.00-1.00. Values closer to 1.00 indicate problems that discriminate well between higher and lower performing students and a value closer 0.00 does not discriminate well (Towns, 2014), indicating that both higher and lower performing students are able to solve the problem correctly. Appendix 5 gives a more detailed explanation of how difficulty and discrimination values were calculated.

#### 4.5.1.2 Questionnaire responses

Questionnaire items regarding students' perceived difficulty and confidence are reported as medians due to the ordinal nature of the data. It also should be noted here that although decimal values in relation to perceived difficulty and confidence do not represent options in the Likert format question, they are used as rounding to the next integer, would change the category of the rating. For example, rounding a confidence value of 3.5 to 4.0 would indicate a median rating of 'somewhat confident' and therefore may be misleading. This method will be used throughout this analysis.

#### 4.5.1.3 Statistical tests

Two-tailed Spearman's rank correlations were used to determine the strength and significance of any relationships present between the variables in the pre-intervention assessment (n=4). Spearman's rank correlations were used as the data was found to be non-parametric (Cohen *et al*., 2018). Because the data was a mixture of continuous and ordinal data, Spearman's rank correlations can be used to compare each these data types to one another, especially when the ordinal data contains five or more levels (Khamis, 2008). An alternative method for correlating ordinal-ordinal, and ordinal-continuous data is the Kendall rank correlation, but this would not have been suitable for continuous-continuous data sets (Khamis, 2008, p.160). Furthermore, the performance of a Spearman's rank correlation is comparable to that of a Kendall rank correlation (Khamis, 2008; Zar, 2010).

### 4.5.2 Qualitative data analysis

#### 4.5.2.1 Student interviews and open-ended questionnaire questions

Analysis of data from the pre-intervention phase was used to gain a more in-depth understanding of students' perceptions of WPs, what they found challenging about them and why. Analysis of the TE and VR phases focussed more on students' perceived impact of the interventions. Responses to the open-ended items of the questionnaire were collated and combined with interview data for analysis.

During student interviews, audio recordings of students' responses were taken and transcribed by myself. Following the advice that "transcriptions need to be constructed so that they are useful for your analysis, but not overwhelming" (Ingram and Elliott, 2019, p. 190), transcriptions were 'cleaned up' before analysis. This process was used to make the transcripts more easily readable whilst maintaining participants' actual words (Carlson, 2016; Kohler Riessman, 1993).

Member checking was carried out to ensure the accuracy of the interview data (Birt *et al*., 2016; Carlson, 2010). As transcripts are partial representations of the data collected (Ingram and Elliott, 2019), the transcribed interviews were returned to participants to allow them to check and confirm the content of the interview and to further reduce the potential for researcher bias (Birt *et al*., 2016). Additionally, as verbatim transcripts may contain grammatical errors and significant use of filler words, some participants may become embarrassed when reading their own words. Therefore, 'cleaned up' transcripts were used to preserve their "dignity and voice" (Carlson, 2016, p. 1112).

Interview transcripts were analysed using thematic analysis as outlined by Braun and Clarke (2006, 2021). The purpose of the coding process was to aid the generation of pertinent themes from the data. An inductive approach to identifying codes and themes was employed, with themes being developed at the semantic level enabling a rich description of the whole data set (Braun & Clarke, 2006, 2021). Therefore, all of the interview data was transcribed. This allowed me to become more familiar with the data, enhancing the reliability of coding and theme development.

Appendices 6 and 7 show examples of a coded transcript and code book.

#### 4.5.2.2 Word problem solving strategy usage

Assessment of strategy usage is difficult to specify comprehensively and score accurately by the use of a rubric (Bisson *et al*., 2016). Therefore, it was assessed using the principles of comparative judgement. Responses to WPs were compared to one another to provide judgements that are reliable despite being derived from subjective, holistic judgements (Jones & Davies, 2024). Due to the small sample size, responses were placed in rank order as suggested by Bramley (2007). Ranking was performed separately by myself and the co-teacher of the class then discussed until an agreement was reached. Although the ranking exercise is inherently subjective with each judge using their own interpretation of 'best', it is in line with similar studies into assessing problem solving in mathematics (e.g., Holmes *et al*., 2017).

Once responses had been ranked, they were grouped together based on their level of strategy usage, and descriptors for these categories were created. In some cases, student responses included elements from different categories. Therefore, a holistic judgement was made of the student's approach, and a 'best-fit' judgement was applied. Appendix 8 gives a summary of the strategy usage categories and descriptors.

## 4.6 Collaboration

In relation to this study myself and the co-teacher of this class collaborated closely in the design of the interventions. Examples of the collaboration points included choosing the questions to be used in each phase of the study and finalising the suggested approaches for each intervention strategy. Additionally, the co-teacher of the class would also use the intervention strategies in their lessons with the class.

Collaboration with students was a key element of this study. A small group of Year 13 students were used to trial aspects of the study, and their feedback was used to inform and refine materials used in the study phases. The participants of the study also played an important role in deciding the nature of the interventions based on the findings from the pre-intervention phase, making the subsequent interventions bespoke to them as a group.

## 4.7 Ethical considerations

This study was granted ethical approval by the Central University Research Ethics Committee (Ethics Approval Reference: EDUC_C1A_23_339) and carried out in accordance with The British Education Research Association (BERA) Ethical Guidelines for Educational Research, fourth edition (2018). The Headteacher was informed and gave permission to carry out the study within the school. Additionally, the safeguarding procedures of the school were followed at all times. As the study used some existing assessment data, a Data Sharing Agreement was drawn up between the university and the school to allow the use existing student work and assessment data.

The study was introduced to students in a chemistry lesson and the purpose of the interviews was clearly explained. Students were provided with information regarding the study, including an opt-out form and an explanation of how to opt-out. Although participants were all classed as competent youths, information regarding the study was sent to parents informing them of the study along with details of how to opt-out of the study. It was made clear that participation in the study optional, and that students may opt-out at any stage without needing to give a reason. It was also stressed to students that all data would be anonymised using a pseudonym, and they would not be identifiable from the information they provided. All students are given female pseudonyms as there was only one male student in the class. Therefore, using a male name may have made this student identifiable. Informed consent was gained from participants based upon mutual trust (BERA, 2018). All participants fully understood the nature of the research without any coercion or duress to participate. All participants gave verbal and written consent to their involvement in the study, including written consent to audio recordings of interviews.

An important issue that requires careful, and significant consideration is the unequal power relationship between researcher and participant. This asymmetric relationship (Farrimond, 2017) is reinforced by the unequal power between teacher and students (Ingram & Elliott, 2019). Moreover, the authority of the teacher persists in all of their interactions, no matter how a teacher may seek to share this authority with students (Buzzelli & Johnston, 2001). To mitigate this, the advice offered by Ingram and Elliott (2019, pp. 165-166) was followed, through:

- Gaining informed consent.
- Interviewing students in pairs or small groups-unless they wished to be interviewed individually.
- Interviews took place in a neutral space, rather than a classroom.
- Emphasising that I was interested in their opinions and views and would not take anything they said personally.
- Reassuring students that there are no ‘right’ or ‘wrong’ answers or behaviours.

## 4.8 Limitations of the study

Although serious thought and consideration was taken to reduce the number of limitations during the design of the study, there still some limitations inherent within the methods chosen.

Whilst the direct effect of the TE intervention may be able to be estimated in relation to the pre-intervention phase, it is most likely not possible to do this for the VR intervention. However, due to the DBR-inspired approach to practitioner research used in this study, each study phase informs the next. Therefore, some elements of the previous phase are taken through into the next phase, and so cannot be easily separated.

The effect of practise cannot be ignored in relation to students' performance in the post-intervention assessments. This may lead to higher scores during the post-intervention assessment through the effect of practise. However, this practise could be considered as a free-choice practice condition, which was found by Pachman *et al* (2013) to have less of an impact that deliberate practice. Therefore, the effects of this may not have been as influential as expected. However, there was no way of measuring this, so the possible effect of this should be noted.

Additionally, as students were actively encouraged to practise the intervention strategies outside of lessons and provided with the worked answers to aid independent study, this could have introduced the worked example effect (Sweller *et al*., 2019). Furthermore, if present the worked example effect may have been advantageous to lower prior knowledge learners (Leppink *et al*., 2014), but disadvantaged higher prior knowledge learners through the ‘expertise reversal effect’ (Kalyuga *et al*., 2003). Again, this was not measured and so the extent to the effect of worked examples is unknown but important to note.

# 5 Findings and Discussion

## 5.1 Which types of chemistry word problem do students find the most challenging?

### 5.1.1 Findings

The findings from the pre-intervention show that problems containing a higher word count, and more solution steps were answered less successfully (see Table 3). The mean proportion of attempted solution steps also shows a similar trend in relation to word count and the number of solution steps. Calculated difficulty values also broadly agree, with the exception of Pre-2 where there were no correct answers due to a misconception regarding ratios that was addressed in the following lesson. Students' perceived difficulty and confidence ratings also decrease as the word count and number of solution steps increase. This indicates that students find these problems more difficult and feel less confident answering them. Problem Pre-3 would appear to be the only question that discriminated well between higher and lower achieving students. This is due to, Pre-1 being completed correctly by all students, whereas Pre-2 and Pre-3 were not completed correctly by any students.

Table 3: Summary information for the Pre-intervention questions

| Question | Solution Steps | Word Count | Mean Attempt | Total correct answers | Correct answer-higher | Correct answer-lower | Difficulty | Discrimination | Median difficulty rating[1] | Median confidence rating[2] |
|---|---|---|---|---|---|---|---|---|---|---|
| Pre-1 | 2 | 11 | 1.00 | 8 | 2 | 2 | 1.00 | 0.00 | 4.00 | 5.00 |
| Pre-2 | 5 | 21 | 0.48 | 0 | 0 | 0 | 0.00 | 0.00 | 4.00 | 3.50 |
| Pre-3 | 9 | 97 | 0.78 | 2 | 2 | 0 | 0.25 | 1.00 | 2.00 | 2.50 |
| Pre-4 | 13 | 191 | 0.40 | 0 | 0 | 0 | 0.00 | 0.00 | 1.00 | 1.50 |

Relationships between problem variables were investigated using two-tailed Spearman's rank correlations (n=4). Figure 4 shows a summary of these correlation values. Due to the small sample size, the statistical significance of most of these correlations was found to be greater than the alpha value of *p=0.05* (unless otherwise stated) and therefore not statistically significant. However, as Cohen *et* al. (2018) state, "statistical significance is not the same as educational significance" (p. 743). Therefore, the correlations found between variables within this sample, could be viewed as educationally significant within the context of this study.

---

[1] On a 1-6 scale, lower numbers represent higher difficulty, and higher numbers represent lower difficulty.

[2] On a 1-6 scale, lower numbers represent lower confidence, and higher numbers represent higher confidence.

Figure 4: Colour coded correlation matrix for the pre-intervention assessment

| | Solution Steps | Median Difficulty Rating | Word Count | Median Confidence Rating | Mean Attempted Steps | Mean Correct Steps | Mean Success Ratio |
|---|---|---|---|---|---|---|---|
| Solution Steps | 1.00 | | | | | | |
| Median Difficulty Rating | -1.00** | 1.00 | | | | | |
| Word Count | 1.00** | -1.00** | 1.00 | | | | |
| Median Confidence Rating | -0.95 | 0.95 | -0.95 | 1.00 | | | |
| Mean Attempted Steps | -0.80 | 0.80 | -0.80 | 0.95 | 1.00 | | |
| Mean Correct Steps | -0.80 | 0.80 | -0.80 | 0.95 | 1.00** | 1.00 | |
| Mean Success Ratio | -0.80 | 0.80 | -0.80 | 0.95 | 1.00** | 1.00** | 1.00 |

**Correlation is significant at the 0.01 level (2-tailed).

Correlation coefficients vary from ±0.80 to ±1.00, indicating large to very large correlations (Cohen *et al*., 2018, p. 772). In summary, the following relationships were found:

- Problems with more solution steps and a higher word count were perceived as more difficult, and perceived confidence was lower.
- Students felt less confident when the perceived difficulty of the problem was higher.
- Students attempted more steps of a problem that contained fewer solution steps, and a lower word count, and when the perceived difficulty of the problem was lower.

- More solution steps were correct in problems with fewer solution steps, less words, and lower perceived difficulty. Additionally, students felt more confident answering the problem and more problem steps were attempted.
- Success ratio increased when problems contained fewer solution steps and fewer words, when perceived difficulty was lower, and when perceived confidence was higher.
- When more steps of a problems were attempted, more of them were completed correctly, and success ratio increased.

Analysis of qualitative data identified five key themes further supporting the findings discussed above. *Finding and organising information* was prevalent among all students and mainly related to a higher word count. "The phrasing and then the different values, but then you don't necessarily need some of them, so if you don't get the question in the first place you won't figure out which one to use" – Helena, and "It's easy to make a mistake by using too much information or using information you don't need" – Julia.

*Managing stress and anxiety*. Louisa detailed her frustrations with the amount of irrelevant or superfluous information in a question, “throwing things at us that are completely useless, you're going to over-think how much it's applicable and it's just gonna stress you out even more”. Furthermore, Sofia outlines "I get confused and if I get confused, I just panic and the whole question was wrong just because I've panicked”.

*Simple vs. complex problems*. Students found questions containing more solution steps, and a higher word count the most challenging. When asked which types of problems she found the least challenging, Helena stated, “ones that explicitly say what you have to find out-it just gives you the things that you need [values and information]”. Conversely, complex problems were viewed as those that contained more words and solution steps, which students found to be more problematic. Margot explained what she felt constituted a more complex problem:

> *“Titration questions, there are quite a lot of steps to think about, things like conversions or if they use different types of formula for the same thing [substance] you’ve got to remember to divide or times things-just quite a lot going on so they’re the trickiest”*

*Structured vs. unstructured problems*. Imogen gave a useful insight into her experience of unstructured problems, “it's confusing to figure out how to get from the information to the answer that you need, there's lots of steps in between that you're not told specifically that you have to do”.

*Confidence*. Feelings of confidence were very closely tied to the type of problem being answered, “for relatively simple calculations I’m confident, if it gets more complicated, I get less confident” – Irini. Zara indicated that her confidence was lower when the number of words and solution steps increased, “you've got to do different steps, the unit conversions and how many words there are, it’s confusing how many words there are”.

### 5.1.2 Discussion

Students' performance and confidence were higher in relation to more simple problems containing fewer words and fewer solution steps. This may be explained by the fact that in these cases the problem syntax generally follows the solution pathway (Strohmaier *et al*., 2022). In this way, the ICL of the problem is likely to be lower as there is less element interactivity (Chandler & Sweller, 1991; van Gog *et al*., 2011). Moreover, the ECL experienced by students is also likely to be lower, as there is very little or no irrelevant information (Kalyuga, 2011). For the simplest problem (Pre-1), all of the above would seem reasonable as all students were successful. Furthermore, the existence of students' pre-existing schema in relation to the problem may have meant that the process of solving this problem was mostly automated. Therefore, optimising GCL and requiring fewer WM resources (Sweller *et al*., 1998, 2019, van Merriënboer & Sweller, 2005).

For more complex problems the opposite is likely true. Fuchs *et al*. (2020), argue that that WM resources involved in solving WPs differ from and are more numerous than those associated with calculations. It could be viewed that in more complex WPs, ICL will be higher due to increased element interactivity (more information and solution steps), as may ECL due to the presence of more irrelevant or superfluous information (Kalyuga, 2011). Additionally, this type of problem is more likely to be less similar to those that students encounter in their daily lessons (Verschaffel *et al*., 2020).

Therefore, existing schemas in the LTM may not match the problem and new solving solutions are sought through the Randomness as Genesis Principle (Sweller *et al.*, 2019). This may in turn lead to the use of less effective problem-solving methods and increase ECL (Sweller *et al.*, 2019, van Merriënboer & Sweller, 2005). In this case, it may be more likely that the limited capacity of the WM is exceeded, and cognitive overload occurs (Overton & Potter, 2011) leading to decreased performance. This may also explain why students perceived these problems as less difficult and account for the strong negative correlations between word count and perceived difficulty, and solution steps and perceived difficulty.

Findings from this study would seem to support the above, particularly in relation to linguistic factors. Students identified the number of words in the problem and the presence of irrelevant information, as specific sources of difficulty, being strongly linked to challenges with identifying the information needed to solve the problem. More linguistically complex problems may require more WM resources in relation to comprehension. This is supported by Jaffe and Bolger (2023) who state, "from a cognitive load perspective, all facets of the linguistic component can influence problem solving" (p. 105). Therefore, the development of a situational model for the problem may take a significant amount of time, possibly exceeding the 41% of total solution time suggested by Leiss *et* al. (2019). The findings of this study support those of Walkington *et al*. (2018), that increased word count is negatively correlated with WP solving accuracy. Additionally, word count was very strongly correlated with the number of solution steps ($r_s$= 1.00, p=<0.01). However, this is likely due to the argument of Daroczy *et al*. (2015) that multiple-step problems also become more difficult linguistically as they “usually contain more phrases and semantic distractors” (p. 348).

Further contributing to the complexity of problem comprehension, is the presence of irrelevant or superfluous information, with the qualitative analysis agreeing with the findings of Muth (1992). Students often cited this in relation to feelings of confusion, anxiety and panic, further reinforcing students’ perceptions of complex chemistry WPs as difficult and feelings of lower confidence. This could be especially important from a self-efficacy perspective, particularly when considering self-efficacy across the dimensions of facet-specificity, level of difficulty, and strength (Bandura, 1997; Street *et al*., 2017; Street *et al*., 2022). Lower self-efficacy in relation to complex problems may be a function students’ perceived lower confidence, and high difficulty. This decreased self-efficacy may then be linked to particular aspects of chemistry WPs such as complex titration problems, evidenced by Margot in the previous section. However, self-efficacy may be preserved or even enhanced in relation to more simple problems where students enjoy success, substantiating self-efficacy and enhancing motivational outcomes (Schunk & DiBenedetto, 2020). This argument may be especially important given that self-efficacy can vary across and within domains (Schunk, 1991), and domain-linked self-efficacy can be debilitating (Yeo & Neal, 2006).

Students often mentioned the increased number of steps involved in more complex problems as a source of difficulty, supporting the findings of Peng *et al*. (2016) that the multiple-step nature of WPs draw upon significant WM resources. The findings are also in line with elements of the Johnstone-El-Bana model, which when considered in isolation, an increasing number of problem steps does relate to lower performance. However, it is much more likely that there are numerous other factors that contribute to this such as students' understanding of chemistry concepts and comprehension (Tsaparlis, 2021b; Daroczy *et al*., 2015).

A particular challenge identified by students that is related to, but not exclusive to the multiple-step nature of the problems, is unit conversions. In complex stoichiometry problems such as Pre-3, unit conversions are required in order to correctly solve the problem. In fact, only one student correctly converted the units for temperature and pressure for Pre-3. Although a necessary step within the problem solution, they may be best considered as examples of implicit information. The presence of implicit information has been shown to negatively affect WP performance, especially when less effective problem-solving strategies such as direct translation are used (Heggarty *et* al., 1995). In relation to chemistry WPs, unit conversions could be argued to require domain-specific comprehension, technical reading ability and domain-specific knowledge as per the assertions of Leiss *et al*. (, 2019) and Vilenius-Tuohimaa *et al*. (2008). Furthermore, unit conversions are likely to increase ICL and require significant WM resources to process.

Based on a combination the analysis of the pre-intervention phase of the study, the reasons outlined above, and consultation with students. The focus of the subsequent intervention phases was assisting students with identifying and organising information in complex titration problems such as Pre-4. These specific areas of difficulty seemed to stem from managing the WM demands of the complex linguistic factors and multiple-step nature of these problems. After careful consideration of the pre-intervention findings the strategies of TE and the use of VRs were chosen. These strategies were selected as it was believed that they would best meet the needs of the students in this study.

## 5.2 How do taught interventions impact students' chemistry word problem performance, perceived difficulty and confidence?

### 5.2.1 Findings

Performance in titration problems between the pre-and-post intervention phases of the study, shows improvement in all of the measures outlined in Table 4. This improvement is larger in problems with fewer words and solution steps, agreeing with the findings from the pre-intervention phase. A similar trend is seen with respect to students' difficulty and confidence ratings, and calculated difficulty and discrimination values.

Table 4: Summary of mean performance in the two intervention phases[3]

| Question | Solution Steps | Word Count | Mean Steps Attempted | Mean Steps Completed Correctly | Mean Success Ratio | Median Difficulty Rating | Median Confidence Rating | Calculated Difficulty | Calculated Discrimination |
|---|---|---|---|---|---|---|---|---|---|
| **Pre-4** | **13** | **191** | **0.40** | **0.13** | **0.20** | **1.00** | **1.50** | **0.00** | **0.00** |
| TE-1 | 7 | 74 | 1.00 | 0.86 | 0.86 | 4.00 | 4.00 | 0.75 | 0.50 |
| TE-2 | 11 | 109 | 0.91 | 0.78 | 0.84 | 3.00 | 3.00 | 0.63 | 1.00 |
| **TE-3** | **12** | **196** | **0.75** | **0.51** | **0.61** | **2.00** | **2.00** | **0.25** | **1.00** |
| VR-1 | 9 | 81 | 0.91 | 0.85 | 0.90 | 4.00 | 4.00 | 0.75 | 1.00 |
| **VR-2** | **10** | **187** | **0.90** | **0.88** | **0.97** | **3.50** | **3.00** | **0.88** | **0.50** |
| *VR-3* | *15* | *273* | *0.92* | *0.62* | *0.65* | *2.00* | *2.50* | *0.38* | *0.50* |

For this analysis it is most useful to compare questions Pre-4, TE-3 and VR-2 as these questions are classed as complex titration questions. These questions are also comparable in relation to their number of words and solution steps. However, problem VR-3 will be discussed separately due to its significantly increased number of words and solution steps. As can be seen in Table 4, performance in complex titration problems has shown a significant increase in each of the three measures of performance: mean proportion of steps attempted, mean number of steps completed correctly, and mean success ratio. This improvement in all areas of student performance indicates that the TE and VR interventions had positive impacts, enabling students to complete these problems more successfully.

---

[3] Question codes beginning with TE represent the text editing intervention phase, and VR represents the visual representations intervention phase.

In relation to perceived difficulty and confidence, a small increase in both was evident. The median difficulty rating for Pre-4 was 1.50 (extremely unconfident), increasing slightly to 2.00 (unconfident) for TE-3 and again to 3.00 for VR-2 (somewhat unconfident). A similar trend can be seen in relation to students' perceived difficulty of the problems, median difficulty was rated as 1.00 (extremely difficult) for Pre-4, 2.00 (difficult) for TE-3, and 3.50 (somewhat difficult) for VR-2.

Calculated difficulty values show a similar trend to students' perceived difficulty, showing that more students achieved correct answers, indicating that the same type of problem became less difficult. Problem TE-3 discriminated well between higher and lower achieving students (1.00), but for VR-2 and VR-3 the value decreased to 0.50 for both problems. This is due to more students from the lower and medium achievement groups achieving a correct answer, in relation to students from the higher achievement group.

Problem VR-3 showed an increase in the mean number of attempted steps (0.92) in comparison to the other complex titration problems in the study. However, the mean proportion of correctly completed steps (0.62), and the success ratio (0.65) was lower than for VR-2, but higher in comparison to both Pre-4 and TE-3. Perceived difficulty and confidence medians (2.00 and 2.50 respectively) were also lower than those for VR-2, and similar to those for TE-3, but higher than Pre-4. Again, the calculated difficulty

(0.38) indicates that the problem may not have actually been as difficult as students perceived it to be as three students achieved a correct answer.

Thematic analysis after each of the intervention phases, with students indicating that each strategy was able to serve a different purpose, but neither was thought to be sufficient by itself. The theme of *Strategy utility* reflected this. In relation to TE students felt that it focussed their attention on the problem through the 'Target' step, with Julia commenting "it helped me to identify what the target was and how to find it through working out the stuff I didn't know". Students also found the 'identify' step of the strategy useful in helping them to find the relevant information from the question, "it helps you organise the information into what you are looking for, and what information you already know, and allows me to make sense of the excessive information that I am given" - Imogen. Many students felt that these first two steps of the strategy were the most useful, with the following steps of list, pair, categorise and calculate being less so, "the pairing step was a bit repetitive and wasn't as useful as the other steps, the listing step can sometimes be hard to make clear if the question contains two volumes for a substance [dilutions]" – Margot. When asked what she found least useful about this strategy, Irini stated "the pairing and categorising steps". A further, problem with the utility of TE was that for some students it created more confusion, "it gave me more stuff to remember which made the question more confusing" – Zara.

Within the same theme for VR, most students perceived the greatest benefit to be organising information clearly, “I was able to see the information more clearly” – Louisa. Other students found this particularly useful in relation to specific aspects of complex titration problems “it helps me organise my thoughts and the work, especially when the question involves a dilution” – Imogen. However, VRs were not without their disadvantages. Some students found it to be time consuming, “it takes a lot of time, and it is not always practical” – Julia. Margot specifically referenced this concern in relation to a test situation when talking about drawing process-related diagrams “it can take a long time, and in an exam, it might be a waste of time”. Some of the same concerns raised about TE were also present within VRs, especially in relation to creating further confusion and difficulties, “there were too many steps, so I ended up forgetting and mixing up a lot of them” – Zara.

*Strategies impact confidence* was mostly found to be in relation to a perception of increased confidence. Findings from the quantitative analysis found that there was a modest increase in students’ confidence between the pre-intervention and post-intervention phases of the study. Many students stated that the interventions increased their confidence buy a small amount. In relation to TE, Helena stated, “I think it made me feel a little bit more confident, but not much to be honest”. With respect to TE, increased feelings of confidence were linked to feeling more able to begin solving the problem, “I think in a way, it helped me gain footing at the beginning of the question to get me started” – Imogen. This feeling was also shared by Margot, “it helped me be more confident because it helps identify what was needed in the question”. Students that continued to find complex titration problems challenging also shared a similar sentiment, “you have a strategy and then you're a bit more confident, but they’re still hard “ – Helena.

In contrast to increased confidence, some students felt that using a specific strategy led to feelings of being overwhelmed, “it’s way too overwhelming, thinking about the whole strategy throughout the question” – Sofia. In this way, this theme of overlaps significantly with the theme of strategy utility.

### 5.2.2 Discussion

As the findings above show, there were clear improvements in performance for to students’ solving of chemistry WPs such as complex titration problems. These improvements are thought to be most likely to the TE and VR interventions reducing the cognitive load associated with completing complex chemistry WPs.

In relation to TE, students felt that it enabled them to make sense of the problem statement more easily, feeling more able to begin solving the problem. Students specifically referenced that the Target and Identify steps were particularly useful in differentiating between the useful and irrelevant information. In this way, this strategy may assist with reducing the ECL associated with irrelevant information (Kayluga, 2011), freeing up more WM to deal with other aspects of the problem (Sweller *et al.*, 2019). In turn, the WM demands associated with the development of a situational model may have been reduced, which is required for problem comprehension (Leiss *et al.*, 2019). This explanation would be in line with other studies into the effect of interventions to assist students in coping with irrelevant information in relation to WM demands (e.g., Fuchs *et al.*, 2020).

In relation to the VR intervention, most students agreed that it enabled them to organise the information extracted from the problem more effectively. The most prominent VR used was a table, with students citing that this allowed them to see clearly what information was missing, informing the next steps in their problem-solving process. This would agree with Ayabe *et al*. (2022) and Zahner and Corter (2010) that empty cells in a table allow solvers to check for missing information. In this way, the use of tables in relation to chemistry WP solving may allow for the activation of existing schema from the LTM or facilitate generation of new schema (Novick & Hurley, 2001; Zahner & Corter, 2010). Therefore, the deployment of relevant schema, is able to effectively bypass the WM (Sweller *et al.*, 2019), and optimise GCL enabling more efficient and effective problem solving (van Merriënboer & Sweller, 2005). Furthermore, students' improved performance when using VRs agrees with the findings of previous studies such as Beitzel and Staley (2015).

The use of strategies such as TE and an effective VR such as a table, seem to reduce the cognitive load related to problem comprehension and the processing of numerical information of chemistry WPs. This may then allow the allocation of more WM resources to tackle the other steps involved in these problems (Sweller *et al*, 1998, 2019) including implicit information such as unit conversions. However, it is important to address students concerns such as less useful elements of the strategies. These elements may have increased ECL as they could be viewed as a less effective problem-solving strategy (Sweller *et al*., 2019; van Merriënboer & Sweller, 2005). Although it would seem that most students decided that they would not use these elements of the strategy, therefore it may not have affected cognitive load. However, some students such as Zara and Sofia indicated trying to remember the steps of the strategies sometimes left them feeling more confused. Therefore, trying to remember all of the steps of the strategy in addition to attempting to solving solve a complex problem, may have resulted in cognitive overload. This may go some way to explaining lower and inconsistent performance during the study.

Although students perceived confidence did improve modestly during the study, it indicated that students on the whole did not feel particularly confident by the end of the VR intervention, despite their increasing performance. This was also evident in interview data, with students stating that they felt slightly more confident “but they’re still hard” - Helena. Considering self-efficacy again (see Section 5.1.2) across the dimensions of facet-specificity, level of difficulty, and strength (Bandura, 1997; Street *et al*., 2017; Street *et al*., 2022), students may still associate this type of problem (facet) with being particularly difficult and therefore their confidence remains low. Furthermore, students have been observed to underestimate their self-efficacy despite improved performance and is observed more in examinations (Talsma *et al*., 2019). Although the post-intervention assessments were not tests, it is likely that students viewed them as such and may underestimate their self-efficacy, reporting this as lower confidence in the questionnaire.

## 5.3 How do students use taught interventions to solve word problems?

### 5.3.1 Findings and discussion

Student responses to the complex titration problems were analysed and categorised in accordance with the descriptors outlined in Appendix 8. Table 5 shows each student's strategy use for the problems, along with the proportion of problem steps attempted and completed correctly. Whilst, completing the steps of the problem is not a pre-requisite for assessing higher strategy usage, more developed and coherent strategies often resulted in a higher proportion of attempted and correctly completed steps.

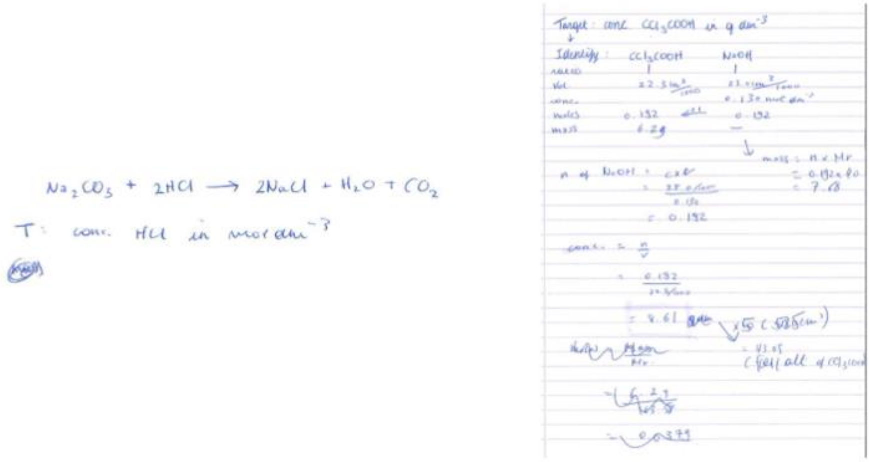
Table 5: Summary of strategy use per student for complex titration problems

| Student | Pre-4 | | | TE-3 | | | VR-2 | | | VR-3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Attempted Steps | Correct Steps | Strategy Use | Attempted Steps | Correct Steps | Strategy Use | Attempted Steps | Correct Steps | Strategy Use | Attempted Steps | Correct Steps | Strategy Use |
| Margot | 0.69 | 0.46 | High | 1.00 | 1.00 | Very High | 1.00 | 1.00 | Very High | 1.00 | 0.53 | Very High |
| Julia | 0.62 | 0.15 | Low | 0.75 | 0.17 | Medium | 1.00 | 1.00 | Very High | 1.00 | 1.00 | Very High |
| Sofia | 0.15 | 0.00 | Low | 0.83 | 0.42 | Medium | 0.80 | 0.70 | Medium | 1.00 | 0.40 | High |
| Louisa | 0.46 | 0.08 | Low | 0.92 | 0.58 | High | 1.00 | 1.00 | Very High | 1.00 | 1.00 | Very High |
| Imogen | 0.54 | 0.15 | Medium | 0.92 | 0.50 | Medium | 1.00 | 1.00 | High | 0.93 | 0.53 | High |
| Irini | 0.62 | 0.15 | High | 1.00 | 1.00 | Very High | 1.00 | 1.00 | Very High | 1.00 | 0.87 | Very High |
| Zara | 0.00 | 0.00 | Very Low | 0.42 | 0.42 | Low | 0.70 | 0.70 | Medium | 0.73 | 0.40 | Low |
| Helena | 0.15 | 0.00 | Very Low | 0.17 | 0.00 | Low | 0.70 | 0.60 | Medium | 0.67 | 0.20 | Medium |

Overall, an improvement in all students' strategy usage can be seen in addition to improvements in the proportion of problem steps attempted and completed correctly. In order to understand how students use taught interventions to develop their word problem solving strategies, it is necessary to examine student responses in more detail. Only the written portions of the students' problem responses are shown in the examples. However, Appendices 9-12 show the full problems with responses, including any annotations that students made to the problem text whilst completing the problems.

Helena's problem-solving strategy developed considerably throughout the study. In relation to strategy usage, her responses began as very low (Pre-4) and low (TE-3 see Figure 5). The response to TE-3 shows a small amount of evidence of TE use in identifying the target of the question. When asked to talk through her approaches to Pre-4 and TE-3, she stated that, "I didn't know how to start and didn't understand the question, I just randomly filled in some of the question" (Pre-4), and "I didn't understand the question or know where to start" (TE-3).

Figure 5: Helena’s responses to the problems TE-3 (left) and VR-3 (right)



Helena’s response to VR-3 (Figure 5), shows a much more developed strategy. There is usage of terminology from TE, limited to the ‘Target’ and ‘Identify’ steps to “help find the important information” – Helena. Additionally, an appropriate VR in the form of a table is used to organise information about relevant substances, “it [table] made the known information clear, helping me to understand what to work out next” – Helena. Furthermore, the VR-3 response also uses illustrations to help her understand the procedural aspects of the question. When asked to outline how she approached the questions, she replied “a combination of both [TE and VR], because to me, that is the best way to understand the question, but the tables helped me to work out what I needed to work out next”. For each of these problems, the proportion of problem steps attempted and completed correctly increased, especially in relation to attempt.

The response to VR-3 was categorised as having a medium strategy usage due to the fact there was still some evidence of trial-and-error in the later stages of the response. When Helena reflected on the question, she identified the cause as “not being able to work out the total volume”.

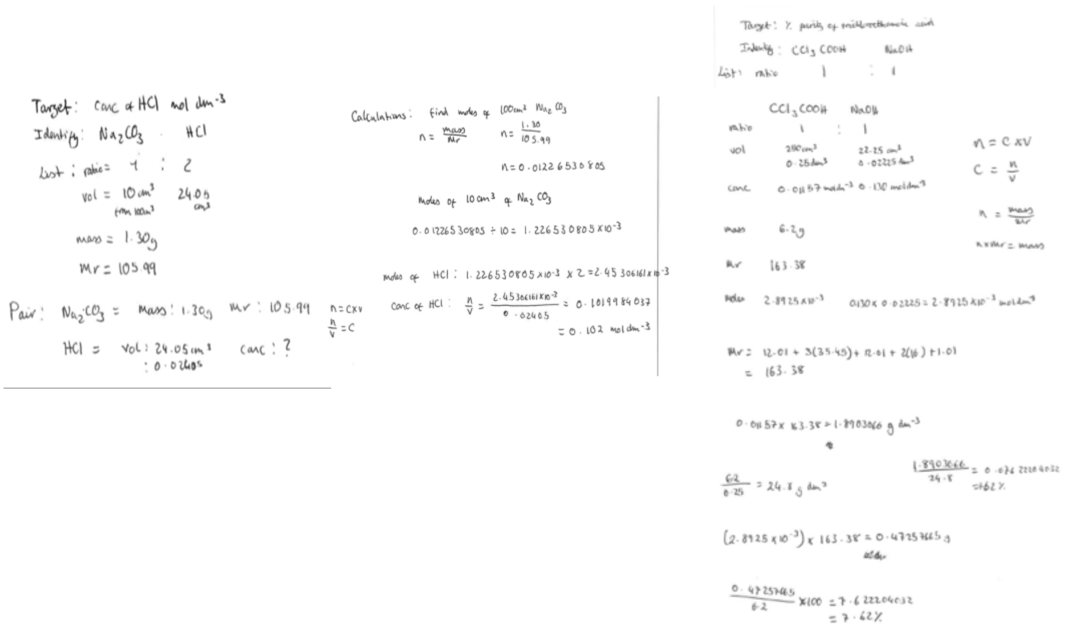
The evolution of Helena's chemistry WP solving strategy is a clear example of how students' strategies can improve due to taught interventions. From initially having little to no strategy, to using a more coherent strategy that enables them to become more successful in solving complex titration problems. The improvements in Helena's strategy use and performance indicate that she is beginning to gain expertise through familiarity with this type of complex problem (Gulacar *et al*., 2013; Tsaparlis, 2021b). Her later responses such as VR-3 would indicate that she is able to use her increased familiarity with the problem type and apply learned schema in the form of elements from TE to identify relevant information, combining this with another schema in relation to a table to organise information. Although this was not assessed, it may also be possible that the two schemas are combined into a single complex schema, that can be handled as a single element in the WM (van Merriënboer & Sweller, 2005). In either case, Helena's improved performance is likely due to her use of schema transferred from LTM, freeing up capacity in the WM (Sweller *et al*., 2019) to deal with the other demands of the problem.

Furthermore, Helena's transition from novice to developing expertise can be further evidenced using categories proposed by Overton *et al*. (2013). The response to TE-3 would sit firmly with the novice category as there is very little evidence of her ability to define the problem or of a structured approach. It is likely that based on the description of her approach, there was a lack of related schemas to draw on, and her

WM was overloaded resulting in cognitive overload (Overton & Potter, 2011). However, her response to VR-3, shows more aspects of expertise, being far more structured with a more evident understanding of the problem. Having a clear conceptual understanding of the relevant chemistry seemed to be the source of Helena's eventual error (Gulacar *et al.*, 2013), being unable to link the values that she had calculated to chemistry-specific knowledge. In this way, the argument of Tsaparlis (2021b), that not all steps of multiple-step problems are equally difficult is also supported. Helena was also able to locate the step of the problem that led to her error in relation to the final volume. This ability to evaluate the problem would indicate that in the process of gaining expertise, there may have also been development in metacognitive skills.

Margot's responses for the same problems (Figure 6) provide insights into how students with existing strategies are able to further develop their chemistry WP solving strategies.

Figure 6: Margot’s responses to the problems TE-3 (left) and VR-3 (right)



Target: conc of HCl mol $dm^{-3}$

Identify: $Na_2CO_3$ HCl

List: ratio = 1 : 2

vol = 10 $cm^3$ (from 100$cm^3$) 24.05 $cm^3$

mass = 1.30g

Mr = 105.99

Pair: $Na_2CO_3$ = mass: 1.30g Mr: 105.99

HCl = vol: 24.05$cm^3$ : 0.02405 conc: ?

$n = c \times v$

$\frac{n}{v} = c$

Calculations: find moles of 100$cm^3$ $Na_2CO_3$

$n = \frac{mass}{Mr}$ $n = \frac{1.30}{105.99}$

$n = 0.0122 6530 805$

moles of 10$cm^3$ of $Na_2CO_3$

$0.01226530805 \div 10 = 1.226530805 \times 10^{-3}$

moles of HCl: $1.226530805 \times 10^{-3} \times 2 = 2.45306161 \times 10^{-3}$

conc of HCl: $\frac{n}{v} = \frac{2.45306161 \times 10^{-3}}{0.02405} = 0.1019984037$

$= 0.102$ mol $dm^{-3}$

Target: % purity of trichloroethanoic acid

Identify: $CCl_3COOH$ NaOH

List: ratio 1 : 1

| | $CCl_3COOH$ | NaOH |
|---|---|---|
| ratio | 1 : | 1 |
| vol | 250$cm^3$<br>0.25$dm^3$ | 22.25 $cm^3$<br>0.02225 $dm^3$ |
| conc | 0.01157 mol$dm^{-3}$ | 0.130 mol$dm^{-3}$ |
| mass | 6.2g | |
| Mr | 163.38 | |
| moles | $2.8925 \times 10^{-3}$ | $0.130 \times 0.02225 = 2.8925 \times 10^{-3}$ mol$dm^{-3}$ |

$n = c \times V$

$c = \frac{n}{V}$

$n = \frac{mass}{Mr}$

$n \times Mr = mass$

$Mr = 12.01 + 3(35.45) + 12.01 + 2(16) + 1.01$

$= 163.38$

$0.01157 \times 163.38 = 1.8903066$ g $dm^{-3}$

$\frac{6.2}{0.25} = 24.8$ g $dm^{-3}$

$\frac{1.8903066}{24.8} = 0.07622204032$

$= 7.62\%$

$(2.8925 \times 10^{-3}) \times 163.38 = 0.47257665$ g

$\frac{0.47257665}{6.2} \times 100 = 7.622204032$

$= 7.62\%$

Margot’s response to Pre-4 was categorised as having high strategy usage, making effective use of a table to organise information. Whilst the proportion of steps attempted in this instance was high (0.69), the proportion of steps completed correctly was lower (0.46). When asked to describe her approach to the problem, she stated, “I underlined the information and used the formulas that I’ve learned”. This combination of high strategy use and low performance may indicate the use of previously learned schema that are not successfully transferred to a similar yet more complex problem. Unsuccessfully applying this existing schema may have led to an increase in ECL as she attempted to generate a solution strategy (Sweller *et al.*, 2019; van Merriënboer & Sweller, 2005). Additionally, this may have possibly caused her to forget important chemistry-specific knowledge (Tsaparlis, 2021b). Margot’s response here shows some elements of expert behaviour in the form of a structured and logical approach (Overton *et al.*, 2013).

However, following the TE intervention, there is clear evidence of strictly adhering to the strategy, completely abandoning her previous use of a table to organise information.  Margot’s approach appears to use TE as a schema as this approach was used in all of the problems in the TE post-intervention assessment.  When describing her approach to this question, she simply stated “text editing”.  Although the TE format was strictly adhered to, she indicated that she found some of the steps “repetitive” and “probably not necessary”.  This may indicate that some students will use taught strategies exactly as they are presented, simply because it has been taught.  This last point would be in line with the findings of Goulet-Lyle *et al*. (2020) who observed that many students will follow a taught strategy closely, especially when it is supported by the socio-mathematical culture of the classroom (such as teacher endorsement of a solution strategy).

Following the VR intervention, it is interesting that Margot reverts to the use of a table to organise information (see Figure 6).  Additionally, the rigid adherence to the TE steps is no longer seen.  Instead, the first three steps of TE are used (‘Target’, ‘identify’ and ‘List’), and then combined with a table to organise the information before completing subsequent calculations.  The overall approach is organised and methodical, with relevant information identified and used to assist in problem solving.  Margot described her approach as in these problems as, “a mixture of text editing and visual representations, because it helps me to understand the questions and organise the information clearly”. Furthermore, links are made between parts of the question to aid the problem-solving process. For VR-2 and VR-3, strategy usage was classed as very high. All problem steps were attempted, but the proportion of steps completed correctly differed between VR-2 (1.00) and VR-3 (0.53). The same values were found in relation to the success ratio for each problem. Upon closer analysis, the reason for the lower performance in VR-3 was due to an intermediate value being calculated incorrectly. However, all subsequent steps completed with this incorrect value were performed correctly. This may either support the arguments of Gulacar *et al*. (2013) related to incomplete chemistry-specific knowledge, or of Tsaparlis (2021b) that increased problem steps increase the chances of error. However, looking for closely at the response and the responses to previous problems it is more likely to be due to the latter argument. All of Margot’s responses were classed as very high for strategy usage and shared nearly all of the elements of expertise suggested by Overton *et al*. (2013).

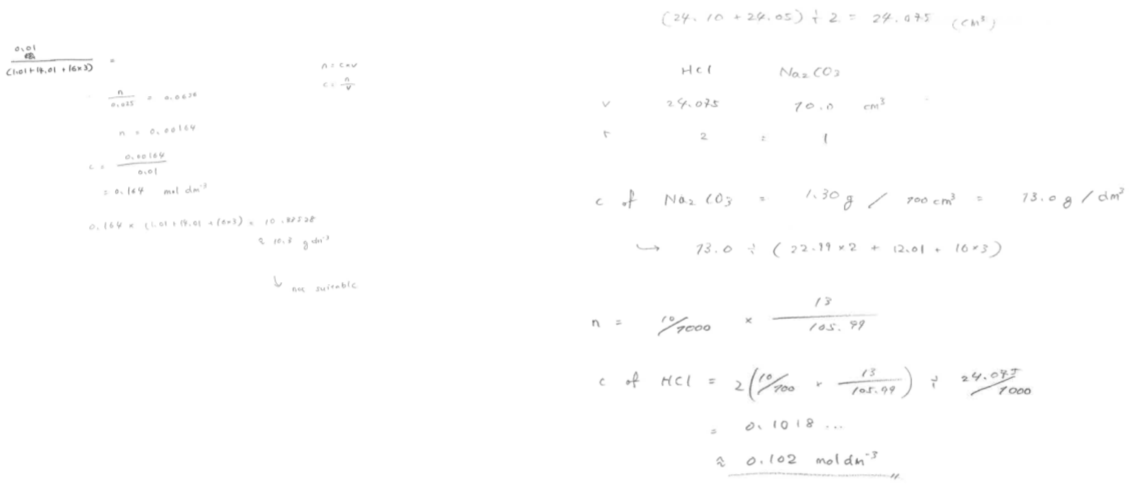
Students acting with agency to combine aspects from the two interventions after the VR phase of the study, was evident in four-out-of-eight responses to VR-2 and VR-3. Indicating that some students sought to adapt taught strategies, ‘cherry picking’ what they felt worked best for them. This may be considered as an example of strategy flexibility, which is defined by Star and Newton (2009) as knowledge of multiple types of solution strategies in addition to the ability and tendency to choose the most appropriate strategy in relation to the specific problem. However, the combination of elements from different strategies may be best understood by considering the elements of strategy flexibility offered by Star and Rittle-Johnson (2008) of: knowledge and use of multiple strategies, and knowledge and use of efficient strategies. In this way, it would seem that these students' knowledge and use of multiple strategies enabled them to act with agency, selecting what they believed to be the most efficient elements. Additionally, this may indicate developing expertise related to metacognitive knowledge, as experts have been observed to choose the most appropriate strategy in relation to particular problems (Star & Newton, 2009).

In all of the cases where elements from the TE and VR strategies were combined, the original language of the taught intervention was retained, as can be seen in the responses of both Helena and Margot. Retaining the original, familiar language of each individual schema, may assist in combining them into a single more complex schema. This may be especially relevant given that new sources of information may combined with existing schema through assimilation (Piaget, 1999). Thus, this new schema is able to be deployed during problem solving, freeing up WM resources to deal with other elements of complex problems.

Two students (Louisa and Zara) preferred to use a strategy based solely upon constructing a table to organise information. For these students, this strategy was in all of the study assessments. For Louisa, this strategy proved to be relatively effective especially in relation to problems VR-2 and VR-3 where all steps were attempted and completed correctly. However, for Zara this strategy produced inconsistent results in terms of strategy usage and the proportion of problem steps attempted and completed correctly (see Table 5). Reasons for their staunch table usage, seem to be rooted in the fact that it is a familiar strategy with Louisa commenting that both TE and VRs "restricted the natural flow" in relation to solving chemistry WPs. Whereas Zara felt that remembering the steps used in each of the interventions "made the questions even more difficult". This apparent resistance to using an alternative strategy or adapting a new strategy may be explained by the reluctance to 'let go' of existing schemas that have previously been successful. Resistance to taught strategies has also been observed by Goulet-Lyle *et al*. (2020), especially in cases when students had used a strategy unsuccessfully. Additionally, when using existing problem-solving strategies, there may be a chance of getting at least *some* of the question correct. This point could prove especially important in relation to test situations, where gaining some marks from tried-and-tested strategies may be preferable to the possibility of scoring nothing when trying a new strategy.

Perhaps most striking and unexpected was the finding that two students (Irini and Julia) independently developed novel strategies over the course of the study that were not evident in the pre-intervention phase. These strategies showed high levels of agency, bearing very little resemblance to any of the strategies used in the study phases. Figure 7 shows the development of Irini's more mathematical strategy.

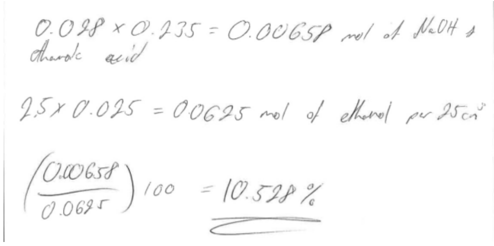
Figure 7: Irini’s Pre-4 response (left) and TE-3 response (right)



Irini found problem Pre-4 challenging, describing her approach at this point as “trial-and-error”, although it meets many of the criteria of high strategy usage. Irini’s approach here shares many commonalities with that of Margot, such as the use of existing schema, leading to a high attempt (0.62) but only 0.15 steps being completed correctly. Similarly to Margot, the eventual failure of the strategy may have been due to increased ECL related to a trial-and-error approach (van Merriënboer & Sweller, 2005). However, her response to TE-3 shows a very different approach (Figure 7). Whilst there is no evidence of TE, all of the pertinent information from the question has been identified and organised in a mathematical manner. This strategy was also observed in her responses to VR-2 and VR-3.

The use of this strategy is successful (all steps complete correctly) with the exception of VR-3, where all steps were attempted and 0.87 were completed correctly. This was due to a miscalculation of an intermediate value that seemed to be related to a chance error rather than a misconception. When asked about this approach, Irini described how she liked to, “shorten the steps”. The strategy employed by Julia, also condenses the problem into a smaller number of steps. Julia’s correct response to problem VR-2 in Figure 9 demonstrates this ability very clearly. Here Julia uses only three steps rather than the ten individual problem steps that may be required.

Figure 8: Julia’s response to VR-2



$0.028 \times 0.235 = 0.00658$ mol of NaOH & ethanoic acid

$2.5 \times 0.025 = 0.0625$ mol of ethanol per $25 cm^3$

$\left(\frac{0.00658}{0.0625}\right) 100 = \underline{\underline{10.528\%}}$

Approaches that show high instances of step-skipping and are consistent with the problem-solving approaches of experts. Irini and Julia’s novel strategies are likely due to the development of a new, complex schema, linking conceptual and procedural knowledge (Ngu & Phan, 2022). From a problem-solving perspective, the examples discussed in this section are of great interest and show how students’ problem-solving strategies develop over time in relation to developing expertise in chemistry WPs.

### 5.3.2 Chemistry word problem solving profiles

Based on the analysis of student responses to the problems in each intervention phase, it became apparent students could be categorised into four profiles. The suggested problem-solving profiles in relation to chemistry WPs are outlined in Table 6 along with their specific characteristics. The profiles are based upon my professional experience, observations from this study and the categories suggested by Overton *et al*. (2013). They are also informed by Gulacar *et al*. (2019) who examined the behaviour of experts' problem-solving strategies in stoichiometry problems.

Table 6: Summary of chemistry WP solving profiles

| | **Profile** | | | |
|---|---|---|---|---|
| | **Innovative Expertise** | **Localised Proficiency** | **Aspirational Beginner** | **Opportunistic Novice** |
| Strategy Usage | Very High | High – Very High | Low-Medium | Very Low |
| Performance | High | High | Low-Medium | Very Low-Low |
| Attempt | High | High | Medium-High | Low |
| Agency of strategy | High | Low-Medium | Low-Medium | Low |
| Step-skipping | Medium-High | Low-Medium | Low | Low |
| Example Response | Irini TE-3<br>(Appendix 11, Fig. 17) | Margot TE-3<br>(Appendix 10, Fig. 14) | Helena VR-3<br>(Appendix 9, Fig. 12) | Sofia Pre-4<br>(Appendix 12, Fig. 20) |

Students that demonstrate innovative expertise are generally very high performing and motivated. Their strategies are developed with agency, either by devising their own, or by combining elements from different taught strategies and retaining their original language. There is also evidence of increased levels of step-skipping. This category shows many attributes associated with strategy flexibility (Star & Newton, 2009; Star &

Rittle-Johnson, 2008), and those of experts in relation to problem solving in chemistry (Gulacar *et al.*, 2019; Ngu & Phan, 2022; Overton *et al.*, 2013).

Localised proficiency refers to the highly effective use of a taught strategy that has been learned fluently and employed correctly. This enables students to demonstrate a high level of proficiency, but only in specific contexts-it is localised to certain types of problem. However, these problem-solving approaches may not be transferred effectively to solve unfamiliar problems. Strategies are very structured indicating some aspects of expertise. However, within the context of an unfamiliar problem, students may not be able to clearly define the problem. In this way, these students may display attributes of experts and novices observed by Overton *et al*. (2013) and Gulacar *et al*. (2019) dependent upon the problem type.

Aspirational beginners are characterised by their strong engagement with strategy usage and resilience. Their attempt level is generally high, but steps are often completed incorrectly. Additionally, errors are mainly related to lower levels of chemistry-specific understanding that are associated with novices, such as errors in unit conversions (Gulacar *et al.*, 2019). Students may also show elements of developing expertise such as a structured strategy (Overton *et al*. 2013).

Opportunistic novices do not tend to have a coherent approach to problem solving and tend to rely on equations and formulae that they have previously learned, employing them in an opportunistic manner.  Formulae may be written out correctly from memory, but incorrect values are often substituted into them.  This category shares many commonalities with the novice category suggested by Overton *et al*. (2013).

These profiles are intentionally broad and multifaceted, and there will be variation within each of them.  For example, the final problem-solving strategies of Margot, Irini and Julia could all be categorised as innovative expertise.  However, Irini and Julia may be considered as demonstrating this more proficiently due to them acting with more agency in the development of their novel strategies, and use of more expert skills such as step-skipping.  Whilst Margot has acted with agency to combine elements of TE and VR to develop a problem-solving approach, she has retained the original language of the taught interventions.  Nevertheless, her agency and development of her own approach should be recognised.  Whilst it is also understood that there will be overlap between each of the profiles and that other profiles will exist, it is believed that they offer a useful framework for understanding students' problem-solving approaches in relation to chemistry WPs.

# 6 Conclusions

## 6.1 Which types of chemistry word problem do students find the most challenging?

Findings from the pre-intervention phase of the study showed that stoichiometry problems with a higher word count and more solution steps were the most challenging. The source of students' difficulties was mainly related to the amount of information in the problem, and the presence of irrelevant and implicit information. Additionally, an increased number of solution steps were also found to negatively affect performance. The number of solution steps were also found to be strongly correlated with the problem word count. Many of the challenges experienced by students are most likely explained by the increased cognitive load of these problems, rather than difficulties related to chemistry-specific concepts. Furthermore, interview data suggests that for the students in this study, the complexity of the linguistic factors of the problems were the greatest source of difficulty. It was also observed that most students did not employ a coherent strategy when attempting more complex WPs.

## 6.2 How do taught interventions impact students' chemistry word problem performance, perceived difficulty and confidence?

The two intervention phases of the study, TE and VR were found to improve performance. Based upon analysis of the qualitative data, increased performance is thought to be due to taught strategies enabling students to develop schema in relation to then problem type. These schemas are believed to assist problem solving by reducing the cognitive load associated with the problem.

Although students did perceive these problems to become less difficult over the course of the study, they were still perceived as difficult. A similar trend was also seen in relation to confidence, which improved by a modest amount indicating that most students still felt less confident when solving more complex chemistry WPs. It is suggested that this perceived low confidence and high difficulty are a related to their self-efficacy with respect to complex chemistry WPs.

## 6.3 How do students use taught interventions to solve chemistry word problems?

Students' problem-solving strategy generally became more developed and organised during the course of the study. The improvements in problem-solving approach were accompanied by an overall improvement in performance, particularly in relation to the number of problem steps attempted. Interestingly, six-out-of-eight students in the study used the taught interventions to develop their own strategy. Four of the new strategies were based upon a combination of the TE and VR interventions, and two students developed novel strategies displaying many characteristics associated with expertise.

## 6.4 Implications for practice and further research

Considering mathematical problems in chemistry as complex WPs equipped me with a deeper theoretical understanding of the challenges that students face in this area of chemistry. However, it is the combination of this theoretical understanding with how students perceive these types of problem that has had the greatest impact on my practice. For example, for the students in this study, the main challenge they faced was sufficiently comprehending the problem so that they felt able to attempt it. This is likely to be true in many chemistry classrooms. Therefore, the findings of this study are likely to be of use to teachers whose students are facing similar challenges.

Observing the development of students' strategies has highlighted the importance of teaching a range of different problem-solving strategies in relation to chemistry WPs. By teaching a range of strategies it may not only expose students to alternative ways of working, but also enable them to exercise agency in relation to these strategies. However, it is acknowledged that the time constraints of different curricula may be a limiting factor with respect to this.

The chemistry WP solving profiles developed as part of this study could potentially be of use in understanding how selected facets of problem-solving attributes and behaviours may compliment or oppose one another. This may assist teachers in identifying individual or co-existing attributes and how they may relate to one another, informing support or interventions on an individual, group or whole-class basis. It would be interesting to further develop and refine this framework with the eventual aim of understanding how and why students may move through different profiles, and how the development of more favourable attributes can be fostered.

However, a missing yet important element of the proposed profiles is self-efficacy, especially when considered across the three dimensions of facet-specificity, strength and level of difficulty (Bandura, 1997; Street *et al.*, 2017; Street *et al.*, 2022). Although perceived confidence and difficulty were measured in this study, improve by a small amount, including these may offer a skewed view.  Related to the findings in relation to perceived confidence and difficulty, it would be interesting to explore how and why students' self-efficacy in relation chemistry WPs changes over time through a longitudinal study.

# 7 References

Anderson, G. L., & Herr, K. (1999). The New Paradigm Wars: Is There Room for Rigorous Practitioner Knowledge in Schools and Universities? *Educational Researcher*, *28*(5), 12–40. https://doi.org/10.3102/0013189X028005012

Anderson, R. C. (1977). The Notion of Schemata and the Educational Enterprise: General Discussion of the Conference. *Schooling and the Acquisition of Knowledge*, 415–432. https://doi.org/10.4324/9781315271644-33/NOTION-SCHEMATA-EDUCATIONAL-ENTERPRISE-GENERAL-DISCUSSION-CONFERENCE-RICHARD-ANDERSON

Ardura, D., Zamora, Á., & Pérez-Bitrián, A. (2021). The role of motivation on secondary school students' causal attributions to choose or abandon chemistry. *Chemistry Education Research and Practice*, *22*(1), 77–92. https://doi.org/10.1039/d0rp00168f

Ayabe, H., Manalo, E., & Vries, E. de. (2022). Problem-appropriate diagram instruction for improving mathematical word problem solving. *Frontiers in Psychology*, *13*, 992625–992625. https://doi.org/10.3389/fpsyg.2022.992625

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bandura, A. (1997). *Self-efficacy: the exercise of control*. W.H. Freeman.

Baumgartner, E., Bell, P., Brophy, S., Hoadley, C., Hsi, S., Joseph, D., Orrill, C., Puntambekar, S., Sandoval, W., & Tabak, I. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189X032001005

Beerepoot, M. T. P. (2023). Formative and Summative Automated Assessment with Multiple-Choice Question Banks. *Journal of Chemical Education*, *100*(8), 2947–2955. https://doi.org/10.1021/acs.jchemed.3c00120

Beitzel, B. D., & Staley, R. K. (2015). The Efficacy of Using Diagrams When Solving Probability Word Problems in College. *The Journal of Experimental Education*, *83*(1), 130–145. https://doi.org/10.1080/00220973.2013.876232

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? *Qualitative Health Research*, *26*(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Bisson, M.-J., Gilmore, C., Inglis, M., & Jones, I. (2016). Measuring Conceptual Understanding Using Comparative Judgement. *International Journal of Research in Undergraduate Mathematics Education*, *2*(2), 141–164. https://doi.org/10.1007/s40753-016-0024-3

Blessing, S. B., & Anderson, J. R. (1996). How People Learn to Skip Steps. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(3), 576–598. https://doi.org/10.1037/0278-7393.22.3.576

Bond, T. G. (2020). *Applying the Rasch model: fundamental measurement in the human sciences* (4th edition). Yan, Z. & Heene, M. (Eds.). Routledge.

Boonen, A. J. H., van Wesel, F., Jolles, J., & van der Schoot, M. (2014). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. *International Journal of Educational Research*, *68*, 15–26. https://doi.org/10.1016/j.ijer.2014.08.001

Bramley, T. (2007). Paired Comparison Methods. In: Newton, P., Baird, J., Goldstein, H., Patrick, H. and Tymms, P. (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 246–300). Qualifications and Curriculum Authority. Availaible from: https://assets.publishing.service.gov.uk/media/5a80d75940f0b62305b8d734/2007-comparability-exam-standards-i-chapter7.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). *Thematic analysis: a practical guide* (Clarke, V. Ed.). SAGE.

British Educational Research Association [BERA] Fourth Edition. (2018). *Ethical Guidelines For Educational Research*. https://www.bera.ac.uk/researchers-resources/publications/ethical- guidelines-for-educational-research-2018

Britton, S., New, P. B., Sharma, M. D., & Yardley, D. (2005). A case study of the transfer of mathematics skills by university students. *International Journal of Mathematical Education in Science and Technology*, *36*(1), 1–13. https://doi.org/10.1080/00207390412331271401

Buzzelli, C., & Johnston, B. (2001). Authority, power, and morality in classroom discourse. *Teaching and Teacher Education*, *17*(8), 873–884. https://doi.org/10.1016/S0742-051X(01)00037-3

Campbell, K. H. (2013). A call to action: why we need more practitioner research. *Democracy & Education*, *21*(2).

Carlson, J. A. (2010). Avoiding traps in member checking. *Qualitative Report*, *15*(5), 1102–1113. https://doi.org/10.46743/2160-3715/2010.1332

Champagne, M. V. (2014) *The Survey Playbook: How to Create the Perfect Survey*. CreateSpace Independent Publishing Platform. Available from: www.createspace.com.

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chandrasegaran, A. L., Treagust, D. F., Waldrip, B. G., & Chandrasegaran, A. (2009). Students' dilemmas in reaction stoichiometry problem solving: Deducing the limiting reagent in chemical reactions. *Chemistry Education Research and Practice*, *10*(1), 14–23. https://doi.org/10.1039/b901456j

Chen, L., & Fouladi, R. T. (2022). Correcting Bias in Extreme Groups Design Using a Missing Data Approach. *Psychological Methods*. https://doi.org/10.1037/met0000508

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Clarke, J. (2005) *Calculations in AS/A Level Chemistry* (7th Ed). Harlow: Longman.

Clark, T. M., Dickson-Karn, N. M., & Anderson, E. (2022). Strategies Undergraduate Students Use to Solve a Volumetric Analysis Problem before and after Instruction. *Journal of Chemical Education*, *99*(11), 3644–3653. https://doi.org/10.1021/acs.jchemed.2c00515

Cochran-Smith, M., & Lytle, S. L. (2009). Teacher Research as Stance. In: *The SAGE Handbook of Educational Action Research*. SAGE Publications Ltd. (pp. 39–49) https://doi.org/10.4135/9780857021021.n5

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (Eighth edition.). London; New York: Routledge.

Cooper, G., & Sweller, J. (1987). Effects of Schema Acquisition and Rule Automation on Mathematical Problem-Solving Transfer. *Journal of Educational Psychology*, *79*(4), 347–362. https://doi.org/10.1037/0022-0663.79.4.347

Cooper, M. M., & Stowe, R. L. (2018). Chemistry Education Research - From Personal Empiricism to Evidence, Theory, and Informed Practice. In: *Chemical Reviews* (Vol. 118, Issue 12, pp. 6053–6087). American Chemical Society. https://doi.org/10.1021/acs.chemrev.8b00020

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (V. L. Plano Clark, Ed.; Third edition.). Sage.

Csíkos, C., Szitányi, J., & Kelemen, R. (2012). The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students. *Educational Studies in Mathematics*, *81*(1), 47–65. https://doi.org/10.1007/s10649-011-9360-z

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H. C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, *6*(APR), 129726. https://doi.org/10.3389/FPSYG.2015.00348/BIBTEX

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the Semantic Structure of Word Problems on Second Graders' Eye Movements. *Journal of Educational Psychology*, *82*(2), 359–365. https://doi.org/10.1037/0022-0663.82.2.359

Depaepe, F., de Corte, E., & Verschaffel, L. (2010). Teachers' metacognitive and heuristic approaches to word problem solving: Analysis and impact on students' beliefs and performance. *ZDM - International Journal on Mathematics Education*, *42*(2), 205–218. https://doi.org/10.1007/s11858-009-0221-5

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method* (Fourth edition.).

Dinkelman, T. (2003). Self-Study In Teacher Education: A Means And Ends Tool For Promoting Reflective Teaching. *Journal of Teacher Education*, *54*(1), 6–18. https://doi.org/10.1177/0022487102238654

Edexcel (2014) *GCE A level Chemistry Paper 2 June 2014*. Pearson Education.

Edexcel (2015) *GCE A level Chemistry Paper 1 Sample Assessment Material*. Pearson Education. Available from: https://qualifications.pearson.com/content/dam/pdf/A%20Level/Chemistry/2015/Specification%20and%20sample%20assessments/A-level-Chemistry-2015-SAMs.pdf

Edexcel (2016) *GCE A level Chemistry Paper 1 June 2016*. Pearson Education. Available from: https://qualifications.pearson.com/content/dam/pdf/A-Level/Chemistry/2015/Exam-materials/8CH0_01_que_20160525.pdf

Edexcel (2017) *GCE AS level Chemistry Paper 1 June 2017*. Pearson Education. Available from: https://qualifications.pearson.com/content/dam/pdf/A-Level/Chemistry/2015/Exam-materials/8CH0_01_que_20170526.pdf

Edexcel (2019) *GCE A level Chemistry Paper 2 June 2019*. Pearson Education. Available from: https://qualifications.pearson.com/content/dam/pdf/A-Level/Chemistry/2015/Exam-materials/9CH0_02_que_20190612.pdf

Edexcel (2020) *GCE AS level Chemistry Paper 1 May 2020*. Pearson Education. Available from: https://qualifications.pearson.com/content/dam/pdf/A-Level/Chemistry/2015/Exam-materials/8ch0-01-que-20220518.pdf

Efklides, A. (2011). Interactions of Metacognition With Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Englert, C. S., Culatta, B. E., & Horn, D. G. (1987). Influence of Irrelevant Information in Addition Word Problems on Problem Solving. *Learning Disability Quarterly*, *10*(1), 29–36. https://doi.org/10.2307/1510752

Ericsson, K. A. (2006). The influence of experience and deliberate practiceon the development of superior expert performance. In: Ericsson, K. A., Charness, N., Feltovich, P. J., and Hoffman, R. R. (Eds.) *The Cambridge Handbook of Expertise and Expert Performance*.  New York, NY: Cambridge University Press, pp. 683–704.

Ericsson, K. A. (2018). The differential influence of experience, practice, and deliberate practice on the development of superior individual performance of experts. In: *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press. pp. 745–769 https://doi.org/10.1017/9781316480748.038

Ericsson, K. A. (1993). *Protocol analysis: verbal reports as data* (Simon, H. A. Ed; Revised edition.). MIT Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and Exceptional Performance: Evidence of Maximal Adaptation to Task Constraints. *Annual Review of Psychology*, *47*(1), 273–305. https://doi.org/10.1146/annurev.psych.47.1.273

Farrimond, H. (2013) *Doing Ethical Research*. Basingstoke, UK: Palgrave Macmillan.

Fink, A. (2002). *How to ask survey questions* (2nd ed.). SAGE

Fisher, J. E., Guha, A., Heller, W., & Miller, G. A. (2020). Extreme-Groups Designs in Studies of Dimensional Phenomena: Advantages, Caveats, and Recommendations. *Journal of Abnormal Psychology (1965)*, *129*(1), 14–20. https://doi.org/10.1037/abn0000480

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *The American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Friedman, H. H. and Amoo, T. (1999) Rating the rating scales. *Journal of Marketing Management*, 9 (3), pp. 114– 23.

Fuchs, L., Fuchs, D., Seethaler, P. M., & Barnes, M. A. (2020). Addressing the role of working memory in mathematical word-problem solving when designing intervention for struggling learners. *ZDM*, *52*(1), 87–96. https://doi.org/10.1007/s11858-019-01070-8

Furió, C., Azcona, R., Ratcliffe, M., & Guisasola, J. (2000). Difficulties in teaching the concepts of ‘amount of substance’ and ‘mole’. *International Journal of Science Education*, *22*(12), 1285–1304. https://doi.org/10.1080/095006900750036262

Galletta, A. (2012). *Mastering the semi-structured interview and beyond: from research design to analysis and publication*. New York University Press.

Gernsbacher, M. A., Soicher, R. N., & Becker-Blease, K. A. (2020). Four Empirically Based Reasons Not to Administer Time-Limited Tests. *Translational Issues in Psychological Science*, *6*(2), 175–190. https://doi.org/10.1037/tps0000232

Glen, S. 2000. “The Dark Side of Purity or the Virtues of Double-Mindedness?” In: Simons, H. and Usher, R. (Eds.), *Situated Ethics in Educational Research*. New York, NY: Routledge Falmer. pp. 12-21

Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four ... or is it two? *Memory (Hove)*, *12*(6), 732–747. https://doi.org/10.1080/09658210344000530

Greene, J. C., & Hall, J. N. (2010). Dialectics and Pragmatism: Being of Consequence. In: SAGE Handbook of Mixed Methods in Social & Behavioral Research. In A. M. Tashakkori & C. B. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed.). SAGE Publications, Inc. https://doi.org/10.4135/9781506335193

Goulet-Lyle, M.-P., Voyer, D., & Verschaffel, L. (2020). How does imposing a step-by-step solution method impact students' approach to mathematical word problem solving? *ZDM*, *52*(1), 139–149. https://doi.org/10.1007/s11858-019-01098-w

Gulacar, O., Cox, C., & Fynewever, H. (2021). Deconstructing the Problem-solving Process: Beneath Assigned Points and Beyond Traditional Assessment. In G. Tsaparlis (Ed.), *Problems and Problem Solving in Chemistry Education: Analysing Data, Looking for Patterns and Making Deductions* (p. 0). The Royal Society of Chemistry. https://doi.org/10.1039/9781839163586-00068

Gulacar, O., Overton, T. L., Bowman, C. R., & Fynewever, H. (2013). A novel code system for revealing sources of students' difficulties with stoichiometry. *Chemistry Education Research and Practice*, *14*(4), 507–515. https://doi.org/10.1039/c3rp00029j

Gulacar, O., Tan, A., Cox, C. T., Bloomquist, J., Jimmy, O., & Cao, N. (2019). Analyzing Characteristics of Experts in the Context of Stoichiometric Problem-Solving. *Education Sciences*, *9*(3), 219. https://doi.org/10.3390/EDUCSCI9030219

Hegarty, M., & Kozhevnikov, M. (1999). Types of Visual-Spatial Representations and Mathematical Problem Solving. *Journal of Educational Psychology*, *91*(4), 684–689. https://doi.org/10.1037/0022-0663.91.4.684

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of Arithmetic Word Problems: A Comparison of Successful and Unsuccessful Problem Solvers. *Journal of Educational Psychology*, *87*(1), 18–32. https://doi.org/10.1037/0022-0663.87.1.18

Herr, Kathryn., & Anderson. G. L. (2005). *The action research dissertation a guide for students and faculty* (G. L. Anderson, Ed.). SAGE.

Hertzum, M., Hansen, K. D., & Andersen, H. H. K. (2009). Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology*, *28*(2), 165–181. https://doi.org/10.1080/01449290701773842

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, *57*(3), 207–220. https://doi.org/10.1080/00461520.2022.2079128

Hoban, R. A., Finlayson, O. E., & Nolan, B. C. (2013). Transfer in chemistry: a study of students' abilities in transferring mathematical knowledge to chemistry. *International Journal of Mathematical Education in Science and Technology*, *44*(1), 14–35. https://doi.org/10.1080/0020739X.2012.690895

Holmes, S. D., He, Q., & Meadows, M. (2017). An investigation of construct relevant and irrelevant features of mathematics problem-solving questions using comparative judgement and Kelly's Repertory Grid. *Research in Mathematics Education*, *19*(2), 112–129. https://doi.org/10.1080/14794802.2017.1334576

Hurley, S. M., & Novick, L. R. (2010). Solving problems using matrix, network, and hierarchy diagrams: The consequences of violating construction conventions. *Quarterly Journal of Experimental Psychology (2006)*, *63*(2), 275–290. https://doi.org/10.1080/17470210902888908

Ingram, J., & Elliott, V. (2019). Transcription Decisions. In *Research Methods for Classroom Discourse* (pp. 189–202). Bloomsbury Academic. https://doi.org/10.5040/9781350072695

Ivankova, N. V. (2015). Mixed Methods Applications in Action Research: From Methods to Community Action. In: *Mixed Methods Applications in Action Research: From Methods to Community Action*. SAGE Publications, Inc. https://doi.org/10.4135/9781071909843

Jaffe, J. B., & Bolger, D. J. (2023). Cognitive Processes, Linguistic Factors, and Arithmetic Word Problem Success: a Review of Behavioral Studies. *Educational Psychology Review*, *35*(4), 105. https://doi.org/10.1007/s10648-023-09821-6

Jerman, M., & Rees, R. (1972). Predicting the Relative Difficulty of Verbal Arithmetic Problems. *Educational Studies in Mathematics*, *4*(3), 306–323. https://doi.org/10.1007/BF00302580

Jia, X., Li, W., & Cao, L. (2019). The role of metacognitive components in creative thinking. *Frontiers in Psychology*, *10*, 2404–2404. https://doi.org/10.3389/fpsyg.2019.02404

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A Comparison of Single and Multiple Strategy Instruction on Third-Grade Students' Mathematical Problem Solving. *Journal of Educational Psychology*, *99*(1), 115–127. https://doi.org/10.1037/0022-0663.99.1.115

Johnston, P. R., Watters, D. J., Brown, C. L., & Loughlin, W. A. (2016). An investigation into student perceptions towards mathematics and their performance in first year chemistry: introduction of online maths skills support. *CHEMISTRY EDUCATION RESEARCH AND PRACTICE*, *17*(4), 123–1214. https://doi.org/10.1039/c6rp00175k

Johnstone, A. H., & El-Banna, H. (1984). New Stars for the Teacher to Steer by? *Journal of Chemical Education*, *61*(10), 847–849. https://pubs.acs.org/sharingguidelines

Johnstone, A. H., & El-Banna, H. (1986). Capacities, demands and processes–a predictive model for science education. *Education in chemistry*, *23*(3), 80-84

Johnstone A. H., (1993), Introduction. In: Wood ,C. and Sleet, R. (Ed.), *Creative Problem Solving in Chemistry*, London: The Royal Society of Chemistry.

Jones, I., & Davies, B. (2024). Comparative judgement in education research. *International Journal of Research and Method in Education*, *47*(2), 170–181. https://doi.org/10.1080/1743727X.2023.2242273

Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, *23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, *30*(1), 17–24. https://doi.org/10.1037/h0057123

Khamis, H. (2008). Measures of association: How to choose? In *Journal of Diagnostic Medical Sonography*, 24(3), pp. 155–162). https://doi.org/10.1177/8756479308317006

Kintsch, W., & Greeno, J. G. (1985). Understanding and Solving Word Arithmetic Problems. *Psychological Review*, *92*(1), 109–129. https://doi.org/10.1037/0033-295X.92.1.109

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, *14*(4), 511–550. https://doi.org/10.1016/0364-0213(90)90008-K

Krawitz, J., & Schukajlow, S. (2020). When Can Making a Drawing Hinder Problem Solving? Effect of the Drawing Strategy on Linear Overgeneralizations and Problem Solving. *Frontiers in Psychology*, *11*, 506–506. https://doi.org/10.3389/fpsyg.2020.00506

Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 236–313). Emerald.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and Novice Performance in Solving Physics Problems. *Science (American Association for the Advancement of Science)*, *208*(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

Lazonby, J. N., Morris, J. E., & Waddington, D. J. (1985). The Mole: Questioning Format Can Make a Difference. *Journal of Chemical Education*, *12*(1), 60–61. https://pubs.acs.org/sharingguidelines

Leiss, D., Plath, J., & Schwippert, K. (2019). Language and Mathematics - Key Factors influencing the Comprehension Process in reality-based Tasks. *Mathematical Thinking and Learning*, *21*(2), 131–153. https://doi.org/10.1080/10986065.2019.1570835

Leiss, D., Schukajlow, S., Blum, W., Messner, R., & Pekrun, R. (2010). The Role of the Situation Model in Mathematical Modelling—Task Analyses, Student Competencies, and Teacher Interventions. *Journal Für Mathematik-Didaktik*, *31*(1), 119–141. https://doi.org/10.1007/s13138-010-0006-y

Lepik, M. (1990). Algebraic Word Problems: Role of Linguistic and Structural Variables. *Educational Studies in Mathematics*, *21*(1), 83–90. https://doi.org/10.1007/BF00311017

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P. M., & van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32–42. https://doi.org/10.1016/j.learninstruc.2013.12.001

Levin, B. B., & Rock, T. C. (2003). The effects of collaborative action research on preservice and experienced teacher partners in professional development schools. *Journal of Teacher Education*, *54*(2), 135–149. https://doi.org/10.1177/0022487102250287

Low, R., & Over, R. (1990). Text Editing of Algebraic Word Problems. *Australian Journal of Psychology*, *42*(1), 63–73. https://doi.org/10.1080/00049539008260106

Low, R., Over, R., Doolan, L., & Michell, S. (1994). Solution of Algebraic Word Problems Following Training in Identifying Necessary and Sufficient Information within Problems. In *Source: The American Journal of Psychology*, 107(3), 423-439. https://www.jstor.org/stable/1422882

Marion, S. B., Reynolds, J. A., Schmid, L., Carter, B. E., Willis, J. H., Mauger, L., & Thompson, R. J. (2023). Beyond Content, Understanding What Makes Test Questions Most Challenging. *Bioscience*, *73*(3), 229–235. https://doi.org/10.1093/biosci/biad007

McKenney, S. E. (2012). *Conducting educational design research* (Reeves, T. C.  Ed.; 1st edition). Routledge. https://doi.org/10.4324/9780203818183

Menter, I. (2016). *A guide to practitioner research in education* (D. L. Elliot, M. Hulme, J. Lewin, & K. Lowden, Eds.). SAGE.

Mertler, C. A. (2017). Action Research: Improving Schools and Empowering Educators. In: *Action Research: Improving Schools and Empowering Educators*. SAGE Publications, Inc. https://doi.org/10.4135/9781483396484

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Morizot, J., Ainsworth, A. T., & Reise, S. P. (2007). Toward Modern Psychometrics Application of Item Response Theory Models in Personality Research. In Robins, R. W., Fraley, R. C., & Krueger, R. F. (Eds.). (2007). *Handbook of research methods in personality psychology*. Guilford Publications. pp. 407-423

Muth, K. D. (1992). Extraneous information and extra steps in arithmetic word problems. *Contemporary Educational Psychology*, *17*(3), 278–285. https://doi.org/10.1016/0361-476X(92)90066-8

Myers, J. A., Witzel, B. S., Powell, S. R., Li, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). *A Meta-Analysis of Mathematics Word-Problem Solving Interventions for Elementary Students Who Evidence Mathematics Difficulties*, *92*(5), 695-742.https://doi.org/10.3102/00346543211070049

Ngu, B. H., Low, R., & Sweller, J. (2002). Text editing in chemistry instruction. *Instructional Science*, *30*(5), 379–402. https://doi.org/10.1023/A:1019833014623

Ngu, B. H., & Phan, H. P. (2022). Developing Problem-Solving Expertise for Word Problems. *Frontiers in Psychology*, *13*, 725280–725280. https://doi.org/10.3389/fpsyg.2022.725280

Ngu, B. H., & Yeung, A. S. (2013). Algebra word problem solving approaches in a chemistry context: Equation worked examples versus text editing. *The Journal of Mathematical Behavior*, *32*(2), 197–208. https://doi.org/10.1016/j.jmathb.2013.02.003

Novick, L. R., & Hurley, S. M. (2001). To Matrix, Network, or Hierarchy: That Is the Question. *Cognitive Psychology*, *42*(2), 158–216. https://doi.org/10.1006/cogp.2000.0746

Opdenacker, C., Fierens, H., Brabant, H. Van, Sevenants, J., Spruyt, J., Slootmaekers, P. J., & Johnstone, A. H. (1990). Academic performance in solving chemistry problems related to student working memory capacity. *International Journal of Science Education*, *12*(2), 177–185. https://doi.org/10.1080/0950069900120206

Overton, T., & Potter, N. M. (2011). Investigating students' success in solving and attitudes towards context-rich open-ended problems in chemistry. *CHEMISTRY EDUCATION RESEARCH AND PRACTICE*, *12*(3), 294–302. https://doi.org/10.1039/C1RP90036F

Overton, T., Potter, N., & Leng, C. (2013). A study of approaches to solving open-ended problems in chemistry. *CHEMISTRY EDUCATION RESEARCH AND PRACTICE*, *14*(4), 468–475. https://doi.org/10.1039/c3rp00028a

Pachman, M., Sweller, J., & Kalyuga, S. (2013). Levels of knowledge and deliberate practice. *Journal of Experimental Psychology. Applied*, *19*(2), 108–119. https://doi.org/10.1037/a0032149

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A Meta-Analysis of Mathematics and Working Memory: Moderating Effects of Working Memory Domain, Type of Mathematics Skill, and Sample Characteristics. *Journal of Educational Psychology*, *108*(4), 455–473. https://doi.org/10.1037/edu0000079

Phan, H. P., & Ngu, B. H. (2021). Introducing the Concept of Consonance-Disconsonance of Best Practice: A Focus on the Development of 'Student Profiling'. *Frontiers in Psychology*, *12*, 557968–557968. https://doi.org/10.3389/fpsyg.2021.557968

Piaget, J. (1999). *The child's conception of physical causality*. Taylor & Francis Group.

Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM*, *52*(1), 33–44. https://doi.org/10.1007/s11858-019-01118-9

Ponte, P. (2005). A Critically constructed concept of action research as a tool for the professional development of teachers. *Journal of In-Service Education*, *31*(2), 273–296. https://doi.org/10.1080/13674580500200279

Pool, J., & Laubscher, D. (2016). Design-based research: is this a suitable methodology for short-term projects?. *Educational Media International EMI.*, *53*(1), 1–11. https://doi.org/info:doi/

Preacher, K. J., Rucker, D. D., MacCallum, R. C., & Nicewander, W. A. (2005). Use of the Extreme Groups Approach: A Critical Re-examination and New Recommendations. *Psychological Methods*, *10*(2), 178–192. https://doi.org/10.1037/1082-989X.10.2.178

Quintero, A. H. (1983). Conceptual Understanding in Solving Two-Step Word Problems with a Ratio. *Journal for Research in Mathematics Education*, *14*(2), 102–112. https://doi.org/10.2307/748578

Randles, C. A., & Overton, T. L. (2015). Expert: Vs. novice: Approaches used by chemists when solving open-ended problems. *Chemistry Education Research and Practice*, *16*(4), 811–823. https://doi.org/10.1039/c5rp00114e

Raviolo, A., Farré, A. S., & Schroh, N. T. (2021). Students' understanding of molar concentration. *Chemistry Education Research and Practice*, *22*(2), 486–497. https://doi.org/10.1039/d0rp00344a

Redish, E. F., & Kuo, E. (2015). Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology. *Science & Education*, *24*(5–6), 561–590. https://doi.org/10.1007/s11191-015-9749-7

Reid, N., & Yang, M. J. (2002). Open-ended problem solving in school chemistry: A preliminary investigation. *International Journal of Science Education*, *24*(12), 1313–1332. https://doi.org/10.1080/09500690210163189

Reisslein, J., Sullivan, H., & Reisslein, M. (2007). Learner Achievement and Attitudes under Different Paces of Transitioning to Independent Problem Solving. *Journal of Engineering Education (Washington, D.C.)*, *96*(1), 45–56. https://doi.org/10.1002/j.2168-9830.2007.tb00914.x

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, *38*(1), 15–22. https://doi.org/10.1207/S15326985EP3801_3

Riessman, C. K. (1993). *Narrative analysis*. Los Angeles: Sage.

Schunk, D. H. (1991). Self-Efficacy and Academic Motivation. *Educational Psychologist*, *26*(3–4), 207–231. https://doi.org/10.1080/00461520.1991.9653133

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, *60*. https://doi.org/10.1016/j.cedpsych.2019.101832

Scott, F. J. (2012). Is mathematics to blame? An investigation into high school students' difficulty in performing calculations in chemistry. *Chemistry Education Research and Practice*, *13*(3), 330–336. https://doi.org/10.1039/c2rp00001f

Silver, E. A. (1979). Student Perceptions of Relatedness among Mathematical Verbal Problems. *Journal for Research in Mathematics Education*, *10*(3), 195–210. https://doi.org/10.2307/748807

Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM*, *41*(5), 557–567. https://doi.org/10.1007/s11858-009-0185-5

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction*, *18*(6), 565–579. https://doi.org/10.1016/j.learninstruc.2007.09.018

Street, K. E. S., Malmberg, L. E., & Stylianides, G. J. (2017). Level, strength, and facet-specific self-efficacy in mathematics test performance. *ZDM - Mathematics Education*, *49*(3), 379–395. https://doi.org/10.1007/S11858-017-0833-0/FIGURES/2

Street, K. E. S., Malmberg, L.-E., & Stylianides, G. J. (2022). *Changes in students' self-efficacy when learning a new topic in mathematics: a micro-longitudinal study*. https://doi.org/10.1007/s10649-022-10165-1

Strohmaier, A. R. (2020). *When reading meets mathematics. Using eye movements to analyze complex word problem solving*. [Doctoral dissertation, Technical University of Munich]. Technical University of Munich. https://doi.org/10.14459/2020md1521471

Strohmaier, A. R., Reinhold, F., Hofer, S., Berkowitz, M., Vogel-Heuser, B., & Reiss, K. (2022). Different complex word problems require different combinations of cognitive skills. *Educational Studies in Mathematics*, *109*(1), 89–114. https://doi.org/10.1007/s10649-021-10079-4

Strömdahl, H., Tullberg, A., & Lybeck, L. (1994). The qualitatively different conceptions of 1 mol. *International Journal of Science Education*, *16*(1), 17–26. https://doi.org/10.1080/0950069940160102

Stylianou, D. A., & Silver, E. A. (2004). The Role of Visual Representations in Advanced Mathematical Problem Solving: An Examination of Expert-Novice Similarities and Differences. *Mathematical Thinking and Learning*, *6*(4), 353–387. https://doi.org/10.1207/s15327833mtl0604_1

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., & Cooper, G. A. (1985). The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. *Cognition and Instruction*, *2*(1), 59–89. https://doi.org/10.1207/s1532690xci0201_3

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. In *Educational Psychology Review*, 31(2), 261–292). Springer New York LLC. https://doi.org/10.1007/s10648-019-09465-5

Tang, H., Kirk, J., & Pienta, N. J. (2014). Investigating the effect of complexity factors in stoichiometry problems using logistic regression and eye tracking. *Journal of Chemical Education*, *91*(7), 969–975. https://doi.org/10.1021/ed4004113

Tarmizi, R. A., & Sweller, J. (1988). Guidance During Mathematical Problem Solving. *Journal of Educational Psychology*, *80*(4), 424–436. https://doi.org/10.1037/0022-0663.80.4.424

Terao, A., Koedinger, K. R., Sohn, M.-H., Qin, Y., Anderson, J. R., & Carter, C. S. (2004). An fMRI study of the Interplay of Symbolic and Visuo-spatial Systems in Mathematical Reasoning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *26*(26).

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How Does Chunking Help Working Memory? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(1), 37–55. https://doi.org/10.1037/xlm0000578

The Education Endowment Foundation. (2021). *Cognitive science approaches in the classroom: a review of the evidence*. https://d2tic4wvo1iusb.cloudfront.net/production/documents/guidance/Cognitive_science_approaches_in_the_classroom_-_A_review_of_the_evidence.pdf?v=1724924081

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, *91*(9), 1426–1431. https://doi.org/10.1021/ed500076x

Tricot, A., & Sweller, J. (2014). Domain-Specific Knowledge and Why Teaching Generic Skills Does Not Work. *Educational Psychology Review*, *26*(2), 265–283. https://doi.org/10.1007/s10648-013-9243-1

Tsaparlis, G. (2021a). Introduction – The Many Types and Kinds of Chemistry Problems. In: G. Tsaparlis (Ed.), *Problems and Problem Solving in Chemistry Education: Analysing Data, Looking for Patterns and Making Deductions* (pp. 1–14). The Royal Society of Chemistry. https://doi.org/10.1039/9781839163586-00001

Tsaparlis, G. (2021b). It Depends on the Problem and on the Solver: An Overview of the Working Memory Overload Hypothesis, Its Applicability and Its Limitations. In: G. Tsaparlis (Ed.), *Problems and Problem Solving in Chemistry Education: Analysing Data, Looking for Patterns and Making Deductions* (p. 0). The Royal Society of Chemistry. https://doi.org/10.1039/9781839163586-00093

Uminski, C., Hubbard, J. K., & Couch, B. A. (2023). How Administration Stakes and Settings Affect Student Behavior and Performance on a Biology Concept Assessment. *CBE - Life Sciences Education*, *22*(2), ar27–ar27. https://doi.org/10.1187/cbe.22-09-0181

van Garderen, D., & Scheuermann, A. M. (2015). Diagramming Word Problems: A Strategic Approach for Instruction. *Intervention in School and Clinic*, *50*(5), 282–290. https://doi.org/10.1177/1053451214560889

van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology*, *36*(3), 212–218. https://doi.org/10.1016/j.cedpsych.2010.10.004

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Verschaffel, L., Depaepe, F., & Van Dooren, W. (2014). Word problems in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education*. pp. 641–645. Dordrecht: Springer.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: a survey. *ZDM*, *52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, *28*(4), 409–426. https://doi.org/10.1080/01443410701708228

Vondrová, N., Novotná, J., & Havlíčková, R. (2019). The influence of situational information on pupils' achievement in additive word problems with several states and transformations. *ZDM*, *51*(1), 183–197. https://doi.org/10.1007/s11858-018-0991-8

Walkington, C., Clinton, V., & Shivraj, P. (2018). How Readability Factors Are Differentially Associated With Performance for Students of Different Backgrounds When Solving Mathematics Word Problems. *American Educational Research Journal*, *55*(2), 362–414. https://doi.org/10.3102/0002831217737028

Walkington, C., Clinton, V., & Sparks, A. (2019). The effect of language modification of mathematics story problems on problem-solving in online homework. *Instructional Science*, *47*(5), 499–529. https://doi.org/10.1007/s11251-019-09481-6

Wang, R., & Krosnick, J. A. (2020). Middle alternatives and measurement validity: a recommendation for survey researchers. *International Journal of Social Research Methodology*, *23*(2), 169–184. https://doi.org/10.1080/13645579.2019.1645384

Waterman, A., & Miller, M. (2022). *Working Memory: A Practical Guide for Teachers*. https://caer.org.uk/wp-content/uploads/CAER-Working-Memory-Guidance.pdf

Westbroek, H., Janssen, F., Mathijsen, I., & Doyle, W. (2022). Teachers as researchers and the issue of practicality. *European Journal of Teacher Education*, *45*(1), 60–76. https://doi.org/10.1080/02619768.2020.1803268

Wong, V. (2023) Science and mathematics: The mathematical demands of science. In: Dillon, J., & Watts, M. (Eds.). (2022). *Debates in Science Education* (2nd ed.). London; New York: Routledge. pp. 239-252 https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.4324/9781003137894

Yeo, G. B., & Neal, A. (n.d.). An Examination of the Dynamic Relationship Between Self-Efficacy and Performance Across Levels of Analysis and Levels of Specificity. *Journal of Applied Psychology*, *91*(5), 1088–1101. https://doi.org/10.1037/0021-9010.91.5.1088

Zahner, D., & Corter, J. E. (2010). The Process of Probability Problem Solving: Use of External Visual Representations. *Mathematical Thinking and Learning*, *12*(2), 177–204. https://doi.org/10.1080/10986061003654240

Zar, J. H. (2010). *Biostatistical analysis* (5th ed.). Prentice-Hall/Pearson.

[return]

## Appendix 1: Text editing intervention final strategy approach

After consultation with the Year 13 trial group and departmental colleagues, the final suggested approach for the text editing intervention can be seen below.

Figure 9: Adapted text editing suggested approach.

**Text editing: suggested approach**

Step 1: Target (what are you being asked to find out?).

Step 2: Identify all substances from the question.

Step 3: List the information that you have about each substance.

Step 3: Pair the information about each substance.

Step 4: Categorise the information as complete, missing, irrelevant (CMI). Decide if the information that you have for each substance is complete, missing, or irrelevant.

Step 5: Calculate. Use the complete information to calculate values (moles, concentrations, etc)

When used with the trial group, the following changes were made:

- Step 5: calculate, was changed to include the extra information in brackets, as the Year 13 students felt that it would be useful to have a reminder about what they could try and use the information to calculate.
- Bold type for **all** in Step 2 was included as the students felt that all of the information needed to be considered, and the bold type would help to emphasise this.

Following the feedback from the trial group, the suggested approach was shared with departmental colleagues for their thoughts and no further changes were made.

[return]

## Appendix 2: Visual representations intervention final strategy approach

Both my colleagues and the Year 13 trail group thought that the overall premise of the "Draw-it" (van Garderen & Scheuermann, 2015) strategy was good, but it contained too many steps and too much information. Additionally, the students thought that it would be useful to include some overlap with the TE strategy such as thinking about what the question is asking to be worked out, if there is any irrelevant information, identifying the different substances, and what is known about them. Departmental colleagues also thought that in Step 3: Organise and Construct, it would be useful to have a reminder about what kind of VRs could be used.

Furthermore, feedback from student questionnaires and interviews from the TE intervention was taken into account in the design of this intervention. A number of students felt that there was significant repetition in the TE strategy, particularly in the steps of identifying, listing and pairing. Therefore, these steps were condensed into the *identify* step.

Figure 10: Visual representations suggested approach

**Visual representations: suggested approach**

Step 1: Read
- Read the question carefully.
- What are you being asked to work out?
- Is there any irrelevant information?

Step 2: Identify
- Identify the substances; what substances are there in the question?
- What values do these substances have (volume, mass, concentration, etc)?

Step 3: Organise and Construct
- How can you represent information in a visual way (diagram) to help you understand it? This could be a table, a graph, a flow diagram, an illustration, etc.
- What steps will you need to take to calculate the answer?

Step 4: Evaluate
- Does your diagram(s) match the question?
- Is your diagram(s) missing anything?
- Will your diagram allow you to work out the answer?
- Can you simplify your diagram? Is there an easier way to solve the problem?

Step 5: Calculate
- Use your diagram(s) to calculate an answer.
- Did you go through all the steps you planned to?
- Check your answer; does it make sense in relation to the question?

Figure 11: Original “Draw-It” strategy (van Garderen & Scheuermann, 2015, p.285)

[return]

## Appendix 3: Post-problem questionnaire

How difficult did you find this question?

| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

[return]

# Appendix 4: Interview guide examples

Pre-intervention interview guide

| |
|---|
| 1. How confident do you feel answering calculations questions in chemistry? Why |
| |
| 2. Do you find calculation questions difficult? Why? |
| |
| 3. What types of calculation question do you find least challenging? Why? |
| |
| 4. What types of calculation question do you find the most challenging? Why? |
| |
| 5. How do you approach calculation questions? Do you use a strategy? |
| |
| 6. What do you feel would help you be more confident when answering calculations question in chemistry? |
| |
| Other notes |

Post-intervention interview guide

| |
|---|
| 1. Do you feel that this strategy has affected your confidence in answering calculations questions in chemistry? How/Why? |
| |
| 2. Did you use this strategy to help you answer the questions? Why/How? |
| |
| 3. Did you find this strategy easy to use? Why? |
| |
| 4. Did you use all the steps of the strategy? Why? |
| |
| 5. How do you feel the strategy could be made better/more useful? |
| |
| 6. What do you feel would help you be more confident when answering calculations question in chemistry? |
| |
| Other notes |

[return]

## Appendix 5: Difficulty and discrimination

In order to determine the difficulty of the problems, the proportion of students answering the question correctly was calculated (Beerepoot *et* al., 2023; Kelley, 1939; Marion *et al*., 2023; Towns, 2014; Uminski *et al*., 2023). As a general guide, difficulty values above 0.75 can be classified as easy, between 0.25 and 0.75 as average, and below 0.25 as difficult (Towns, 2014). This method for calculating question difficulty is mostly used in relation to multiple-choice problems (Beerepoot *et* al., 2023; Kelley, 1939; Marion *et al*., 2023; Towns, 2014; Uminski *et al*., 2023), which posed a potential issue in relation to the problems used in this study. Due to the nature of the problems used in this study, an answer was classified as correct when the proportion of correct steps completed was equal to, or greater than 0.70. This allowed for a distinction to be made between students that demonstrated higher and lower levels of achievement in multiple-step problems. This distinction is in line with previous studies conducted into assessing student performance and understanding, such as Reisslein *et al*. (2007, p. 49).

Discrimination is a measure of how discriminating performance in a problem is between high-performing and low-performing students (Kelley 1939). Discrimination can be calculated by various methods such as those using Rasch analysis, Item Response Theory (IRT), and the extreme group method. The latter of these options was chosen due to Rasch analysis and IRT requiring larger sample sizes of approximately 30-300 respondents to provide reliable estimates (Bond *et al*., 2020; Morizot *et al*., 2007). Therefore, due to the small sample size of this study (n=8), the extreme group method originally outlined by Kelley (1939) was used. Although there are limitations when using an extreme groups method, such as the introduction of bias when excluding portions of a sample (Chen & Fouladi, 2022; Fisher *et al*., 2020), it was felt that this method was appropriate as it would allow the effect of any intervention to be found if it exists, and it being well suited to the exploratory stage of research (Preacher *et al*., 2005).

In this study, the discrimination score per question was based on students' responses to each individual question. It was calculated using the difference between the number of correct responses by the students within the highest 27% of students (n=2) and the students scoring in the lowest 27% (n=2), divided by the number of students that compose 27% of the respondents (n=2) (Kelley, 1939; Marion *et al*., 2023). The students within the lowest and highest 27% of respondents were based on the pre-intervention assessment scores and kept constant throughout the study. This allowed for comparison during the different study phases. The highest performing students were Margot and Irini, and the lowest performing students were Zara and Helena.

[return]

## Appendix 6: Coded interview transcript example

Interviewer: How confident do you feel you felt answering calculation questions in chemistry?

NeEm LC. Fi-Int.

Participant: Not really that confident because they were quite…I thought quite overwhelming, there is quite a lot of information within the questions you have to pick out, and you had to pick out which ones needed to be applied for the question because sometimes they give you information which wasn't necessary.

HInt. Ir-Int.

Interviewer: Why did you feel that you found calculation questions difficult?

PD,Ht.

Participant: I don't think that I don't think I found it difficult, but I wouldn't be easy either if there were more the medium style question, I need to think about it more than the other ones and probably put a bit more effort in and then more likely to be the ones that get wrong. They required more thought

MEff. LC/KSEf.

Interviewer: What types of calculations you find least challenging?

PS-Sim/PS-SQ

Participant: Probably just simple [calculations] what is the moles, concentration [where] we have to do this simple formula that you can follow. Whereas the titration ones they end up with quite a lot of steps you have to go through and then maybe have to change things.

PT-Ti MS

Interviewer: What do you find the most challenging questions?

PT-Ti MS MS-Con

Participant: Titration questions, there are quite a lot of steps to think about ,things like conversions or if they use different types of formula for the same thing [substance] you got remember to divide or times things-just quite a lot going on so they're the trickiest.

MEft

Int: Would you say that the maths is difficult?

MnD MS MnD.

Participant: No, it's easy, it's just knowing what to do and which type of maths to use in a way for the questions remembering to do all the steps the maths itself for me at least is simple, it's just timsing and dividing, but it's remembering what things you have to times and things you have to divide which is the harder part.

Interviewer: OK, so the actual maths itself isn't a problem, so what them difficult? You said lots of steps, is there anything else about those types of questions like titration or back titration apart from the number of steps so conversions dilutions anything else in there that makes it tricky?

PS-WC.

Participant: That question likes to use a lot of words and kind of make it quite a long question we have to read and really pick out the information which is also I feel what makes them more difficult, as a simple one it's normally like here's the concentration here's the volume figure out the moles, whereas this one they'll give you a large question that you must read and find the information.

PS-SQ

Interviewer: Did you have a particular strategy or approach that you used for calculation start problem?

Str-VR

Participant: Sometimes a table, I'm gonna put my information there and see what formulas I could use to help find out the missing gaps in the table probably want to do and then maybe highlight stuff within the question to kind of focus for what I'm using.

Str-IF

Interviewer: OK and then what do you think would have helped you be more confident in calculations?

Str-Pr

Str-Diff.

Participant: maybe practising them or having a different strategy maybe that would make them easier but definitely more practise.

Str-Pr

[return]

# Appendix 7: Example code book

| **Code** | **Description** |
|---|---|
| LC | Refers to any indication of feelings of lower confidence |
| HC | Refers to any indication of feelings of higher confidence |
| NeEm | Refers to any indication of negative emotions |
| PoEm | Refers to any indication of positive emotions |
| Hi-Inf | Refers to higher amounts of information in a problem |
| Lo-Inf | Refers to lower amounts of information in a problem |
| Ir-Inf | Refers to irrelevant information in problems |
| Fi-Inf | Refers to difficulties in finding relevant information in a question that can be used to answer it. |
| HiMe-Eff | Refers to indications of high mental effort needed |
| LoMe-Eff | Refers to indications of low mental effort needed |
| PS-Sim | Refers to problem structure-simple problem structure |
| PS-Com | Refers to problem structure-complex problem structure |
| PS-SP | Refers problems structure-structured problems |
| PS-UP | Refers problems structure-unstructured problems |
| PS-HWC | Refers to mentions of problem structures-higher word count/information |
| PS-LWC | Refers to mentions of problem structures-lower word count/information |
| L-Sef | Refers to indications of low self-efficacy |
| H-Sef | Refers to indications of high self-efficacy |
| MS | Refers to multiple steps of a problem |
| PT-Ti | Refers to a specific problem type-titration problem |
| MS-Conv | Refers to unit conversions in multiple-step calculations |
| MND | Refers to feelings that the maths in the problem is not difficult |
| Str-VR | Refers to the use of a visual representation strategy such as diagram |
| Str-IF | Refers to the use of an information finding strategy such as highlighting |
| Str-Pr | Refers to using practise as a strategy |
| Str-Diff | Refers to thoughts of using a different strategy to what had been used. |

[return]

# Appendix 8: Strategy usage categories and descriptors

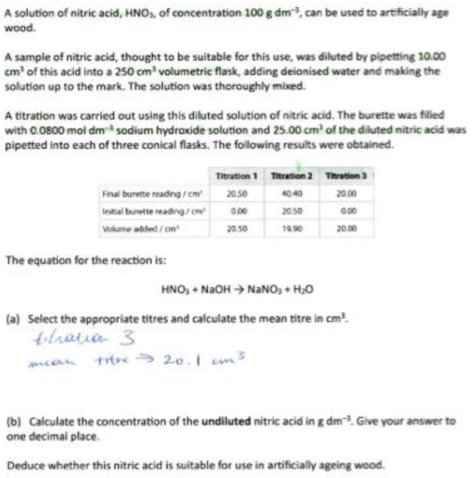
Table 7: Summary of strategy usage categories and descriptors

| **Strategy Usage** | **Description of general observations** |
|---|---|
| 1 – None/Very low | No or very little evidence of any form of strategy usage. May be limited to highlighting of text in the problem text, some of which may be relevant. Copying of information from the problem verbatim. Often results in no attempt to answer the question. |
| 2 – Low | Identifies (in some way) some relevant problem text, may be evidence of removal of irrelevant information (crossing out), some attempt to organise information, often in the form of writing a relevant equation/formula with some attempt at calculation (correct or incorrect). Problem solution is incomplete with few steps attempted. Trial-and-error with no coherence and no clear evaluation in relation to the problem. Copying of information from the problem verbatim. |
| 3 – Medium | Identification of most of the relevant information (highlighting or direct use in organisation method) but there are mistakes/inconsistencies, may be evidence of removal of irrelevant information (crossing out), there is an obvious attempt to organise information, most calculation steps are attempted but not all correct due to mistakes in identifying relevant information. There may still be evidence of a trial-and-error strategy-especially in the later stages of calculation. |
| 4 – High | Majority of relevant information is identified with few mistakes (highlighting/underlining or direct use in organisation method), information is generally organised before calculations take place, majority/all calculation steps attempted, many of which are correct, may be some evidence of evaluation of sub-steps/final answers in relation to the question but this is inconsistent. |
| 5 – Very High | All relevant information is identified and used correctly in calculations, information is organised and structured, evidence of evaluation of sub-steps/final answers in relation to the question, usually evidence of self-correction of errors, all steps completed with no/very few errors. |

[return]

# Appendix 9: Helena's full strategy use

Figure 12: Helena's full responses to the problems Pre-4 (left) and TE-3 (right)



A solution of nitric acid, $HNO_3$, of concentration 100 g dm$^{-3}$, can be used to artificially age wood.

A sample of nitric acid, thought to be suitable for this use, was diluted by pipetting 10.00 cm$^3$ of this acid into a 250 cm$^3$ volumetric flask, adding deionised water and making the solution up to the mark. The solution was thoroughly mixed.

A titration was carried out using this diluted solution of nitric acid. The burette was filled with 0.0800 mol dm$^{-3}$ sodium hydroxide solution and 25.00 cm$^3$ of the diluted nitric acid was pipetted into each of three conical flasks. The following results were obtained.

| | Titration 1 | Titration 2 | Titration 3 |
|---|---|---|---|
| Final burette reading / cm$^3$ | 20.50 | 40.40 | 20.00 |
| Initial burette reading / cm$^3$ | 0.00 | 20.50 | 0.00 |
| Volume added / cm$^3$ | 20.50 | 19.90 | 20.00 |

The equation for the reaction is:

$$HNO_3 + NaOH \rightarrow NaNO_3 + H_2O$$

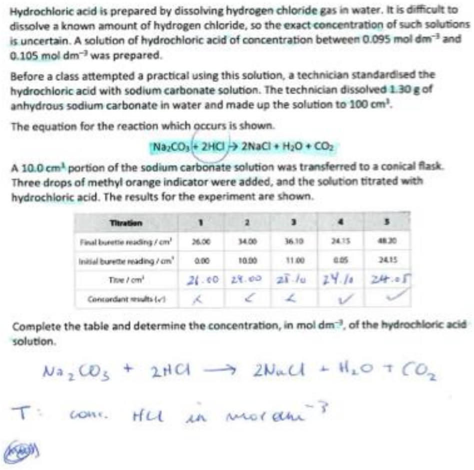
(a) Select the appropriate titres and calculate the mean titre in cm$^3$.

titration 3

mean titre ⇒ 20.1 cm$^3$

(b) Calculate the concentration of the **undiluted** nitric acid in g dm$^{-3}$. Give your answer to one decimal place.

Deduce whether this nitric acid is suitable for use in artificially ageing wood.



Hydrochloric acid is prepared by dissolving hydrogen chloride gas in water. It is difficult to dissolve a known amount of hydrogen chloride, so the exact concentration of such solutions is uncertain. A solution of hydrochloric acid of concentration between 0.095 mol dm$^{-3}$ and 0.105 mol dm$^{-3}$ was prepared.

Before a class attempted a practical using this solution, a technician standardised the hydrochloric acid with sodium carbonate solution. The technician dissolved 1.30 g of anhydrous sodium carbonate in water and made up the solution to 100 cm$^3$.

The equation for the reaction which occurs is shown.

$$Na_2CO_3 + 2HCl \rightarrow 2NaCl + H_2O + CO_2$$

A 10.0 cm$^3$ portion of the sodium carbonate solution was transferred to a conical flask. Three drops of methyl orange indicator were added, and the solution titrated with hydrochloric acid. The results for the experiment are shown.

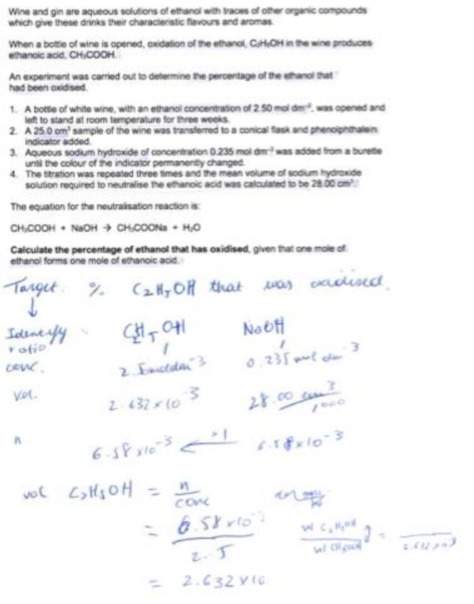
| Titration | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Final burette reading / cm$^3$ | 26.00 | 34.00 | 36.10 | 24.15 | 48.20 |
| Initial burette reading / cm$^3$ | 0.00 | 10.00 | 11.00 | 0.05 | 24.15 |
| Titre / cm$^3$ | 26.00 | 24.00 | 25.10 | 24.10 | 24.05 |
| Concordant results (✓) | ✗ | ✗ | ✗ | ✓ | ✓ |

Complete the table and determine the concentration, in mol dm$^{-3}$, of the hydrochloric acid solution.

$$Na_2CO_3 + 2HCl \longrightarrow 2NaCl + H_2O + CO_2$$

T: conc. HCl in mol dm$^{-3}$

Figure 13: Helena's full responses to the problems VR-2 (left) and VR-3 (right)



Wine and gin are aqueous solutions of ethanol with traces of other organic compounds which give these drinks their characteristic flavours and aromas.

When a bottle of wine is opened, oxidation of the ethanol, $C_2H_5OH$ in the wine produces ethanoic acid, $CH_3COOH$.

An experiment was carried out to determine the percentage of the ethanol that had been oxidised.

1. A bottle of white wine, with an ethanol concentration of 2.50 mol dm$^{-3}$, was opened and left to stand at room temperature for three weeks.
2. A 25.0 cm$^3$ sample of the wine was transferred to a conical flask and phenolphthalein indicator added.
3. Aqueous sodium hydroxide of concentration 0.235 mol dm$^{-3}$ was added from a burette until the colour of the indicator permanently changed.
4. The titration was repeated three times and the mean volume of sodium hydroxide solution required to neutralise the ethanoic acid was calculated to be 28.00 cm$^3$.

The equation for the neutralisation reaction is:

$$CH_3COOH + NaOH \rightarrow CH_3COONa + H_2O$$

**Calculate the percentage of ethanol that has oxidised**, given that one mole of ethanol forms one mole of ethanoic acid.

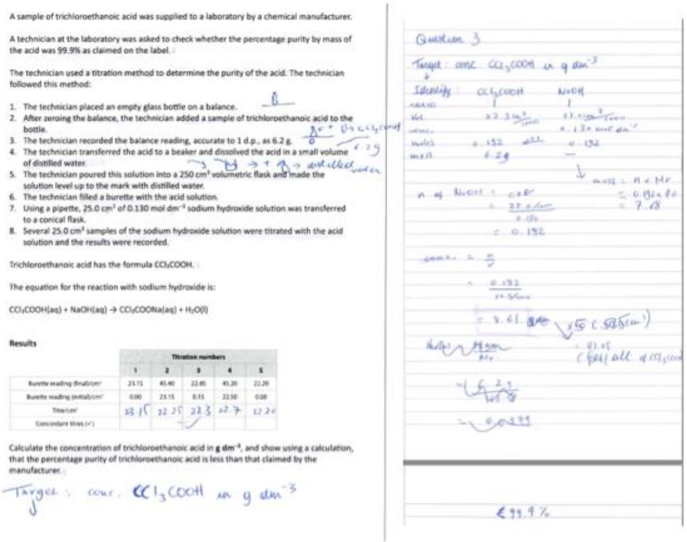
Target: % $C_2H_5OH$ that was oxidised

↓

| | | |
|---|---|---|
| Identify: | $C_2H_5OH$ | NaOH |
| ratio | 1 | 1 |
| conc. | 2.5 mol dm$^{-3}$ | 0.235 mol dm$^{-3}$ |
| vol. | $2.632 \times 10^{-3}$ | $\frac{28.00\ cm^3}{1000}$ |
| n | $6.58 \times 10^{-3}$ $\xleftarrow{\times 1}$ | $6.58 \times 10^{-3}$ |

$$vol\ C_2H_5OH = \frac{n}{conc} = \frac{6.58 \times 10^{-3}}{2.5} = 2.632 \times 10$$

$$\frac{vol\ C_2H_5OH}{vol\ CH_3COOH} = \frac{}{2.632 \times 10^{-3}}$$



A sample of trichloroethanoic acid was supplied to a laboratory by a chemical manufacturer.

A technician at the laboratory was asked to check whether the percentage purity by mass of the acid was 99.9% as claimed on the label.

The technician used a titration method to determine the purity of the acid. The technician followed this method:

1. The technician placed an empty glass bottle on a balance.
2. After zeroing the balance, the technician added a sample of trichloroethanoic acid to the bottle.
3. The technician recorded the balance reading, accurate to 1 d.p., as 6.2 g.
4. The technician transferred the acid to a beaker and dissolved the acid in a small volume of distilled water.
5. The technician poured this solution into a 250 cm$^3$ volumetric flask and made the solution level up to the mark with distilled water.
6. The technician filled a burette with the acid solution.
7. Using a pipette, 25.0 cm$^3$ of 0.130 mol dm$^{-3}$ sodium hydroxide solution was transferred to a conical flask.
8. Several 25.0 cm$^3$ samples of the sodium hydroxide solution were titrated with the acid solution and the results were recorded.

6.2 g

→ + distilled water

Trichloroethanoic acid has the formula $CCl_3COOH$.

The equation for the reaction with sodium hydroxide is:

$$CCl_3COOH(aq) + NaOH(aq) \rightarrow CCl_3COONa(aq) + H_2O(l)$$

**Results**

| | Titration numbers | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Burette reading (final)/cm$^3$ | 23.15 | 45.40 | 22.45 | 45.20 | 22.20 |
| Burette reading (initial)/cm$^3$ | 0.00 | 23.15 | 0.15 | 22.50 | 0.00 |
| Titre/cm$^3$ | 23.15 | 22.25 | 22.3 | 22.7 | 22.2 |
| Concordant titres (✓) | | | ✓ | | |

Calculate the concentration of trichloroethanoic acid in **g dm$^{-3}$**, and show using a calculation, that the percentage purity of trichloroethanoic acid is less than that claimed by the manufacturer.

Target: conc. $CCl_3COOH$ in g dm$^{-3}$

Question 3

Target: conc $CCl_3COOH$ in g dm$^{-3}$

Identify: $CCl_3COOH$ NaOH

mass = n × Mr

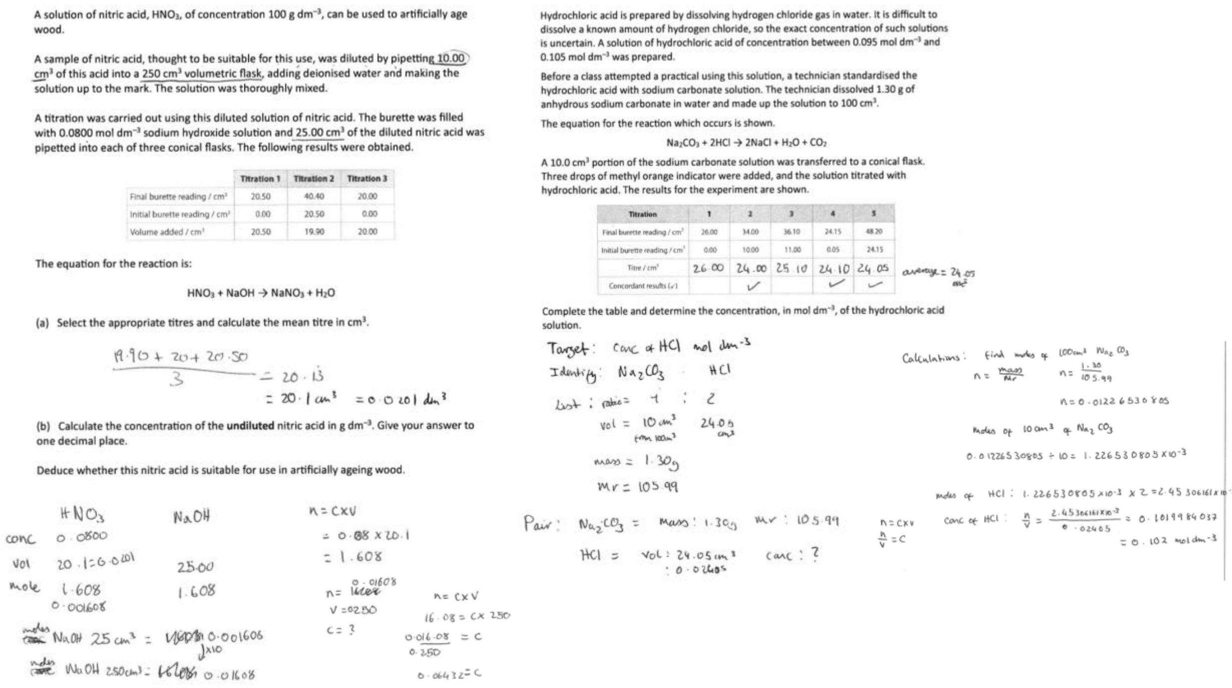
= 0.192

< 99.9 %

[return]

# Appendix 10: Margot’s full strategy use

Figure 14: Margot’s full responses to the problems Pre-4 (left) and TE-3 (right)



A solution of nitric acid, $HNO_3$, of concentration 100 g $dm^{-3}$, can be used to artificially age wood.

A sample of nitric acid, thought to be suitable for this use, was diluted by pipetting 10.00 $cm^3$ of this acid into a 250 $cm^3$ volumetric flask, adding deionised water and making the solution up to the mark. The solution was thoroughly mixed.

A titration was carried out using this diluted solution of nitric acid. The burette was filled with 0.0800 mol $dm^{-3}$ sodium hydroxide solution and 25.00 $cm^3$ of the diluted nitric acid was pipetted into each of three conical flasks. The following results were obtained.

| | Titration 1 | Titration 2 | Titration 3 |
|---|---|---|---|
| Final burette reading / $cm^3$ | 20.50 | 40.40 | 20.00 |
| Initial burette reading / $cm^3$ | 0.00 | 20.50 | 0.00 |
| Volume added / $cm^3$ | 20.50 | 19.90 | 20.00 |

The equation for the reaction is:

$$HNO_3 + NaOH \rightarrow NaNO_3 + H_2O$$

(a) Select the appropriate titres and calculate the mean titre in $cm^3$.

$\frac{19.90 + 20 + 20.50}{3} = 20.13$

$= 20.1\ cm^3 = 0.0201\ dm^3$

(b) Calculate the concentration of the **undiluted** nitric acid in g $dm^{-3}$. Give your answer to one decimal place.

Deduce whether this nitric acid is suitable for use in artificially ageing wood.

| | $HNO_3$ | NaOH |
|---|---|---|
| conc | 0.0800 | |
| vol | 20.1=0.0201 | 25.00 |
| mole | 1.608<br>0.001608 | 1.608 |

n = C×V

= 0.08 × 20.1

= 1.608

n = ~~1.608~~ 0.01608

V = 0.250

c = ?

n = C×V

16.08 = C× 250

$\frac{0.016.08}{0.250}$ = C

0.06432 = C

moles NaOH 25 $cm^3$ = ~~1.608~~ 0.001608 ↓×10

moles NaOH 250$cm^3$ = ~~1.608~~ 0.01608

Hydrochloric acid is prepared by dissolving hydrogen chloride gas in water. It is difficult to dissolve a known amount of hydrogen chloride, so the exact concentration of such solutions is uncertain. A solution of hydrochloric acid of concentration between 0.095 mol $dm^{-3}$ and 0.105 mol $dm^{-3}$ was prepared.

Before a class attempted a practical using this solution, a technician standardised the hydrochloric acid with sodium carbonate solution. The technician dissolved 1.30 g of anhydrous sodium carbonate in water and made up the solution to 100 $cm^3$.

The equation for the reaction which occurs is shown.

$$Na_2CO_3 + 2HCl \rightarrow 2NaCl + H_2O + CO_2$$

A 10.0 $cm^3$ portion of the sodium carbonate solution was transferred to a conical flask. Three drops of methyl orange indicator were added, and the solution titrated with hydrochloric acid. The results for the experiment are shown.

| Titration | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Final burette reading / $cm^3$ | 26.00 | 34.00 | 36.10 | 24.15 | 48.20 |
| Initial burette reading / $cm^3$ | 0.00 | 10.00 | 11.00 | 0.05 | 24.15 |
| Titre / $cm^3$ | 26.00 | 24.00 | 25.10 | 24.10 | 24.05 |
| Concordant results (✓) | | ✓ | | ✓ | ✓ |

average = 24.05 $cm^3$

Complete the table and determine the concentration, in mol $dm^{-3}$, of the hydrochloric acid solution.

Target: conc of HCl mol $dm^{-3}$

Identify: $Na_2CO_3$ HCl

List: ratio = 1 : 2

vol = 10 $cm^3$ from 100$cm^3$ 24.05 $cm^3$

mass = 1.30g

Mr = 105.99

Pair: $Na_2CO_3$ = mass: 1.30g Mr: 105.99

HCl = vol: 24.05 $cm^3$ : 0.02405 conc: ?

n = C×V

$\frac{n}{V}$ = C

Calculations: find moles of 100$cm^3$ $Na_2CO_3$

n = $\frac{mass}{Mr}$ n = $\frac{1.30}{105.99}$

n = 0.0122 6530 805

moles of 10 $cm^3$ of $Na_2CO_3$

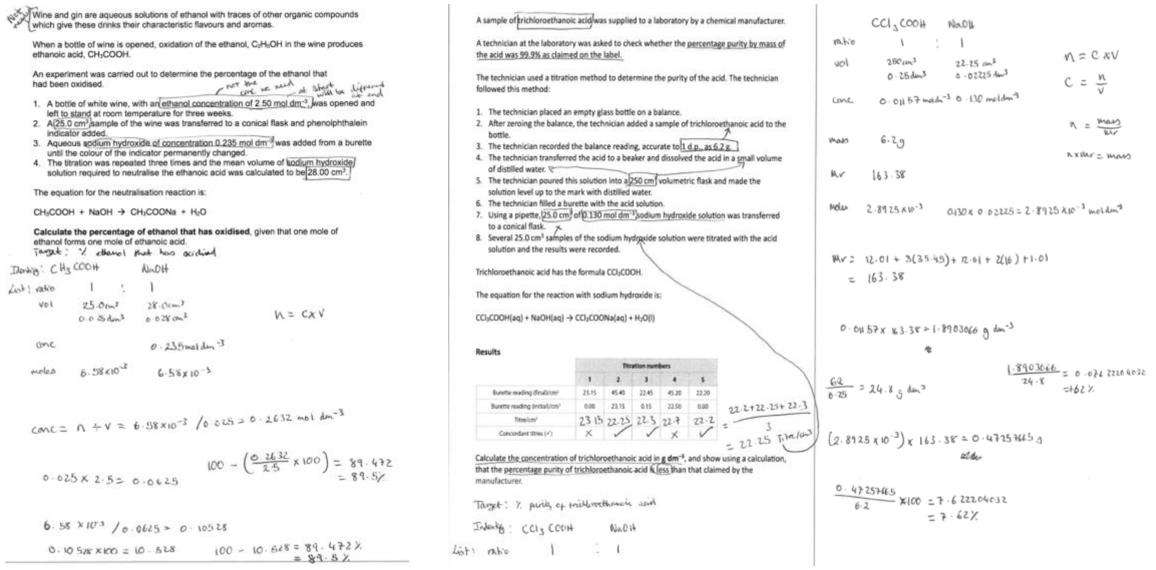
0.01226530805 ÷ 10 = 1.226530805 $\times 10^{-3}$

moles of HCl: 1.226530805 $\times 10^{-3}$ × 2 = 2.45306161 $\times 10^{-3}$

conc of HCl: $\frac{n}{V} = \frac{2.45306161 \times 10^{-3}}{0.02405}$ = 0.1019984037

= 0.102 mol $dm^{-3}$

Figure 15: Margot’s full responses to the problems VR-2 (left) and VR-3 (right)



Wine and gin are aqueous solutions of ethanol with traces of other organic compounds which give these drinks their characteristic flavours and aromas.

When a bottle of wine is opened, oxidation of the ethanol, $C_2H_5OH$ in the wine produces ethanoic acid, $CH_3COOH$.

An experiment was carried out to determine the percentage of the ethanol that had been oxidised.

not the conc we need

at start will be different at end

1. A bottle of white wine, with an ethanol concentration of 2.50 mol $dm^{-3}$, was opened and left to stand at room temperature for three weeks.
2. A 25.0 $cm^3$ sample of the wine was transferred to a conical flask and phenolphthalein indicator added.
3. Aqueous sodium hydroxide of concentration 0.235 mol $dm^{-3}$ was added from a burette until the colour of the indicator permanently changed.
4. The titration was repeated three times and the mean volume of sodium hydroxide solution required to neutralise the ethanoic acid was calculated to be 28.00 $cm^3$.

The equation for the neutralisation reaction is:

$$CH_3COOH + NaOH \rightarrow CH_3COONa + H_2O$$

**Calculate the percentage of ethanol that has oxidised**, given that one mole of ethanol forms one mole of ethanoic acid.

Target: % ethanol that has oxidised

Identify: $CH_3COOH$ NaOH

| | $CH_3COOH$ | NaOH |
|---|---|---|
| List: ratio | 1 : | 1 |
| vol | 25.0$cm^3$<br>0.025$dm^3$ | 28.0$cm^3$<br>0.028$cm^3$ |
| conc | | 0.235mol $dm^{-3}$ |
| moles | 6.58$\times 10^{-3}$ | 6.58$\times 10^{-3}$ |

n = C×V

conc = n ÷ V = 6.58$\times 10^{-3}$ /0.025 = 0.2632 mol $dm^{-3}$

0.025 × 2.5 = 0.0625

$100 - \left(\frac{0.2632}{2.5} \times 100\right) = 89.472$

= 89.5%

6.58 $\times 10^{-3}$ /0.0625 = 0.10528

0.10528 ×100 = 10.528

100 − 10.528 = 89.472%

= 89.5%

A sample of trichloroethanoic acid was supplied to a laboratory by a chemical manufacturer.

A technician at the laboratory was asked to check whether the percentage purity by mass of the acid was 99.9% as claimed on the label.

The technician used a titration method to determine the purity of the acid. The technician followed this method:

1. The technician placed an empty glass bottle on a balance.
2. After zeroing the balance, the technician added a sample of trichloroethanoic acid to the bottle.
3. The technician recorded the balance reading, accurate to 1 d.p., as 6.2 g.
4. The technician transferred the acid to a beaker and dissolved the acid in a small volume of distilled water.
5. The technician poured this solution into a 250 $cm^3$ volumetric flask and made the solution level up to the mark with distilled water.
6. The technician filled a burette with the acid solution.
7. Using a pipette, 25.0 $cm^3$ of 0.130 mol $dm^{-3}$ sodium hydroxide solution was transferred to a conical flask.
8. Several 25.0 $cm^3$ samples of the sodium hydroxide solution were titrated with the acid solution and the results were recorded.

Trichloroethanoic acid has the formula $CCl_3COOH$.

The equation for the reaction with sodium hydroxide is:

$$CCl_3COOH(aq) + NaOH(aq) \rightarrow CCl_3COONa(aq) + H_2O(l)$$

**Results**

| | Titration numbers | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Burette reading (final)/$cm^3$ | 23.15 | 45.40 | 22.45 | 45.20 | 22.20 |
| Burette reading (initial)/$cm^3$ | 0.00 | 23.15 | 0.15 | 22.50 | 0.00 |
| Titre/$cm^3$ | 23.15 | 22.25 | 22.3 | 22.7 | 22.2 |
| Concordant titres (✓) | × | ✓ | ✓ | × | ✓ |

$= \frac{22.2 + 22.25 + 22.3}{3}$

= 22.25 Titre/$cm^3$

Calculate the concentration of trichloroethanoic acid in g $dm^{-3}$, and show using a calculation, that the percentage purity of trichloroethanoic acid is less than that claimed by the manufacturer.

Target: % purity of trichloroethanoic acid

Identify: $CCl_3COOH$ NaOH

List: ratio 1 : 1

| | $CCl_3COOH$ | NaOH |
|---|---|---|
| ratio | 1 : | 1 |
| vol | 250$cm^3$<br>0.25$dm^3$ | 22.25 $cm^3$<br>0.02225 $dm^3$ |
| conc | 0.01157 mol$dm^{-3}$ | 0.130 mol$dm^{-3}$ |
| mass | 6.2g | |
| Mr | 163.38 | |
| moles | 2.8925$\times 10^{-3}$ | 0.130× 0.02225 = 2.8925$\times 10^{-3}$ mol$dm^{-3}$ |

n = C×V

C = $\frac{n}{V}$

n = $\frac{mass}{Mr}$

n×Mr = mass

Mr = 12.01 + 3(35.45) + 12.01 + 2(16) + 1.01

= 163.38

0.01157× 163.38 = 1.8903066 g $dm^{-3}$

$\frac{6.2}{0.25}$ = 24.8 g $dm^{-3}$

$\frac{1.8903066}{24.8}$ = 0.07622204032

= 7.62%

$(2.8925 \times 10^{-3}) \times 163.38$ = 0.47257665 g

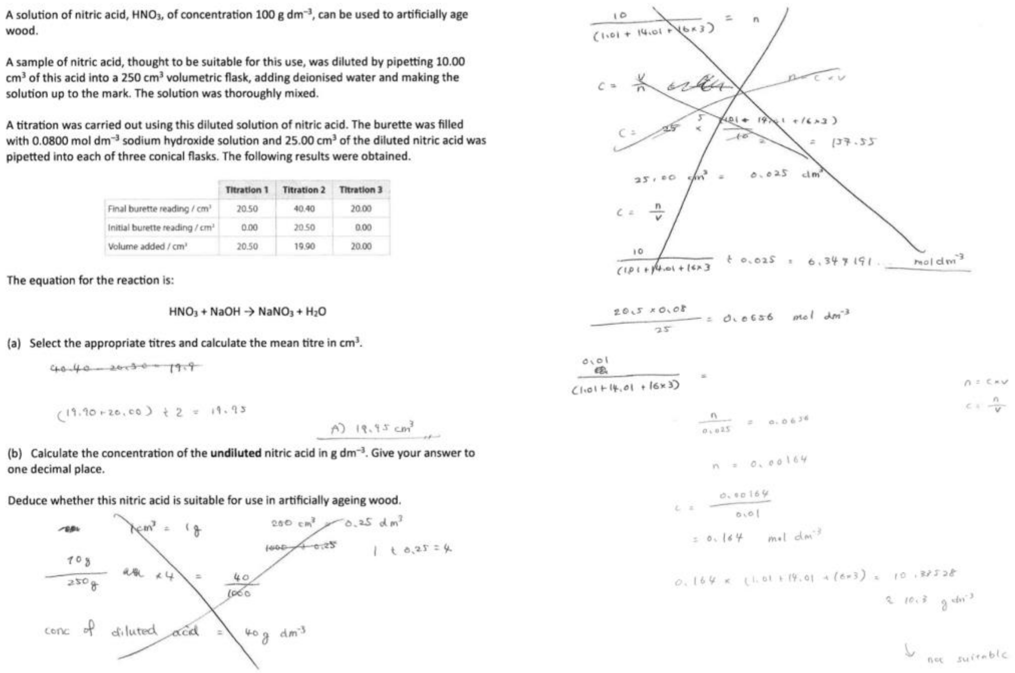
$\frac{0.47257665}{6.2}$ ×100 = 7.622204032

= 7.62%

[return]

# Appendix 11: Irini's full strategy use

Figure 16: Irini's full Pre-4 response



A solution of nitric acid, $HNO_3$, of concentration 100 g $dm^{-3}$, can be used to artificially age wood.

A sample of nitric acid, thought to be suitable for this use, was diluted by pipetting 10.00 $cm^3$ of this acid into a 250 $cm^3$ volumetric flask, adding deionised water and making the solution up to the mark. The solution was thoroughly mixed.

A titration was carried out using this diluted solution of nitric acid. The burette was filled with 0.0800 mol $dm^{-3}$ sodium hydroxide solution and 25.00 $cm^3$ of the diluted nitric acid was pipetted into each of three conical flasks. The following results were obtained.

| | Titration 1 | Titration 2 | Titration 3 |
|---|---|---|---|
| Final burette reading / $cm^3$ | 20.50 | 40.40 | 20.00 |
| Initial burette reading / $cm^3$ | 0.00 | 20.50 | 0.00 |
| Volume added / $cm^3$ | 20.50 | 19.90 | 20.00 |

The equation for the reaction is:

$$HNO_3 + NaOH \rightarrow NaNO_3 + H_2O$$

(a) Select the appropriate titres and calculate the mean titre in $cm^3$.

~~40.40 – 20.30 = 19.9~~

(19.90 + 20.00) ÷ 2 = 19.95

A) 19.95 $cm^3$

(b) Calculate the concentration of the **undiluted** nitric acid in g $dm^{-3}$. Give your answer to one decimal place.

Deduce whether this nitric acid is suitable for use in artificially ageing wood.

~~1 $cm^3$ = 1 g~~

~~$\frac{10 g}{250 g}$ × 4 = $\frac{40}{1000}$~~

~~250 $cm^3$ = 0.25 $dm^3$~~

~~1000 ÷ 0.25~~

1 ÷ 0.25 = 4

~~conc of diluted acid = 40 g $dm^{-3}$~~

~~$\frac{10}{(1.01 + 14.01 + 16 \times 3)}$ = n~~

~~c = $\frac{V}{n}$~~ ~~n = c × v~~

~~c = $\frac{25}{10}$ × (1.01 + 14.01 + 16 × 3) = 157.55~~

~~25.00 $cm^3$ = 0.025 $dm^3$~~

c = $\frac{n}{V}$

~~$\frac{10}{(1.01 + 14.01 + 16 \times 3}$ ÷ 0.025 = 6.347191 mol $dm^{-3}$~~

$\frac{20.5 \times 0.08}{25}$ = 0.0656 mol $dm^{-3}$

$\frac{0.01}{(1.01 + 14.01 + 16 \times 3)}$ =

n = c × v

c = $\frac{n}{V}$

$\frac{n}{0.025}$ = 0.0656

n = 0.00164

c = $\frac{0.00164}{0.01}$

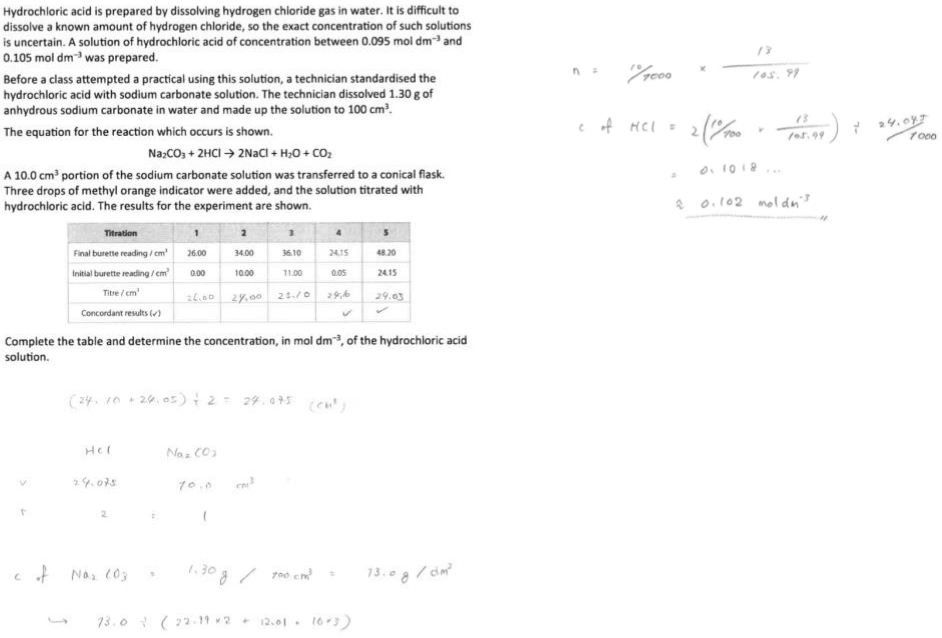
= 0.164 mol $dm^{-3}$

0.164 × (1.01 + 14.01 + (16 × 3) = 10.33528

≈ 10.3 g $dm^{-3}$

↳ not suitable

Figure 17: Irini's full TE-3 response



Hydrochloric acid is prepared by dissolving hydrogen chloride gas in water. It is difficult to dissolve a known amount of hydrogen chloride, so the exact concentration of such solutions is uncertain. A solution of hydrochloric acid of concentration between 0.095 mol $dm^{-3}$ and 0.105 mol $dm^{-3}$ was prepared.

Before a class attempted a practical using this solution, a technician standardised the hydrochloric acid with sodium carbonate solution. The technician dissolved 1.30 g of anhydrous sodium carbonate in water and made up the solution to 100 $cm^3$.

The equation for the reaction which occurs is shown.

$$Na_2CO_3 + 2HCl \rightarrow 2NaCl + H_2O + CO_2$$

A 10.0 $cm^3$ portion of the sodium carbonate solution was transferred to a conical flask. Three drops of methyl orange indicator were added, and the solution titrated with hydrochloric acid. The results for the experiment are shown.

| Titration | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Final burette reading / $cm^3$ | 26.00 | 34.00 | 36.10 | 24.15 | 48.20 |
| Initial burette reading / $cm^3$ | 0.00 | 10.00 | 11.00 | 0.05 | 24.15 |
| Titre / $cm^3$ | 26.00 | 24.00 | 25.10 | 24.10 | 24.05 |
| Concordant results (✓) | | | | ✓ | ✓ |

Complete the table and determine the concentration, in mol $dm^{-3}$, of the hydrochloric acid solution.

(24.10 + 24.05) ÷ 2 = 24.075 ($cm^3$)

| | HCl | $Na_2CO_3$ |
|---|---|---|
| v | 24.075 | 10.0 $cm^3$ |
| r | 2 : | 1 |

c of $Na_2CO_3$ = 1.30 g / 100 $cm^3$ = 13.0 g / $dm^3$

↳ 13.0 ÷ (22.99 × 2 + 12.01 + 16 × 3)

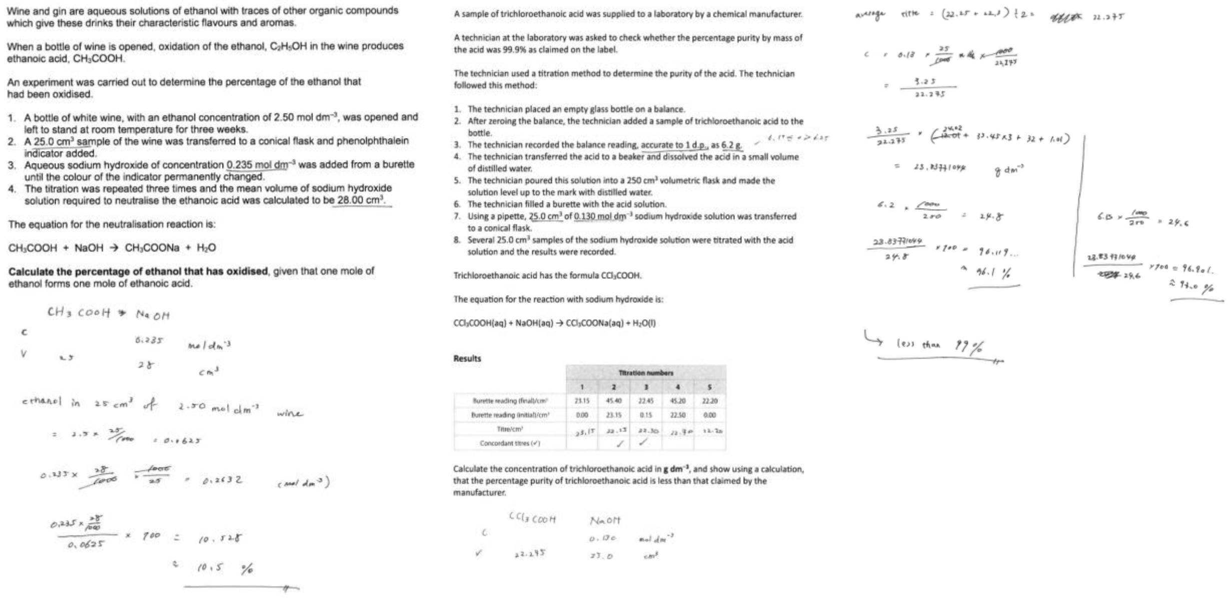
n = $\frac{10}{1000}$ × $\frac{13}{105.99}$

c of HCl = $2\left(\frac{10}{100} \times \frac{13}{105.99}\right) \div \frac{24.075}{1000}$

= 0.1018 ...

≈ 0.102 mol $dm^{-3}$

Figure 18: Irini's full responses to the problems VR-2 (left) and VR-3 (right)



Wine and gin are aqueous solutions of ethanol with traces of other organic compounds which give these drinks their characteristic flavours and aromas.

When a bottle of wine is opened, oxidation of the ethanol, $C_2H_5OH$ in the wine produces ethanoic acid, $CH_3COOH$.

An experiment was carried out to determine the percentage of the ethanol that had been oxidised.

1. A bottle of white wine, with an ethanol concentration of 2.50 mol dm$^{-3}$, was opened and left to stand at room temperature for three weeks.
2. A 25.0 cm$^3$ sample of the wine was transferred to a conical flask and phenolphthalein indicator added.
3. Aqueous sodium hydroxide of concentration 0.235 mol dm$^{-3}$ was added from a burette until the colour of the indicator permanently changed.
4. The titration was repeated three times and the mean volume of sodium hydroxide solution required to neutralise the ethanoic acid was calculated to be 28.00 cm$^3$.

The equation for the neutralisation reaction is:

$CH_3COOH + NaOH \rightarrow CH_3COONa + H_2O$

**Calculate the percentage of ethanol that has oxidised**, given that one mole of ethanol forms one mole of ethanoic acid.

$CH_3COOH + NaOH$

c    0.235 mol/dm$^{-3}$

V  25    28 cm$^3$

ethanol in 25 cm$^3$ of 2.50 mol dm$^{-3}$ wine

$= 2.5 \times 25/1000 = 0.0625$

$0.235 \times \frac{28}{1000} \div \frac{1000}{25} = 0.2632$ (mol dm$^{-3}$)

$\frac{0.235 \times \frac{28}{1000}}{0.0625} \times 100 = 10.528$

$\approx 10.5\ \%$

---

A sample of trichloroethanoic acid was supplied to a laboratory by a chemical manufacturer.

A technician at the laboratory was asked to check whether the percentage purity by mass of the acid was 99.9% as claimed on the label.

The technician used a titration method to determine the purity of the acid. The technician followed this method:

1. The technician placed an empty glass bottle on a balance.
2. After zeroing the balance, the technician added a sample of trichloroethanoic acid to the bottle.
3. The technician recorded the balance reading, accurate to 1 d.p., as 6.2 g. $6.15 \leq x \geq 6.25$
4. The technician transferred the acid to a beaker and dissolved the acid in a small volume of distilled water.
5. The technician poured this solution into a 250 cm$^3$ volumetric flask and made the solution level up to the mark with distilled water.
6. The technician filled a burette with the acid solution.
7. Using a pipette, 25.0 cm$^3$ of 0.130 mol dm$^{-3}$ sodium hydroxide solution was transferred to a conical flask.
8. Several 25.0 cm$^3$ samples of the sodium hydroxide solution were titrated with the acid solution and the results were recorded.

Trichloroethanoic acid has the formula $CCl_3COOH$.

The equation for the reaction with sodium hydroxide is:

$CCl_3COOH(aq) + NaOH(aq) \rightarrow CCl_3COONa(aq) + H_2O(l)$

**Results**

| | Titration numbers | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Burette reading (final)/cm$^3$ | 23.15 | 45.40 | 22.45 | 45.20 | 22.20 |
| Burette reading (initial)/cm$^3$ | 0.00 | 23.15 | 0.15 | 22.50 | 0.00 |
| Titre/cm$^3$ | 23.15 | 22.25 | 22.30 | 22.70 | 22.20 |
| Concordant titres (✓) | | ✓ | ✓ | | |

Calculate the concentration of trichloroethanoic acid in **g dm$^{-3}$**, and show using a calculation, that the percentage purity of trichloroethanoic acid is less than that claimed by the manufacturer.

$CCl_3COOH$    $NaOH$

c        0.130 mol dm$^{-3}$

V  22.275    25.0 cm$^3$

average titre = (22.25 + 22.3) ÷ 2 = 22.275

$c = 0.13 \times \frac{25}{1000} \times \frac{1000}{22.275}$

$= \frac{3.25}{22.275}$

$\frac{3.25}{22.275} \times (\overset{24.02}{\cancel{12.01}} + 35.45 \times 3 + 32 + 1.01)$

$= 23.83771049$ g dm$^{-3}$

$6.2 \times \frac{1000}{250} = 24.8$

$\frac{23.83771049}{24.8} \times 100 = 96.119...$

$\approx 96.1\ \%$

$6.15 \times \frac{1000}{250} = 24.6$

$\frac{23.83771049}{24.6} \times 100 = 96.901.$

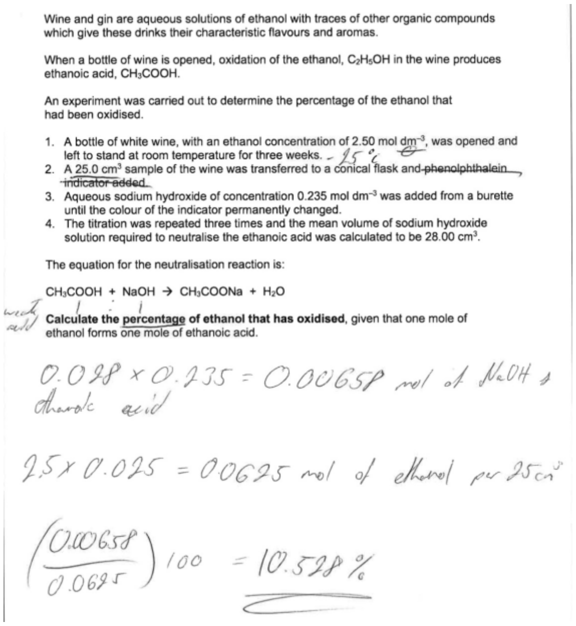
$\approx 97.0\ \%$

↳ less than 99 %

[return]

# Appendix 12: Julia's and Sofia's full strategy use

Figure 19: Julia's full response to VR-2



Wine and gin are aqueous solutions of ethanol with traces of other organic compounds which give these drinks their characteristic flavours and aromas.

When a bottle of wine is opened, oxidation of the ethanol, $C_2H_5OH$ in the wine produces ethanoic acid, $CH_3COOH$.

An experiment was carried out to determine the percentage of the ethanol that had been oxidised.

1. A bottle of white wine, with an ethanol concentration of 2.50 mol dm$^{-3}$, was opened and left to stand at room temperature for three weeks. - 25°C
2. A 25.0 cm$^3$ sample of the wine was transferred to a conical flask and ~~phenolphthalein indicator added~~.
3. Aqueous sodium hydroxide of concentration 0.235 mol dm$^{-3}$ was added from a burette until the colour of the indicator permanently changed.
4. The titration was repeated three times and the mean volume of sodium hydroxide solution required to neutralise the ethanoic acid was calculated to be 28.00 cm$^3$.

The equation for the neutralisation reaction is:

$CH_3COOH + NaOH \rightarrow CH_3COONa + H_2O$

weak acid

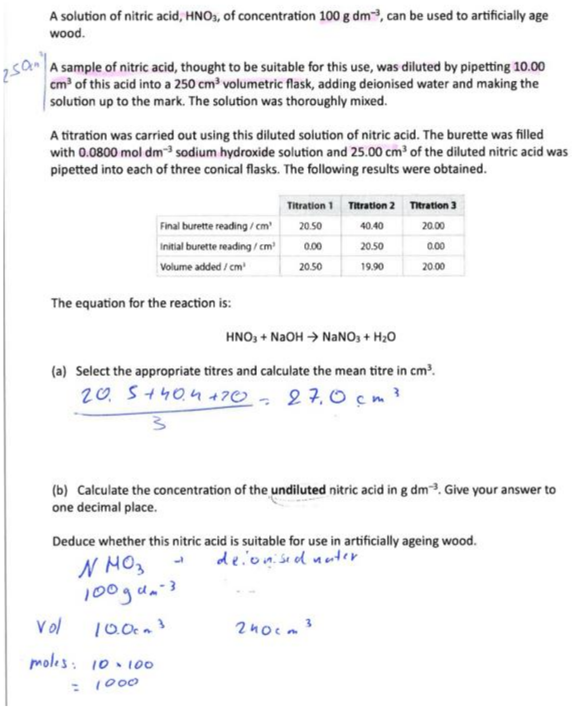
**Calculate the percentage of ethanol that has oxidised**, given that one mole of ethanol forms one mole of ethanoic acid.

0.028 × 0.235 = 0.00658 mol of NaOH & ethanoic acid

2.5 × 0.025 = 0.0625 mol of ethanol per 25cm$^3$

$\left(\frac{0.00658}{0.0625}\right) 100 = 10.528\%$

Figure 20: Sofia's full response to Pre-4



A solution of nitric acid, $HNO_3$, of concentration 100 g dm$^{-3}$, can be used to artificially age wood.

250cm$^3$

A sample of nitric acid, thought to be suitable for this use, was diluted by pipetting 10.00 cm$^3$ of this acid into a 250 cm$^3$ volumetric flask, adding deionised water and making the solution up to the mark. The solution was thoroughly mixed.

A titration was carried out using this diluted solution of nitric acid. The burette was filled with 0.0800 mol dm$^{-3}$ sodium hydroxide solution and 25.00 cm$^3$ of the diluted nitric acid was pipetted into each of three conical flasks. The following results were obtained.

| | Titration 1 | Titration 2 | Titration 3 |
|---|---|---|---|
| Final burette reading / cm$^3$ | 20.50 | 40.40 | 20.00 |
| Initial burette reading / cm$^3$ | 0.00 | 20.50 | 0.00 |
| Volume added / cm$^3$ | 20.50 | 19.90 | 20.00 |

The equation for the reaction is:

$HNO_3 + NaOH \rightarrow NaNO_3 + H_2O$

(a) Select the appropriate titres and calculate the mean titre in cm$^3$.

$\frac{20.5+40.4+20}{3} = 27.0\ cm^3$

(b) Calculate the concentration of the **undiluted** nitric acid in g dm$^{-3}$. Give your answer to one decimal place.

Deduce whether this nitric acid is suitable for use in artificially ageing wood.

$NHO_3$ → deionised water

100g dm$^{-3}$

Vol 10.0cm$^3$ 240cm$^3$

moles: 10 × 100

= 1000

[return]

## Appendix 13: Pre-intervention assessment

# Pre-intervention assessment

**Instructions**

1. You should attempt each question, even if you don't feel like you can complete it.
2. After you answer each question, you should complete the evaluation-please be as detailed and honest as you can in the evaluation.
3. You should show all of your working.
4. If you have any questions, please ask-this is not a test.
5. You may use a calculator and a Data Booklet to help you.
6. The blank space after each question is there for you to write your answer and show your working. You do not have to fill the entire space.

Acknowledgements:

1. Questions 2 and 3 taken from Edexcel GCE A level Chemistry Paper 1 June 2016, p2 and p10.
2. Question 4 taken from: Edexcel GCE AS level Chemistry Paper 1 May 2020, p 16.

## Question 1

Calculate the number of moles in 4.60g of ethanol, $C_2H_5OH$.

**Evaluation: Question 1**

How difficult did you find this question?

| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 2

What is the number of ions in 9.53 g of magnesium chloride, $MgCl_2$?

[Avogadro constant = $6.02 \times 10^{23}$ $mol^{-1}$]

**Evaluation: Question 2**

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 3

When malachite, $Cu_2CO_3(OH)_2$ is heated to approximately 300 °C, water, carbon dioxide and copper(II) oxide are formed.

The equation for this decomposition is: $Cu_2CO_3(OH)_2 \rightarrow 2CuO + CO_2 + H_2O$

Calculate the maximum volume of carbon dioxide that could be produced when 0.810 g of malachite is thermally decomposed.

Assume that the gas is collected at a temperature of 25 °C and 101 kPa pressure.
Give your answer to an appropriate number of significant figures and state the units.

[The ideal gas equation is $pV = nRT$. Gas constant (R) = 8.31 J $mol^{-1}$ $K^{-1}$]

**Evaluation: Question 3**

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

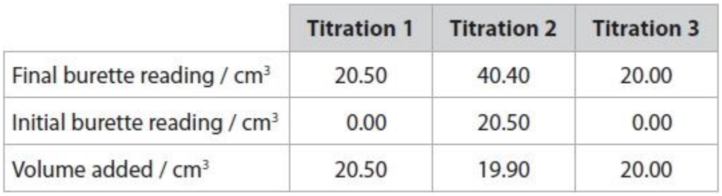
How did you approach this question? Did you use any particular strategies?

## Question 4

A solution of nitric acid, $HNO_3$, of concentration 100 g dm$^{-3}$, can be used to artificially age wood.

A sample of nitric acid, thought to be suitable for this use, was diluted by pipetting 10.00 cm$^3$ of this acid into a 250 cm$^3$ volumetric flask, adding deionised water and making the solution up to the mark. The solution was thoroughly mixed.

A titration was carried out using this diluted solution of nitric acid. The burette was filled with 0.0800 mol dm$^{-3}$ sodium hydroxide solution and 25.00 cm$^3$ of the diluted nitric acid was pipetted into each of three conical flasks. The following results were obtained.



| | Titration 1 | Titration 2 | Titration 3 |
|---|---|---|---|
| Final burette reading / cm$^3$ | 20.50 | 40.40 | 20.00 |
| Initial burette reading / cm$^3$ | 0.00 | 20.50 | 0.00 |
| Volume added / cm$^3$ | 20.50 | 19.90 | 20.00 |

The equation for the reaction is:

$HNO_3 + NaOH \rightarrow NaNO_3 + H_2O$

(a) Select the appropriate titres and calculate the mean titre in cm$^3$.

(b) Calculate the concentration of the undiluted nitric acid in g dm$^{-3}$. Give your answer to one decimal place.

Deduce whether this nitric acid is suitable for use in artificially ageing wood.

Extra space for your answer.

**Evaluation: Question 4**

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

[return]

## Appendix 14: Text Editing post-intervention assessment

# Post-Text editing assessment

### Instructions

1. You should attempt each question, even if you don't feel like you can complete it.
2. After you answer each question, you should complete the evaluation-please be as detailed and honest as you can in the evaluation.
3. You should show all of your working.
4. If you have any questions, please ask-this is not a test.
5. You may use a calculator and a Data Booklet to help you.
6. The blank space after each question is there for you to write your answer and show your working.  You do not have to fill the entire space.

Acknowledgements:

1. Question 2 adapted from Edexcel GCE A level Chemistry Paper 2 June 2014.
2. Question 3 taken from: Edexcel GCE AS level Chemistry Paper 1 May 2017.

**Question 1**

25.0cm$^3$ of sodium hydroxide solution, NaOH with an unknown concentration was placed in a conical flask and titrated with a hydrochloric acid solution, HCl with a concentration of 0.0125 mol dm$^{-3}$ added from a burette. 23.5 cm$^3$ of the hydrochloric acid solution was needed to neutralise the sodium hydroxide solution.

The equation for the reaction is:
$NaOH + HCl \rightarrow NaCl + H_2O$

Calculate the concentration of the sodium hydroxide solution in mol dm$^{-3}$.

## Evaluation: Question 1

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 2

Brand **X** is unlike many conventional toilet cleaners in that it does not contain bleach, but instead contains hydrochloric acid, HCl. The label states that the toilet cleaner contains 9 g of HCl per 100 $cm^3$ of the toilet cleaner.

An industrial technician was given the task of checking the validity of this statement. Using a 25.0 $cm^3$ portion of the toilet cleaner, the technician carried out a titration using 2.50 mol $dm^{-3}$ sodium hydroxide solution, NaOH. 24.55 $cm^3$ of the sodium hydroxide solution was required to neutralise the hydrochloric acid in the toilet cleaner solution.

Calculate the mass of HCl present in 100 $cm^3$ of the toilet cleaner.

## Evaluation: Question 2

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 3

Hydrochloric acid is prepared by dissolving hydrogen chloride gas in water. It is difficult to dissolve a known amount of hydrogen chloride, so the exact concentration of such solutions is uncertain. A solution of hydrochloric acid of concentration between 0.095 mol dm$^{-3}$ and 0.105 mol dm$^{-3}$ was prepared.

Before a class attempted a practical using this solution, a technician standardised the hydrochloric acid with sodium carbonate solution. The technician dissolved 1.30 g of anhydrous sodium carbonate in water and made up the solution to 100 cm$^3$.

The equation for the reaction which occurs is shown.

$$Na_2CO_3 + 2HCl \rightarrow 2NaCl + H_2O + CO_2$$

A 10.0 cm$^3$ portion of the sodium carbonate solution was transferred to a conical flask. Three drops of methyl orange indicator were added, and the solution titrated with hydrochloric acid. The results for the experiment are shown.

The figure originally presented here cannot be made freely available via ORA because of copyright.

Complete the table and determine the concentration, in mol dm$^{-3}$, of the hydrochloric acid solution.

**Extra space for your answer.**

## Evaluation: Question 3

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

[return]

Appendix 15: Visual representations post-intervention assessment

# Post-Visual representations assessment

## Instructions

1. You should attempt each question, even if you don't feel like you can complete it.
2. After you answer each question, you should complete the evaluation-please be as detailed and honest as you can in the evaluation.
3. You should show all of your working.
4. If you have any questions, please ask-this is not a test.
5. You may use a calculator and a Data Booklet to help you.
6. The blank space after each question is there for you to write your answer and show your working.  You do not have to fill the entire space.

Acknowledgements:

1. Question 1 taken from: Clarke, J. (2005) *Calculations in AS/A Level Chemistry* (7th Ed).  Harlow, Longman. p 81.
2. Question 2 adapted from Edexcel GCE A level Chemistry Paper 2 June 2019, Question 6.
3. Question 3 taken from: Edexcel GCE A level Chemistry Paper 1 Sample Assessment Material, Question 8.

**Question 1**

Lime water is calcium hydroxide solution. In an experiment to find the concentration of calcium hydroxide in lime water, 25.0cm$^3$ of lime water was placed in a conical flask and titrated with a hydrochloric acid solution, HCl with a concentration of 0.0400 mol dm$^{-3}$. 18.8 cm$^3$ of the hydrochloric acid solution was needed to neutralise the lime water.

The equation for the reaction is:

$Ca(OH)_2 + 2HCl \rightarrow CaCl_2 + 2H_2O$

Calculate the concentration of calcium hydroxide solution in **g dm$^{-3}$**.

## Evaluation: Question 1

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 2

Wine and gin are aqueous solutions of ethanol with traces of other organic compounds which give these drinks their characteristic flavours and aromas.

When a bottle of wine is opened, oxidation of the ethanol, $C_2H_5OH$ in the wine produces ethanoic acid, $CH_3COOH$.

An experiment was carried out to determine the percentage of the ethanol that had been oxidised.

1. A bottle of white wine, with an ethanol concentration of 2.50 mol dm$^{-3}$, was opened and left to stand at room temperature for three weeks.
2. A 25.0 cm$^3$ sample of the wine was transferred to a conical flask and phenolphthalein indicator added.
3. Aqueous sodium hydroxide of concentration 0.235 mol dm$^{-3}$ was added from a burette until the colour of the indicator permanently changed.
4. The titration was repeated three times and the mean volume of sodium hydroxide solution required to neutralise the ethanoic acid was calculated to be 28.00 cm$^3$.

The equation for the neutralisation reaction is:

$CH_3COOH + NaOH \rightarrow CH_3COONa + H_2O$

**Calculate the percentage of ethanol that has oxidised**, given that one mole of ethanol forms one mole of ethanoic acid.

## Extra space for your answer

## Evaluation: Question 2

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

## Question 3

A sample of trichloroethanoic acid was supplied to a laboratory by a chemical manufacturer.

A technician at the laboratory was asked to check whether the percentage purity by mass of the acid was 99.9% as claimed on the label.

The technician used a titration method to determine the purity of the acid. The technician followed this method:

1. The technician placed an empty glass bottle on a balance.
2. After zeroing the balance, the technician added a sample of trichloroethanoic acid to the bottle.
3. The technician recorded the balance reading, accurate to 1 d.p., as 6.2 g.
4. The technician transferred the acid to a beaker and dissolved the acid in a small volume of distilled water.
5. The technician poured this solution into a 250 cm$^3$ volumetric flask and made the solution level up to the mark with distilled water.
6. The technician filled a burette with the acid solution.
7. Using a pipette, 25.0 cm$^3$ of 0.130 mol dm$^{-3}$ sodium hydroxide solution was transferred to a conical flask.
8. Several 25.0 cm$^3$ samples of the sodium hydroxide solution were titrated with the acid solution and the results were recorded.

Trichloroethanoic acid has the formula $CCl_3COOH$.

The equation for the reaction with sodium hydroxide is:

$CCl_3COOH(aq) + NaOH(aq) \rightarrow CCl_3COONa(aq) + H_2O(l)$

**Results**

Calculate the concentration of trichloroethanoic acid in **g dm$^{-3}$**, and show using a calculation, that the percentage purity of trichloroethanoic acid is less than that claimed by the manufacturer.

**Space for your answer.**

## Evaluation: Question 3

How difficult did you find this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Difficult | Difficult | Somewhat Difficult | Somewhat Easy | Easy | Extremely Easy |

How confident did you feel answering this question?

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Extremely Unconfident | Unconfident | Somewhat Unconfident | Somewhat Confident | Confident | Extremely Confident |

What did you find least challenging about this question?

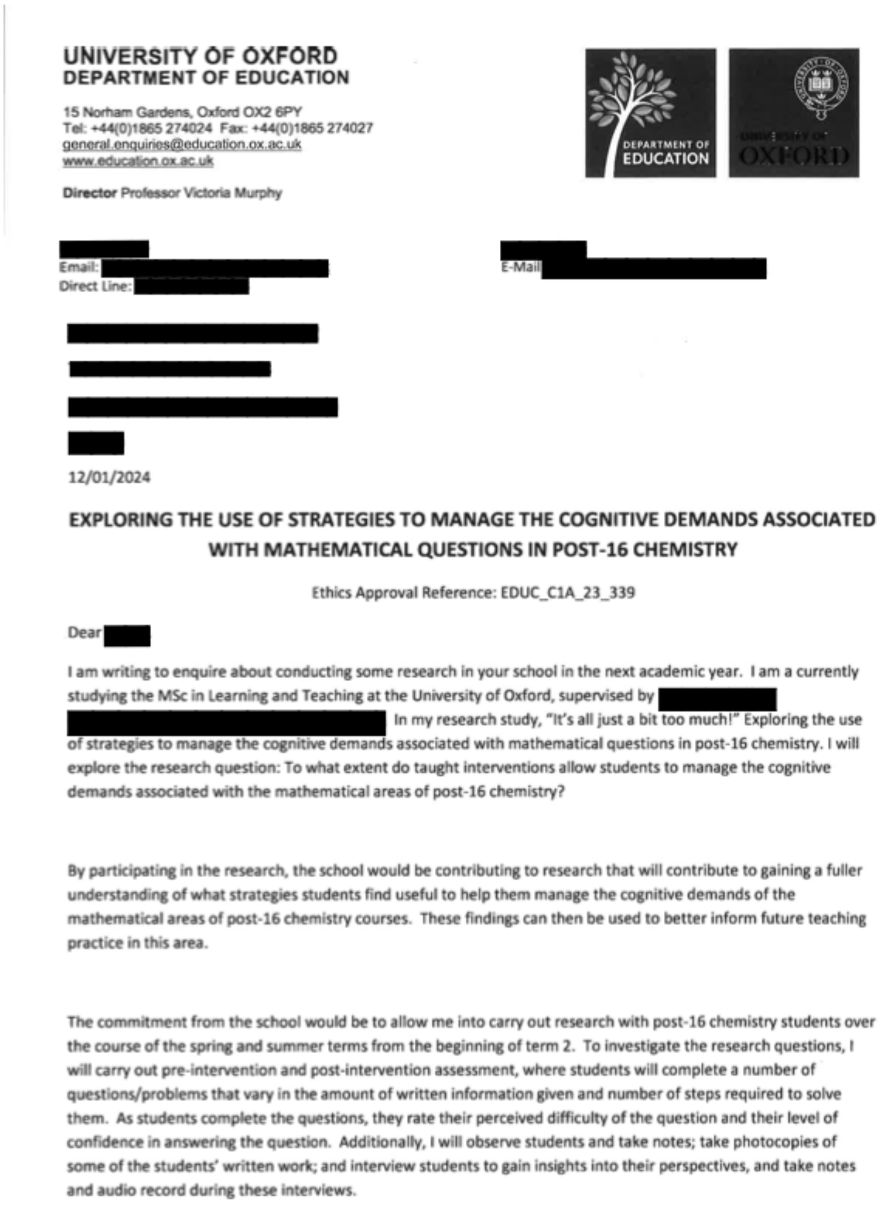
What did you find most challenging about this question?

How did you approach this question? Did you use any particular strategies?

[Return to TOC]

## Appendix 16: Headteacher permission letter

It should be noted that the title of the study was changed after ethical approval had been granted. This change took place through the official University channels and was approved. The school were also informed of the change of the study title. All of the following documentation contains the old study title.



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

[redacted]
Email: [redacted]
Direct Line: [redacted]

[redacted]
E-Mail [redacted]

[redacted]

12/01/2024

**EXPLORING THE USE OF STRATEGIES TO MANAGE THE COGNITIVE DEMANDS ASSOCIATED WITH MATHEMATICAL QUESTIONS IN POST-16 CHEMISTRY**

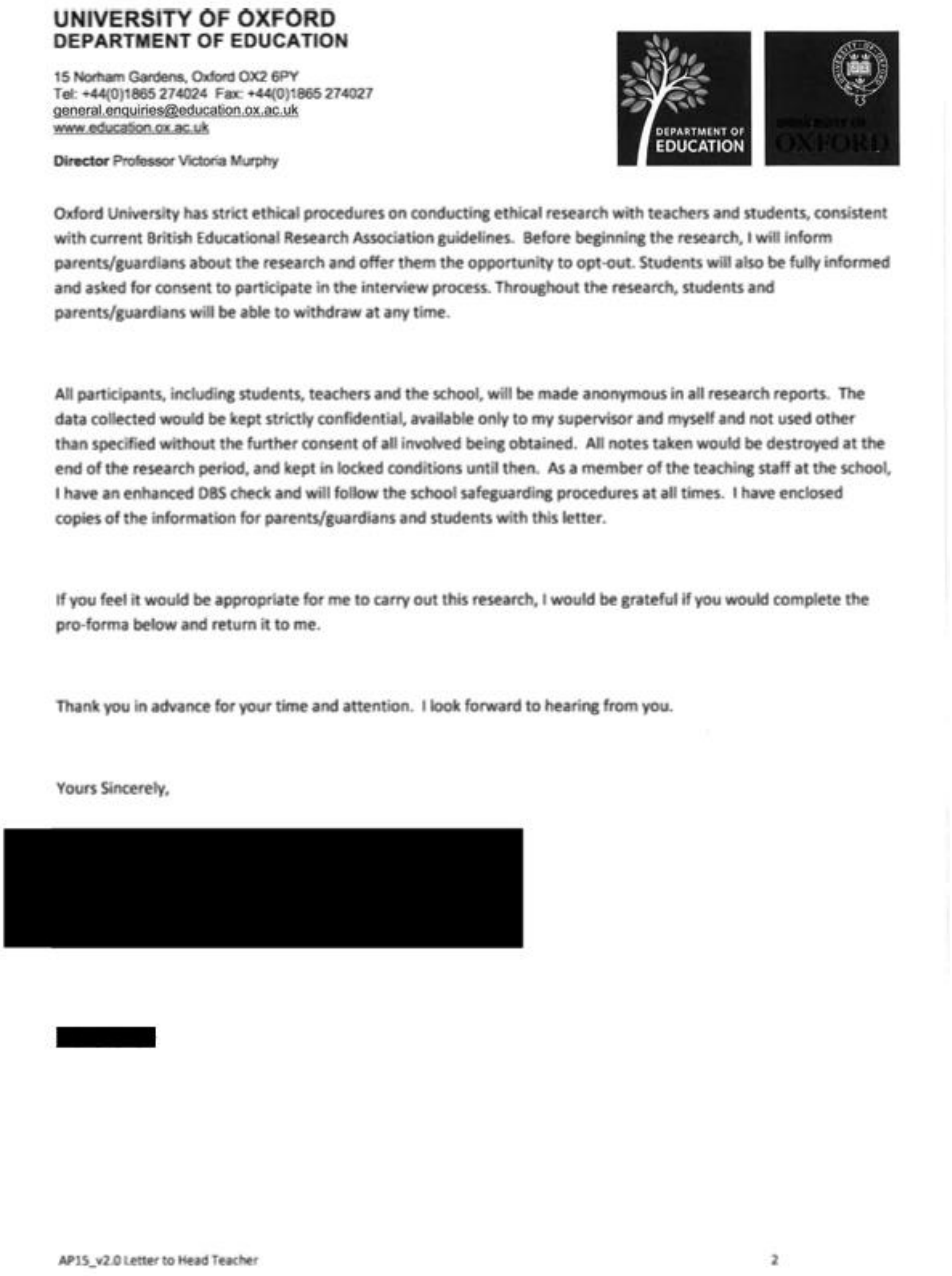
Ethics Approval Reference: EDUC_C1A_23_339

Dear [redacted]

I am writing to enquire about conducting some research in your school in the next academic year. I am a currently studying the MSc in Learning and Teaching at the University of Oxford, supervised by [redacted] In my research study, "It's all just a bit too much!" Exploring the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry. I will explore the research question: To what extent do taught interventions allow students to manage the cognitive demands associated with the mathematical areas of post-16 chemistry?

By participating in the research, the school would be contributing to research that will contribute to gaining a fuller understanding of what strategies students find useful to help them manage the cognitive demands of the mathematical areas of post-16 chemistry courses. These findings can then be used to better inform future teaching practice in this area.

The commitment from the school would be to allow me into carry out research with post-16 chemistry students over the course of the spring and summer terms from the beginning of term 2. To investigate the research questions, I will carry out pre-intervention and post-intervention assessment, where students will complete a number of questions/problems that vary in the amount of written information given and number of steps required to solve them. As students complete the questions, they rate their perceived difficulty of the question and their level of confidence in answering the question. Additionally, I will observe students and take notes; take photocopies of some of the students' written work; and interview students to gain insights into their perspectives, and take notes and audio record during these interviews.



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

Oxford University has strict ethical procedures on conducting ethical research with teachers and students, consistent with current British Educational Research Association guidelines. Before beginning the research, I will inform parents/guardians about the research and offer them the opportunity to opt-out. Students will also be fully informed and asked for consent to participate in the interview process. Throughout the research, students and parents/guardians will be able to withdraw at any time.

All participants, including students, teachers and the school, will be made anonymous in all research reports. The data collected would be kept strictly confidential, available only to my supervisor and myself and not used other than specified without the further consent of all involved being obtained. All notes taken would be destroyed at the end of the research period, and kept in locked conditions until then. As a member of the teaching staff at the school, I have an enhanced DBS check and will follow the school safeguarding procedures at all times. I have enclosed copies of the information for parents/guardians and students with this letter.

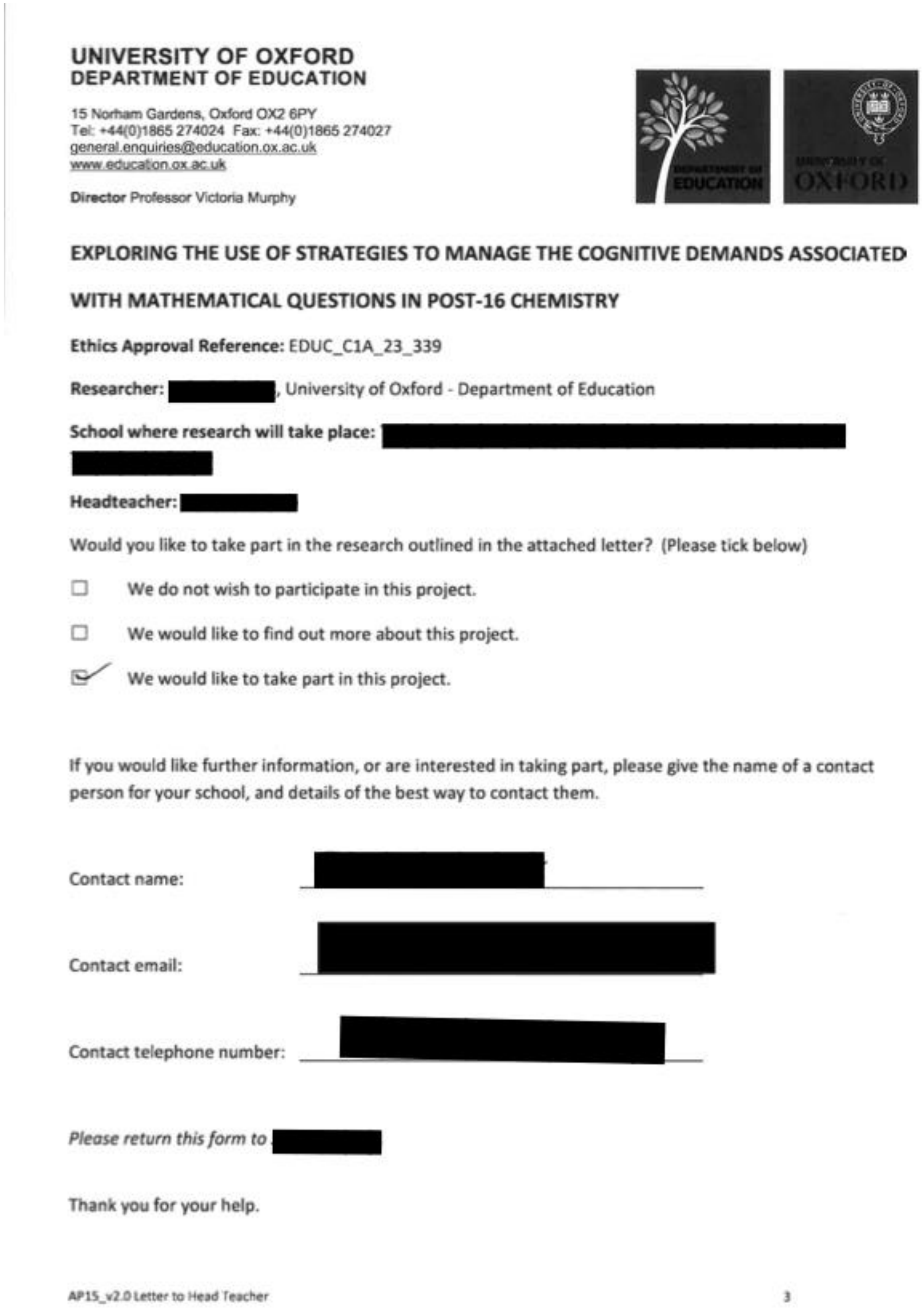
If you feel it would be appropriate for me to carry out this research, I would be grateful if you would complete the pro-forma below and return it to me.

Thank you in advance for your time and attention. I look forward to hearing from you.

Yours Sincerely,

[redacted]

[redacted]



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

## EXPLORING THE USE OF STRATEGIES TO MANAGE THE COGNITIVE DEMANDS ASSOCIATED WITH MATHEMATICAL QUESTIONS IN POST-16 CHEMISTRY

**Ethics Approval Reference:** EDUC_C1A_23_339

**Researcher:** [redacted], University of Oxford - Department of Education

**School where research will take place:** [redacted]

**Headteacher:** [redacted]

Would you like to take part in the research outlined in the attached letter? (Please tick below)

☐ We do not wish to participate in this project.

☐ We would like to find out more about this project.

☑ We would like to take part in this project.

If you would like further information, or are interested in taking part, please give the name of a contact person for your school, and details of the best way to contact them.

Contact name: [redacted]

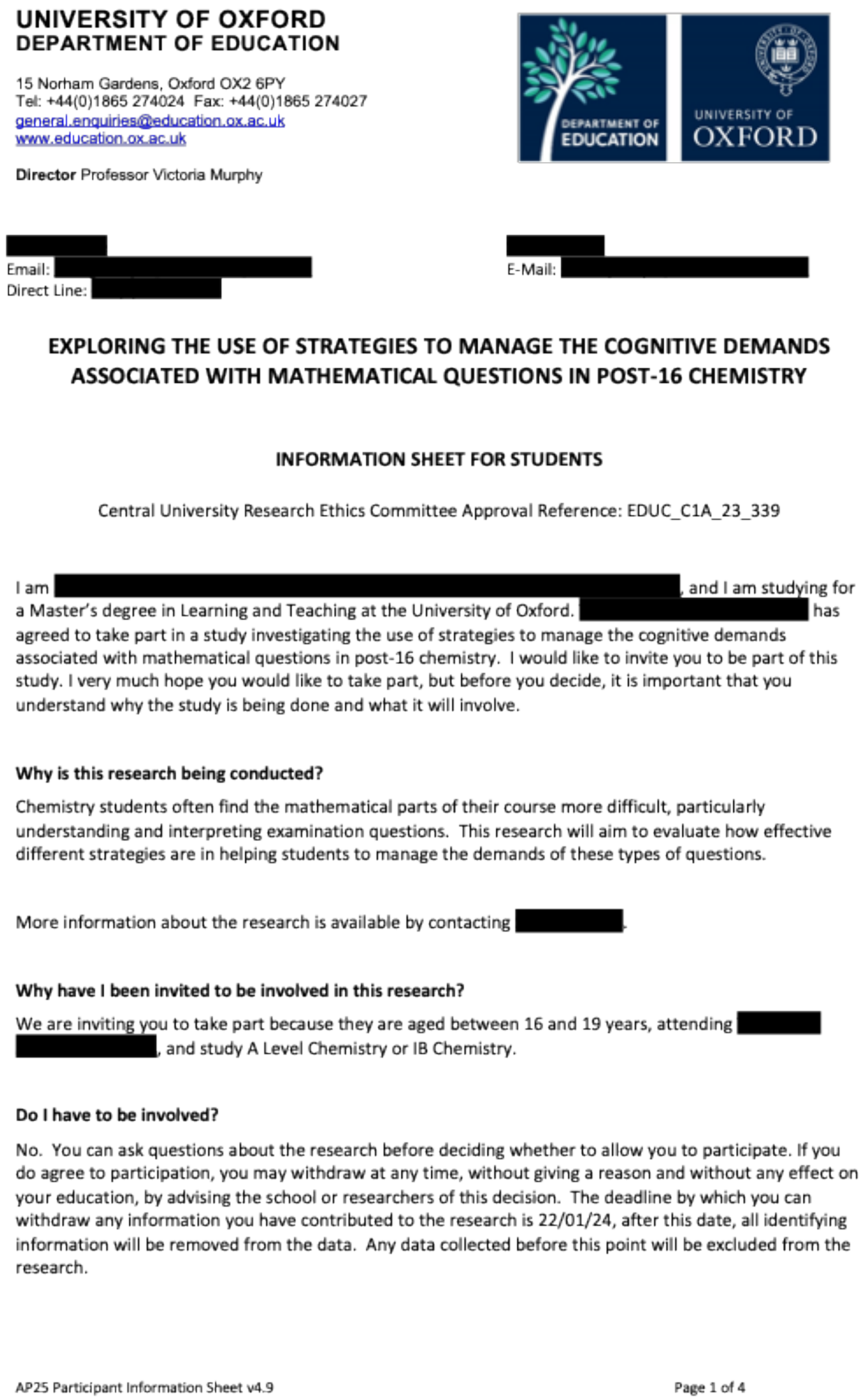
Contact email: [redacted]

Contact telephone number: [redacted]

*Please return this form to* [redacted]

Thank you for your help.

# Appendix 17: Student information sheet



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024  Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

[redacted]
Email: [redacted]
Direct Line: [redacted]

[redacted]
E-Mail: [redacted]

## EXPLORING THE USE OF STRATEGIES TO MANAGE THE COGNITIVE DEMANDS ASSOCIATED WITH MATHEMATICAL QUESTIONS IN POST-16 CHEMISTRY

### INFORMATION SHEET FOR STUDENTS

Central University Research Ethics Committee Approval Reference: EDUC_C1A_23_339

I am [redacted], and I am studying for a Master's degree in Learning and Teaching at the University of Oxford. [redacted] has agreed to take part in a study investigating the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry. I would like to invite you to be part of this study. I very much hope you would like to take part, but before you decide, it is important that you understand why the study is being done and what it will involve.

**Why is this research being conducted?**

Chemistry students often find the mathematical parts of their course more difficult, particularly understanding and interpreting examination questions. This research will aim to evaluate how effective different strategies are in helping students to manage the demands of these types of questions.

More information about the research is available by contacting [redacted].

**Why have I been invited to be involved in this research?**

We are inviting you to take part because they are aged between 16 and 19 years, attending [redacted], and study A Level Chemistry or IB Chemistry.

**Do I have to be involved?**

No. You can ask questions about the research before deciding whether to allow you to participate. If you do agree to participation, you may withdraw at any time, without giving a reason and without any effect on your education, by advising the school or researchers of this decision. The deadline by which you can withdraw any information you have contributed to the research is 22/01/24, after this date, all identifying information will be removed from the data. Any data collected before this point will be excluded from the research.

If you are not involved in the research, you will still be taught the strategies to manage the demands of mathematical questions in chemistry, but no data will be collected from you.

**What will happen if I take part?**

- The research will take place between January and September 2024, and data collection will take place between January 2024 and May 2024.
- You will be asked to complete a series of mathematical chemistry questions and rate them in terms of their difficulty. This will take place during the school day.
- You will be taught different strategies to help them approach mathematical questions in chemistry and asked to evaluate how useful you found them. This will take place during the school day.
- You will be asked to take part in a paper questionnaire, that will take place during the school day.
- You may be asked to take part in small group interviews to gain a better understanding of their views and experiences. This will take place during the school day and last approximately 30 minutes. These may be audio recorded.
- Examples of your work from the current academic year, and direct quotations from interview and questionnaire responses may also be used as illustrative examples.
- Existing assessment data will be used to compare achievement in the mathematical areas of chemistry in relation to other areas of chemistry (subject to a data sharing agreement).
- Any data collected in the research will have identifying information removed as soon as possible after collection.

**What are the possible disadvantages and risks in taking part?**

There are no intended or perceived risks of being involved in the research project. You will not be identifiable from the data collected.

**Are there any benefits in taking part?**

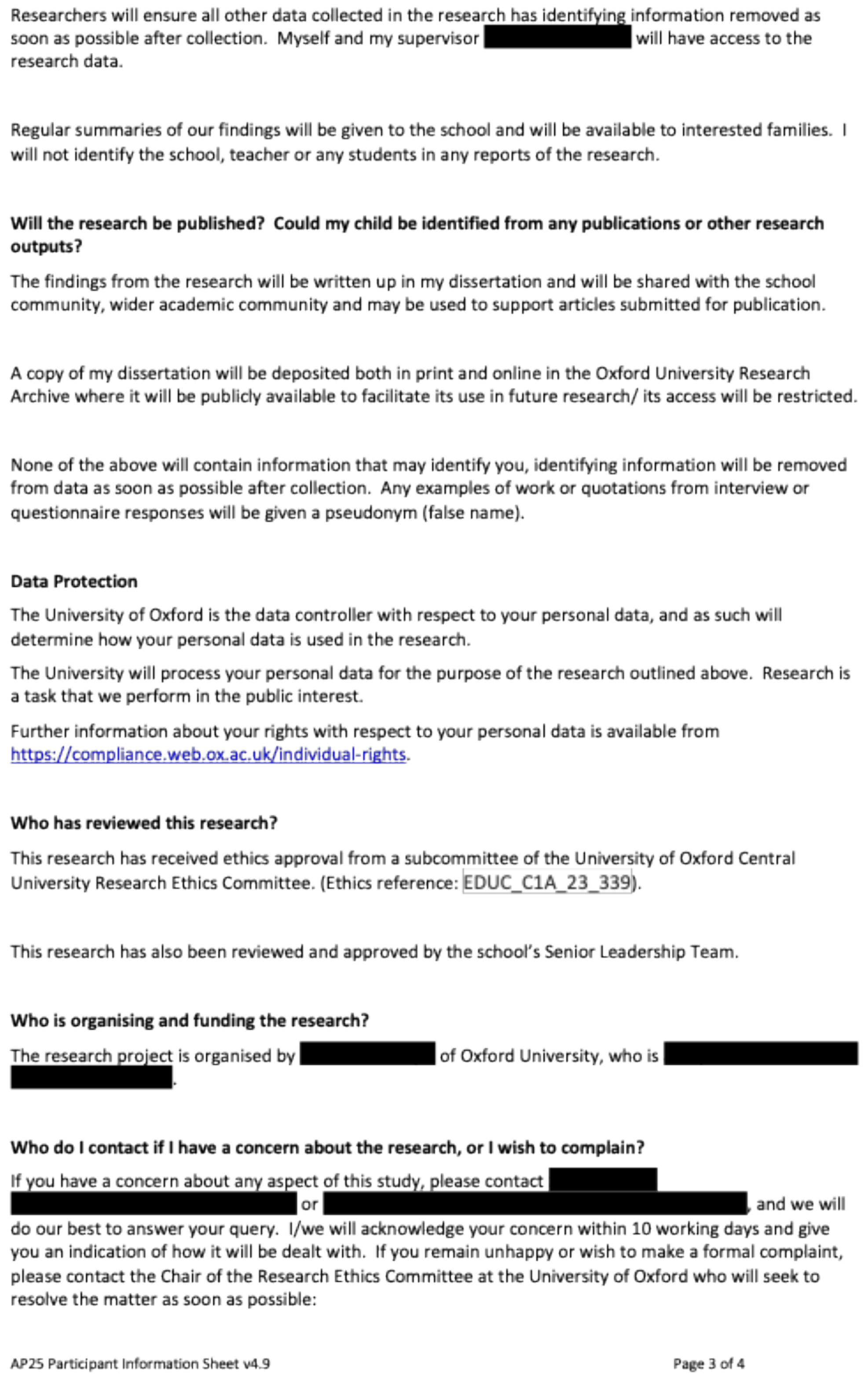
By taking part in this research you will be contributing to our understanding of what methods are more effective in managing the demands of answering mathematical questions in chemistry.

**What information will be collected and why is the collection of this information relevant for achieving the research objectives?**

The personal data that is being collected in this study, will be you name, in addition to whether you and your parent(s)/guardian(s) have given you permission to take part. Examples of work from the current academic year, and direct quotations from interview and questionnaire responses may also be used as illustrative examples. Interviews may be audio recorded. Existing assessment data will be used to compare achievement in the mathematical areas of chemistry in relation to other areas of chemistry (subject to a data sharing agreement). Once the data collection stage is completed, all research data will be given a pseudonym (false name), and you will not be identifiable.

Identifiable data (including consent forms) will be stored securely as scanned copies on the University of Oxford's OneDrive for Business, paper copies will be shredded. Other research data will be stored for 3 years after publication or public release of the work of the research.

Opt-out forms will be retained by the school for the duration of the research, and for as long as the school determines appropriate after research activities have concluded at the school.



Researchers will ensure all other data collected in the research has identifying information removed as soon as possible after collection. Myself and my supervisor [redacted] will have access to the research data.

Regular summaries of our findings will be given to the school and will be available to interested families. I will not identify the school, teacher or any students in any reports of the research.

**Will the research be published? Could my child be identified from any publications or other research outputs?**

The findings from the research will be written up in my dissertation and will be shared with the school community, wider academic community and may be used to support articles submitted for publication.

A copy of my dissertation will be deposited both in print and online in the Oxford University Research Archive where it will be publicly available to facilitate its use in future research/ its access will be restricted.

None of the above will contain information that may identify you, identifying information will be removed from data as soon as possible after collection. Any examples of work or quotations from interview or questionnaire responses will be given a pseudonym (false name).

**Data Protection**

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research.

The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest.

Further information about your rights with respect to your personal data is available from https://compliance.web.ox.ac.uk/individual-rights.

**Who has reviewed this research?**

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference: EDUC_C1A_23_339).

This research has also been reviewed and approved by the school's Senior Leadership Team.

**Who is organising and funding the research?**

The research project is organised by [redacted] of Oxford University, who is [redacted].

**Who do I contact if I have a concern about the research, or I wish to complain?**

If you have a concern about any aspect of this study, please contact [redacted] [redacted] or [redacted], and we will do our best to answer your query. I/we will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Chair, Social Sciences & Humanities Interdivisional Research Ethics Committee; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB

### What should I do next?

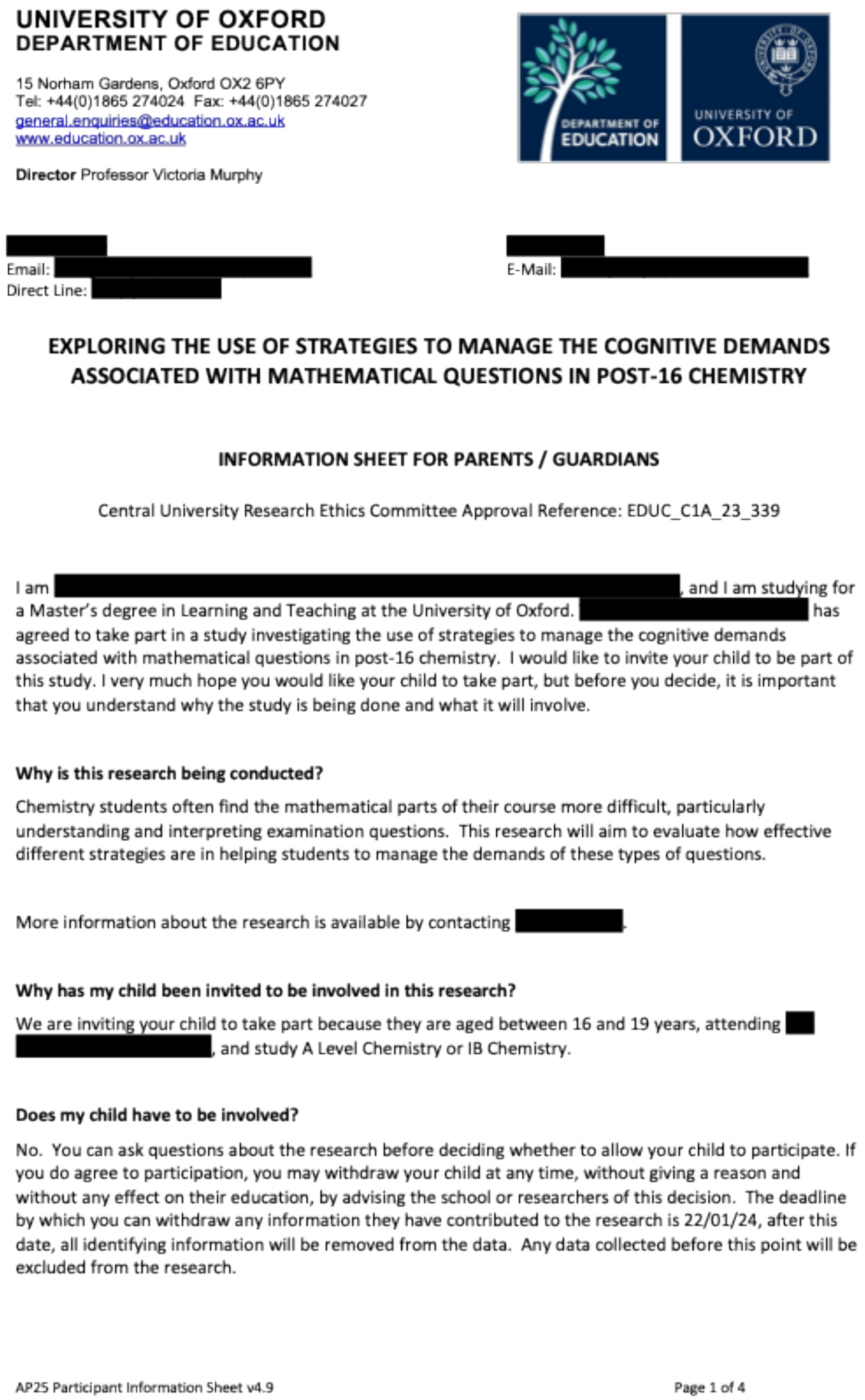
If you would not like to take part in this study, please complete the opt-out form and return it to [redacted] [redacted]. Please remember that you may withdraw at any time, without giving a reason, by notifying the researcher.

### Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

[redacted]
E-Mail: [redacted]

# Appendix 18: Parent/Guardian information sheet



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

[redacted]
Email: [redacted]
Direct Line: [redacted]

[redacted]
E-Mail: [redacted]

## EXPLORING THE USE OF STRATEGIES TO MANAGE THE COGNITIVE DEMANDS ASSOCIATED WITH MATHEMATICAL QUESTIONS IN POST-16 CHEMISTRY

### INFORMATION SHEET FOR PARENTS / GUARDIANS

Central University Research Ethics Committee Approval Reference: EDUC_C1A_23_339

I am [redacted], and I am studying for a Master’s degree in Learning and Teaching at the University of Oxford. [redacted] has agreed to take part in a study investigating the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry. I would like to invite your child to be part of this study. I very much hope you would like your child to take part, but before you decide, it is important that you understand why the study is being done and what it will involve.

**Why is this research being conducted?**

Chemistry students often find the mathematical parts of their course more difficult, particularly understanding and interpreting examination questions. This research will aim to evaluate how effective different strategies are in helping students to manage the demands of these types of questions.

More information about the research is available by contacting [redacted].

**Why has my child been invited to be involved in this research?**

We are inviting your child to take part because they are aged between 16 and 19 years, attending [redacted], and study A Level Chemistry or IB Chemistry.

**Does my child have to be involved?**

No. You can ask questions about the research before deciding whether to allow your child to participate. If you do agree to participation, you may withdraw your child at any time, without giving a reason and without any effect on their education, by advising the school or researchers of this decision. The deadline by which you can withdraw any information they have contributed to the research is 22/01/24, after this date, all identifying information will be removed from the data. Any data collected before this point will be excluded from the research.

If you are not involved in the research, you will still be taught the strategies to manage the demands of mathematical questions in chemistry, but no data will be collected from you.

**What will happen if I take part?**

- The research will take place between January and September 2024, and data collection will take place between January 2024 and May 2024.
- You will be asked to complete a series of mathematical chemistry questions and rate them in terms of their difficulty. This will take place during the school day.
- You will be taught different strategies to help them approach mathematical questions in chemistry and asked to evaluate how useful you found them. This will take place during the school day.
- You will be asked to take part in a paper questionnaire, that will take place during the school day.
- You may be asked to take part in small group interviews to gain a better understanding of their views and experiences. This will take place during the school day and last approximately 30 minutes. These may be audio recorded.
- Examples of your work from the current academic year, and direct quotations from interview and questionnaire responses may also be used as illustrative examples.
- Existing assessment data will be used to compare achievement in the mathematical areas of chemistry in relation to other areas of chemistry (subject to a data sharing agreement).
- Any data collected in the research will have identifying information removed as soon as possible after collection.

**What are the possible disadvantages and risks in taking part?**

There are no intended or perceived risks of being involved in the research project. You will not be identifiable from the data collected.

**Are there any benefits in taking part?**

By taking part in this research you will be contributing to our understanding of what methods are more effective in managing the demands of answering mathematical questions in chemistry.

**What information will be collected and why is the collection of this information relevant for achieving the research objectives?**

The personal data that is being collected in this study, will be you name, in addition to whether you and your parent(s)/guardian(s) have given you permission to take part. Examples of work from the current academic year, and direct quotations from interview and questionnaire responses may also be used as illustrative examples. Interviews may be audio recorded. Existing assessment data will be used to compare achievement in the mathematical areas of chemistry in relation to other areas of chemistry (subject to a data sharing agreement). Once the data collection stage is completed, all research data will be given a pseudonym (false name), and you will not be identifiable.

Identifiable data (including consent forms) will be stored securely as scanned copies on the University of Oxford's OneDrive for Business, paper copies will be shredded. Other research data will be stored for 3 years after publication or public release of the work of the research.

Opt-out forms will be retained by the school for the duration of the research, and for as long as the school determines appropriate after research activities have concluded at the school.

Researchers will ensure all other data collected in the research has identifying information removed as soon as possible after collection. Myself and my supervisor [redacted] will have access to the research data.

Regular summaries of our findings will be given to the school and will be available to interested families. I will not identify the school, teacher or any students in any reports of the research.

**Will the research be published? Could my child be identified from any publications or other research outputs?**

The findings from the research will be written up in my dissertation and will be shared with the school community, wider academic community and may be used to support articles submitted for publication.

A copy of my dissertation will be deposited both in print and online in the Oxford University Research Archive where it will be publicly available to facilitate its use in future research.

None of the above will contain information that may identify your child, identifying information will be removed from data as soon as possible after collection. Any examples of students' work or quotations from interview or questionnaire responses will be given a pseudonym (false name).

**Data Protection**

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your child's personal data is used in the research.

The University will process your child's personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest.

Further information about your rights with respect to your personal data is available from https://compliance.web.ox.ac.uk/individual-rights.

**Who has reviewed this research?**

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference: EDUC_C1A_23_339).

This research has also been reviewed and approved by the school's Senior Leadership Team.

**Who is organising and funding the research?**

The research project is organised by [redacted] of Oxford University, who is [redacted].

**Who do I contact if I have a concern about the research, or I wish to complain?**

If you have a concern about any aspect of this study, please contact [redacted] or [redacted], and we will do our best to answer your query. I/we will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Chair, Social Sciences & Humanities Interdivisional Research Ethics Committee; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB

**What should I do next?**

If you would not like your child to take part in this study, please ask them to complete the opt-out form and return it to [redacted]. Please remember that your child may withdraw at any time, without giving a reason, by notifying the researcher.

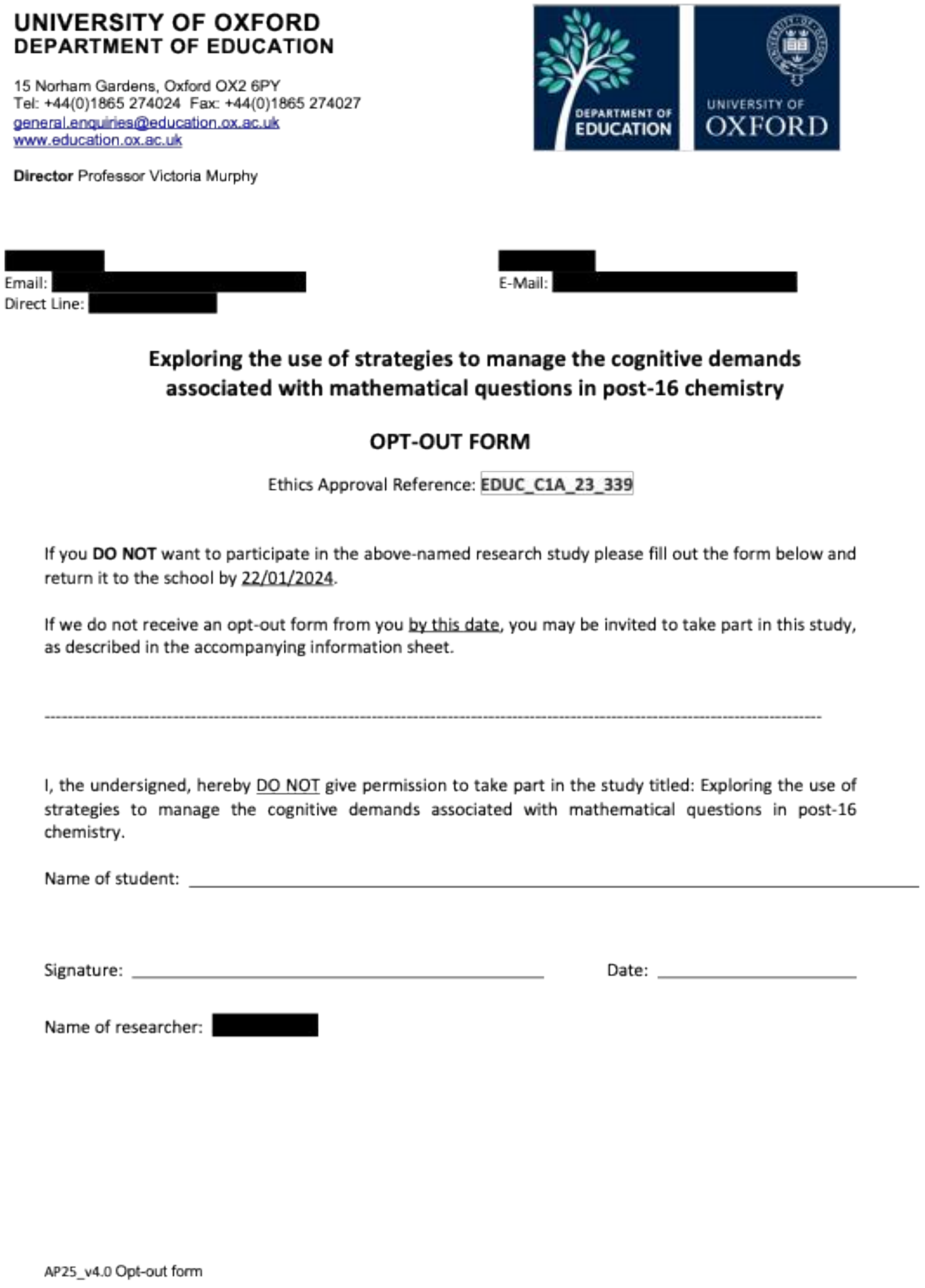
**Further Information and Contact Details**

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

[redacted]
E-Mail: [redacted]

[Return to TOC]

# Appendix 19: Study opt-out form



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

[redacted]
Email: [redacted]
Direct Line: [redacted]

[redacted]
E-Mail: [redacted]

**Exploring the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry**

**OPT-OUT FORM**

Ethics Approval Reference: EDUC_C1A_23_339

If you **DO NOT** want to participate in the above-named research study please fill out the form below and return it to the school by 22/01/2024.

If we do not receive an opt-out form from you by this date, you may be invited to take part in this study, as described in the accompanying information sheet.

-----------------------------------------------------------------------------------------------------------

I, the undersigned, hereby DO NOT give permission to take part in the study titled: Exploring the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry.

Name of student: ________________________________________

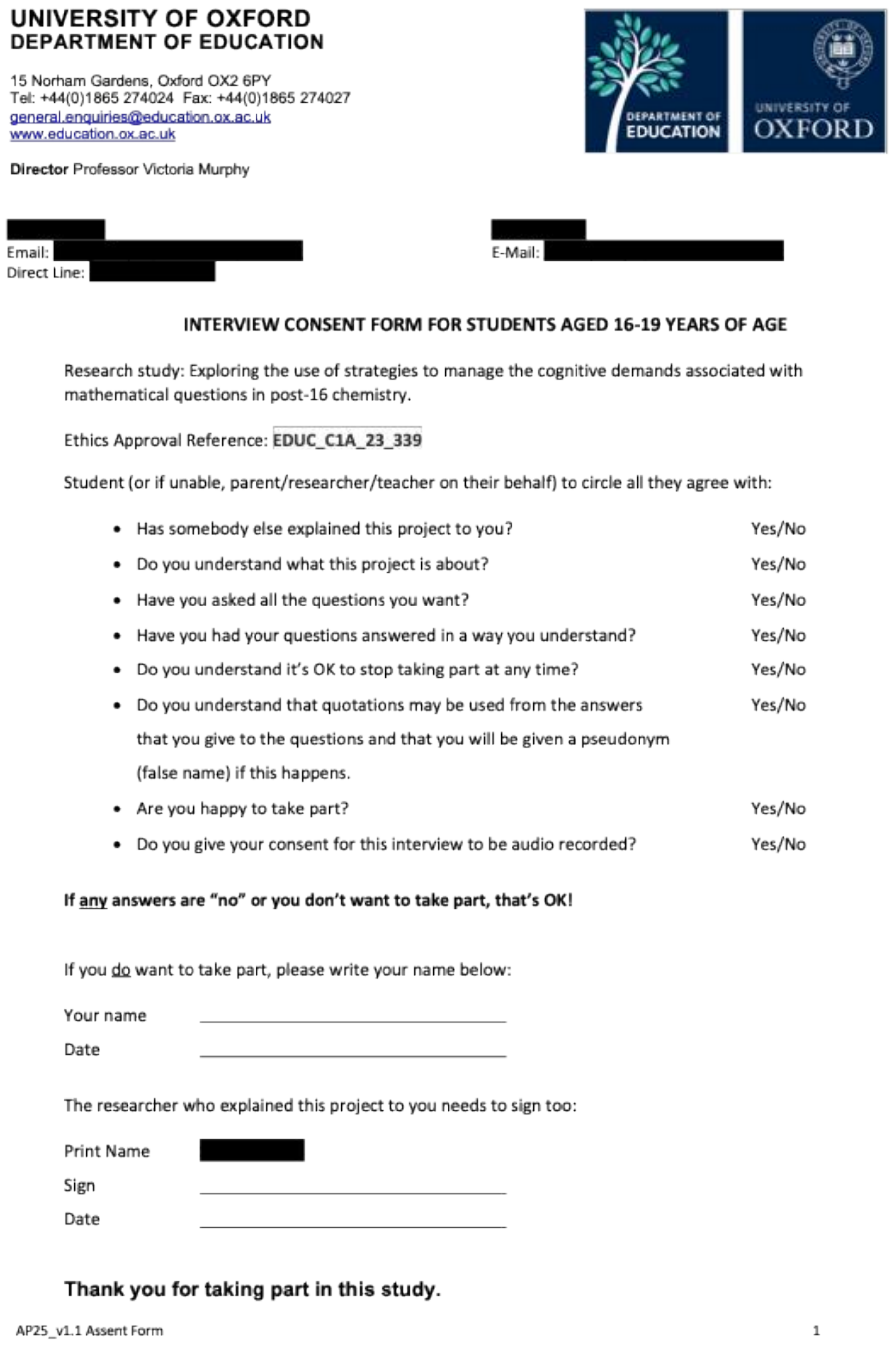
Signature: ____________________________ Date: ______________

Name of researcher: [redacted]

AP25_v4.0 Opt-out form

[Return to TOC]

# Appendix 20: Interview consent form



**UNIVERSITY OF OXFORD**
**DEPARTMENT OF EDUCATION**

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

**Director** Professor Victoria Murphy

[redacted]
Email: [redacted]
Direct Line: [redacted]

[redacted]
E-Mail: [redacted]

**INTERVIEW CONSENT FORM FOR STUDENTS AGED 16-19 YEARS OF AGE**

Research study: Exploring the use of strategies to manage the cognitive demands associated with mathematical questions in post-16 chemistry.

Ethics Approval Reference: **EDUC_C1A_23_339**

Student (or if unable, parent/researcher/teacher on their behalf) to circle all they agree with:

| | |
|---|---|
| • Has somebody else explained this project to you? | Yes/No |
| • Do you understand what this project is about? | Yes/No |
| • Have you asked all the questions you want? | Yes/No |
| • Have you had your questions answered in a way you understand? | Yes/No |
| • Do you understand it's OK to stop taking part at any time? | Yes/No |
| • Do you understand that quotations may be used from the answers that you give to the questions and that you will be given a pseudonym (false name) if this happens. | Yes/No |
| • Are you happy to take part? | Yes/No |
| • Do you give your consent for this interview to be audio recorded? | Yes/No |

**If any answers are "no" or you don't want to take part, that's OK!**

If you do want to take part, please write your name below:

Your name ______________________________

Date ______________________________

The researcher who explained this project to you needs to sign too:

Print Name [redacted]

Sign ______________________________

Date ______________________________

**Thank you for taking part in this study.**

AP25_v1.1 Assent Form 1