2019

Statistical analysis of the effect of lesson and learning styles on the learning and retention of information of secondary school chemistry students

Barry, L.

Barry, L. (2019) 'Statistical analysis of the effect of lesson and learning styles on the learning and retention of information of secondary school chemistry students', The Plymouth Student Scientist, 12(1), p. 305-344.
http://hdl.handle.net/10026.1/14690

The Plymouth Student Scientist
University of Plymouth

Statistical analysis of the effect of lesson and learning styles on the learning and retention of information of secondary school chemistry students

Lauren Barry

Project Advisor: Dr Roy Lowry, Earth and Environmental Sciences, Plymouth University, Drake Circus, Plymouth, PL4 8AA

Abstract

The study is an interdisciplinary study as it links to both chemistry and also social science. The project investigated the learning and recall of information of year 9 students at Richard Lander School. In total 85 students participated in this study. Two top set ability groups and two bottom set ability groups were selected by the school to participate in either a non-practical or practical lesson. The students all completed a pre-lesson test to assess their knowledge of the lesson subject (chromatography) prior to the lesson. This was used to establish that the top/bottom set ability was the same for each group (1) regardless of whether or not they were participating in the practical or non-practical lesson. The students also completed a post lesson, retention of information and VARK test. The results of these tests were analysed for statistically significant differences using SPSS software to perform relevant t-tests.

No significant statistical difference was found in the learning of information for the top set ability students as a result of the lesson style (practical vs non-practical). A significant statistical difference was found for the bottom set as a result of the lesson style (practical vs non-practical) with the practical lesson style yielding higher test scores. From these results it was concluded that practical work enhances the learning of information for low ability students, but provides little benefit for high ability students. The results of the VARK comparisons showed very little conclusive evidence, however students with a multimodal learning approach appear to learn better than any other students regardless of whether the lesson is practical or not. Both the individual learning style of the students and the style of lesson were found to have the greatest effect on students from the bottom set.

Introduction

The study is a combination of both social science ideas and methodologies and chemistry aimed at investigating how the presentation of ideas in a lesson can affect the students learning. The study of science is widely deemed as an important component of education. It is said to influence the way students think, learn, problem solve and evaluate information across a wide range of subjects and often in everyday life (2).

This particular study focuses on how beneficial practical work is to learners of different abilities and learning styles. Undoubtedly practical work has been proven to improve the practical skill levels of students (3). However, as the majority of secondary school students do not consider a career in chemistry, the cultivation of these skills is not necessarily as important as the understanding of the subject. With respect to the learning of chemistry many studies demonstrate that there is still a place for practical work and that as students learn in a variety of ways the removal of chemistry practicals could hinder the ability of some students to understand the content of the curriculum. This study hopes to investigate how beneficial practical lessons are and to whom they are most beneficial to.

The work of Piaget

Chemistry is taught as part of the core subject of science in secondary schools. The teaching of science in secondary schools has been the subject of many studies within teaching. One key idea in the learning of science is “constructivism” (4), which was pioneered by Jean Piaget. Piaget believed that children were not less intelligent but that they thought differently from adults and used active learning processes to make observations and learn about the world. In this way children were actual thinking like a scientists (5). Piaget believed in four stages of development, The Sensorimotor Stage (birth -2 years), The Preoperational Stage (2-7 years), The Concrete Operational Stage (7-11years) and The Formal Operational Stage (12 years and up)(6). Piaget’s work was and is still considered to revolutionary, Georg W. Oesterdiekhoff states that, “Jean Piaget (1896–1980) is known as the greatest scholar in the entire history of developmental psychology and possibly scientifically more relevant than Freud, Binet, Lewin, Skinner, Eysenck,” (7).

Although the core beliefs of Piaget’s work have been largely well received there have been questions raised about his stages not being accurate as some children have been found to develop earlier than Piaget suggested (8). Other inconsistencies were drawn from Piaget’s belief that children learnt by physically exploring their world, this proved to not always be true as children whose physical abilities had been impaired (due to being born with either missing or poorly functioning limbs) where seen to have normal cognitive development (9) home. Although Piaget’s work has been criticised over the years it still has a huge amount of relevance within teaching and learning, particularly across the sciences. This is due to science having practical applications and in lesson time students are given an opportunity to learn from experiences they physically witness.

The process of chromatography

Chromatography is a key analytical method used within chemistry. For the purposes of this experiment a very simple model was used, but it does have much more complicated and wider reaching applications. The chromatography method used was pen ink separation via chromatography paper. This method of separation was subject to a large volume of work by Friedrich Goppelsroeder who has been referred to as the “pioneer of paper chromatography" (10). Goppelsroeder's work centred around the separation of commercially used dyes to determine the components of pure dyes. An example of this was his separation of picric acid from curcuma dye and from indigo (11). Teaching chromatography in secondary education is relevant as students wishing to continue to study will encounter chromatography not only in chemistry applications, but also in many other areas of study including but not limited to biology, forensics and pharmacology. The principles taught at secondary school set the basis of understanding for the widely used practices of Gas chromatography (GC) and High-Performance Liquid Chromatography (HPLC).

A key principle of chromatography is the idea of retention/retardation factors (R_f). If an experiment is exactly replicated the R_f values of a compound will remain the same. If different mobile and or stationary phases are used then the compound may be seen to have a different R_f value (12). The method for calculating the R_f can be seen below in equation 1.

Equation 1

$$R_f = \frac{\textit{The migration of distance of compound}}{\textit{The migration distance of the solvent front}}$$

The units for distance must be the same on the numerator and the denominator. The solvent front is the distance the mobile phase has moved. Both distances are measured from the baseline.

The Importance of teaching

The British educational system has the ability to cover a wide range of educational needs, this is particularly apparent within in science as there is scope for both written work and practical learning. Practical experiments give students the opportunity to learn from what they see and do during lesson time.

Even with the wide range of mediums available to teach sciences students can feel disconnected with the subject matter, this has been demonstrated in a number of studies. In findings from a questionnaire by the Oxford, Cambridge and RSA examinations board (OCR) it can be seen that chemistry and physics faired considerably worse than biology,

"51% of teenagers think science lessons are boring, confusing or difficult" (13).

Students were also asked if they would study science if it became non-compulsory, to which the reply was,

"45% said they would take biology, 32% chemistry, 29% physics and 19% combined science"(13).

In a separate study by the Royal Society of Chemistry (RSC) there were a lot of positive findings about the general public's feelings on chemists (14). Unfortunately, although the public believe chemists make a difference to their world and consider them to be honest, approachable and enthusiastic 28% of the public believe them to also be boring. This backs up the OCR findings that 51% of students find science as a whole boring and as chemistry is a subject within science it is highly likely that this statement also applies to chemistry.

As much of the general public's knowledge of chemistry will come from their perception formed during their school years it is fair to assume that the widely held belief that chemistry is boring starts at school. It can therefore be construed that chemistry lessons are or at least have been unable to stimulate and engage students. In a 1988 study of students aged 17 very few students seemed to have enjoyed or engaged with the science lessons they were taught at school (15). An important question for consideration here is "what can be done to improve public perceptions of science?"

Unfortunately, currently the best way of assessing how engaged students are is by their exam results. There are obvious flaws in this process (mainly students could be extremely engaged and excited by a subject but not necessarily take on the relevant information needed to pass an exam or students have found the lesson appealing but due to other learning needs perform poorly exams). The alternative is to survey the students opinion for their opinion of the lesson and how engaging they found it (16), however this tells us little about what content has actually been learnt.

With this current system in mind research has identified that there is a strong correlation with the quality of teaching and the results attained by students. A key argument for this is that the countries in which students perform the best have the best teachers as they recruit the very top percent of their graduates for teaching. For example, South Korea recruits the top 5% of graduates (17) and in 2014 was ranked the top educational system in the world (18). Finland employs the top 10% of graduates (19) and is ranked top 5. Not only is there a similarity in that both education systems perform extremely well and have a proportion of the top graduates there is also a decrease in the quality of the education system as the quality of the graduate is diluted. In 2015 the UK's most successful subject (science) was placed 15th (19) in the world standings. The government has many strategies to improve this score, one such strategy is to reduce the teacher work load to allow for teacher to, "concentrate on teaching and their own development"(20, 21).This would hopefully improve the quality of teaching, however the majority of the policies that the government has proposed appear to focus on the infrastructure of the schools as opposed to the teaching (22). In a separate report in 2013 the differences between the English education system and those of the highest achieving systems at key stage 3 were highlighted as,

"Compared to the average for other countries, England has younger teachers and head teachers, fewer modern language teachers, more autonomous schools, significantly greater numbers of teaching assistants and administrative and managerial staff in schools, and teachers reporting longer total working hours on average but not face-to-face teaching hours" (22).

The interesting point to note here is the large amount of time and money invested in employing managerial and administrative staff, whilst also having teaching assistants taking on more work to a higher standard than they may be comfortable with.

Relevant work to the study

Origins of VARK

VARK stands for Visual, Aural, Read/write, and Kinaesthetic learning styles. The development of the VARK learning styles and questionnaire (23) was started by Neil Fleming, the first article he wrote alluding to learning styles in this sense was “Not Another inventory” (24). In the conclusion of the article Fleming suggest that some of the many positive outcomes from this type of learning and the questionnaire are empowering students to reflect upon their sensory preferences and to modify their study methods accordingly. One of Fleming’s main inspirations appears to be Dunn and Dunn (25). The work by Dunn and Dunn centred on the concept that by adapting lesson styles to suit individual needs the students learnt information more quickly and were less prone to disruptive behaviour (26). The validity of the method used to create this concept has been called into question (27).

As a result of his research Neil Fleming went on to produce the VARK questionnaire. The questionnaire was designed to highlight each student’s favoured learning style so that the learner would be aware of methods of learning which could benefit them and also teachers would be able to tailor lessons to suit individual learning needs.

Recent work on evaluating the use of learning styles has shown that in 2014 90% of teachers in the UK believed that students learn better when there learning styles are taken into consideration (28). However, the reliability of the research on learning styles has often been called in to question(29, 30). This is due to the bias present in Neil Flemings work, particularly due to the fact that he is self-published and both Neil Fleming and Dunn and Dunn can be seen in the work to reference work that they themselves have undertaken in order to validate their own methods. Criticism has also been drawn on the misconceptions of learning styles as it can be seen that while learners may have a specific learning style this will in fact vary from task to task (29).

The impact of practical work

Experimentation is core part of the majority of scientific research and has therefore been used to teach science to students in order for the students to gain an understanding of both analytical thinking and practical skills.

The necessity of the use of practical working within science based education has been upheld by the national science teachers association (NSTA) which actively endorses the use of practical work within lessons (31).

In a study conduct by Abrahams and Millar students aged 11-16 participated in 25 practical lessons to assess whether or not the practical lessons helped the students to understand relevant scientific theories. The study concluded that the expectations of the teachers were that by visualising a scientific experiment the students would then be able to uses this experience and apply it to their theory work. However, for the majority of the students this was not the case and without prompting, parallels between the work in the lab and the area of study were not made. The teachers were seen to expect the learning of theory concepts to emerge organically of their own accord during the practical processes. The study concluded that there was very little value to practical work pursued in this manner (32).

Social science concepts

There are many approaches to social science studies of the four broad approaches outlined by Porta and Keating this study has been designed to incorporate a post-positivist approach. The bases of which is outlined in table 1.

Table 1: Issues for consideration when applying a post-positivist approach

Issue	Post-positivist ideas
Does social reality exist?	Objective, critical realism
Is reality knowable?	Yes, but not easy to capture
Relationship between the scholar and his/her object	Knowledge is influenced by the scholar; deductive procedures
Forms of knowledge	Probabilistic law

The core principle that post-positivist research works from is that social science is similar to other physical sciences and as such the researcher can describe and analyse it (33). The post-positivist approach is carried out in natural conditions (34) (e.g. the classroom as opposed to laboratory conditions).

This study was modelled on concepts put forward by Heritier in particular the acquisition of “generalizable knowledge the respect to the phenomena in question” (33).The phenomena in question is that of memory and retention of information of KS3 chemistry students. The experiment was developed with reference to the “casual explanation”, which is used to describe studies that are based on a number of assumptions about the way the world works. For a study to be considered casual the cause in question must generate an effect (35). In the case of this study the cause under consideration is practical work and the effect is the learning and retention of information. Studies must follow set rules on reliability and replicability which refers to any steps taken being possible to replicate and the outcomes reassessed. In order to evaluate the validity of the study the effectiveness of the measurements used in the study to deliver relevant results is analysed (33). Prognostication must also be considered by using existing studies to draw conclusions. When designing this method, the reliability and replicability, validity and prognostication were all considered.

The necessity of this study

<u>The current academic position</u>

The current position on the use of practical work is that it is beneficial to all students as it teaches them a set of skills not available through solely participating in non- practical lessons. A study which compiled information on the usefulness of laboratory work within in scientific learning found that,

“the assumption that laboratory experiences help students understand materials, phenomena, concepts, models, and relationships, almost independent of the nature of the laboratory experience, continues to be widespread in spite of sparse data from carefully designed and conducted studies”(36).

Studies also concluded that learning styles have a great influence on the understanding and retention of information and that without practical lessons students who learn best visually and kinaesthetically are not given a chance to excel.
Although there are a large number of studies on the relevance of practical work within science there are also many gaps partly due to a huge range in external factors, whether this is individual students, teachers, facilities, schools or the syllabus.

The gap in the research

There is a gap in the current work where although learning styles are often assessed via a questionnaire the effects of different lesson styles on students learning and the compatibility of the lesson style with the student's regards learning style are not compared. Another reason for performing this study is that although the learning and retention of information for students participating in practical lessons has been compared to non-practical lessons there is variation on the non-practical lesson style and very few studies of this nature have been carried out in the UK.

Aims and objectives

Aim
The aim of this project was to investigate the effect of lesson styles and learning styles on the learning and retention of information
Objectives:

- Identify a practical suitable for participation and non-practical explanation.
- Trial the chosen practical in the university lab prior to incorporating it in lessons.
- Carry out non-practical and practical style lesson with secondary school students of the same ability.
- Identify a method for testing learning and recall of the lesson.
- Use statistics to see if there is a significant difference in the learning and retention. of information with regards to lesson style and learner style.

Materials and Methods

Ethical approval

Prior to beginning any work with the students' ethical approval was sought and granted from the University of Plymouth. The experiment was conducted within the stipulations of the ethical approval guidelines.

Study groups

The study groups used were selected by the school where the experiment took place. All the study groups were part of year 9 as this year was deemed to have more time available as they were not yet undergoing exam preparation for GCSE's. The lower years were discounted as their practical experience was limited. The students all have two science lessons a week in which they are taught biology, physics and chemistry. The chosen groups had varying abilities, two of the groups were top sets of A/B ability (9L1, 9L2) and two sets were bottom sets of very low ability with some learning difficulties (9L5,9L6). From these groups 9R6 and 9L2 were selected to do the non-practical lesson due to timetabling.

As all the lessons had to be taught by the same teacher to ensure they were as similar as possible. It was not possible to teach every lesson on the same day. The lessons were taught at the times shown in table 2.

Table 2: Timetabling of study groups

Group	**Lesson type**	**Class set**	**Date**		
			4/12/17	8/12/17	11/12/17
			Period		
9R5	Practical	Bottom	3		3
9L1	Practical	Top	5		1
9R6	Non-practical	Bottom		2	
9L2	Non-practical	Top		3	

As shown in the timetable it was necessary for the practical groups to have two lessons in order to cover the work as the set up and carrying out of the practical took up lesson time.
To assess long term retention of information all the groups undertook the same test on the week beginning 26th March.

Subject of study

The subject taught was chromatography; this was also selected by the school, due to it being a short practical which therefore allowed for knowledge tests to be carried out at the beginning of the lesson. Chromatography is taught in the school as part of the C2: Elements, compounds and mixtures module, specifically C2.1: Separation techniques with the OCR exam board (37).

As this was not the student's first experience of chromatography it was necessary to conduct a knowledge test at the beginning of each lesson to establish the level of knowledge students had prior to the lesson (38). The same test was conducted at the end of the lesson. The lesson was planned with reference to ideas from the RSC chromatography practical (39). Although many of the students had some experience of using chromatography to separate dyes, the lesson plan built on this by teaching the calculation of R_f values (40) which had not previously been taught.

Practical experiment

The practical selected for the students to perform was separating inks using chromatography and calculating the R_f values (40).The students were then asked if any of their R_f values from the "suspects" matched an R_f value of a pen ink found on some "evidence" (ransom note). For this reason, the practical was carried out in the laboratories at Plymouth University prior to the students taking part in the practical in order to provide the students with the $R_{f\ value}$ from the ransom note. The trial was also used to ensure the safety of the experiment and the reliability of the results to ensure the students would be able to attain useful data. Whilst these trials were conducted a video of the practical was also developed to demonstrate the results for the classes not taking part in the practical. No COSHH assessments were needed as the apparatus used for the experiment was non-hazardous.

Reagents and equipment

The pens selected were a Pilot wytebord marker, a Tiger permanent marker and a stabilo point 88. Deionised water was used as a solvent and also to wash all glassware prior to use. The chromatography paper used was Whatman 100x300mm sheets catalogue number 3001-845. A thermometer was used to assess any change in temperature.

Procedure

A pencil line was drawn across the chromatography paper approximately 2.5 cm from the bottom of the paper. At equidistant intervals the 3 inks were dotted along the line using 5 dots of ink on the same spot for each ink. The paper was the attached to a piece of dowel using a paper clip and measured up against the beaker in order to approximate the amount of solvent needed. The solvent was then poured into the beaker and the paper placed in the solvent so that the solvent touched the paper but did not touch the paper line. Care was taken to ensure the paper did not touch the sides of the beaker. Once the paper was in place a timer was started. The experiment was continued until the inks were not seen to separate further. The paper was then removed and left to dry before measurements for the R_f value calculations were made.

The lesson

Learning outcomes

As the practical was determined by the school, the lesson was planned around the practical and using the OCR GCSE gateway science chemistry A specification(37).The learning outcomes for both the practical and non-practical lesson are outlined in table 3.

Table 3: Learning outcomes for all lessons

Learning outcomes
Calculate R/F values
Principle of chromatography separation
Identify immiscible and miscible liquids
Calculate mean

Practical Lesson Trial

The practical work was trialled at a Royal Society of Chemistry event. The event saw 6 groups of 15 to 20 students partake in a 30min chromatography practical. During which they used the same pens as were used in the experiment practical and calculated R_f values. Due to the nature of the timetable the students were not able to wait for their chromatography paper to dry prior to calculating R_f values. The students were given a handout with some instructions and tables to fill out. Due to time constraints the students did not calculate the means of the R_f values. On average the practical took between 20-25mins.

Lesson planning

Prior to planning the lesson, the information necessary for the students to learn and the needs of the study were both considered (41).This was done in accordance with the integrated course design (41) and is demonstrated in figure 1.

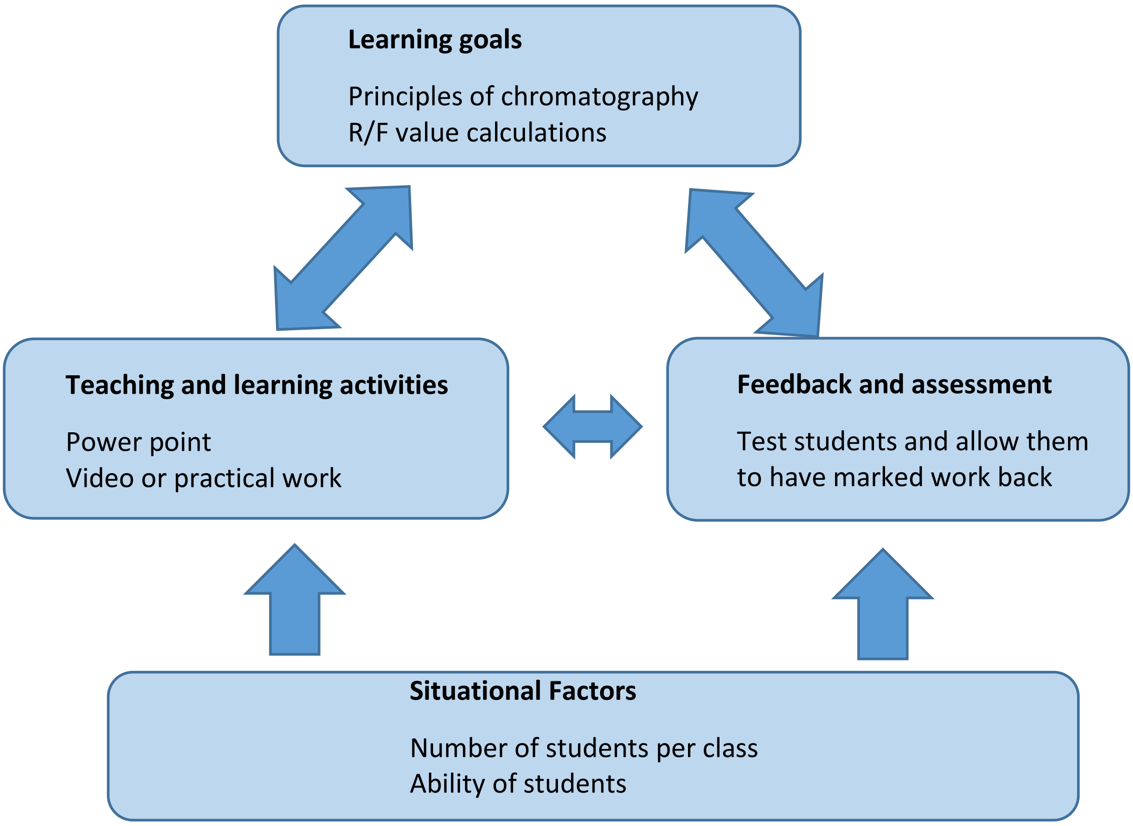

Figure 1: Flow chart of key concepts for lesson following the integrated course design

The lessons were both structured using the same initial PowerPoint slides with small adaptations for the practical and non-practical lesson. PowerPoint was selected as a learning tool for all classes as it has been shown to significantly benefit learners (42). During the course of both lessons, the students were given opportunities to ask and answer questions both in front of the class and also individually as they worked. Understanding was monitored using a short 10 question multiple choice test. The same test was used to assess their knowledge prior to engaging in the lesson. The learning objectives laid out in the table 3 were adhered to. The activities the students participated in were designed to challenge misconceptions and stimulate learning and questioning, during this process the questions also allowed the chance for participants to make mistakes (41). Throughout the lesson's progression was monitored with a variety of activities including crosswords and calculations.

Practical lesson 1

At the beginning of the lesson the students completed a VARK test (23) and a knowledge test . The students all signed their consent and were informed their results would not be counted towards other work they were doing with the school. As many of the students were concerned, they would not know the answers as they had not been taught the lesson, they were reassured that no prior knowledge was expected of them. Prior to the students carrying out the practical they were first given and overview of the concepts of chromatography before being shown a picture of the apparatus set up all via a power point. Initially the students were shown a slide demonstrating some uses of chromatography in order to generate interest before learning about the principles of its use as a separation technique via diagrams. To generate more interest in the subject the practical was presented in the form of a forensic problem (43).The students were asked to use chromatography and R_f values to deduce which pen ink was responsible for a ransom note. To give clarity to the experiment the students were shown two diagrams one of the appropriate ways to set up their chromatography paper and a second demonstrating all of the apparatus at their disposal. The students then carried out their experiment in groups of 2 or 3. Once the practical had been running for approx. 15 mins the practical was stopped, the paper was initialled by the students in pencil, hung to dry and the lesson cleared away. The students where then asked which pen, they thought was responsible for writing the note. A question which they were not yet able to answer. The majority of students wrongly deduced that the ink which moved the most was responsible for the note. This provided an opportunity for the students to question themselves on their understanding and a brief introduction on R/F values to be given prior to the start of the second lesson.

Practical lesson 2

At the beginning of lesson two the students were asked to collect their pieces of chromatography paper and shown a slide on how to calculate R_f values. The students were then given a handout of the tables they were required to fill out and shown a slide of the handout with an explanation of how to calculate the mean. The students first calculated their R_f values with supervision and the opportunity to ask for help. Once this was completed, they were asked to swap results with two other groups in order to calculate the mean.

The students were then shown a slide revealing the R_f value of the ink responsible for the ransom note and asked which pen had the same value and was therefore responsible for the note. Once all the students had verified the correct pen responsible the importance of taking a mean and reliability of their results was discussed. An opportunity to ask questions was then given. The students all worked quickly and were engaged with the subject* and so had time to complete a crossword designed to help with revision of the subject. Finally, the students were asked to complete the same knowledge test they started the lesson with. This allowed for their understanding of the lesson to be monitored.

* The top set students also had enough time to complete a work sheet on gas chromatography which was not relevant to their knowledge test.

Non practical lesson

The students completed the same tests as in the practical lesson 1.(38) The students worked through essentially the same power point as used for the practical, but rather than completing the practical they were shown a video recorded in the University of Plymouth Laboratories using the same pen inks. After watching the video, the students were asked which pen was responsible for writing the note and again the majority answered incorrectly that the pen which moved the most must be responsible. This presented an opportunity for a conversation about R_f values. They were then shown a slide recapping the information that was available before moving onto the slide on calculating R_f values. The students were then shown a slide including tables with values for the R_f calculations already included as they had non data due to not completing the practical. The table on the slide were given out to each student as a handout and the students then did the necessary calculations. The R_f value of the pen responsible for the note was again revealed and the students asked to identify which pen was responsible for the note. Once the class had correctly identified the pen responsible the students completed, he same crossword as the practical group before re-taking the knowledge test.

Final test

To assess the students' retention of information all the students re-took the same knowledge test after 19 weeks. The long delay on the completion of the final test was due to complications with timetabling at the school due to staff illness and is discussed later

Results, Statistical analysis and Interpretation

Results

Once all the teaching sessions, pre, post, retention tests and the VARK tests were completed the tests were marked. Each student was allocated identifiers in order to retain anonymity. The identifier was applied to all their results for each test and the name removed. This allowed for the individual results of the student to be grouped together. Students that did not complete all of the tests due to absence were removed from the results. This section comprises of a "statistical overview" and "Analysis and interpretation of results". The data generated by this study is analysed both graphically and with the use of t-tests.

Statistical overview

Once all the teaching sessions, pre, post, retention tests and the VARK tests were completed the tests were marked. Each student was allocated identifiers in order to retain anonymity. The identifier was applied to all their results for each test and the name removed. This allowed for the individual results of the student to be grouped together. Students that did not complete all of the tests due to absence were removed from the results.

The sample sizes were not consistent across the experiment as the students in the low ability groups had smaller class sizes. The majority of the analysis was performed using the t-test in order to determine whether two sets of data are different from one another. There are a variety of t-tests available, but the focus of this study was on the paired t-test (44)when comparing tests from the same class as the sample sizes were equal and the independent t-test(44) when comparing results from separate classes which were therefore of different sample size.

For all statistical analysis SPSS software was used. The software uses Levene's test (45)to analyse the variance of the absolute deviation scores of the groups. For equal variance the spread of the data from the average value must be similar. A value for the Levene test over 0.05 confirms equal variance. Two sets of data are provided by the software and the results of the Levene test used to determine whether equal variance was found and as a result the relevant data was selected. For all of the results in this study the Levene test results were above 0.05 which indicated that the compared data was of equal variance. The Levene's test results have been included in the relevant tables as part of the results. For the t-tests a P value was generated using SPSS software. The P value is a significance level which indicates how likely the sample result is if the population means are equal. If the P value was below 0.05, the results were deemed statistically different. If the P value was above 0.05 than the results were considered not statistically different (46). All statistical tests were completed with 95% confidence unless otherwise stated. In order to clarify how the software was used an unaltered annotated example output has been included in figure 2. Not all the data contained in the output was necessary and the study focuses on the highlighted values for the Levene's test and the t-test p-values.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pre-test results	Equal variances assumed	.003	.956	-1.100	40	.278	-.61905	.56263	-1.75617	.51807
	Equal variances not assumed			-1.100	39.86	.278	-.61905	.56263	-1.75628	.51819

This is the p-value for the Levene's test. As it is above 0.05 equal variance is assumed and therefore only the data highlighted in red is applicable

This is the p-value for the t-test which is above 0.05 indicating that the data is not statistically significantly different

Figure 2: Example of SPSS output for the comparison of post-test results between top set students who participated in a practical lesson and top set students who participated in a non-practical lesson

The relative standard deviation (RSD) of the results was included in the relevant tables. RSD is a standardized measure of dispersion of a probability distribution or frequency distribution. The larger the value the greater the dispersion of the data. A small value indicates high precision. The equation for the calculation of the RSD can be seen below in equation 2.

Equation 2

$$RSD = \sigma \div \mu \times 100$$

Where σ = the standard deviation (stdev)

μ = the mean

Standard deviation is often used within chemistry research to identify and remove any outliers. This is calculated by multiplying the standard deviation by 3 and then any values which are the mean +/- 3 stdev are removed (47). The results of this study produced no outliers due to the large standard deviations (which coincidently the presence of outliers will increase). However, if outliers were present they would not have been removed as the study is post-positivist and therefore critical realism (table 1) must be used to conclude that any students which may appear as outliers are still relevant as people are individuals capable of reacting to the same stimuli and producing a variety of different results.

Analysis of results

Comparison of the results from the learning of top set ability students taught practically or non-practically

The mean data for the practical lesson was 4.67 and for the non-practical lesson 5.29 (out of a possible 10) shown in table 4. This suggests that the ability of both groups is similar, however further statistical analysis is shown below in table 5 proving this to be the case. The RSD for the practical and non-practical lesson is shown as 37.9% and 35.5% respectively in table 4. Both values are quite high however it is to be expected with a study; as the study is a post-positivist, social science model all the data is considered valuable and as such outliers have not been excluded as students by their nature are individuals and will not perform differently for a variety of reasons due to many external circumstances which are outlined later.

Table 4: mean comparison between 9L1 and 9L2 pre-test results

Lesson type	Number of students	Mean score	stdev	RSD (%)
Practical	21	4.67	1.77012	37.9
Non-Practical	21	5.29	1.87464	35.5

To analysis the data for significant statistical differences SPSS was used to generate the data shown in table 5. Equal variance was assumed due to the value of 0.956 for the

Levene test being greater than 0.05 for the pre-test results. The P value shown for the pre-test results was 0.278 which is above 0.05 suggests there is no statistical difference in the knowledge and understanding of the subject between classes 9L1 and 9L2 prior to under taking the lesson. This result allows for the post- test and VARK tests of the top set practical and non-practical groups to be compared as prior to starting the lesson their knowledge of the subject was statistically the same.

Table 5: Independent t-test results for comparison of means between 9L1 and 9L2 pre, post and retention test results.

Test	Pairing	Levene's Test for Equality of Variances	Significance level (P)	Statistically significant difference?
Pre-test	9L1-9L2	0.956	0.278	No
Post-test		0.102	0.501	No
Retention		0.602	0.464	No

The mean data for the post-test results in table 6 shows 9L1 to have a mean 8.44 and 9L2 a mean of 8.71 out of 10. This suggests that the understanding of the lesson content in both groups is similar. Levene's test for the post-test results (table 5) shows a value of 0.102 which is above 0.05 and therefore the variances can be considered equal as the variances are equal the relevant value for P was 0.501 which is greater than 0.5 indicating no significantly different data. This data therefore suggests that there is no statistically significant difference between the mean scores of class 9L1 (practical lesson) and 9L2 (non-practical lesson). The relative standard deviation (table 6) was much smaller for both lesson styles in the post-test results suggesting less spread on the data and therefore suggesting most of the student's results were close to the mean. This was not seen to be the case earlier with the pre-test results in table 4.

Table 6: Mean Comparison between 9L1 and 9L2 post-test results

Lesson Type	Number of students	Mean	stdev	RSD (%)
Practical	16	8.44	1.36473	16.2
Non-Practical	21	8.71	1.00712	11.6

The final test the students underwent was a retention of information test to see how the learning style affected their long-term retention of information. Table 7 shows the mean for the practical group is 6.6 out of 10 for the practical group and 7.14 out of 10 for the non-practical. This suggests that the non- practical group retained a similar amount of information as the practical group the possible reasons for this are discussed in "Which lesson style is best?". In table 7 The RSD is seen to be 24.8% and 40.2% for the practical and non-practical retention test results respectively. This has increased from the

results of the post-test and for the non-practical lesson it is greater than the pre-test. Suggesting a wide variety of results for the retention of information when a non-practical lesson style is adopted.
To assess whether this is an accurate statement an independent t-test was again performed on the data and the results shown in table 6. The P value for the retention test results in table 6 was 0.464 which is greater than 0.05, therefore there is no significant difference between the retention of information of the students with regards to the two lesson styles. This result is discussed further in “The effects of lesson style”.

Table 7: t-test comparison between 9L1 and 9L2 retention test results

Lesson Type	Number of students	Mean	stdev	RSD (%)
Practical	20	6.60	1.63514	24.8
Non-practical	21	7.14	2.86855	40.2

The means of the pre-test results are lower than the retention test results so a paired t-test was done to see if the students had retained any more knowledge than they had prior to the lesson. The results are shown in table 8.

Table 8: paired t-test results for the comparison between pre-test and retention results for both 9L1 and 9L2

Lesson Type	Pairing	stdev	Significance level (P)	Statistically significant difference?
Practical	Pre-test - retention	1.69442	0.000	Yes
Non-practical	Pre-test - retention	3.33188	0.002	Yes

Both the practical lesson and the non-practical lesson show P values below 0.05 for the comparison of the test score from the pre-test and the retention test, therefore it can be inferred that there is a statistically difference in the knowledge the students had before they participated in the lesson and at the time of undertaking the retention test. This is discussed further in “The effects of lesson style”.

Figure 3 shows a graphical representation of these means. It can be seen that there is very little difference in mean scores between the two lesson styles and although the non-practical group appear to do best overall as they also scored better on the pre-test it is possible that they already have some understanding of the subject prior to taking part in the lesson. Although any advantage the non-practical group had was not found to be statistically significant (table 10).

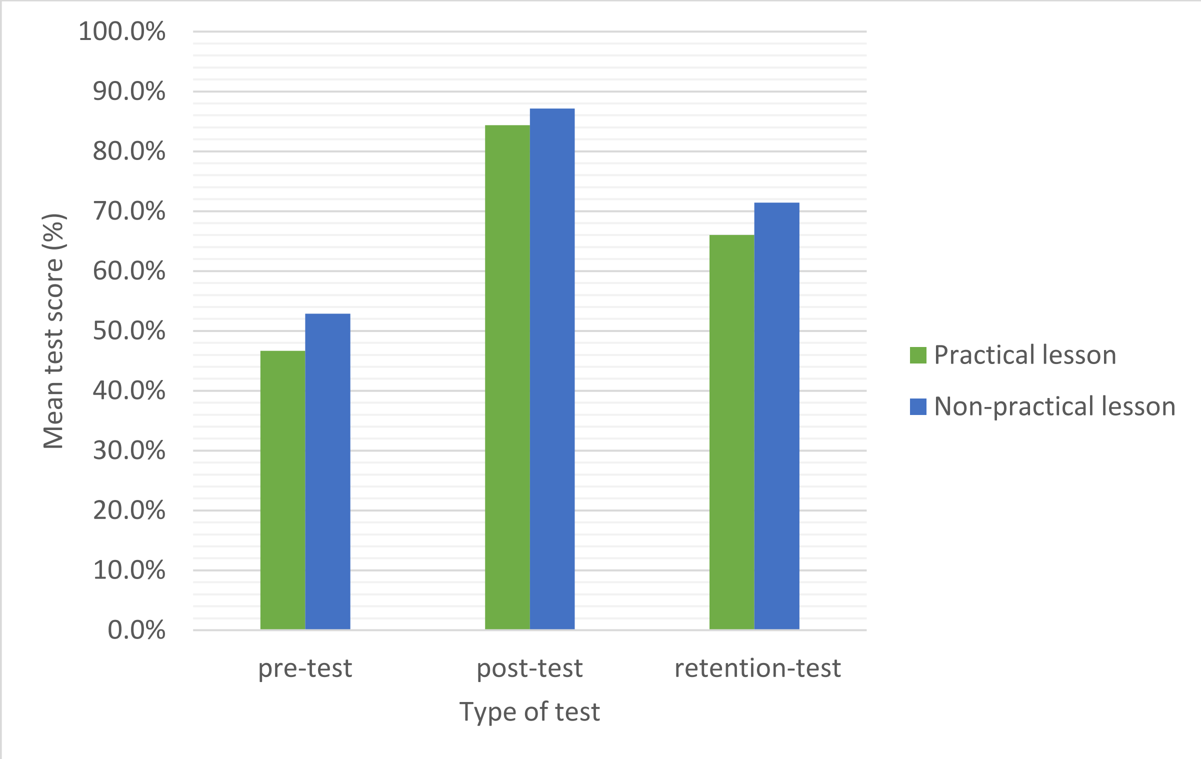

Figure 3: Bar chart comparing mean data from the pre, post and retention tests for 9L1 and 9L2

Rationalising top set discrepancies using VARK learning styles

VARK results for top set practical

By comparing the pre and post-test scores of 9L1 (figure 4) it can be seen that students 11 and 22 appeared to gain no understanding from the lesson as they have a lower score than prior to engaging with the practical lesson. As this is a top set the students are all assumed to be of a similar ability.

A possible explanation for this lack of understanding is that the lesson style was not comparable to their learning style. This is discussed further in "The compatibility of learning and lesson styles". All the students also undertook a VARK test to assess how they learnt best and the results for students 11, 22 and 23 (figure 5).

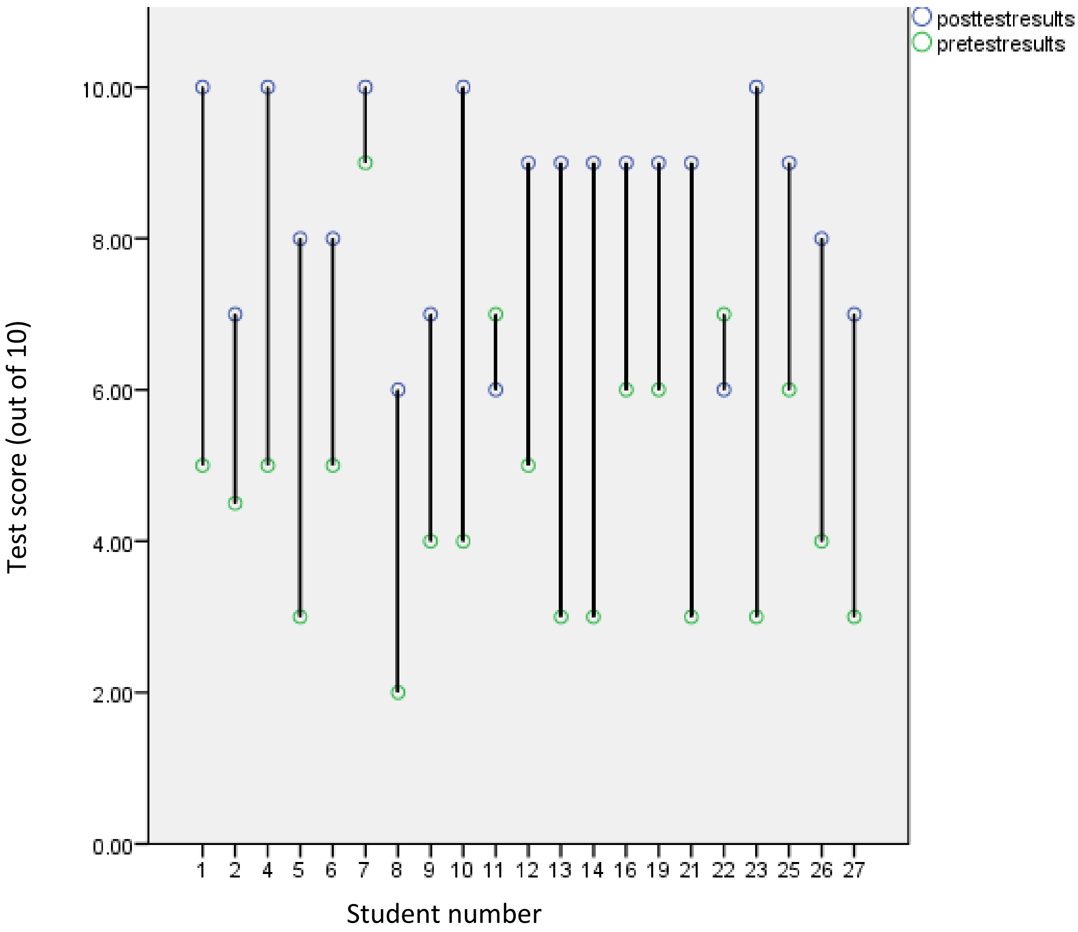

Figure 4: Scatter graph showing a comparison of 9L1 pre and post-tests scores

The graph (figure 5) shows the VARK tests results from students 11, 22, and 23. Students 11 and 22 both did poorly in the post-test whereas student 23 has been displayed for comparison as this student showed the greatest improvement from their pre-test to post-test scores as shown in figure 4. All three students picked different learning styles for different situations laid out in the VARK test. From the results it can be seen that student 11 preferred aural and kinaesthetic learning, student 22 strongly favoured kinaesthetic learning and student 23 had a slight preference for reading as a learning style although they rate all the other learning style highly and of a similar value to their learning. From these results it would be expected that student 22 would learn best from the practical lesson style. However, in spite of favouring the kinaesthetic method of learning both students 11 and 22 did not appear to learn from the lesson as their marks decreased. Whereas student 23 achieved full marks after the lesson despite having very little understanding at the start and favouring a learning style least suited to practical work. Possible reasons for this result are discussed in “The compatibility of learning and lesson styles”.

<u>VARK results for top set non-practical</u>

Unlike the results for the practical lesson every student who participated in the non-practical lesson from top set achieved some improvement between their pre-test and post-test scores. The lowest post-test score was from student 4 shown below in figure 6. The largest improvement was seen by students 5 and 24.

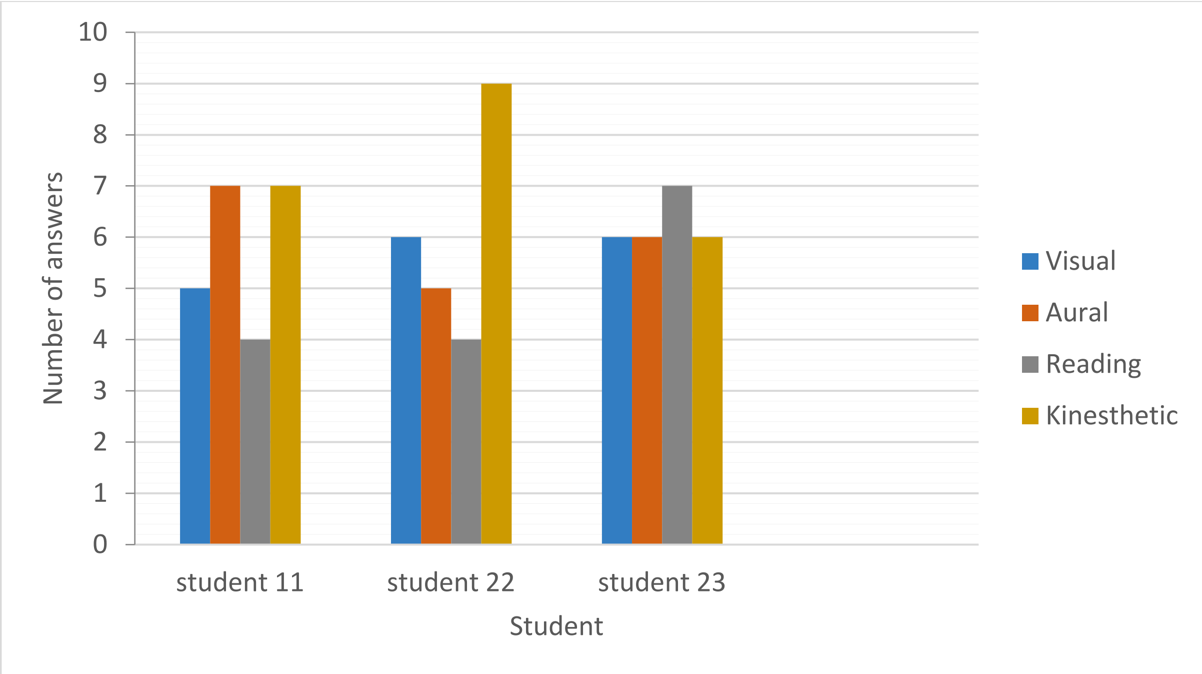

Figure 5: Bar chart showing VARK results for students 11, 22 and 23 in class 9L1

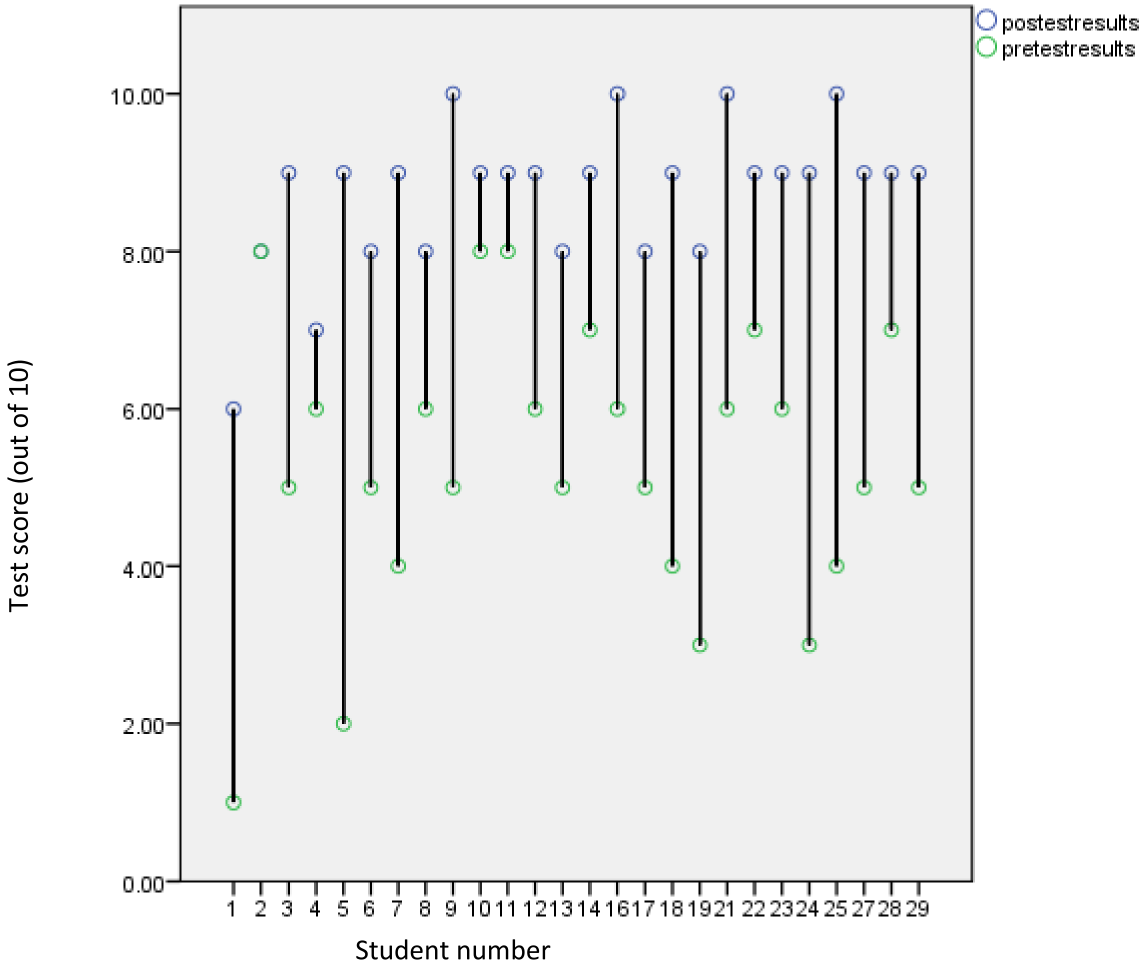

Figure 6: Scatter graph showing a comparison of 9L2 pre and post-test scores

The VARK test results of students 4, 5 and 24 have been compared in figure 7.

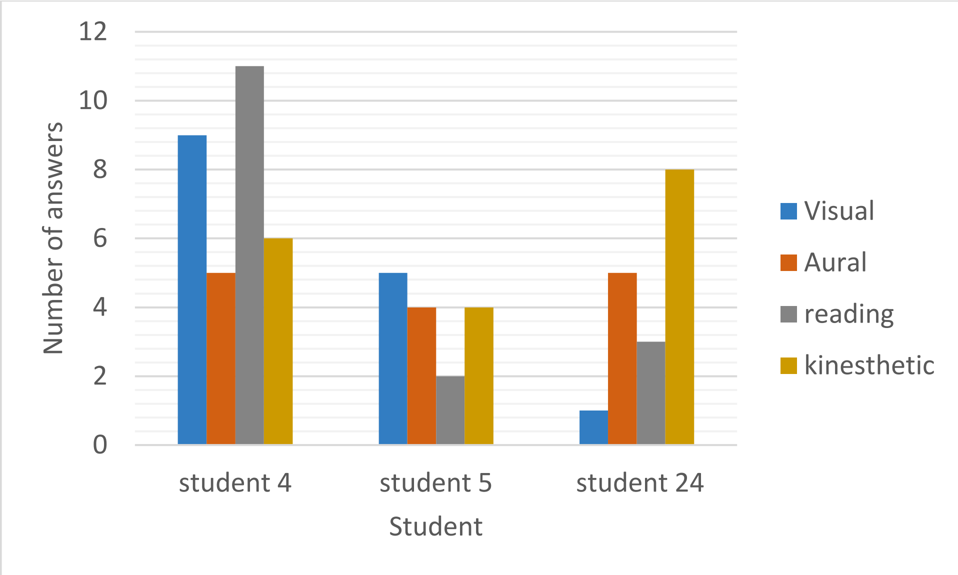

Figure 7: Bar chart showing VARK results for students 4, 5 and 24 in class 9L2

Student 4 favours a reading and visual learning style, which should be most suitable to a non-practical lesson, however they were seen to make the least improvement. Students 5 and 24 saw the most improvement but have different learning styles; student 5 appears to prefer visual and aural learning whereas student 24 prefers a kinaesthetic learning style above all others. From the VARK results it would be expected that student 4 would fair significantly better than student 24 however the opposite is true in this example. Student 24 and student 5 did improve the same amount, but student 24 attained full marks (10/10) in the post-test.

Comparison of the results from the learning of bottom set ability students taught practically or non-practically

The mean score for the bottom sets practical group (9R5) was 3.20 and the mean for the non-practical group (9R6) was slightly lower at 3.00, shown in table 9. This suggests that the ability of both groups is similar, however further statistical analysis (table 9) proved this to be the case. The relative standard deviation (RSD) for the practical and non-practical lesson is shown as 50.4 % and 28.9% respectively below in table 9. As with the RSD for the top set pre-test the RSD values for bottom set are high for both lesson types although the RSD of the practical lesson pre-test is particularly high a 50.4%. As mentioned previously large RSD values are to be expected due to the nature of the study.

Table 9: mean comparison between 9R5 and 9R6 pre-test results

Lesson type	Number of students	Mean	stdev	RSD (%)
Practical	15	3.20	1.61245	50.4
Non-practical	7	3.00	0.86603	28.9

From the data shown in table 10 equal variance is assumed due to the value of 0.195 being greater than 0.05 for the pre-test results. The P value of 0.736 is above 0.05 which suggest there is no statistical difference in the knowledge and understanding of the subject between classes 9R6 and 9R5 prior to under taking the lesson.

Table 10: Independent t-test results for comparison of means between 9R5 and 9R6 pre, post and retention test results.

Test	Pairing	Levene's Test for Equality of Variances	Significance level (P)	Statistically significant difference?
Pre-test	9R5-9R6	0.195	0.736	No
Post-test		0.216	0.002	Yes
Retention		0.769	0.605	No

The mean score for the post-test results shows 9R5 to have a mean score of 6.29 and 9R6 a mean of 3.7 out of 10 (table 11). This suggests that the understanding of the lesson content in both groups is different. Further statistical analysis is shown below in table 12 proving this to be the case. The relative standard deviation for the post-test results (practical 23.7% and non-practical 55.6%) are seen to be nearly the reverse of those from the pre-test (practical 50.4% and non-practical 25.9%). This indicates that the practical lesson style group saw a much smaller spread of data as well as a higher mean score for the post-test result whereas the non-practical lesson saw a higher mean score for the post-test result, but also a larger spread of data than in their pre-test results.

Table 11: Mean comparison between 9R5 and 9R6 post-test results

Lesson type	Number of students	Mean	stdev	RSD (%)
Practical	15	6.29	1.490	23.7
Non-practical	7	3.70	2.058	55.6

For the post-test results table 10 shows a Levene's test value of 0.216 which is above 0.05 and therefore the variances can be considered equal as the variances are equal the relevant value for P is 0.002 this is less than 0.5 indicating significantly different data. This data therefore suggests that there is a statistically significant difference between the mean scores of class 9R5 (practical lesson) and 9R6 (non-practical lesson) for the post-test. As the mean score achieved by the 9R5 students was greater at 6.29 than 3.70 scored by 9R6 it can be inferred that the practical lesson style improved the test scores of the students as prior to beginning the lesson the abilities of both classes (9R5 and 9R6) were found to be statistically the same as shown previously in table 10. This is discussed further in “Which lesson style is best?”.

The final test the students underwent was a retention of information test to see how the learning style affected their long-term retention of information. Table 12 shows the mean for the practical group is 3.09 out of 10 for the practical group and 3.57 out of 10 for the non-practical. This suggests that the non- practical group retained more information than the practical group. In table 12 the RSD’s for both practical and non-practical lesson styles once again show a large spread of data with values of 62.2% and 50.8% respectively.

Table 12: mean comparison between 9R5 and 9R6 retention test results

Lesson type	Number of students	Mean	stdev	RSD (%)
Practical	15	3.09	1.92117	62.2
Non-practical	7	3.57	1.81265	50.8

To assess whether this is an accurate statement an independent t-test was performed on the data and table 10 shows that there is no significant difference between the retention of information between the two lesson styles on the classis’s retention of information as the P value of 0.605 is greater than 0.05. The means are also similar to those of the pre-test results so a paired t-test was done to see if the students had retained any more knowledge than they had prior to the lesson. The results are shown in table 13.

Table 13: paired t-test results for the comparison between pre-test and retention results for both the practical lesson and the non-practical lesson

Lesson type	Pairing	stdev	Significance level (P)	Statistically significant difference?
Practical	Pre – retention test	1.60128	0.502	No
Non-practical		1.90238	0.457	No

Both the practical lesson and the non-practical lesson show P values above 0.05 (table 13), therefore in can be inferred that there is no statistically difference in the knowledge the students had before they participated in the lesson and at the time of undertaking the retention test. This is discussed further in “The effects of lesson style”.

Figure 8 shows a graphical representation of the mean score for the pre, post and retention test completed by bottom set both practical and non-practical. It can be seen that there is very little difference in mean scores between the two lesson styles prior to beginning the lesson (pre-test). Whereas the results of the post-test show a much higher mean for the practical group indicating better understanding at the end of the lesson.

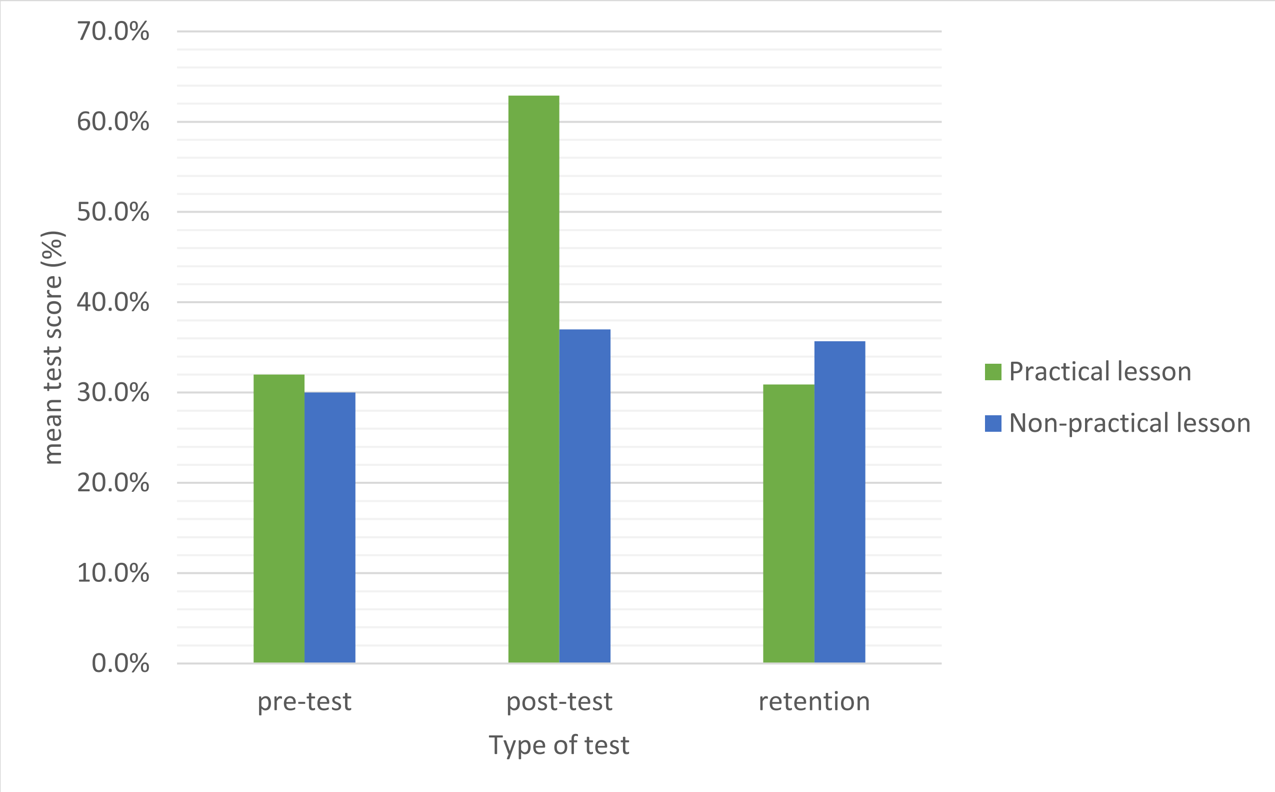

Figure 8: Bar chart comparing mean data from the pre, post and retention tests for 9R5 and 9R6

Rationalising bottom set discrepancies using VARK learning styles

VARK results for bottom set practical

Figure 9 is a scatter graph which matches the pre and post-test scores of the individual students from the bottom set practical group together.

The graph (figure 9) shows that student 11 appeared to gain no understanding from the lesson as they have a lower score than prior to engaging with the practical lesson. To assess if this is due to their learning style not being suitable for the type of lesson their VARK score was assessed along with the VARK scores of two students who improved the most out of the group. For interest students 2 and 13 were selected as both improved the same amount (by 4 marks) but had different levels of understanding (student 2 had a final score of 5, student 13 a score of 8). The VARK questionnaire results for students 11, 2 and 13 can be seen in figure 10.

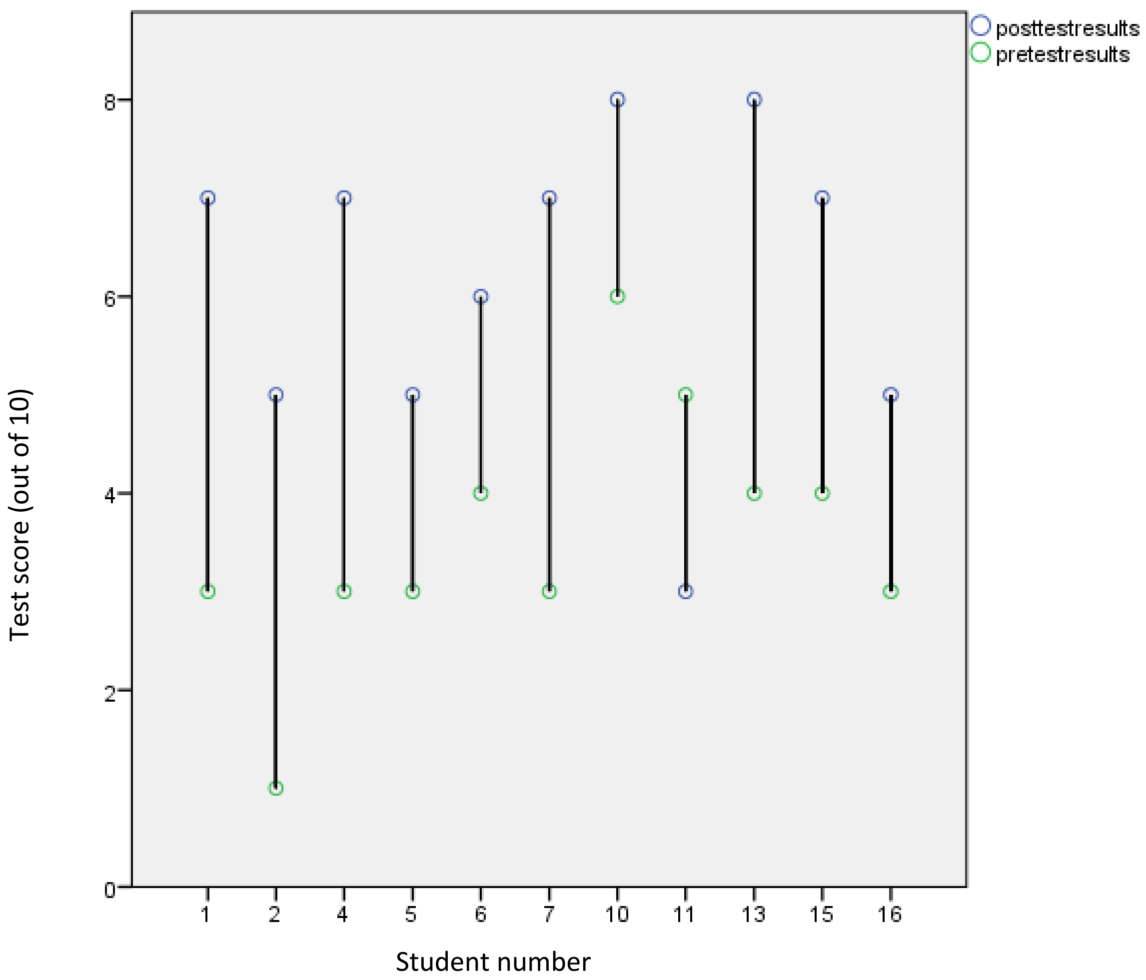

Figure 9: Scatter graph showing a comparison of 9R5 pre and post-test scores

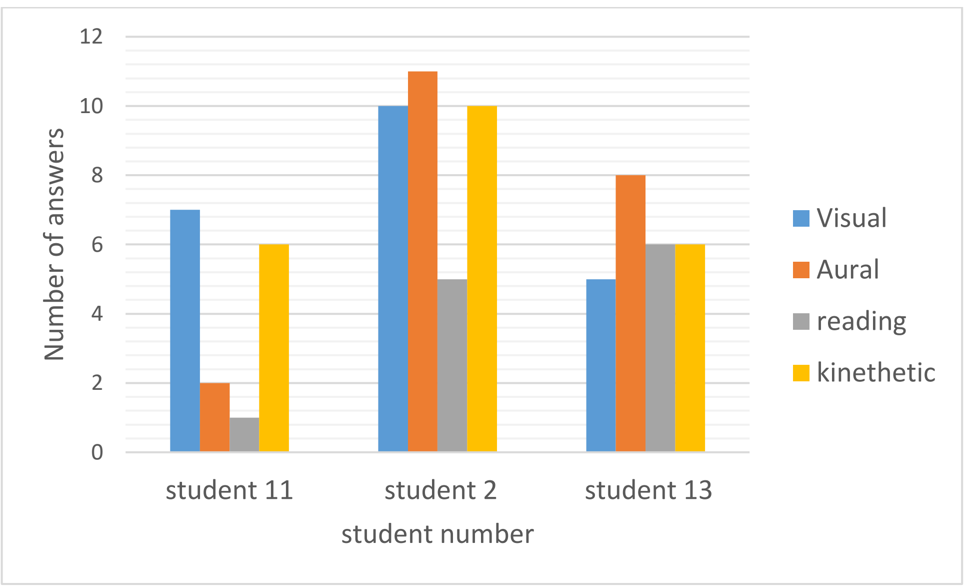

Figure 10: Bar chart showing VARK results for students 11, 2, 13 in class 9R5

The graph (figure 10) shows student 11 had the least understanding of the lesson content, but appears to have a preference for visual and kinaesthetic learning which would suggest that the practical lesson would be his preferred lesson style. Students 2 and 13 are two students that demonstrated the highest increase in understanding after the lesson (although there were other students how increased by the same amount). The VARK results for students 2 and 13 both show a preference for aural learning which should not have been and advantage to them during the practical lesson structure. This is discussed further in "The compatibility of learning and lesson styles".

VARK results for bottom set non-practical

Figure 11 is a scatter graph comparing the pre and post-test result of the bottom set non-practical group.

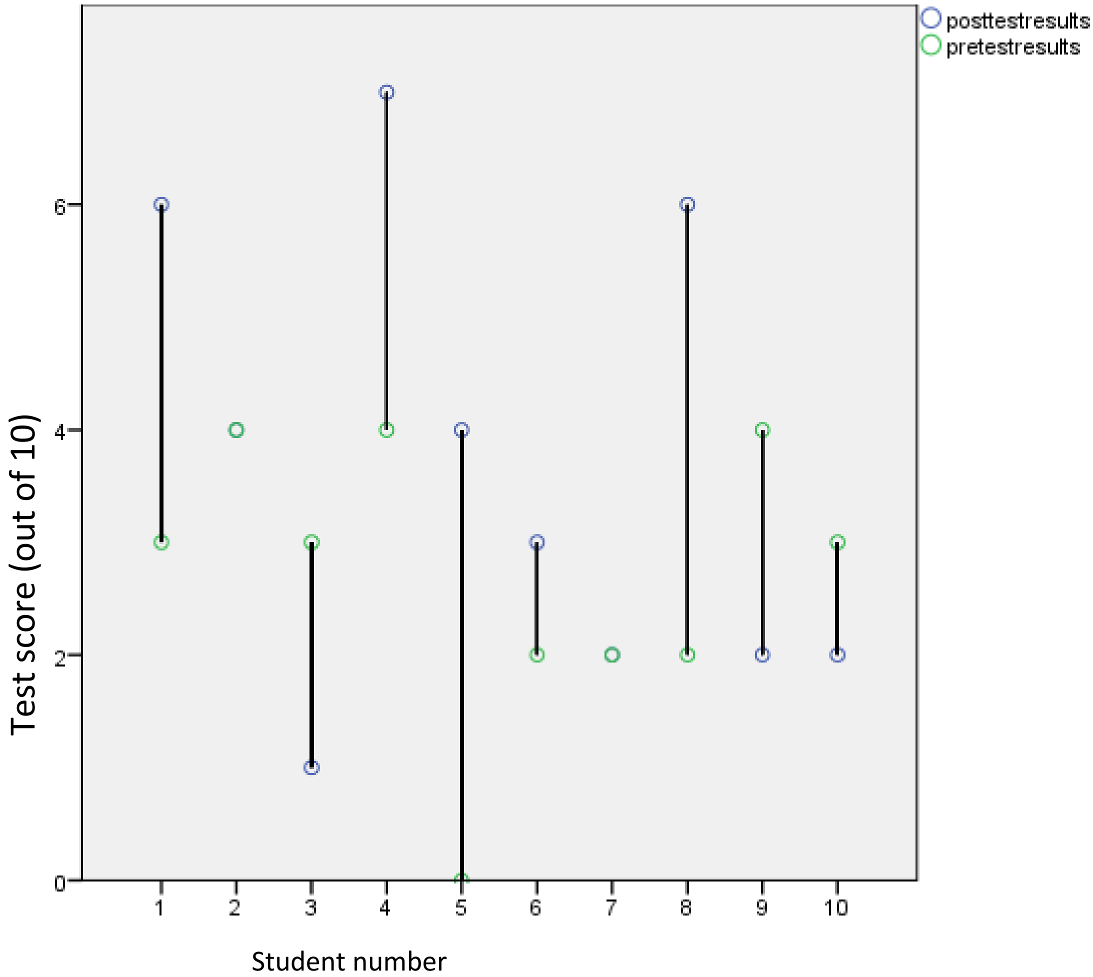

Figure 11: Scatter graph showing a comparison of 9R6 pre and post-test scores

The graph in figure 11 shows that for students 2 and 7 show the same mark was attained in both tests and therefore no improvement was made. Students 9 and 10 both appear to have less understanding after the lesson. The largest improvement was made by student

5 and the highest mark was from student 4. In figure 12 students 4, 5, 9 and 10 VARK results were compared, as their results show the largest contrast.

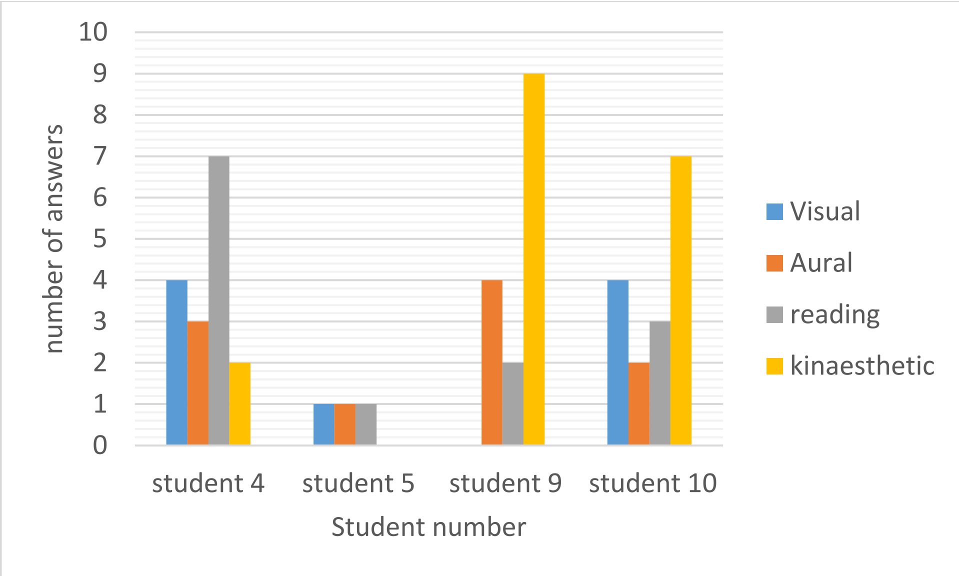

Figure 12: Bar chart showing VARK results for students 4, 5, 9, and 10 in class 9R6

Figure 12 shows that students 9 and 10 were strongly kinaesthetic learners and in this respect the non-practical lesson style should not have suited their learning style the least. Students 4 and 5 score kinaesthetic learning as the least helpful method for themselves which indicates that the non-practical lesson style is most beneficial for their learning. Student 5 was considered by the teaching staff to be the least able due to suffering from severe autism.

Comparison of class average VARK scores

Figure 13 shows the average VARK score of each class. The score for each learning varies considerably as students were allowed to choose multiple answers to questions. Therefore, displaying the information this way does not provide the clearest picture.

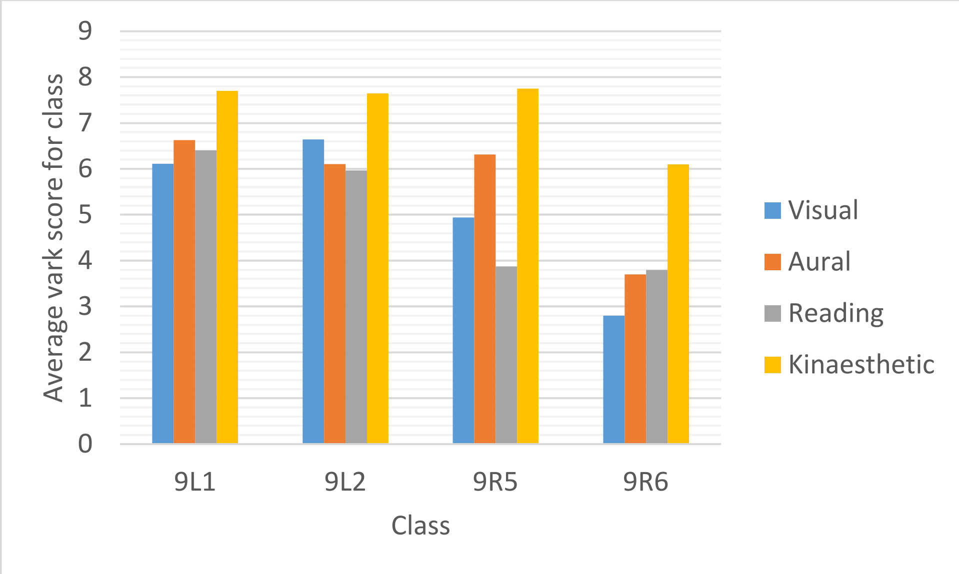

Figure 13: Bar chart of mean VARK learning style score for each class

To make better use of this data the ratios of the VARK scores from the class averages have been compared below in figure 14. The ratios were calculated by taking the smallest value for each class and dividing the VARK class results by this number the data for this is shown in table 14

Table 14: The mean VARK score for each class and the calculated ratio used to plot the graph shown in figure 14

Learning style	**9L1**		**9L2**		**9R5**		**9R6**	
	Mean score	Ratio	Mean score	Ratio	Mean score	Ratio	Mean score	Ratio
Visual	6.11	1	6.64	1.11	4.94	1.27	2.8	1
Aural	6.63	1.08	6.11	1.02	6.31	1.63	3.7	1.32
Reading	6.41	1.05	5.96	1	3.88	1	3.8	1.36
Kinaesthetic	7.70	1.26	7.64	1.28	7.75	2	6.1	2.18

The bar chart (figure 14) shows that all the classes appear to prefer a kinaesthetic form of learning, but the strength of this preference is seen to vary with the ability of the class. The result for the two top sets (9L1 and 9L2) show fairly similar results for all the learning styles indicate that they are comfortable learning information across all the learning styles and therefore can be considered to be multi-modal learners.

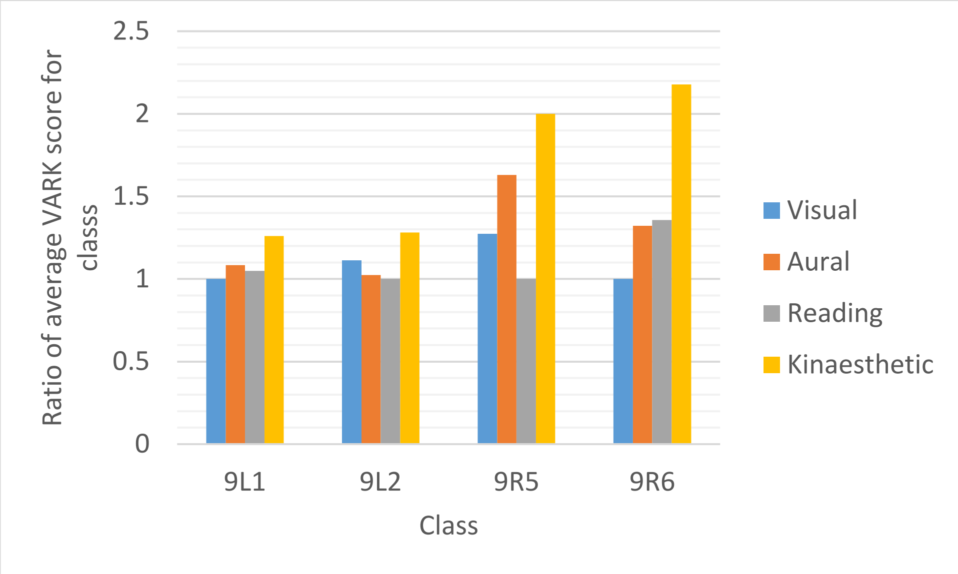

Figure 14: Bar chart of VARK learning style ratio for each class

The two bottom sets (9R5 and 9R6) both show less variety in their learning style preference with visual and reading learning style appearing the least popular of learning styles. This is discussed further in “The effects of learning styles”.

Comparison of post-test score with relation to lesson and learning style compatibility

For the two different lesson types two different learning styles were preferable. Students participating in the practical lesson with either a visual or kinaesthetic approach to learning should learn better than those with an aural or reading learning style and vice versa for the non-practical group. To investigate this, students who were predominately either visual or kinaesthetic learners were grouped together and students who favoured aural or reading learning styles were also grouped together. An independent t-test was then performed. Table 15 shows the results from 9L1 the top set practical group.

Table 15: Mean post-test results comparison between visual/Kinaesthetic and Aural/ reading learning style for 9L1

Learning style	**Lesson type**	**Number of students**	**Mean**	**stdev**	**RSD (%)**
Visual/kinaesthetic	Practical	10	8.10	1.37032	16.95
Aural/ reading		11	8.72	1.27208	14.59

The mean test score for both groupings of the learning style appears to be similar seen at 8.10 for visual/kinaesthetic learners and 8.72 for Aural/reading learners (table 15). The results for the Aural/reading learners is higher than that for the visual/kinaesthetic learners which is not the expected result as the lesson style should have been more beneficial to kinaesthetic/visual learners (discussed further in "The effects of learning styles"). This test was repeated for the top set non-practical group and the results can be seen in table 16.

Table 16: Mean post-test results comparison between visual/Kinaesthetic and Aural/ reading learning styles for 9L2

Learning style	Lesson type	Number of students	Mean	stdev	RSD (%)
Visual/kinaesthetic	Non-practical	18	8.89	0.90025	10.13
Aural/ reading		9	8.56	1.01379	11.84

Table 16 shows the reverse of table 15 with the mean score being slightly higher for the visual/kinaesthetic students being slightly higher at 8.89, this again goes against the expected result. To see if the results for the mean scores seen in both table 15 and 16 were statistically different an independent t-test was performed with the results shown in table 17.

Table 17: Independent t-test results for the comparison between visual/Kinaesthetic and Aural/ reading post-test results for top set practical group and non-practical group

Lesson style	Pairing	stdev	Significance level (P)	Statistically significant difference?
Practical	Visual/kinaesthetic-Aural/ reading	0.57654	0.290	No
Non-practical		0.38297	0.392	No

For both the practical and non-practical top set lesson the p-value was seen to be greater than 0.05 (0.2690 for the practical and 0.392 for the non-practical). This suggest there was no statistically significant difference between the test scores for either lesson type with regards to learner style. Indicating that there was no statistically significant difference between the post-test scores based on the compatibility between the students learning style and the lesson style (further discussion in 4.5)

The same analysis of results was carried out for the bottom set students. The results for the mean post-test score comparison between visual/Kinaesthetic and Aural/ reading learning style for the top set practical group can be seen in table 18.

Table 18: Mean comparison between visual/Kinaesthetic and Aural/ reading post-test results for top set practical group

Learning style	Lesson type	Number of students	Mean	stdev	RSD (%)
Visual/kinaesthetic	Practical	6	6.50	1.76068	27.08
Aural/ reading		5	5.80	1.30384	22.48

The mean results in table 18 show that the students with Visual or kinaesthetic learning style achieved a higher mean sore of 6.50 whereas students with an aural or reading learning style scored slightly less with a mean of 5.80.
This test was replicated below in table 18 for the bottom set non- practical lesson.

Table 19: Mean comparison between visual/Kinaesthetic and Aural/ reading post-test results for bottom set practical group

Learning style	Lesson type	Number of students	Mean	Std. Deviation	RSD (%)
Visual/kinaesthetic	Non-practical	6	3.00	1.78885	59.6
Aural/ reading		3	5.00	2.64575	52.9

The results for the bottom set non-practical group shown in table 19 show that the means appear to be very different with the visual and kinaesthetic students scoring very poorly (3.00) compared to the Aural and reading students (5.00). This was the expected result as the lesson style was more accessible to Aural and reading learners. The statistical significance of this was assessed in table 20 and no statistically significant difference was found. Indicating that again there was no statistically significant difference between the post-test scores based on the compatibility between the students learning style and the lesson style.

Table 20: Independent t-test results for the comparison between visual/Kinaesthetic and Aural/ reading post-test results for bottom set practical group and non-practical group

Lesson type	Pairing	Std. Deviation	Significance level (P)	Statistically significant difference?
Practical	Visual/kinaesthetic- Aural/ reading	0.95316	0.481	No
Non-practical		1.46385	0.214	No

Discussion

This study aims to use the results previously discussed addressing key points relative to the study. Before discussing these points, the key hypothesis investigated are outlined below. The interesting areas of the study were selected as “The effects of lesson style”, “Which lesson style is best?”, “The effects of learning styles” and “The compatibility of learning and lesson styles”.

Hypotheses

This study investigated three hypotheses (table 21), the null hypotheses were accepted or rejected as a result of the t-tests. If $H_1 = H_0$ the null hypothesis was accepted and the hypothesis rejected, If $H_1 \neq H_0$ the null hypothesis was rejected and the alternative hypothesis (H_1) accepted.

Table 21: The null hypotheses investigated in this study and whether they were excepted or rejected

Null Hypotheses (H_0)		**Top set**	**Bottom set**
1	There was no significant difference between the students’ knowledge for either non-practical lessons or practical lessons prior to the lesson	$H_1 = H_0$	$H_1 = H_0$
2	There was no significant difference in the information learnt between non-practical lessons and practical lessons	$H_1 = H_0$	$H_1 \neq H_0$
2	There was no significant difference between the retention of new information from either non-practical lessons or practical lessons	$H_1 = H_0$	$H_1 = H_0$

As null hypothesis 1 was accepted this allows for the results of the same ability students in either the practical or non-practical groups to be compared as they are statistically the same for both top and bottom set. For the bottom set hypothesis 2 the alternative hypothesis was accepted as there was a significant difference in the information learnt between non-practical lessons and practical lessons

The effects of lesson style

On learning information

The results of the t-tests (table 10) show us that for students of high ability (9L1 and 9L2) the lesson type had very little effect on their learning and retention of information. This was demonstrated as both the practical group (9L1) and the non-practical group (9L2) showed levels of improvement which were statistically the similar for both the post lesson and retention tests.

The t-tests for students of low ability (table 15, 9R5 and 9R6) show that lesson style had an effect on their learning as the levels of improvement seen in the post-test results were proven to be statistically significantly different. As the results for the students participating in the practical lesson style (9R5) were higher it can be inferred that the lower ability students learn better practically. A possible explanation for this is that practical work is known to enhance motivation. The increase of motivation is likely to be minimal in a class that is already highly motivated, therefore the practical lesson benefited the top students the least in this respect as the students are already considered highly motivated. The improvement seen in the bottom set as a result of the practical work could be due to a large increase in motivation.(48)

On retention of information

For all the students (both top and bottom set) none results for the retention of information were found to be significantly different as a result of the lesson style (tables 10 and 15). Therefore, this study suggests that the lesson style (either practical or non-practical) has no effect on the retention of information. These results appear counter intuitive as if the students learnt the information better practically as was the case for bottom set then it would be expected that they would also retain the information better. There are a number of explanations for this result. Firstly, due to complications with the school the time between the final lesson and the retention test was 19 weeks. This length of time without any revision is probably too long as without revision most students will have forgotten the majority of the lesson content within 15 minutes of leaving the lesson (49) . However, the top set classes did see some improvement when their mean pre-test results were compared to their mean post-test results (table 10). There was also a difference in the way this test was carried out. Rather than having the same teacher who carried out all the other tests this test (due to the constraints of the school) was carried out by four different teachers who would all have given different introductions to the test and expressed different levels of its importance. Teachers have a great effect on the success of a lesson and can account for up to 30% of the variance of student achievement (50) Finally, the school also choose to change the format of the test by printing it in a much smaller format which may have affected some of the student's ability as size of font can affect the perception of how difficult a test maybe (51). Due to all of these external factors it is likely that the results are not as reliable as those for the post-test.

Which lesson style is best?

The results of the study show that while top set ability classes can improve without practical work the use of practical work is beneficial to bottom set ability students (tables 10 and 15). It is also worth mentioning that at no point throughout the study did practical work provide statistically significant lower test scores.

A study by Jerome Thompson & Kola Soyibon(52) investigated, *"whether the use of the combination of lecture, teacher demonstrations, class discussion and student practical work in small groups significantly improved the experimental subjects' attitudes to chemistry and understanding of electrolysis more than their control group counterparts who were not exposed to practical work".*

The study concluded that this experimental teaching method was significantly more successful than the control group and therefore validates the results seen by the bottom set (9R5 and 9R6) where the practical lesson produced statistically higher test results. The students in the Thompson investigation participated in a pre and post-lesson tests similar to the method used in this study. However, the Thompson study doesn't stipulate the general ability of the students it selected. A possible reason for the result provided for the top set students in this study is that they are of a higher academic ability than the students in the Thompson study and therefore practical work has less of an impact on their learning as they have already adapted to learn in a number of different styles (multimodal learners).

A large proportion of work on education in science suggests that students learn best by "doing" (48). However through the results of this experiment it can be seen that a more accurate statement is that " some students learn best by doing" and as practical work is time consuming to ensure all learners reach their full potential a more pragmatic approach to the use of practical lesson should be adopted so as students who learn best this way have the majority of their lessons practical based , but for those who show no preference practical's are only necessary when learning a specific practical skill (e.g. a titration) and not as a method of teaching theory. An explanation into the shortcomings of practical work has been given by Dearden who stated,
"a teaching method which genuinely leaves things open for discovery also necessarily leaves open the opportunity for not discovering them" (53).

This statement goes some way to explain why different studies produce different findings on the benefits of practical work and suggests that the most important factor is possibly the teacher and the way the lesson is structured.

The effects of learning Style

The results of the VARK questionnaire show that the majority of learners have a preference for learning kinaesthetically regardless of ability as shown in figure 14. This is in line with the findings for first year medical students for which a kinaesthetic learning style was the most prevalent in uni-modal learners (54) (those which have a strong single preference of learning style)The area in which learning style differs between the high ability students and low ability students is in the difference between how strongly a learning style is favoured. Although students in the top set did marginally prefer kinaesthetic learning, they also exhibited similar competence with all the available learning style whereas low ability students' results suggested that other methods of learning were not as useful to them. This suggests that top set students have the advantage of being able to utilise a larger range of learning styles and is a possible reason for there being no significant statistical difference between the results of the post-test results for the practical and non-practical lessons. Whereas for the bottom set were less capable at using other learning styles and therefore practical learning was statistically significantly more beneficial.

The compatibility of learning style and lesson style.

The learning styles most suited to a practical lesson were visual and kinaesthetic(55), the post-test results for students favouring these learning styles were compared to the post-

test results of students favouring aural and reading learning styles, which should be more beneficial to non-practical work.
The results in tables 18, 19 and 20 show that for top set ability groups there was no significant difference in post-test score for either lesson style regardless of whether or not the students preferred learning style was compatible. These results suggest that the reason for students 11 and 12 performing poorly was not as a result of their learning style being incompatible with the lesson style. It is possible that an external factor is responsible for these results. Tables 21, 22 and 23 show the same to be true for the bottom set. This result opposes the work done by Peacock which suggested that mismatching teaching and learning style affected learning(56). Peacock's work was based on the opinions of students and teachers via interviews and therefore is not necessarily reliable as students who performed badly could possibly infer this as a fault of the teaching rather than a lack of personal application.
Another reason the results may be inconclusive is that the VARK method is extremely hard to validate, Dr Svinicki stated that, *"We found that VARK was hard to validate statistically, including with several modifications we tried and several statistical strategies such as multidimensional scaling. We just couldn't get a good fit with the data. This does not mean that the instrument itself is not valid or desirable, but it shouldn't be used in research; that is not its strength"*(57).

For these tests the favoured learning styles were assigned as the learning style selected most often by the student. In the raw data it can be seen that the majority of students selected answers for all the available learning styles the only difference being the amount of times a learning style was selected by an individual. This indicates that to view a student as only having one learning style is probably inaccurate as the majority tend to display a multimodal approach (in a separate study 63.9% of students were found to be multimodal learners with only 36.1% having one specific learning style (58)) and therefore are able to use their less favoured skills to a similar level of effectiveness. This may explain why no significant difference was found.

Limitations

Due to the nature of the study there are many limitations and therefore it was necessary to discuss the relevant points in the following sections "External factors" and "The study".

External factors

The individuals

Due to the nature of this study which uses students there are a wide number of external factors which may influence the results. Firstly, the individual students will have a wide variation of home lives which can affect learning (59, 60). Their home environments may be a positive place where learning is encouraged, but could equally be a place which causes them stress and inhibits their ability learn. Other external factors related to the individuals that have been proven to affect learning include diet, exercise and sleeping patterns. There is also a possibility that some students may revisit the lesson work whereas other students will not. The student's overall classroom skills were also likely to influence the results as although the lesson was focused on chemistry, students with a low reading ability may have found this hampered their learning experience due to the written test(61). As a result, students with a good scientific understanding may have not been able to gain a test score representative of their ability as all of these factors are

external, they are difficult to control and for the purposes of this study (due to its causal nature) no attempt to do this was made. Hodson stated,
"Each student's array of formal (academic) knowledge and informal (everyday) knowledge, compounded by highly personal experiential and affective elements—what we might call their personal framework of understanding will, in part, have been accumulated under a range of different circumstances, often under different stimuli and sometimes for different purposes" (48)

Classes and lesson style

The classes used were as grouped by the school, this meant that all the class sizes were different with 9R6 and 9R5 being considerably smaller than 9L1 and 9L2. This was due to 9R6 and 5 having more learning needs and as well as the class size affecting the learning of the students some students also had access to a teaching assistant. All these measures are done specifically to help the lower ability students learn so it is likely to have also affected the results of this study.

The time the students had to learn the lesson content also varied from practical to non-practical lesson as due to the nature of practical work time had to be allocated for this to be carried out and as a result students participating in the practical work had double the amount of lesson time when compared to the non-practical lessons. Students of different abilities require different amounts (62) of time to process and learn the same information it is possible that the improvement seen by 9R5 is directly related to the amount of time spent on the subject. Equally the lack of improvement for 9L1 could be due to the students not needing the extra time the practical provided them.

The teacher

A large number of studies have highlighted the teacher as being a key component in a successful practical lesson and stress that ineffective portrayal of scientific ideas due to poor teaching can often lead to students receiving no benefit from practical work, (28, 39). All the classes also had different regular teachers and it is possible that how the teachers introduced to study to the students affected the level interest in the topic.

The study

A big limitation of this study is its size. The study only covers one age group and therefore doesn't recognize that practical work may be more beneficial to students at specific times during their education. It is also small and as it was conducted in rural Cornwall does not portray a true representation of the general population.

Conclusions

The conclusion of this study is that whilst practical work undoubtedly teaches a range of physical skills unavailable through non-practical work it is not always the best approach for the learning of information. In the case of the top set students it can be concluded that there was no benefit seen in the retention and learning of information through practical work as there was no statistically significant difference between the post-test results for the non-practical and practical lesson styles. As the practical lesson was time inefficient (taking double the time to complete) it is in the interest of these students to participate in the non-practical lesson as the same grade will be achieved in less time.

For the lower ability students (bottom set), a greater increase in the learning of information was shown in the post-test result for the practical lesson. The practical lesson produced as mean post-test result of 6.29 whereas the non-practical lesson produced a mean post-test result of 3.7.These results were seen to be statistically different from one another when analysed using t-tests and echo the results of the Thompson (52). From this result it can be inferred that the less able students benefitted from the practical lesson style. Whether it is the process of learning practical that caused this improvement in their learning or the lesson style taking more time and therefore giving students to reflect and build ideas is inconclusive.

The data from the VARK questionnaire proved inconclusive in regards to whether students learn better when the type of lesson matches their learning style as there was no statistically significant difference found for the students' post-test results whose learning style was compatible with the lesson style when compared with the results of students whose learning style were not compatible with the lesson style they participated in. The only real conclusion to be draw was that multimodal learners appear to learn better than all other learners regardless of the type of lesson.

Further work

To produce more reliable results for the application of the conclusion to the population to be plausible a much larger study is need involving a greater number of students of varying ages, ethnicities and genders. The study should also be expanded to include a number of different chemistry practices to try to ascertain if there is some practical's that always produce more positive results when taught either none practically or practically.

To try and build a larger picture of the benefits of laboratory work an investigation in to how it helps to build key laboratory skills could be of interest.

The lesson style chosen for this study are by no means the only types of lesson available to students so comparison between practical work and other leading styles for example practical demonstration of participation would relate well to the material covered in this study.

To investigate how much the time taken to learn the lesson material affected the results a similar study could be produced with an equal amount of time spent on both the practical and non-practical lesson styles.

References

1. Zohar A, Nemet F. Fostering students' knowledge and argumentation skills through dilemmas in human genetics. Journal of research in science teaching. 2002;39(1):35-62.

2. Texas Uo. Importance of science education in schools Texas: University of Texas; 2017 [Available from: https://academicpartnerships.uta.edu/articles/education/importance-of-science-education.aspx. Accessed 09/04/18]

3. Duckett SB, Garratt J, Lowe ND, Nigel D. Key skills: what do chemistry graduates think. University Chemistry Education. 1999;3(1):1-9.

4. Steffe LP, Gale JE. Constructivism in education: Lawrence Erlbaum Hillsdale, NJ; 1995.

5. Beilin H. Piaget's enduring contribution to developmental psychology. Developmental psychology. 1992;28(2):191-204.

6. Piaget J, Cook M. The origins of intelligence in children: International Universities Press New York; 1952.

7. Oesterdiekhoff GW. Is a forgotten subject central to the future development of sciences? Jean Piaget on the interrelationship between ontogeny and history. Personality and Individual Differences. 2016;98:118-26.

8. Bower T. The development of object-permanence: Some studies of existence constancy. Perception & Psychophysics. 1967;2(9):411-8.

9. Chess S. The Plasticity of Human Development: Alternative Pathways. Journal of the American Academy of Child & Adolescent Psychiatry.17(1):80-91.

10. Kritchevsky D. Friedrich Goppelsroeder: Pioneer of paper chromatography. Journal of Chemical Education. 1959;36(4):196.

11. Goppelsroeder F. Ueber eine fluorescirende Substanz aus dem Kubaholze und über Fluorescenzanalyse. Zeitschrift für analytische Chemie. 1868;7(1):195-211.

12. Fanali S. Liquid chromatography Applications. Fanali S, editor. Amsterdam: Amsterdam : Elsevier; 2013.

13. BBC. Science dull and hard,pupils say 2005 [Available from: http://news.bbc.co.uk/1/hi/education/4100936.stm. Accessed 23/01/18]

14. Chemistry Rso. Public attitudes to chemistry 2015 [Available from: http://www.rsc.org/globalassets/04-campaigning-outreach/campaigning/public-attitudes-to-chemistry/public-attitudes-to-chemistry-research-report.pdf. Accessed 23/01/18]

15. Watts M, Ebbutt D. Sixth-formers' views of their science education, 11-16. International Journal of Science Education. 1988;10(2):211-9.

16. Hofstein A, Levy Nahum T, Shore R. Assessment of the Learning Environment of Inquiry-Type Laboratories in High School Chemistry2001. 193-207 p.

17. Education Df. The Importance of teaching. In: Education Df, editor. 2010. p. 1-95.

18. Lepi K. The Top 10 (And Counting) Education Systems In The World: Edudemic; 2014 [Available from: http://www.edudemic.com/learning-curve-report-education/. Accessed 23/01/18]

19. Coughlan S. Pisa tests: UK lags behind in global school rankings. 2016.

20. Education Df. Reducing teacher workload. In: Education Df, editor. 2018.

21. Education Df. Freedom and Autonomy for Schools National Association. 2018.

22. Micklewright J, Jerrim J, Vignoles A, Jenkins A, Allen R, Ilie S, et al. Teachers in secondary schools: evidence from TALIS 2013. 2014:15-6.

23. Fleming ND. The VARK Questionnaire, How Do I Learn Best? : VARK learn limited; 2018 [Available from: http://vark-learn.com/the-vark-questionnaire/. Accessed 10/03/18]

24. Fleming N, D, Mills C. Not Another Inventory,rather a Catalyst for Reflection. To Improve the Academy. 1992;11:137.

25. Fleming ND. Teaching and learning styles : VARK strategies. 2nd ed. ed. Christchurch, N.Z.: Christchurch, N.Z. : N.D. Fleming; 2006. p. 17

26. Dunn R. Understanding the Dunn and Dunn learning styles model and the need for individual diagnosis and prescription. Journal of Reading, Writing, and Learning Disabilities International. 1990;6(3):223-47.

27. Dunn R, Griggs SA, Olson J, Beasley M, Gorman BS. A Meta-Analytic Validation of the Dunn and Dunn Model of Learning-Style Preferences. The Journal of Educational Research. 1995;88(6):353-62.

28. Howard-Jones PA. Neuroscience and education: myths and messages. Nature Reviews Neuroscience. 2014;15(12):817-24.

29. Willingham DT, Hughes EM, Dobolyi DG. The scientific status of learning styles theories. Teaching of Psychology. 2015;42(3):266-71.

30. Leite WL, Svinicki M, Shi Y. Attempted validation of the scores of the VARK: Learning styles inventory with multitrait–multimethod confirmatory factor analysis models. Educational and psychological measurement. 2010;70(2):323-39.

31. Klein SE, Yager RE, McCurdy DW, Tafel M, Perez L, Bybee RW, et al. AN NSTA POSITION STATEMENT: The Laboratory is Vital in Science Instruction in the Secondary School. The Science Teacher. 1982;49(2):20-3.

32. Abrahams I, Millar R. Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. International Journal of Science Education. 2008;30(14):1945-69.

33. Della Porta DK, Micheal. Approaches and Methodologies in the Social Sciences: A Pluralist Perspective. Cambridge: Cambridge University Press; 2008. p. 8, 13-14,23

34. Goldenberg MJ. On evidence and evidence-based medicine: lessons from the philosophy of science. Social science & medicine. 2006;62(11):2621-32.

35. Gerring J, Yesnowitz J. A normative turn in political science? Polity. 2006;38(1):101-33.

36. Hofstein A, Lunetta VN. The laboratory in science education: Foundations for the twenty-first century. Science Education. 2004;88(1):28-54.

37. OCR. Elements, compounds and mixtures. Cambridge2016.

38. Branton R. The effects of teaching style on student learning of DNA. Louisiana: Louisiana State University; 2012. p. 1-67.

39. RSC. Outreach: crime scene chromatography: Royal society of chemistry; 2013 [Available from: http://www.rsc.org/learn-chemistry/resource/res00001607/outreach-crime-scene-chromatography. Accessed 03/02/18]

40. Burrows A. Chemistry 3 : introducing inorganic, organic and physical chemistry. Oxford: Oxford : Oxford University Press; 2009. p. 528

41. Fink L, D. A Self-Directed Guide to Designing Courses for Significant Learning 2003. p. 1-7.

42. Lari FS. The impact of using PowerPoint presentations on students' learning and motivation in secondary schools. Procedia-Social and Behavioral Sciences. 2014;98:1672-7.

43. Hulleman CS, Godes O, Hendricks BL, Harackiewicz JM. Enhancing interest and performance with a utility value intervention. Journal of Educational Psychology. 2010;102(4):880.

44. Starkings S. Quantitative Data Analysis with IBM SPSS 17, 18 & 19: A Guide for Social Scientists by Alan Bryman and Duncan Cramer2012. p. 173-185

45. Gastwirth JL, Gel YR, Miao W. The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. Statistical Science. 2009;24(3):343-60.

46. University KS. SPSS Tutorials: Independent Samples t Test; 2018 [updated 06/03/2018. Available from: https://libguides.library.kent.edu/SPSS/IndependentTTest. Accessed 12/03/18]

47. Boccadoro C. Outliers: Mutual Fund Observer; 2015 [Available from: https://www.mutualfundobserver.com/2015/06/outliers/. Accessed 17/04/18]

48. Hodson D. Learning Science, Learning about Science, Doing Science: Different goals demand different learning methods. International Journal of Science Education. 2014;36(15):2534-53.

49. Bok D. Our underachieving colleges: A candid look at how much students learn and why they should be learning more: Princeton University Press; 2009.

50. Hattie J. Teachers Make a Difference, What is the research evidence? 2003:3.

51. Hyunjin S, Norbert S. If It's Hard to Read, It's Hard to Do: Processing Fluency Affects Effort Prediction and Motivation. Psychological Science. 2008;19(10):986-8.

52. Thompson J, Soyibo K. Effects of Lecture, Teacher Demonstrations, Discussion and Practical Work on 10th Graders' Attitudes to Chemistry and Understanding of Electrolysis. Research in Science & Technological Education. 2002;20(1):25-37.

53. Dearden R. Instruction and learning by discovery. The Concept of Education (International Library of the Philosophy of Education Volume 17): Routledge; 2010. p. 100-14.

54. Kharb P, Samanta PP, Jindal M, Singh V. The learning styles and the preferred teaching—learning strategies of first year medical students. Journal of clinical and diagnostic research: JCDR. 2013;7(6):1089.

55. Dobson JL. Learning style preferences and course performance in an undergraduate physiology class. Advances in Physiology Education. 2009;33(4):308-14.

56. Peacock M. Match or mismatch? Learning styles and teaching styles in EFL. International Journal of Applied Linguistics. 2002;11(1):1-20.

57. Fleming N, Baume D. Learning Styles Again: VARKing up the right tree! Educational Developments. 2006;7(4):4.

58. Baykan Z, Naçar M. Learning styles of first-year medical students attending Erciyes University in Kayseri, Turkey. Advances in Physiology Education. 2007;31(2):158-60.

59. Fehrmann PG, Keith TZ, Reimers TM. Home influence on school learning: Direct and indirect effects of parental involvement on high school grades. The Journal of Educational Research. 1987;80(6):330-7.

60. Nelson G. The relationship between dimensions of classroom and family environments and the self-concept, satisfaction, and achievement of grade 7 and 8 students. Journal of community Psychology. 1984;12(3):276-87.

61. Mji A, Makgato M. Factors associated with high school learners' poor performance: a spotlight on mathematics and physical science. South African journal of education. 2006;26(2):253-66.

62. Bloom BS. Time and learning. American psychologist. 1974;29(9):682.