An Investigation into Utilising ‘Inference Squares’ in Tutoring Pupils to Use the Techniques Required in Answering PISA Questions

Daniel North
14/05/2017

1. Introduction

1.1. Background

This investigation gave multiple PISA (Programme for International Student Assessment) style science questions, as published by PISA (OECD, 2009), to two year 10 classes. Support on how to go about answering the questions (using inference squares) was given to only one of the classes.

The investigation was conducted in a Welsh-medium community comprehensive school in South Wales. Most of the pupils come from the local town and neighbouring villages, but many are from "the underprivileged wards" (Estyn, 2012)beyond the town. However only 9.8% of the student body are entitled to FSM (Free School Meals), as opposed to the Welsh national average of 17.4% (Welsh Government, 2017). Around 80% of pupils are from homes where English is the primary language, but all can "speak Welsh to a first language standard" (Estyn, 2012).

The school offers pupils the opportunity to sit their science GCSE qualifications through the medium of English. In this study, the English medium classes were chosen to take part in the study. These classes were selected because they were further ahead in their GCSE units, which allowed them more freedom to take part in the investigation. It was also decided that comparing both sets within one language medium was far more reliable than comparing results from both English and Welsh medium classes, because the question would then be given in two languages, adding to the uncertainty in the results.

1.2. Aims

The aim of this investigation was to gauge the success of inference squares in supporting pupils answering a PISA-style question. This topic was chosen in light of the 2015 PISA test results. In the latest PISA tests Wales scored significantly lower than the rest of the UK, which Wales' First Minister Carwyn Jones referred to as "uncomfortable reading" (Hume, 2016). The Welsh Assembly Government already has plans for improvement in place, including the new Donaldson curriculum (Donaldson, 2015). While the new curriculum proposed by Donaldson should improve pupil ability in answering PISA-style questions, other methods must still be investigated in preparation for the 2018 PISA tests (Donaldson, 2015; OECD, 2017).

2. Literature Review

2.1. PISA

Organised by the OECD (Organisation for Economic Co-operation and Development), the PISA test is an internationally agreed upon test, where candidates are assessed on their capability in "science, mathematics, reading, collaborative problem solving and financial literacy" (OECD, 2017). It is sat by a random selection of 15-year-old pupils, across many countries, economies and education systems, once every three years. All 35 OECD countries take part in the test, and several other countries/economies opt to take part in each test (Appendix I). The test formally assesses pupil performance and skills, while also collecting data on pupil attitudes and motivations. PISA does not prescribe to one curriculum and draws on elements from curriculums across the world, however it is not constrained by the need to find common denominators (OECD, 2017).

> *"A key objective of PISA is to inform and support education policy decision making within countries. A three-year cycle provides countries with timely information that includes data and analyses to consider the impact of policy decisions and related programs."* – (OECD, 2017)

As PISA is not focused on grading and comparing individual pupils, but assessing the ability of pupils across an entire education system, there is no need for every pupil to sit the same test (OECD, 2017). Therefore, several tests are written for PISA, with each pupil being randomly assigned only one test, but each country/economy collectively completing all tests. This gives the OECD a greater coverage of the ability of pupils and more efficient data on each country and economy involved (OECD, 2017). This is the cause of many criticisms of PISA, as the organisation uses a mathematical model developed by Georg Rasch to calculate the scores of pupils and countries taking part. Few people truly understand what happens within PISA's statistical analysis, but Professor Svend Kreiner who is a former pupil of Rasch, claims that for this model to work properly all PISA questions would have to be of equal difficulty to each other, and in every participating country, to create a comparable (and meaningful) scale (Stewart, 2013b). However, the Rasch model is widely used within educational testing, and fellow mathematician Trevor Bond is of the opinion that "if all items of a test are supposed to measure the same 'thing', why should it then make a difference which items we use when comparing persons?" (Bond and Fox, 2015).

PISA scores its test on a proficiency scale, individually developed for each subject area by the Rasch model (Stewart, 2013b). The subject areas assessed are reading, mathematics and science. The current scales have not changed since the 2006 PISA test, making the past four tests directly comparable (OECD, 2009). The scales are graded at six distinct levels, dependant on the candidate's proficiency in answering the question (OECD, 2009). For example, a pupil lacking the skills needed to answer the simplest questions would be graded as below Level 1, whereas a pupil with those skills would receive a higher level. These levels correspond to a set number of points, which are then averaged for all pupils across the country/economy and used to rank each country/economy involved in that year's study (OECD, 2017). The OECD does not give a collective score for all subject areas combined, but rather it takes the mean score of each subject and uses the score of that years focus subject area in ranking a country. The average score for an OECD country is 500 points, with a standard deviation of 100 points (OECD, 2017). Two in three pupils in OECD countries receive scores between 400 – 600 points (OECD, 2017). It is important to note here that the OECD member countries are like-minded and economically prosperous, and its scale is designed around these; so, when other countries and economies participate in the study they can often be at an advantage or disadvantage to the average scores (OECD, 2008).

> *"Pisa seems to be well on its way to being institutionalised as the main engine in the global accountability juggernaut, which measures, classifies and ranks students, educators and school systems from diverse cultures and countries using the same standardised benchmarks"* – (Meyer and Benavot, 2013)

The neutral wording of the above quote is an excellent example of how the institutionalisation of PISA as a primary force in global accountability can be interpreted in many different ways; is PISA's growing power a positive force in education, or do standardised benchmarks enforce an unfairly narrow, Western-centric, culturally specific criteria for success? There is a serious concern about the true validity of PISA rankings, as justly comparing the reading ability of pupils speaking vastly different languages is difficult, if not realistically impossible. To add to this difficulty PISA must find a fair comparison of simple measures, in over 70 education systems, in a way which all participating countries agree with (OECD, 2017). Michael Goldstein, head of the OECD schools division, admits that there is a risk that the questions used to ensure comparability across counties, can end up becoming the lowest common denominator, as these questions can still be read and answered differently in different countries (Stewart, 2013b; OECD, 2017). This is one of the largest flaws in PISA and the issue taken with many claims against it (Stewart, 2013b).

2.2. The 2015 PISA Test

The most recent PISA test was conducted in 2015, with the results being published on the 6th of December 2016. 71 countries took part in the test, and 13 test papers were written by PISA. It also marked the first time that the test was completed on computers, with an overwhelming majority of the countries opting for the new digital version (OECD, 2017). According to past PISA studies, the ability of pupils in the reading segment of the test is greater when presented as the first item on a test (as it was in the test in 2000). In contrast, the ability in the reading segment is at its worst when it was between the science and mathematics segment (as was mostly the case in the following years). To increase the effectiveness of test design, PISA focuses on one subject area every year, on a nine-year rotation for all three areas (OECD, 2017).

The 2015 test focused on scientific skills and literacy, with experts in OECD countries collaborating to write questions that required the abilities they considered to be the "most important for the future success of students in an increasingly science-based world." (OECD, 2017). The assessment was designed so that 86% of pupils were split into groups and only answered questions on two subject areas (science and reading, science and mathematics, and science and collaborative problem solving). One half of the pupils in each group answered the science questions first, and the other half of each group answered the science questions last. There were also two assessments where reading and mathematics were each the assessment focus, this was done to lower the uncertainty of the results between this test and previous years (OECD, 2015a).

In the 2015 PISA test, Wales once again ranked the lowest of the UK's constituent countries (Hume, 2016). While the UK is grouped together in OECD membership, and sits the test as one entity, each country volunteers a random selection of its pupils to take part in the test. The data is collected and released by PISA in a way that allows the government to break down the results of each individual country (Jerrim and Shure, 2016). Out of the 71 countries that took part in the 2015 test, the UK ranked 15th for science - which was the sole subject used for ranking that year (OECD, 2015a). In contrast to this, had the constituent countries of the UK been considered separately, Wales would have ranked 35th (OECD, 2016b)

With Wales obsessing over its PISA rankings in recent years, it is crucial to state that the OECD themselves warn that PISA rankings should not be overanalysed, as "large variation in single-ranking positions is likely" (Paton, 2013). Despite this regularly expressed warning, the PISA rankings are treated almost like a sports league by the media and politicians alike, with countries comparing rankings on single points, even though this difference is statistically insignificant. The OECD claims no accountability in this behaviour, stating that they cannot control the way countries use the results, but they are the organisation that continues to publish the results of its assessments in the form of crude rankings (Stewart, 2013a). It is important to note that the OECD explains that even significant differences are not necessarily indicative of a difference in the quality of education systems, but rather a result of a combination of many factors in a child's development up to the age of 15 (Tienken, 2014).

2.3. Wales in PISA 2015

In Wales, 140 schools (of 211 schools in total) participated in the test; from those 140 schools, 3,451 pupils (of approximately 182,500 pupils) participated in the test. That equates to 92% of the number of schools that were invited to participate, and 88% of the number of pupils who were 'randomly' selected within those schools (Jerrim and Shure, 2016). The pupil selection process is only quasi-random, as the pupils are selected to represent the current statistics for Wales' education system (Jerrim and Shure, 2016). This is an example of stratified sampling, in which the population is divided into discrete sub-populations corresponding to their values in the data intended to be analysed. After sorting the population into their relevant sub-populations, then a random sample can be taken from each sub-population, ensuring a diverse sample comparable to the true population (Särndal, Swensson and Wretman, 2003)

	2015 Statistics	2015 PISA Sample
% Free School Meal eligible	17.4	18
% achieving Level 2 threshold including English/Welsh and Maths	57.9	59
% Welsh medium instruction	20	17
% schools in Green Support Category	14	19
% schools in Red Support Category	11	9

Fig.1: Table comparing the actual 2015 education statistics for Wales against the 2015 PISA test's random sample (Welsh Government, 2015b; Lloyd and Newby, 2016).

Unfortunately, while this system does benefit from having an accurate representation of the population, it may also leave some large margins of error in the smallest sub-populations. These would primarily come from the Red Support Category schools and the Welsh medium schools, because while it may give a true reflection of the number of these schools in the population, it is also collecting a smaller sample size from these schools to achieve that (Jerrim and Shure, 2016).

Out of the 140 schools that participated in the 2015 PISA test, only 24 (or 17%) were Welsh medium schools (see Fig.1). This is a similar figure to the percentage of 15-year-olds who sat a Welsh First Language GCSE in 2015 (16%, approximately 5,500 pupils), however a further 30% of 15-year-olds sat a Welsh Second Language GCSE (Full Course) the same year. All Welsh pupils (in both Welsh and English medium schools) were given the option to sit either the English or Welsh language version of the PISA test. Of the 3,451 pupils participating, 575 pupils were studying Welsh as a first language GCSE, and 337 of those pupils chose to complete the PISA test in Welsh. There were two additional pupils who were not studying Welsh as a first language that chose to sit the Welsh test, bringing the total numbers to 339 Welsh-medium tests and 3,112 English-medium tests (Jerrim and Shure, 2016).

The data collected on these groups allows for a valuable comparison between the pupils involved. The majority of pupils who sat the Welsh exam were female (52%), whereas they were the minority on the English test (49%). The parents of the pupils who sat the Welsh test were more likely to be university educated (48%) and own more books, compared to the parents of those who sat the English test (where 36% went to university). There was also a significant difference in the number of pupils who spoke Welsh at home, with 38% of the Welsh test pupils doing so as opposed to 2% of the English test pupils (Jerrim and Shure, 2016).

	Test in English	Test in Welsh
Science (overall)	487	466
Mathematics	478	475
Reading	480	455

Fig.2: Table showing the mean scores in each category for pupils who sat the 2015 PISA test in English against pupils who sat the test in Welsh (Jerrim and Shure, 2016).

However, the most important comparison for this investigation is the test results, as can be seen in Fig.2. Apart from Mathematics, where the mean test results are similar, the pupils who sat the test in Welsh appear to be at a significant disadvantage to those who sat the test in English (just above and below the 5% level, where differences become statistically significant). The reason for this is unclear, but the 2016 study by UCL (University College London) noted that the differences in "background characteristics cannot explain the achievement gap between pupils who completed the test in English versus Welsh" (Jerrim and Shure, 2016). This was also the trend in the 2012 test; however, in 2012 the Welsh paper pupils scored significantly higher in mathematics, increasing the evidence that the Welsh medium pupils only suffer significantly in science and reading (National Assembly for Wales, 2013).

	Test in English	Test in Welsh
English medium education	485	N/A
Welsh medium education	495	466

Fig.3: Table comparing the Welsh science (overall) scores in the 2015 PISA test Between English and Welsh medium pupils, and whether they sat the English or Welsh paper. Note: N/A is listed due to insufficient sample size (Jerrim and Shure, 2016).

The Welsh medium pupils who sat the English paper scored significantly higher than the English medium pupils, and therefore even higher than their Welsh paper peers. There is a 29-point difference in mean scores between the Welsh medium peers, with a similar result for reading and no difference observed in mathematics (neither are included in Fig.3 as science is the overall score). The data suggests that many Welsh medium pupils could improve their scores significantly by choosing to sit the test in English (Jerrim and Shure, 2016). This is the kind of data PISA provides that allows countries to improve their education systems.

2.4. Success of PISA

Examples of the successful use of PISA results can be found in multiple countries across the world. In Germany, the PISA test in the year 2000 found that 20% of pupils did not reach the baseline level of reading proficiency, while the performance difference between the richest and poorest pupils were some of the largest amongst OECD countries. In response, Germany underwent an education reform, creating new education standards by which to compare pupil progress and making all-day schools a requirement. In 2015, Germany was above the OECD average for reading proficiency, with much of that improvement taking place in the country's lowest achieving and disadvantaged schools (OECD, 2016a).

Likewise, Brazil was the lowest-scoring country in the 2003 PISA test, with over half of the pupils scoring at or below the baseline proficiency level for mathematics, while fewer than 1% scored at the top level. Brazil consequently set itself the goal of reaching at least the OECD average by 2021. By the 2015 test, Brazil had increased its mathematics score by 34 points (now scoring 377 points), approximately equal to a full average year of schooling; and while it still must improve significantly to reach its goal, it is following the OECD guidelines to do so (OECD, 2016a).

A recent example with similar circumstance to Wales is Scotland. After the 2009 PISA test, Scotland was still behind the OECD average in mathematics, as it had been in previous years. With several reviews into the Scottish curriculum already underway, this result helped the report by Graham Donaldson pass and accepted into forming the new curriculum (Donaldson, 2010). The success of this curriculum is debateable, but it is currently only 4 years old (Kennedy and Doherty, 2012).

Considering PISA's successes, there are many who question its power, and its responsibility to wield it. PISA is the most influential international education study, with the power to make governments pay attention; as can be seen with the so called "PISA shock", experienced by Germany in 2001, Brazil in 2004, and across the UK since 2010. It is somewhat worrying that the organisation built to compare and guide the education systems of countries is now being used to justify large and controversial reforms of the systems it is attempting to measure (Stewart, 2013a). These reforms are often extreme responses to the important but questionable data PISA provides; in 2010 the Welsh Education Minister Leighton Andrews referred to the 2009 PISA result as a "systemic failure" and a "matter of absolute urgency", both overexaggerating the PISA results and ignoring Wales' results in other tests (Darren, 2010).

2.5. The Donaldson Curriculum

This continued through the 2012 PISA test, until in 2015 Graham Donaldson performed a review of the Welsh education system (Donaldson, 2015). This review was requested by the Welsh Assembly Government after a concerning OECD report in 2014, which raised issues such as PISA scores, continuity between Key Stages, and the lack of pupil engagement (OECD, 2014). The Welsh Government accepted the proposals of the Donaldson report, and is currently moving ahead with implementing a new curriculum guided by these proposals (Lewis *et al.*, 2015). Guided by these proposals "a network of pioneer schools" (Welsh Government, 2015a, p.2) is developing strategies of teaching the new curriculum across Wales, and in September 2018 their strategies will be made available across Wales, along with a government support network to assist other schools in making the change. Schools will then have until September 2021 to implement the changes in their curriculum, at which point the assessments will be in place (Welsh Government, 2015a).

> *"The school curriculum in Wales should be defined as including all of the learning experiences and assessment activities planned in pursuit of agreed purposes of education"* – (Donaldson, 2015, p.115)

The above mentioned by the Donaldson report are to help young people develop as ambitious, capable learners; enterprising, creative contributors; ethical, informed citizens; and healthy, confident individuals (Donaldson, 2015, p.29). These aims should prepare young people to be ready to learn throughout their lives; ready to play a full part in life and work; and be ready to lead fulfilling lives as full members of society (Donaldson, 2015, p.29). Similar aims are already stated in the Education Act 2002 (Her Majesty's Government, 2002), but these only provide a general guide to the curriculum aims, and according to Donaldson are unlikely to have a meaningful impact on teaching or learning (Donaldson, 2015, p.22). Recently a trend has emerged across other education systems of building curriculums around a set of key skills and the development of their potential for multi-disciplinary application (Donaldson, 2015). The aims proposed by Donaldson have been constructed so that they can directly influence decisions about the curriculum, its aims, and how they will alter over the many years this curriculum may be in place (Donaldson, 2015). After all, one of the major criticisms of the current curriculum is how it has been reviewed and adjusted several times since the 1988 Education Reform Act, but still largely consists of now outdated aims (Lewis *et al.*, 2015).

One of the aims of the new Donaldson curriculum is to group individual subjects into "areas of learning" (Donaldson, 2015, p.15), as has been the trend in many countries in recent years (including Wales' foundation phase). These areas of learning "should promote and underpin continuity and progression and help to make the structure easier to understand" (Donaldson, 2015, p.39); this is to be achieved by removing the barriers around individual subjects and allowing skills, ideas and themes to flow between them, creating a balance of each area across the pupil's education. By creating these common links between subjects, the application of information pupils learnt in another subject becomes more natural, and this is one of the core themes in PISA questions (Donaldson, 2015; OECD, 2017). It is therefore important that schools (especially 'pioneer schools') begin to develop new teaching techniques that help pupils effectively apply information learnt across many subjects, as new assessment styles (and of course the 2015 PISA test) are likely to require that the pupil can apply what they have learnt elsewhere to the question at hand.

2.6. Inference Squares

One such technique for encouraging the application of transferable skills is known as an inference square (Appendix VI). This technique has been used in history classes for many years, and appears to be a good method for encouraging pupils to consider the problem linearly. It is a technique that the OECD themselves have included in past PISA papers, and is also referenced by the Welsh Government as an example resource (OECD, 2009, cited in Welsh Assembly Government, 2009).

The inference square is a set of three or more concentric squares, and each one asks a deeper question about what can be inferred from the source material. The main example consists of the centre square asking what the source is telling you, the middle square asks what else you can infer from what the source has said, and the outer square asks what questions the source has raised but has left unanswered. The pupil would then answer each of these questions within the boxes. This technique requires the pupil to read through the source carefully, think about what it is saying and extract the relevant information (Huddleston, 2014).

PISA strives to improve the quality of education systems worldwide, for better or worse (OECD, 2017). However, one of the key problems with both the criticism and support of PISA is that there is no universal consensus on what exactly education system quality is or means; it is entirely down to opinions of those forming the curriculums and education policies. Without a clear definition of what constitutes a high-quality education system, it is down to the individual's belief in PISA and their methods as to whether it is possible to use its assessments as a measure of quality (Hanberger, 2014). However, this paper will not be focusing on the importance of PISA tests per se, or on the suitability of PISA as a measure of education quality, but rather on the pupils' ability to answer the questions that they set.

3. Methodology & Ethics of Educational Research

3.1. School Policy

Conducting any kind of research within a school requires a strict adherence to the school's own policies, and a careful consideration of the potential effect on the pupils' learning. Considering that the research involved GCSE classes that were approaching their examinations, there was considerable pressure to finish the topics that must be covered. It was therefore requested that only subject-relevant questions that served to further pupil understanding were used in the investigation. However, "Ysgol X" has no unique or dedicated policies regarding its teachers performing educational research, and therefore only requires that the researcher follows the Data Protection Act (1998) and all ethical guidelines advised by the BERA (British Educational Research Association) (Her Majesty's Government, 1998; BERA, 2011).

3.2. BERA Guidelines

The ethics of educational research is a very complex topic that requires much consideration, as every action that could potentially impact the learning process must be justifiable. The guidelines set forward by BERA are always under review, and attempt to encompass the many different subjects, philosophies, theories and methodologies that are part of the field of educational research (BERA, 2011). All guidelines are set to ensure research is conducted in an environment with ethical respect toward the person(s), knowledge, democratic values, the quality of the research and academic freedom. Similarly, while conducting any research, the researcher has a responsibility to participants, the research community, educational professionals and the public (BERA, 2011). The guidelines relevant to this research are to ensure that all participants consent to the research, that the investigation is open and disclosed with all personal data remaining private, and that all participants are given the right to withdraw; none of the other guidelines factor into research within the bounds of normal school operation (BERA, 2011).

3.3. Data Protection

During this investigation, pupil results for PISA questions were kept digitally in an Excel spreadsheet. This was a potential risk if the data was accessed by another individual. As per BERA guidelines, all data was encrypted within the spreadsheet to ensure confidentiality, which was done using a reference key. Once the PISA questions had been completed and marks calculated, the confidential data was made anonymous by randomising all the codes assigned to pupils; this way all research analysis done on the data was also anonymised (BERA, 2011). The file was stored on the school's internal network, where the only potential access was by teaching staff. The original file was kept on the network after the data was anonymised to ensure it remained accessible in compliance with the Data Protection Act (1998); likewise, the reference key was kept securely off the network (Her Majesty's Government, 1998).

3.4. Adverse Effects

When conducting research in a school setting, a particular concern is the research having an adverse effect, such as being an interruption to pupil progress or the PISA questions proving an ineffective learning method. This could also be a risk to the practicing teacher of the class, as they are then required to play catch up before the examination period. A back-up lesson plan was always prepared in case the PISA question was too difficult to answer. Further to this, it was made clear to all pupils involved that should they feel any detriment to their usual work as a result of these tests, it should be brought to the attention of a teacher (BERA, 2011). All BERA guidelines were followed during this investigation to ensure that any risks towards the pupil, the school and university (through partnership) were both minimal and justifiable.

3.5. Disclosure and Consent

It was important that that the investigation be fully disclosed, with all participants giving their informed consent and being given the right to withdraw. All pupils that were asked to be involved in the study were considered to be of an appropriate age, and have the intellectual capability, to form their own views on the investigation, and therefore offer their own consent to participate (BERA, 2011). The method in which informed consent was provided was to firstly explain the uses and methods of PISA, its importance within the new curriculum, and the aims of this investigation; all this was done verbally during a lesson, and then an information sheet was given to all pupils (Appendix IV). It was made clear that taking part did not require any additional work on their behalf, as all pupils would be required to sit the tests anyway (as a part of normal school operation), and there was no need to revise for the tests. The pupils were given a few days to make their decisions (until the following lesson), before they were then given a consent form to sign; either consenting or not consenting to their data being used in this investigation. These forms, included in Appendix IV & V, outlined the research proposed, and exactly to what the pupils were consenting.

It is crucial when working with minors that all documents are written in suitable language. In this investigation, it was important that both the information sheet and consent form were fully understood by every 15-year-old pupil that read them. Complex academic language is unnecessary to explain the concept of PISA and the aims of this investigation. Many argue that the core tenet of educational research is respect toward the participant, and this is imperative when the participants are minors (Coombs, 1998). Everyone has the right to understand, further enquire or dispute any contract that they enter into; therefore, when your participants are minors, it is pointless to use overly complex language as long as the information is clear (United Nations, 1948). The consent form serves as more than a trigger to begin the research, but also as a form of security for the participants that the research will not overstep any boundaries laid out in the form. By signing the consent form, both parties are entered into a contractual agreement, where the participant has agreed to 'supply' data for research, and the researcher has agreed to commit to the ethical standards proposed within the contract (Panza and Potthas, 2010).

3.6. The Research

During the research itself, all work was done within the bounds of normal school operation. The PISA questions were chosen on topics relevant to what the pupils had been studying as part of the GCSE curriculum, and answered during lessons. The only breaking of normal school operation is that the pupils never had their marks returned to them. That is not to say the pupils weren't allowed to request their own marks, which they have the right to do (BERA, 2011). This was decided because PISA questions are harshly marked, and many pupils received what would be perceived as a low mark in a GCSE paper. Bad grades tend to have a detrimental effect on pupils, so giving them their results for a test they will never formally sit is largely pointless, especially when they will never receive constructive AfL (Assessment for Learning) feedback on how to improve because of the investigation's method of comparison (Children Schools and Families Committee of Her Majesty's Government, 2008). It is worth noting that, usually, pupils would never receive their individual PISA scores (OECD, 2017).

4. Evaluation

4.1. The Investigation

This investigation gave two PISA questions to two year 10 classes. The classes chosen were the two highest ability sets through the English medium pathway; these will be referred to as S1 and S2, for higher and lower respectively. These two sets are both mixed ability classes, and are the only English medium sets undertaking a GCSE in science. This means that the highest ability pupil in S1 is far more scientifically able than the lowest ability pupil in S2, in terms of the GCSE grading system. However, considering the information presented within the UCL study, it was decided that two sets with a broad range of ability would be more comparable than two sets with similar ability but answering in different languages (Jerrim and Shure, 2016).

To assist pupils in answering the PISA questions, they were taught how to use an inference square. The inference square given to them consisted of three questions; "what does the question tell you?", "what do you already know about this?" and "how does this relate to what the question has asked?" (Appendix VI). The method is simple; before answering the question itself, answer each question on the inference square. The theory is that this will help pupils organise their thoughts, and only use the relevant information within a question (Huddleston, 2014). This method of using an inference square to aid in answering the PISA questions was only taught to one class, to allow its effect to be measured. The S2 pupils were chosen for this as they were the ones expected to make the largest increase in performance, and had they only been given the same verbal feedback that S1 received, their performance may not have increased at all.

Each class was given two PISA questions a fortnight apart, which covered six lessons in total; one lesson for each question and four lessons in between. The pupils were not given any assistance or warning with the first question, which was taken as the control question to benchmark each pupil. The first question (Q1) given was titled 'Energy Positive' and focused on the topic of convection in a carbon neutral house, which was set to coincide with the end of the GCSE unit 'Making Use of Energy', which includes heat transfers (Appendix II; OECD, 2009; CBAC (WJEC), 2016). After these were marked, each class received feedback on how they had answered and what the marking scheme required in each answer. The pupils were also allowed to see their papers, which were only marked with a tick for each point the marking scheme required. Over the following four lessons, both classes resumed their usual work, now on the unit 'Features of Waves' (CBAC (WJEC), 2016). During these lessons, S2 was taught how to use an inference square and given several opportunities to do so. The second question (Q2) given was titled 'Ultrasound' and focused on the properties of sound waves and x-rays, again tying into their GCSE work (Appendix III; OECD, 2009). During this test S2 were also given a printed inference square each.

4.2. Results

PISA questions are graded out of only a few marks, with each desired answer usually being either correct (2 marks), partially correct (1 mark) or incorrect (0 marks) (OECD, 2009). The first question was marked out of 9 marks, which offered 1 mark for part one, 4 marks for part two, and 4 marks in total for part three. The second question was marked out of 11 marks, which offered 4 marks for part one, 4 marks for part two, and 3 marks in total for part three. These marks were calculated as percentages, and are listed below in Fig.4. All averages in this investigation are the arithmetic mean.

S1			S2		
Pupil	**Q1 (%)**	**Q2 (%)**	**Pupil**	**Q1 (%)**	**Q2 (%)**
S1A	44	55	S2A	44	82
S1B	33	36	S2B	22	45
S1C	44	45	S2C	22	36
S1D	89	100	S2D	44	36
S1E	78	73	S2E	44	55
S1F	22	36	S2F	33	36
S1G	78	82	S2G	22	27
S1H	33	36	S2H	22	64
S1I	44	36	S2I	22	45
S1J	67	64	S2J	56	73
S1K	44	55	S2K	11	36
S1L	100	91	S2L	44	45
S1M	33	36	S2M	0	27
S1N	22	64	S2N	11	27
S1O	33	45	S2O	33	36
S1P	22	36	S2P	22	64
S1Q	67	55			
S1R	56	55			
S1S	22	55			
S1T	56	73			
S1U	56	36			
S1V	22	45			
S1W	22	36			
Average	47	54	Average	28	46

Fig.4: Table containing the marks of all pupils for both Q1 and Q2, separated by class and including class averages.

Fig.4 gives us the results for both questions of each pupil, separated into their respective sets. The pupils were assigned a random letter, with their set as a prefix, to keep their work confidential. It is firstly evident from these results that S1 did far better in the first question than S2, with the marks awarded being quite varied, but including some very high marks. In S2 all the marks are fairly similar at around 22%, without many higher marks. The highest and lowest marks attained in Q1 by S1 pupils were 100% and 22%, whereas in S2 they are 56% and 0%, so both ends of the pupil ability spectrum (in answering PISA style questions) can be seen in these results. However, it can also be seen that marks improved considerably in the second question. A better visual representation of the difference between the average marks for each class can be seen in Fig.5.

In S1, the average mark is closer to the lowest mark than the highest mark for both questions; this clearly shows that the majority of pupils got marks below this average, with only a few achieving very high marks (which is confirmed through the highest and lowest marks, and Fig.4). This is also true for S2 Q2, however in Q1 the average was closer to the highest mark than the lowest mark, meaning that the majority of pupils scored above the average. This does not mean that S2 did worse in their second test, but rather that the lowest achieving pupils in Q1 made the greatest improvement in Q2. As can be seen in Fig.5, there is a much greater improvement in the marks of S2 than there are in S1.

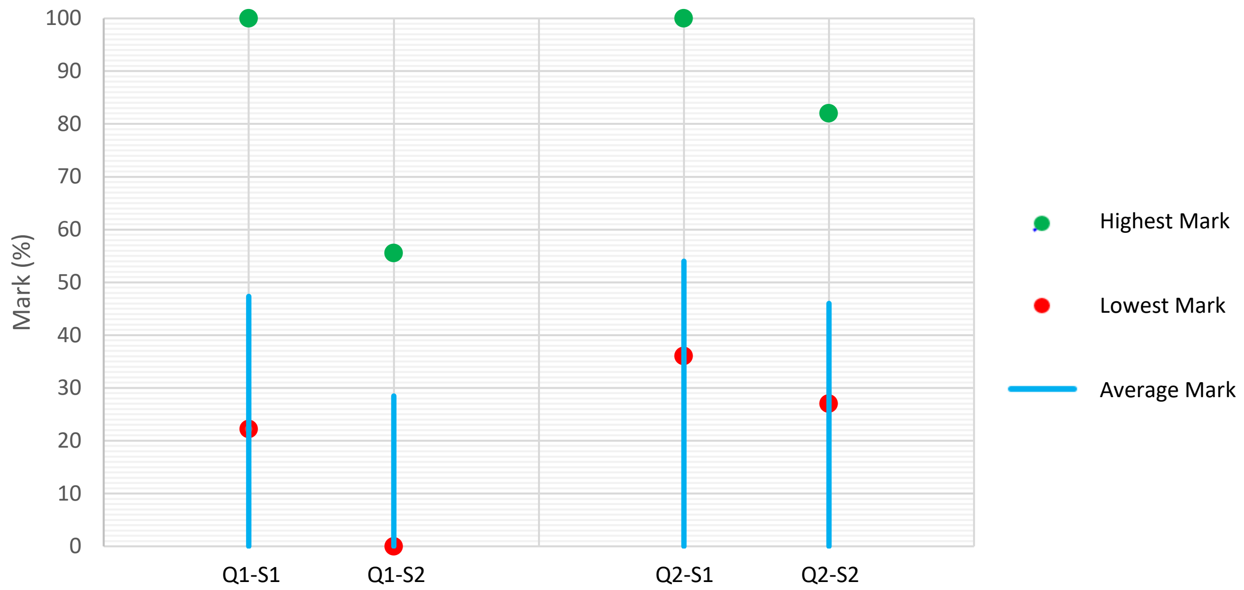

Fig.5: Graph displaying the class averages for each question, including both the highest and lowest mark achieved in that class.

4.3. Improvement

S1		S2	
Pupil	**Improvement (%)**	**Pupil**	**Improvement (%)**
S1A	10	S2A	37
S1B	3	S2B	23
S1C	1	S2C	14
S1D	11	S2D	-8
S1E	-5	S2E	10
S1F	14	S2F	3
S1G	4	S2G	5
S1H	3	S2H	41
S1I	-8	S2I	23
S1J	-3	S2J	17
S1K	10	S2K	25
S1L	-9	S2L	1
S1M	3	S2M	27
S1N	41	S2N	16
S1O	12	S2O	3
S1P	14	S2P	41
S1Q	-12		
S1R	-1		
S1S	32		
S1T	17		
S1U	-19		
S1V	23		
S1W	14		
Average	7	Average	18

Fig.6: Table containing the change in each pupil's mark between Q1 and Q2, including class averages.

The improvement of each pupil is given in Fig.6, and clearly S2 made a far greater increase in performance than S1. The biggest improvement in both sets was a 41% increase in marks, which was attained by one S1 pupil and two S2 pupils. The lowest decrease in performance was -19% by an S1 pupil; it is important to note that while seven S1 pupils received lower marks on the second question, it happened to only one S2 pupil. Overall, the S2 pupils improved their results by an average of 18%, or 11% more marks than the S1 pupils. The varied results of the S1 pupils, which are mostly small increases, brings into question whether there was any statistically significant improvement in performance. However, the large increases in the results of S2 pupils are clearly an improvement from the first test.

There is of course the possibility that Q2 was an easier question than the first. Given that the question had the same number of sub-questions but more marks available, it may have been easier to score marks on the paper; however, PISA was looking for more in-depth answers to those questions, potentially making them harder. It is also important to consider that both questions were chosen for their relevance to recently completed work, this means that for both questions, unlike a normal PISA test, the work was fresh in their memory, possibly narrowing the gap between the relative difficulties of each question. Additionally, if Q2 truly was an easier question, it can be assumed that S1 should have performed far better as the top set class, whereas in reality there was very little improvement at all.

4.4. Gender

	S1		S2	
Average	Male	Female	Male	Female
Q1	51	37	29	28
Q2	58	44	50	42
Improvement	7	7	21	14

Fig.7: Table containing the average marks and improvement for male and female pupils, separated by class.

From Fig.7 we can see that the male pupils on average outperformed the female pupils in each class on both questions. This fits with OECD data on gender difference in science, where male pupils are on average over 10 points ahead of female pupils in the UK on PISA science questions (OECD, 2015b). Answering Q1, all S1 pupils were significantly ahead of S2 pupils, but on Q2 there was only a 2% difference between the female pupils in S1 and S2. The average improvement of S1 pupils is 7% for both male and female pupils, whereas in S2 it is 21% and 14% respectively, meaning they improved two and three times more than the S1 pupils.

4.5. PISA Scores

Unfortunately, the OECD keeps the exact model it uses for calculating the difficulty scales and scores a secret (Stewart, 2013b). This makes it impossible to convert the marks and percentages obtained from this investigation into PISA points directly. However, as the pupils who sat this test were Welsh medium pupils sitting the English medium test, we can assume their average score was similar to that of Wales at 495 points (Fig.3). The average score for the top 10% of pupils in Wales was 602 points in 2015, and 368 points for the bottom 10% of pupils (Jerrim and Shure, 2016). By grouping both sets together, and calculating the average percentage for the top and bottom 10% of pupils, we can equate a percentage to each of these scores.

	Q1		Q2	
Average	Percentage (%)	PISA Score	Percentage (%)	PISA Score
Top 10%	86	602	89	603
All	40	495	51	535
Bottom 10%	14	368	29	450

Fig.8: Table containing the approximated PISA scores equivalent to average marks in both questions. Polynomial equation used $y = -0.04x^2 + 6.84x + 280.78$

With 39 pupils taking part in the investigation, the top and bottom 10% was taken from the four best and worst marks in Q1. Once the average marks were calculated, they were equated with their PISA score analogues and graphed. With only three data points, the graph has high margins of error, but without more PISA data available it is unable to be calculated or improved. A polynomial trendline was then calculated for these three data points, as given in Fig.8. This polynomial equation was then used to calculate the PISA scores equivalent to the Q2 averages.

Of course, one of the large issues with this model is that the questions the pupils answered were different each time. This was previously addressed as one of the issues with PISA's own analysis, though PISA themselves claim that their analysis is sound (Schleicher, 2013; Stewart, 2013b). As the current scales have not changed since 2006, and the OECD average is intended to be 500 points before the test takes place, then it can be assumed all tests since 2006 have been written to the same difficulty (where the questions used in this investigation were released in 2009) (OECD, 2009). PISA's exact methods for judging difficulty are unknown, but to keep the same measurement and average, the difficulties must be relatively similar. However, these difficulties may have been judged per section, with each question being of varying difficulty. As PISA changes the test focus and structure every three years, it is difficult to know with any certainty how the questions used compare to the 2015 test. Therefore, the model proposed above will be used to judge the success of the inference squares, but the results must be treated with caution as no effective error analysis can be done without greater access to PISA's data.

	S1		S2	
	Q1	Q2	Q1	Q2
Average	523	545	446	519

Fig.9: Table containing the approximated PISA scores for each set in both questions, using the polynomial from Fig.8.

Interestingly, using this method, Fig.8 suggests that the top 10% had very little improvement on the test, where the bottom 10% increased by almost 80 points. This idea fits when the top 10% are all S1 and the bottom 10% are all S2 (for Q1). The top and bottom 10% were not calculated for the individual classes, as they would have been the results of only two pupils each, which would have been very unreliable. Fig.9 gives the PISA scores for each set on both questions, and it is once again clear that S2 made the greater improvement. It is possible that S1 was already above the Welsh and UK average with their results in Q1, but it may be improbable, and this brings the reliability of the model into question. However, the Q1 score of S2 is reasonable, approximately 50 points below the Welsh average, making their improvement of 73 points a very impressive achievement.

5. Conclusion

The results of this investigation suggest that that the inference square had a positive impact on pupil ability to answer PISA-style questions. In just two weeks of using an inference square, the pupils of S2 improved their ability in answering PISA questions by an average of 18%. As mentioned above, this is 11% more than the pupils of S1. Such an increase, especially by a lower ability class, is certainly connected to the techniques used. There is still the possibility that Q2 was an easier question, but this remains unlikely, and without conducting further investigation on the difficulty of the questions given, it is difficult to know exactly how difficult each question was. In terms of PISA points, this equates to an improvement of 73 points for S2, compared to 22 points for S1 (Fig.9). This puts the S2 average for Q2 at just 4 points (or 1%) behind the S1 average for Q1, or in other words the lower ability S2 pupils were able to improve to the initial performance level of the S1 pupils. This is a very impressive improvement if true.

There is also the evidence that all but one pupil in S2 improved in the second question. S1 had a greater variation in the second set of results. This could be a result of some pupils getting unusually high marks in the first question and being unable to reproduce work of that quality. However, it is more plausible that the inference squares had a positive impact on the pupils' ability to organise their thoughts. Fig.4 gives the results of all pupils, and when looking through the Q2 results for each class, it is evident that S1 contains varied results, whereas S2 are more concentrated with only a few outliers.

It is possible that using the inference squares is more effective for male pupils. This claim could be supported by the fact that both the male and female pupils improved by the same average percentage in S1, where no inference squares were used. It could also be that the female pupils were already struggling in the subject due to their lower ability in science, and therefore they made less of an improvement even when given a technique to focus their ideas (OECD, 2015b). It could also be argued that the female pupils merely believe that they have a lower ability in science, but unfortunately both possibilities have the same outcome (OECD, 2015b). Additionally, recent studies have shown that girls around 15 years of age lose confidence and suffer from anxiety in STEM subjects, even when they poses a natural ability in them, which is likely an influencing factor (Adams, 2015; OECD, 2015c).

Considering the evidence gathered during this investigation, from the both questions and the analysis done on the results, it is evident that the inference square does have a measurable positive impact on the ability of all pupils to answer PISA-style questions. Providing a tool to help pupils focus that which they already know, in a test which requires the application of skills across multiple subjects, surely has a positive impact on pupil ability to answer those questions. The marks of pupils in both sets make the conclusion clear, and even with an exaggerated model of PISA scores, the effect this would have on pupils during a PISA test is unmistakable. Further investigation could pinpoint the pupils that this would most benefit, but this investigation concludes that the use of an inference square has a positive impact on pupils while answering PISA-style questions, and has the potential to significantly improve their marks.

6. Limitations

Further investigation could significantly improve on the findings of this paper. One such improvement would be to test the male and female pupils of both sets to see if the same trend can be found in the pupils with greater scientific ability. One class is not large enough to provide sufficient data to prove that the inference square is of greater use to male pupils. Similarly, the inference square was only given to a lower ability class in this investigation, and further work should strive to find whether this method is of greater use to a certain demographic. So, by giving the inference squares to a large range of pupils with varying ability, a greater understanding of the squares effect on improvement can be achieved.

Another worthwhile venture would be to investigate the effect of this technique in both English and Welsh medium classes. As a bilingual country, it is important to understand how our teaching methods work in both languages, and how to adapt or create new methods to compensate for any disparity. I propose that this investigation be repeated, but using both the English and Welsh streams in Ysgol X. By once again using both top sets, the same investigation can be run and any improvement compared between the two languages. This could also be merged with investigating the inference square over multiple abilities, by splitting each set in two and giving the inference square to only half the class.

This investigation had to be completed in a very short timeframe, which meant that only two questions could be completed, giving only one data point for improvement and a control. While this particular experiment showed an improvement in marks, the next question could have reversed that entirely. It is very important in terms of reliability that any further research be carried out over a longer period of time, and include many more questions.

PISA's insistence on keeping their methods and data secret, and only releasing information they believe is relevant is unfortunate. If there was more available data, a far more accurate model of how marks compare to PISA points could be created. Alternatively, more questions and/or pupils could give a better estimate of PISA's analysis methods.

7. References

Adams, R. (2015) 'Girls lack self-confidence in maths and science problems, study finds', *The Guardian*. Available at: https://www.theguardian.com/education/2015/mar/05/girls-lack-self-confidence-maths-science-oecd-school-engineering. (Accessed: 15 May 2017)

BERA (2011) 'Ethical Guidelines for Educational', *British Educational Research Association (BERA)*

Bond, T. and Fox, C. M. (2015) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Third. Taylor & Francis.

CBAC (WJEC) (2016) 'WJEC GCSE in SCIENCE (DOUBLE AWARD)'. CBAC (WJEC).

Children Schools and Families Committee of Her Majesty's Government (2008) *Testing and Assessment: Third Report of Session 2007-08, Vol. 2: Oral and Written Evidence*. Her Majesty's Stationery Office.

Coombs, J. R. (1998) 'Educational Ethics: Are We On The Right Track?', *Educational Theory*, 48(4), p. 555.

Darren, E. (2010) '"No excuses" for Pisa flop', *The Times Educational Supplement Cymru*, (334), p. 1. Available at: http://search.proquest.com.ezproxy.uwtsd.ac.uk/docview/847581367/abstract/57C4BD704E35499FPQ/1?accountid=130472. (Accessed: 26 April 2017)

Donaldson, G. (2010) *Teaching Scotland's Future: Report of a review of teacher education in Scotland*.

Donaldson, G. (2015) *Successful Futures*.

Estyn (2012) *A report on Ysgol X*.

Hanberger, A. (2014) 'What PISA Intends To and Can Possibly Achieve', *European Educational Research Journal*, 13(2), pp. 167–180. Available at: http://umu.diva-portal.org/smash/get/diva2:713996/FULLTEXT01.pdf. (Accessed: 24 April 2017)

Her Majesty's Government (1998) *Data Protection Act 1998*, *Her Majesty's Stationery Office*. Available at: http://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpga_19980029_en.pdf. (Accessed: 17 April 2017)

Her Majesty's Government (2002) *Education Act 2002*. Available at: http://www.legislation.gov.uk/ukpga/2002/32/pdfs/ukpga_20020032_en.pdf. (Accessed: 15 May 2017)

Huddleston, J. (2014) *Inference Squares*. Available at: http://jeh651.wixsite.com/jwslearning/inference-squares (Accessed: 16 April 2017).

Hume, C. (2016) 'Pisa: Wales still worst in UK in world education tests', *BBC News*, 6 December. Available at: http://www.bbc.co.uk/news/uk-wales-38208738. (Accessed: 16 April 2017)

Jerrim, J. and Shure, N. (2016) *Achievement of 15-Year-Olds in Wales: PISA 2015 National Report*.

Kennedy, A. and Doherty, R. (2012) *Professionalism and partnership: panaceas for teacher education in Scotland?* Available at: https://pure.strath.ac.uk/portal/files/11697031/JEP_Final_version_accepted_Apr_2012.doc. (Accessed: 26 April 2017)

Lewis, H., Burns, A., Thomas, S., Jones, A., Thomas, A. and Griffiths, J. (2015) 'Welsh Government - Record Of Proceedings 30/06/15', in *Statement: The Welsh Government Response to the Donaldson Review*. Cardiff: Welsh Government.

Lloyd, R. and Newby, A. (2016) *Key Education Statistics Wales 2016*. Available at: http://dera.ioe.ac.uk/26153/1/160427-key-education-statistics-2016-en.pdf. (Accessed: 17 April 2017)

Meyer, H. D. and Benavot, A. (2013) *PISA, Power, and Policy: the emergence of global educational governance*. Symposium Books (Oxford Studies in Comparative Education).

National Assembly for Wales (2013) *Programme for International Student Assessment (PISA) 2012: Research Note*. Available at: http://www.assembly.wales/Research Documents/Programme for International Student Assessment (PISA) 2012 - Research note-09122013-252334/rn13-030-English.pdf. (Accessed: 05 May 2017)

OECD (2008) 'The OECD'. OECD. Available at: http://www.oecd.org/newsroom/34011915.pdf. (Accessed: 25 April 2017)

OECD (2009) *Take the Test: Sample questions from OECD's PISA assessments*. OECD.

OECD (2014) *Improving Schools in Wales: An OECD Perspective*.

OECD (2015a) *PISA 2015 Results in Focus*, *OECD*.

OECD (2015b) *The ABC of Gender Equality in Education*, *OECD Publishing*.

OECD (2015c) 'Trends Shaping Education 2015 Spotlight 7 - Gender Equality', (2014), pp. 1–12.

OECD (2016a) 'How does PISA help shape education reform?' OECD. Available at: https://www.youtube.com/watch?v=-xpOn0OzXEw. (Accessed: 19 April 2017)

OECD (2016b) *Results from PISA 2015: Country Note: United Kingdom*. Available at: https://www.oecd.org/pisa/PISA-2015-United-Kingdom.pdf. (Accessed: 19 April 2017)

OECD (2017) *PISA: FAQ*, *OECD*. Available at: http://www.oecd.org/pisa/pisafaq/ (Accessed: 19 April 2017).

Panza, C. and Potthas, A. (2010) *Ethics For Dummies*. Wiley Publishing Inc.

Paton, G. (2013) 'International school league tables "utterly wrong"', *The Telegraph*, July. Available at: http://www.telegraph.co.uk/education/educationnews/10191157/International-school-league-tables-utterly-wrong.html. (Accessed: 24 April 2017)

Särndal, C. E., Swensson, B. and Wretman, J. (2003) *Model Assisted Survey Sampling*. Springer New York (Springer Series in Statistics).

Schleicher, A. (2013) 'Attacks on Pisa are entirely unjustified', *The Times Educational Supplement Scotland*, (2328), p. 33. Available at: http://search.proquest.com.ezproxy.uwtsd.ac.uk/docview/1437083190/abstract/97FEF500847148A2PQ/1?accountid=130472. (Accessed: 26 April 2017)

Stewart, W. (2013a) 'How Pisa came to rule the world', *The Times Educational Supplement Scotland*, (2346), p. 18. Available at: http://search.proquest.com.ezproxy.uwtsd.ac.uk/docview/1492828680/abstract/F05B68CAA0ED4293PQ/1?accountid=130472. (Accessed: 26 April 2017)

Stewart, W. (2013b) 'Is Pisa fundamentally flawed?', *The Times Educational Supplement Scotland*, (2327), p. 16. Available at: http://search.proquest.com.ezproxy.uwtsd.ac.uk/docview/1433969860/abstract/7930062732F24088PQ/1?accountid=130472. (Accessed: 26 April 2017)

Tienken, C. H. (2014) 'PISA Problems', *AASA Journal of Scholarship and Practice*, 10(4), pp. 4–18.

United Nations (1948) *Universal Declaration of Human Rights*. United Nations. Available at: http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf. (Accessed: 05 May 2017)

Welsh Assembly Government (2009) *Task Sheet 1: Source Square*, *Welsh Government: Learning Wales*. Available at: http://learning.gov.wales/docs/learningwales/publications/130612-module-8-task-sheet-1-en.pdf. (Accessed: 12 May 2017)

Welsh Government (2015a) *A curriculum for Wales – a curriculum for life*.

Welsh Government (2015b) *National School Categorisation System 2015*. Available at: http://gov.wales/docs/dcells/publications/150130-national-school-categorisation-system-en-v2.xlsx. (Accessed: 29 April 2017)

Welsh Government (2017) *Pupil Level Annual School Census*, *StatsWales*. Available at: https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Schools-Census/Pupil-Level-Annual-School-Census. (Accessed: 16 April 2017)

8. Appendices

Appendix I: The OECD Countries

OECD MEMBER COUNTRIES

Australia	www.oecd.org/australia
Austria	www.oecd.org/austria
Belgium	www.oecd.org/belgium
Canada	www.oecd.org/canada
Czech Republic	www.oecd.org/czech
Denmark	www.oecd.org/denmark
Finland	www.oecd.org/finland
France	www.oecd.org/france
Germany	www.oecd.org/germany
Greece	www.oecd.org/greece
Hungary	www.oecd.org/hungary
Iceland	www.oecd.org/iceland
Ireland	www.oecd.org/ireland
Italy	www.oecd.org/italy
Japan	www.oecd.org/japan
Korea	www.oecd.org/korea
Luxembourg	www.oecd.org/luxembourg
Mexico	www.oecd.org/mexico
Netherlands	www.oecd.org/netherlands
New Zealand	www.oecd.org/nz
Norway	www.oecd.org/norway
Poland	www.oecd.org/poland
Portugal	www.oecd.org/portugal
Slovak Republic	www.oecd.org/slovakia
Spain	www.oecd.org/spain
Sweden	www.oecd.org/sweden
Switzerland	www.oecd.org/switzerland
Turkey	www.oecd.org/turkey
United Kingdom	www.oecd.org/uk
United States	www.oecd.org/us

31

Appendix II: Question 1

'Smart' carbon positive energy house

Experts from the Welsh School of Architecture have designed and built the UK's first purpose-built, low-cost energy smart house, capable of exporting more energy to the national electricity grid than it uses.

The house, designed by Professor Phil Jones and his team based at the Welsh School of Architecture, has been built as a prototype to meet tough new targets for zero net carbon emission housing set by UK Government.

Designed and constructed as part of the Wales Low Carbon Research Institute's (LCRI) project, and supported by Swansea University, its unique design combines for the first time reduced energy demand, renewable energy supply and energy storage to create an energy positive house.

"The Welsh and UK Governments – and governments across the EU – have set targets for very low 'nearly zero' energy buildings by 2020, and zero carbon new housing can deliver this and more. This means that as an academic community we have to rise to that challenge and come-up with innovative new ways to build houses of the future," according to Professor Jones, who heads the project.

Zero carbon energy performance involves a combination of reduced energy demand and renewable energy supply, using the electricity grid to import and export energy.

Electrical and thermal storage have also been used to allow energy generated at the house to be used directly by the occupiers.

In order to drastically reduce the energy demand, the house was built with high levels of thermal insulation reducing air leakage and uses an innovative energy efficient design which includes low carbon cement, structural insulated panels external insulated render, transpired solar collectors and low emissivity double glazed aluminium clad timber frame windows and doors.

The south facing roof comprises of glazed solar photovoltaic panels, fully integrated into the design of the building, allowing the roof space below to be naturally lit. This has been designed to reduce the cost of bolting on solar panels to a standard roof.

The house's energy systems combine solar generation and battery storage to power both its combined heating, ventilation, hot water system and its electrical power systems which includes appliances, LED lighting and heat pump. The solar air system preheats the ventilation air which is topped up from a thermal water store.

Question 1 "In order to drastically reduce the energy demand, the house was built with high levels of thermal insulation reducing air leakage…" Which form(s) of thermal transfer does this reduce? Select one answer	
Conduction	
Convection	
Radiation	
Conduction and convection	
Convection and radiation	
Radiation and conduction	

Question 2 What strategies does the house employ to ensure supply of electricity throughout the year?

Question 3

Explain why:

(a) Ice cubes float on water
(b) Ice cubes can cool the entire drink—even the liquid at the bottom of the glass

Appendix III: Question 2

SCIENCE UNIT 26: ULTRASOUND

In many countries, images can be taken of a foetus (developing baby) by ultrasound imaging (echography). Ultrasounds are considered safe for both the mother and the foetus.

The doctor holds a probe and moves it across the mother's abdomen. Ultrasound waves are transmitted into the abdomen. Inside the abdomen they are reflected from the surface of the foetus. These reflected waves are picked up again by the probe and relayed to a machine that can produce an image.

QUESTION 26.1

To form an image the ultrasound machine needs to calculate the *distance* between the foetus and the probe.

The ultrasound waves move through the abdomen at a speed of 1540 m/s. What measurement must the machine make so that it can calculate the distance?

....

....

....

QUESTION 26.2

An image of a foetus can also be obtained using X-rays. However, women are advised to avoid having X-rays of their abdomens during pregnancy.

Why should a woman avoid having her abdomen X-rayed during pregnancy *in particular*?

....

....

....

QUESTION 26.3

Can ultrasound examinations of expectant mothers provide answers to the following questions? Circle "Yes" or "No" for each question.

Can an ultrasound examination answer this question?	Yes or No?
Is there more than one baby?	Yes / No
What colour are the baby's eyes?	Yes / No
Is the baby about the right size?	Yes / No

Appendix IV: Information Sheet

Information on the Research

What is the research?

There is a new curriculum that will begin to be introduce across Wales next year. This curriculum may not affect you, but it will change the way that future pupils are taught and tested. One of the ways this will change is that GCSE paper may end up having questions that are similar to the questions on PISA tests, which unfortunately Wales currently isn't very good at answering.

So, this research aims to find a way to help pupils answer these PISA questions. You will be given a few different example PISA questions that tie into what you are learning, and given part of a lesson to answer them to the best of your ability. They are hard, but you will not be assessed on them. Everyone will be taking part in these questions, but it is up to you whether or not I can use your results. If you give permission for your results to be used, they will be anonymous and compared with the results of classmates, Wales and pupils around the world in a research paper.

What is PISA?

PISA is the Programme for International Student Assessment, an organization which tests school systems around the world. About 90 countries take part in PISA, with the aim being to learn off each other to better ourselves. Unfortunately, Wales is currently 35th on this list. The PISA test is given to a random sample of 15-year-olds around the world every three years, and the last test was in 2015, so you may end up sitting the test next year.

Your rights

- Your results will not be used without your consent.
- If you give consent, all of your results will be anonymised, meaning your name is removed and untraceable.
- You have the right to ask for your results at any time (but be aware that these are not reflective of your ability at answering GCSE style questions, and may be lower than you would expect).
- You have the right to withdraw from the research at any time, and your results will be removed entirely.

Appendix V: Consent Form

Research Consent Form

PISA Questions

Date

March 28, 2017

The aim of this research is to measure the ability of pupils in answering PISA questions, and to find a way of helping them do well. This investigation will not interfere with your normal GCSE work, you do not have to revise, and there won't be any penalty for low marks (though I do ask that you try your best, as the work is on topic). However, this work is not optional, you must still take part in these lessons. This form offers you the choice of consenting to your results being used in this investigation.

By taking part in this investigation you agree to have your results for each question compared with that of classmates, Wales and pupils around the world. All your results will be anonymised, meaning that your name and any other information that is not your mark will be removed and untraceable. You have the right to ask for your results at any time. You also have the right to withdraw from this investigation at any time, and your results will be removed from the investigation completely.

I ask your permission to use your marks to try and find a better way of helping yourselves and others, to answer this style of question. If you are happy for this to happen, please sign below.

I give permission for my results to be used in this investigation, as agreed in the terms above.

Name: ______________________

Signed: ______________________

Date: ______________________

Appendix VI: Inference Square Used