UCS Research Review

Issue 1, January 2022

Table of Contents

Editorial

Emma Taylor

This inaugural issue of the UCS Research Review (University College School's educational research journal), showcases a vast array of innovative and original research carried out by staff at the school. The skills and expertise demonstrated here illustrate the time and attention paid by practitioners to the close consideration of their teaching, as well as the school's commitment to evidence informed policy and practice. I am very grateful for the help and support of the Editorial Board (named below) who have given up their time to review and support with editing the articles that follow. They have also provided invaluable feedback on the layout, design and structure of the journal.

The journal is split into three sections, with the first section dedicated to original research undertaken by staff, some of which consists of abridged versions of papers that have been submitted as part of a Masters being undertaken by the author. Keith Bugler starts us off with an enlightening comparison of the priorities displayed by those students who might be considered intrinsically motivated yet of differering attainment profiles. This is followed by Emma Dell's qualitative exploration of marking; an in-depth study of the marking practices taking place at UCS. The outcomes of this research have since been used to inform the school's assessment policy. Emma Desmond then provides an overview of her research examining the handling of low-level disruption in a year 8 classroom. Her work provides us with a practical artefact or guide for teachers to use based on her findings. The next article, by Amy Holland, consists of a timely overview of some of the key literature examining the impact of single-sex environments on the retention of girls studying A-level Physics.

Maria Pia Maggioni's work forms part of her participation in the Subject Expert Micro-Programme organised by the UCL Institute of Education Confucius Institute for Schools. She provides an analysis of the impact of self-discovery vocabulary learning on pupil outcomes in Mandarin. This is followed by the final article in this section, a piece of research undertaken by Katie Matthews, Emma Taylor and Andrew Quirke into the extent to which low or no stakes testing in the classroom has an impact on pupil outcomes. The findings of this study have since been embedded in classroom practice at UCS as well as within the UCS Revision Toolkit (a guide to effective evidence based revision techniques for students).

In the second section of the journal, Abby Caplin's extended essay provides a thorough analysis of the influence of PISA on global education policy, carefully considering both the advantages and drawbacks of PISA for policy reform. This is then followed by a third section dedicated to reviews of recently published books with a focus on teaching and learning, and I am very grateful to Sophie Bennett, Charlotte Hawes, Laura McGill, Jay Thomson and Kimberley Ward for taking the time to submit their informative and illuminating reviews.

Many thanks to all the contributors for giving up their time to submit articles, essays and reviews. I hope you find their contributions interesting and relevant and that you enjoy reading this first issue of the UCS Research Review.
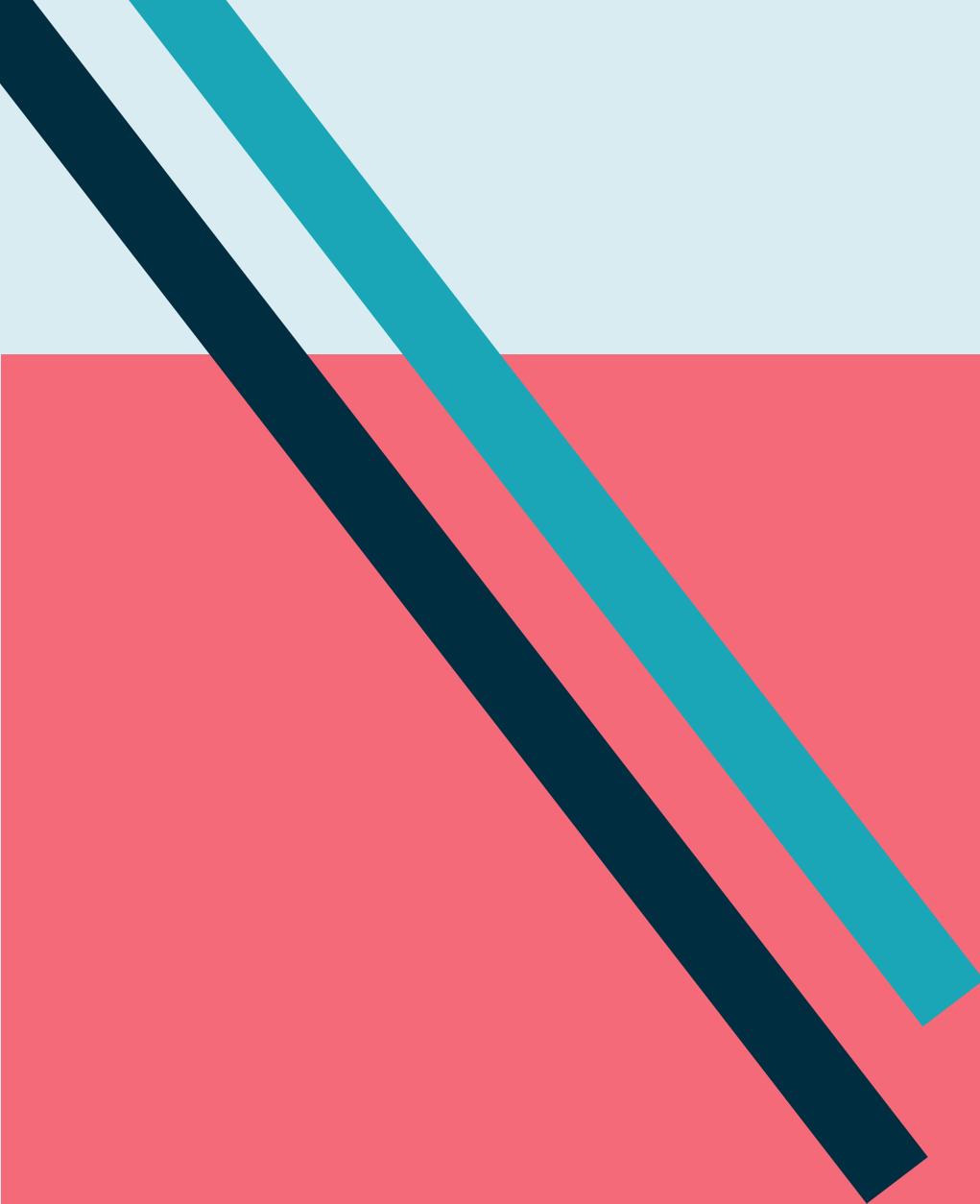
Original Research
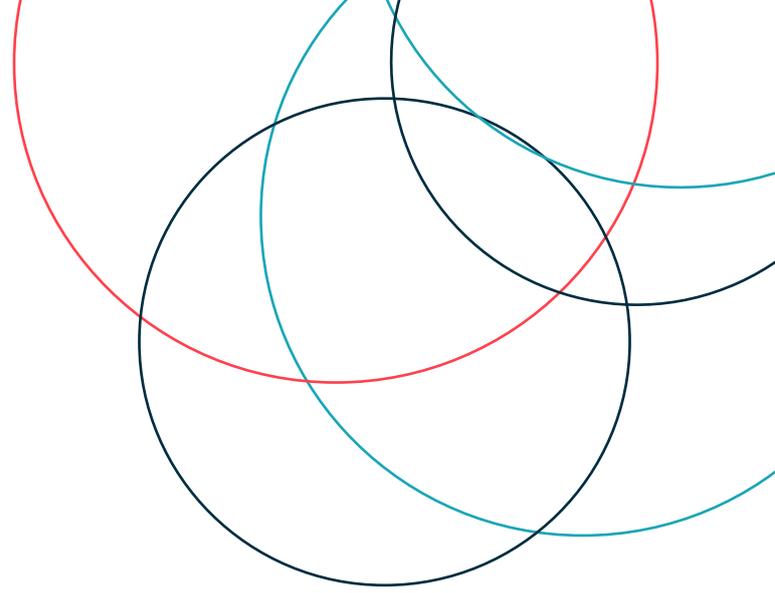
A comparison of the internal priorities displayed by intrinsically motivated students of differing attainment profiles

Keith Bugler

Abstract

A great deal of research has been carried out into the correlation between intrinsically motivated pupils and high educational attainment. The theoretical framework of self-determination theory (SDT) was used to profile students' motivation levels in their chemistry lessons at UCS. This was done through both the pupil's chemistry teacher and form tutor detailing what they believed to be each pupil's motivation type, and pupils self-reporting via the situational motivation survey (SIMS). Subsequently two pupils, both displaying intrinsic motivation traits, but at opposite ends of the attainment spectrum, were interviewed to elucidate their prioritisation of the underlying psychological pillars found within SDT. It was found that the pupil's both displayed high levels of autonomy, perceived competence and relatedness but they prioritised each of these traits very differently. The pupil with higher attainment placed perceived competence as most important, while the lower attaining pupil's source of motivation stemmed from an unusually high sense of autonomy.

Introduction

One way to classify motivation is to divide it into two broad types. Intrinsic motivation is defined by Ryan and Deci as when an individual does a task because they find it "inherently interesting or enjoyable" (2000, p.55). This contrasts with extrinsic motivation where a person is motivated by external forces, such as the offer of a reward or coercion to undertake a task (Cecere, Mancinelli and Mazzanti, 2014). Some examples of extrinsic motivation in an educational setting include: motivating through praise, influencing students through positive student-teacher relationships, and influence or pressure within a peer group.

One widely used model is that of self-determination theory (SDT, see Figure 1), which can be thought of as a taxonomy of motivation, proposed by the academics Deci and Ryan. Self-determination theory states that "people are inherently prone toward psychological growth and integration, and thus toward learning, mastery and connection with others" (Ryan and Deci, 2020 p.1). However, for this to occur individuals must be given appropriate scaffolding. The authors suggest that the three key pillars of this are autonomy, competence and relatedness.

Regulatory Styles | Amotivation | Extrinsic motivation → External Regulation, Introjection, Identification, Integration | Instrinsic motivation

	Amotivation	External Regulation	Introjection	Identification	Integration	Instrinsic motivation
Regulatory Styles	Amotivation	External Regulation	Introjection	Identification	Integration	Instrinsic motivation
Motivation Regulators	Non-relevance Nonintentionality Low perceived competence	Rewards Punishment Compliance	Ego led Approval from self Peer Approval	Conscious valuing of activity Self-endorsement of goals	Hierarchical synthesis of goals Congruence	Interest Enjoyment Inherent satisfaction
Source of motivation	Impersonal - not associated with self	External	Somewhat External	Somewhat Internal	Internal	Internal

Figure 1. An SDT taxonomic model of the differing kinds of motivation as proposed by Ryan and Deci (2000a)

Autonomy is supported by experiences of interest or engagement, but can be undermined by external control. Competence is defined as being specifically linked to mastery, a feeling or sense that an individual can "succeed and grow" (p.1) in completing a task or demonstrating a skill. Relatedness is a "sense of belonging and connectedness" (p.1) which can be built within a respectful and caring environment.

According to SDT, workplace or education practices which encourage autonomy, competence and relatedness will in turn empower individuals and lead to a greater proportion of employees or students displaying intrinsic motivation. This is important as numerous studies (Csikszentmihalyi, Rathunde and Whalen, 1993; Clinkenbeard, 2012; Olszewski-Kubilius, Subotnik and Worrell, 2015) have found a strong correlation between high attainment and intrinsic motivation. Thus educators would in an ideal world want students to be intrinsically motivated within a lesson or learning task.

Methods and Student Motivation Profile Results

Measuring the motivation of a pupil and assigning a taxonomic label is a complex endeavour. Depending on the task or academic subject at hand an individual's motivation may vary depending on the importance they ascribe to it. In addition, a school is not a homogeneous environment, and the core pillars of perceived competence, relatedness and autonomy as identified by SDT will vary from classroom to classroom, affecting the motivation a student feels. Subsequently a peer-reviewed questionnaire, the Situational Motivation Survey (SIMS), was used to ascribe pupil motivation.

Five volunteers from a Lower Remove (aged 13-14) chemistry class were selected to complete the SIMS regarding their views on their chemistry lessons. They were then ascribed a motivation classification according to their SIMS score. Teachers of the students who completed the SIMS were also asked to self-report which motivation style the pupils displayed. Finally, end of topic test scores, assessed homework and student banding data for each student participating was consulted and each participant sorted into an attainment quintile.

Student	Motivation style in chemistry according to SIMS	Attainment data quintile
Student A	Intrinsic Motivation/Very High Order Extrinsic Motivation	1st
Student B	Intrinsic Motivation	2nd
Student C	High Order Extrinsic Motivation	3rd
Student D	High Order Extrinsic/Intrinsic Motivation	2nd
Student E	Intrinsic Motivation	5th

Table 1: Students motivations toward chemistry from SIMS data and their placement in the cohort attainment data

Student	Motivation style in chemistry according to SIMS	Chemistry teacher ascribed motivation	Form tutor ascribed motivation
Student A	Intrinsic Motivation/ Very High Order Extrinsic Motivation	Intrinsic Motivation	High Order Extrinsic Motivation
Student B	Intrinsic Motivation	High Order Extrinsic Motivation	High Order Extrinsic Motivation
Student C	High Order Extrinsic Motivation	High Order Extrinsic Motivation	Lower Order Extrinsic Motivation
Student D	High Order Extrinsic Motivation/Intrinsic Motivation	High Order Extrinsic	Intrinsic Motivation
Student E	Intrinsic Motivation	Intrinsic Motivation	Intrinsic Motivation

Table 2: Comparison of motivation styles as ascribed by SIMS, chemistry teacher and form tutor.
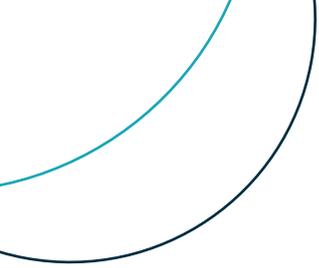
After using the SIMS, attainment ranking and teacher perceptions of motivation style to filter the data, it became clear that there were two suitable data points for comparison. Student A is in the first quintile for attainment and was identified as displaying either intrinsic motivation or a particular type of very high order extrinsic motivation. Student E was identified by teacher perception and SIMS data to be intrinsically motivated and is in the fifth quintile for attainment data. As both Student A and E displayed similar motivation profiles but very different attainment data, the students were invited to interview.

Due to time constraints only two interviews were conducted, one with each student. Interview questions from previous work by Netcoh using SDT concepts were used (2017). A narrative analysis of the interview transcripts was used. For brevity, a condensed series of findings will be discussed.
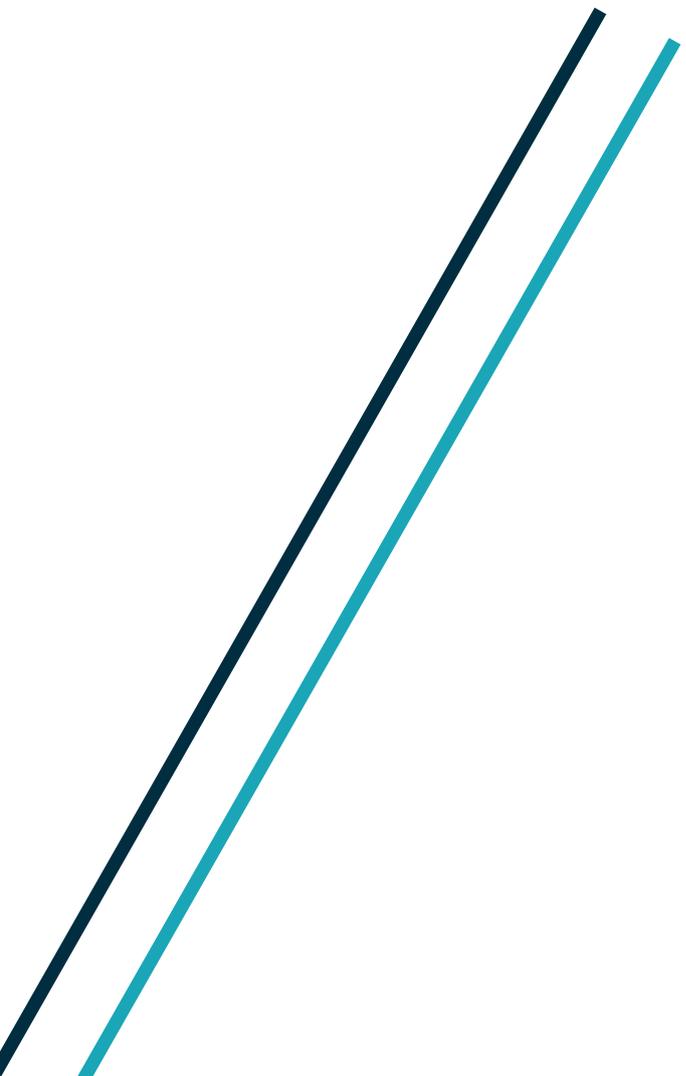
Interview Findings

Student A repeatedly made unprompted references to traits synonymous with the psychological pillars of motivation; relatedness, perceived competence and autonomy. Relatedness and its associated environment were a feature of many of the answers Student A gave to questions. These comments were often spontaneous but Student A would emphasise the value they placed on one-to-one discussions with teachers in subjects they found interesting.

Interviewer: "How does getting a certificate, let's say in front of a whole school assembly, compare to reading about something that interests you?"

Student A: "I would say the reading [is more of a motivator]...but reading a textbook on its own is not as interesting as discussing it with a teacher, if you have a conversation about something you've learned, then that's forever"

Interviewer: "What things can teachers do to help you be motivated?"

Student A: "Well enthusiasm, if they are excited about what they are teaching it definitely makes it more interesting..."

Through narrative analysis it was concluded that while student A displayed high levels of all three psychological pillars, they ranked perceived competence highest and autonomy lowest.

Student E also gave answers suggesting relatively high levels of autonomy, perceived competence and relatedness. Interestingly their test scores relative to their cohort placement was of little interest to Student E. However it was on the theme of autonomy that the most revealing exchanges occurred. Student E believed that at UCS teachers supported

9

flexibility in what pupils do, which could be viewed as a proxy for freedom and autonomy, as demonstrated by the following exchange.

Interviewer: "Do you think you influence what you learn about? Or you have any control over that?"

Student E: "Yes, it's up to me and what I retain."

Interviewer: "So you feel you can influence what you learn at school?"

Student E: "Yes."

Interviewer: "Do you think how good you are at a subject affects your enjoyment and motivation?"

Student E: "I don't think being good at the subject affects how much you enjoy the subject, if you're not very good at a subject but you like it then you can obviously get better at it."

Here Student E shows a very clear sense of autonomy and appears to prioritise it above competence and attainment, with their theory of "you can obviously get better at it" implying possibility for improvement, rather than competence in a domain being fixed or uncontrollable. This internal narrative of autonomy and self-control over their school life was repeatedly emphasised with the following quote summarising ideas expressed throughout the entire interview:

Student E: "I don't do chemistry because I have no choice, I don't have a choice but nobody is forcing me to retain the information so I'm in control."

Conclusions and Implications

Student A: Perceived Competence > Relatedness > Autonomy

Student E: Autonomy >>> Relatedness > Perceived Competence

Figure 2: Proposed hierarchy of the three psychological pillars behind students' motivation

Student A is an archetypal intrinsically motivated student with associated high attainment. Student E however has low attainment, relative to the school and cohort. Despite this Student E displays clear intrinsic motivation, potentially fuelled by a notably high sense of autonomy.

From this work I concluded that students with similar motivation profiles but differing attainment levels do display different prioritisation of the psychological pillars which underpin motivation. The findings from the interview with Student E highlight one potential way that teachers can motivate lower attaining students, by reinforcing a sense that students have control and autonomy over their outcomes, whilst supported in a caring and engaged classroom environment.

If educators wish to build and support the motivation of their students and learners, they should first consider how to make them feel like empowered citizens of the classroom with a sense of autonomy. Furthermore, it would be a fruitful area of research to identify students who display lower attainment and intrinsic motivation, and observe which of the three psychological pillars; autonomy, perceived competence and relatedness, are prioritised. This could be contrasted with higher attaining intrinsically motivated students who may show different prioritisation of the three psychological pillars.
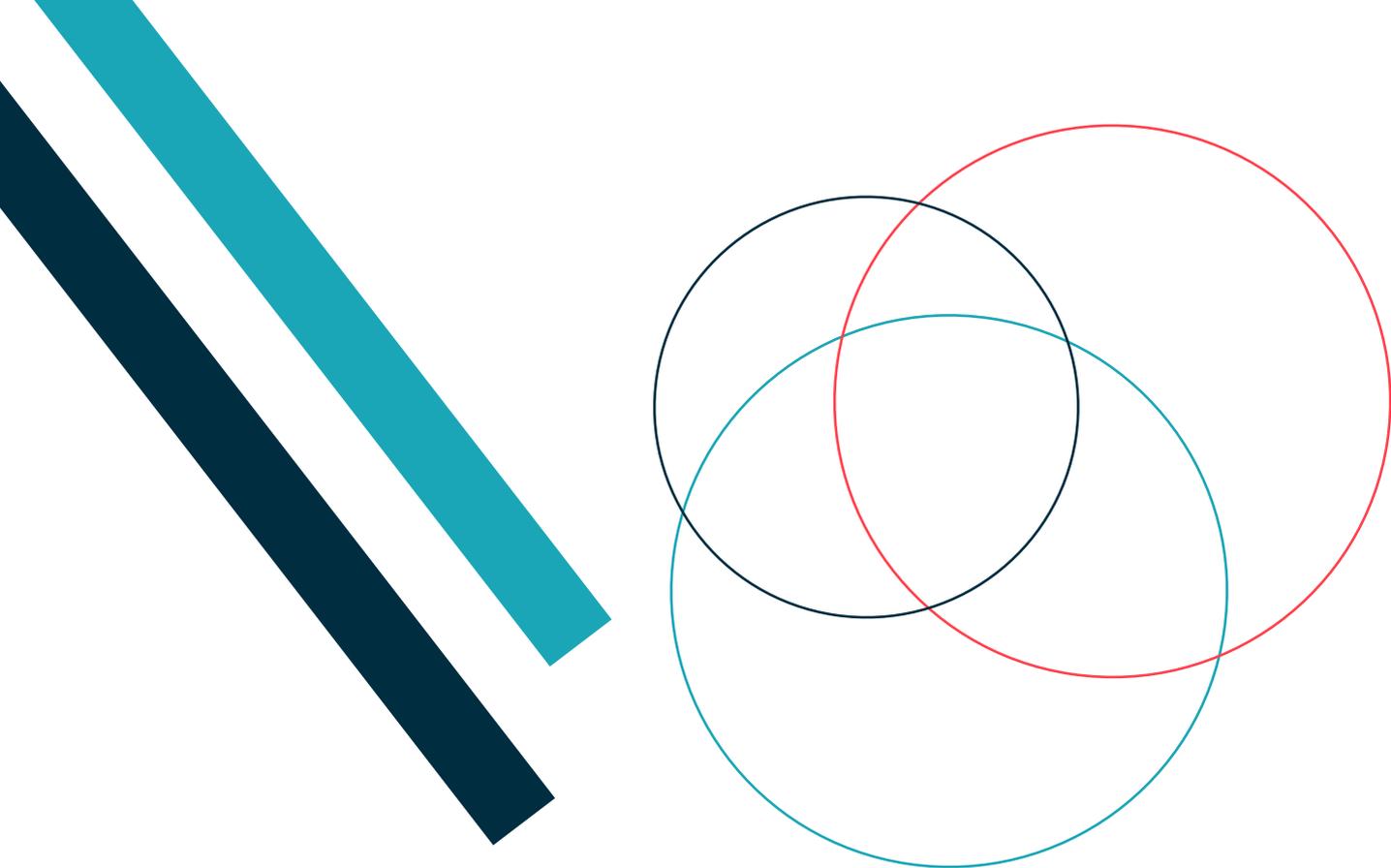
References

Cecere, G., Mancinelli, S. and Mazzanti, M. (2014) 'Waste prevention and social preferences: The role of intrinsic and extrinsic motivations', *Ecological Economics*, 107.

Clinkenbeard, P. R. (2012) 'Motivation and gifted students: Implications of theory and research', *Psychology in the Schools,* 49(7).

Csikszentmihalyi, M., Rathunde, K. and Whalen, S. (1993) 'Talented adolescents: The roots of success and failures', *British Journal of Educational Studies.*

Netcoh, S. (2017) *Students' Experiences with Personalized Learning: An Examination Using Self-Determination Theory.* University of Vermont.

Olszewski-Kubilius, P., Subotnik, R. F. and Worrell, F. C. (2015) 'Antecedent and concurrent psychosocial skills that support high levels of achievement within talent domains', *High Ability Studies,* 26(2).

Ryan, R. M. and Deci, E. L. (2000) 'Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions', *Contemporary Educational Psychology,* 25(1).

Ryan, R. M. and Deci, E. L. (2020) 'Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions', *Contemporary Educational Psychology.*

A qualitative exploration of marking

Emma Dell

Abstract

The purpose of this article is to outline research undertaken into marking practices at UCS. It also reflects on the wider literature related to this topic. Pupils and teachers were surveyed to explore their opinions on effective marking, with focus groups of Middle School and Sixth Form pupils providing further depth. Results demonstrate that pupils highly value specific and actionable marking, where feedback is provided together with opportunities to act on this feedback. In addition, it is important for teachers to consider the emotional investment that pupils have in their work and its marking, and that explaining the purpose of a task enhances its pedagogical value.

Introduction and context

While it is widely acknowledged that assessment sits at the core of teaching and learning, there remains a lack of consensus in what constitutes effective assessment. Written marking, in particular, often has a reputation in teaching for being rigidly prescribed and a burdensome workload, yet its impact on pupil progress is under-researched. In fact, the Education Endowment Foundation 2016 review of written marking found scant existing evidence on the practice, especially at secondary school level, and described an urgent need for further research.[1] Therefore marking is a ripe area for research, especially at UCS with its commitment to evidence-based practice. This study was designed to explore how marking is currently perceived at UCS by pupils and teachers, to discern common themes of best practice, and to determine ways in which marking can promote pupil progress and well-being, while maintaining a reasonable workload for teachers.

Methods and results

In the first stage of the study, a survey was emailed out as a Google Form, to all pupils from the Lower Remove through to the Sixth during the autumn term 2020. The survey recorded the pupil's year, in order for comparisons to be made across the Middle School and Sixth Form, but collected no other identifiable information. The survey started by asking pupils to categorise their most and least helpful forms of marking. Then pupils used a Likert scale to express their level of agreement with statements such as 'marking helps me improve'. Finally, pupils answered free-response questions describing examples of helpful marking. 243 pupils responded. Survey data were analysed to determine most and least helpful forms of marking from a pupil viewpoint (Table 1). Answers to free response questions were manually analysed using inductive coding. In this process, the responses were read through, then codes were created based upon common themes in these responses, such as preference for 'specificity' in marking. These codes were then applied to the data and their frequency was tallied (Table 2).

Most helpful forms of marking	% of pupils who choose this option
Numerical marks	82
Comments	82
Ticks/crosses	57
Model answers	55
Verbal feedback	49

Least helpful forms of marking	% of pupils who choose this option
Peer marking	67
Stickers	64
Smiley faces	62
Commendations	36
Questions	22

Table 1: Top five most and least helpful forms of marking from pupil survey data

Coded responses to the free response question 'describe examples of marking that you find most helpful'	Number of pupil responses containing this code (out of a sample of 144)
Model answers	25
Comments	21
Google Doc Comments	15
Improvements	13
Specific/Detailed	13
Identifying errors	11
Grades plus comments	10
Grades	5

Table 2: Descriptors of most helpful forms of marking from pupil survey data

In parallel to this pupil survey, a survey (also as a Google Form) was emailed out to all teaching staff during the autumn term 2020. This survey was anonymous and collected no identifying information. There were 67 responses. Teachers were asked a similar series of Likert scale questions to the pupils, as well as a free response question on their most effective forms of marking.

Answers to the free response questions were manually analysed using inductive coding as with the pupil survey data. As a control point, a question about the frequency of responding to marking was asked to both teachers and pupils (Table 3). There is a high level of consistency between pupil and teacher responses, providing evidence for the reliability of the survey data.

How often do teachers ask you to respond to their marking?	Pupil response
For every/most piece of work	33%
For some pieces of work	41%
Rarely for pieces of marked work	23%

How often do you ask pupils to respond to your marking?	Teacher response
For every/most piece of work	36%
For some pieces of work	40%
Rarely for pieces of marked work	22%

Table 3: Comparison of pupil and teacher survey data on the frequency of response to marking

In order to further explore the themes revealed in the survey data, two focus groups with the researcher were set up, one with seven pupils from the Remove, and one with six pupils from the Sixth. These focus groups used open-ended questions to explore the pupil's thoughts on marking in greater depth.. Pupils were invited to volunteer for the focus groups via email, and then selected on a first-come-first-served basis. The audio from the focus groups was digitally recorded and then transcribed. During transcription, all identifying details were removed and the recording was then deleted. The transcriptions were analysed through inductive coding, which generated eight key themes, all supported through pupil quotations:

- Pupils value homework being set for a purpose
- Pupils believe homework should be marked
- Pupils value going through marking in an active manner
- Pupils value comments with marks
- Markschemes in isolation are unhelpful; markschemes with guidance are helpful
- Pupils value opportunities to put their feedback into practice
- Sixth pupils view homework as more purposeful and motivational than Remove pupils
- Pupils are not always aware of the educational value of tasks they are set

Discussion and recommendations

The first key takeaway from the study is how much pupils value marking (Figure 1). The majority of pupils surveyed believe that marking helps them improve, and they described many positive aspects of marking. In fact, pupils viewed marking as having a more positive impact than teachers. Additionally, marking was viewed more positively by pupils in the Sixth Form compared to the Middle School. It may be worth considering the types

of assignments and marking that are used with sixth form pupils and how these could be equally applied lower down the school.

Marking helps me improve

Student

Strongly agree
Agree
Neutral
Disagree
Strongly Disagree

238 responses

My marking has a positive impact on sudent progress

Teacher

66 responses

Strongly agree
Sixth form: 49%
Middle School: 30%
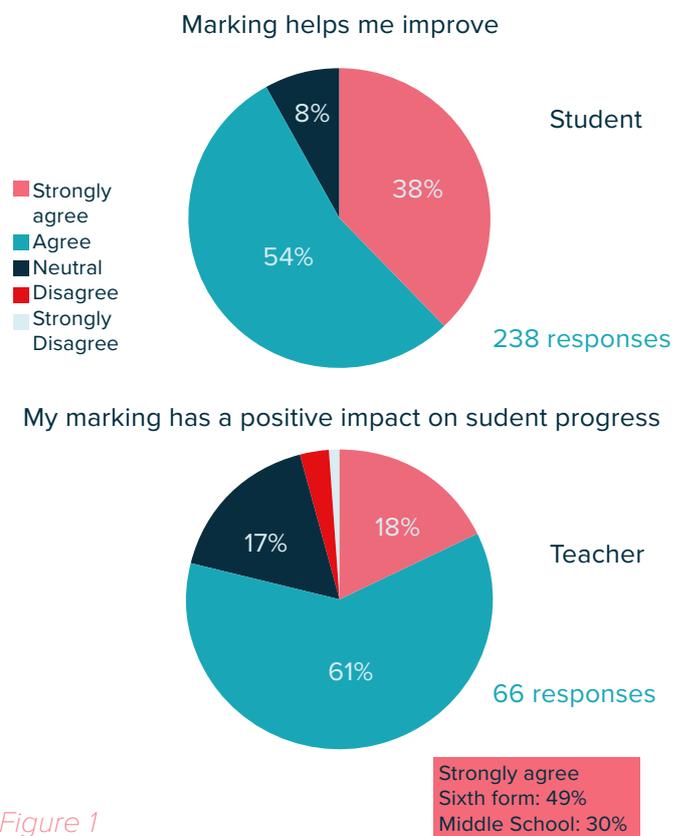
Figure 1

The study also revealed the diverse forms of marking taking place, and that effective marking does not necessarily need to involve red ink in an exercise book. Pupils and teachers described the benefits of verbal feedback in the survey, and this is supported by UCL's Verbal Feedback project which showed improved engagement, performance and pupil-teacher relationships through the use of verbal feedback.[2] Specific, actionable comments are highly appreciated by pupils, but some pupils feel that they receive too much written feedback which can be overwhelming. In survey data, several teachers also agreed that marking should be selective. The evidence supports that written feedback should be specific and elaborative, and provided in manageable units.[3] Effective marking can also take place within lessons. In survey data, pupils frequently mentioned going through model answers as an effective method of marking.
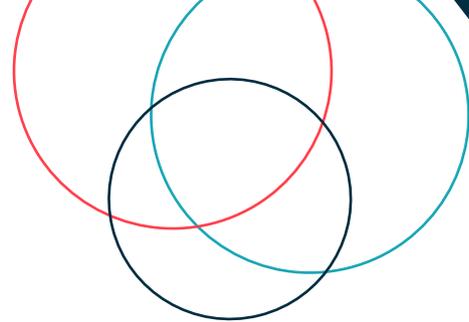
No teacher suggested this as a marking style in the teacher survey. This is perhaps because teachers may not consider this a form of marking, however the data suggests it is viewed as such by pupils and is highly valued. There is an important subtlety here as many pupils dislike exam board markschemes because they find them difficult to use:

"sometimes you look at the markscheme and you don't know how to implement those kind of structures into your writing, and also having feedback from your teachers is a nice supplement to it".

They prefer a teacher-led worked-example since:

"they can articulate it to you in a way that you might understand it better".

Use of model answers for self-assessment in class may also help to ease teacher workload.

One point emphasised throughout the data and the literature is that the most important part of marking is what a pupil does with it. John Hattie describes that feedback can only reliably be measured by how it is received, not how it is given.4 Marking is therefore a process. Pupils were very positive when describing experiences of actively using their marking and making improvements. Table 3 shows that pupils do not currently respond to all pieces of marked work. The evidence suggests that pupils are unlikely to benefit from marking unless some time is set aside to enable students to consider and respond to marking; feedback is only useful if it is used.1 The experience of making progress as a result of acting on feedback is also associated with pleasure:

"if last time you did this wrong, but this time you did it correctly, I think that just gives you a nice feeling. You're sort of going somewhere".

The data also highlighted the importance of the pupil-teacher relationship involved in marking. Marking is one of the areas of teaching that can elicit a highly personal and emotional response in pupils. Pupils view homework and marking as a joint endeavour; a pupil has put effort into an assignment and expects effort on the part of the teacher. There is a strong connection between marking and value: "if they mark it, it has a lot more purpose" and "I don't think there's much point in doing homework if it's not marked". When work is not marked it is "demoralising". Pupils can discern patterns in how teachers mark, and put less effort into a subject when they notice that marking is infrequent. Pupils also value timeliness in marking, especially with tests. It is important to be aware of the emotional investment that pupils have in their marking, especially when their views are not aligned with educational evidence. For instance, evidence shows that it is more effective for teachers to mark selectively rather than every piece of homework, and so pupils should be shown the value in self- and peer-marking. Additionally, 'overlearning' aids academic progress,5 but pupils describe irritation with repetitive practice: "I feel like just doing lots and lots of questions, like in Maths, doesn't really help you". Explaining the purpose of an assignment is therefore critical. Pupils are very keen to receive grades for their work. Evidence shows that it may be more effective to focus on formative comments (or to separate the release of a grade from the release of comments), and so again pupils will need explanation and reassurance as to why a grade is not being given.6,7 Receiving feedback can be a challenging experience, and it is improved by a strong pupil-teacher relationship.8 Pupils will feel more trust in the process when they have a clear understanding of how and why an assignment is being marked.
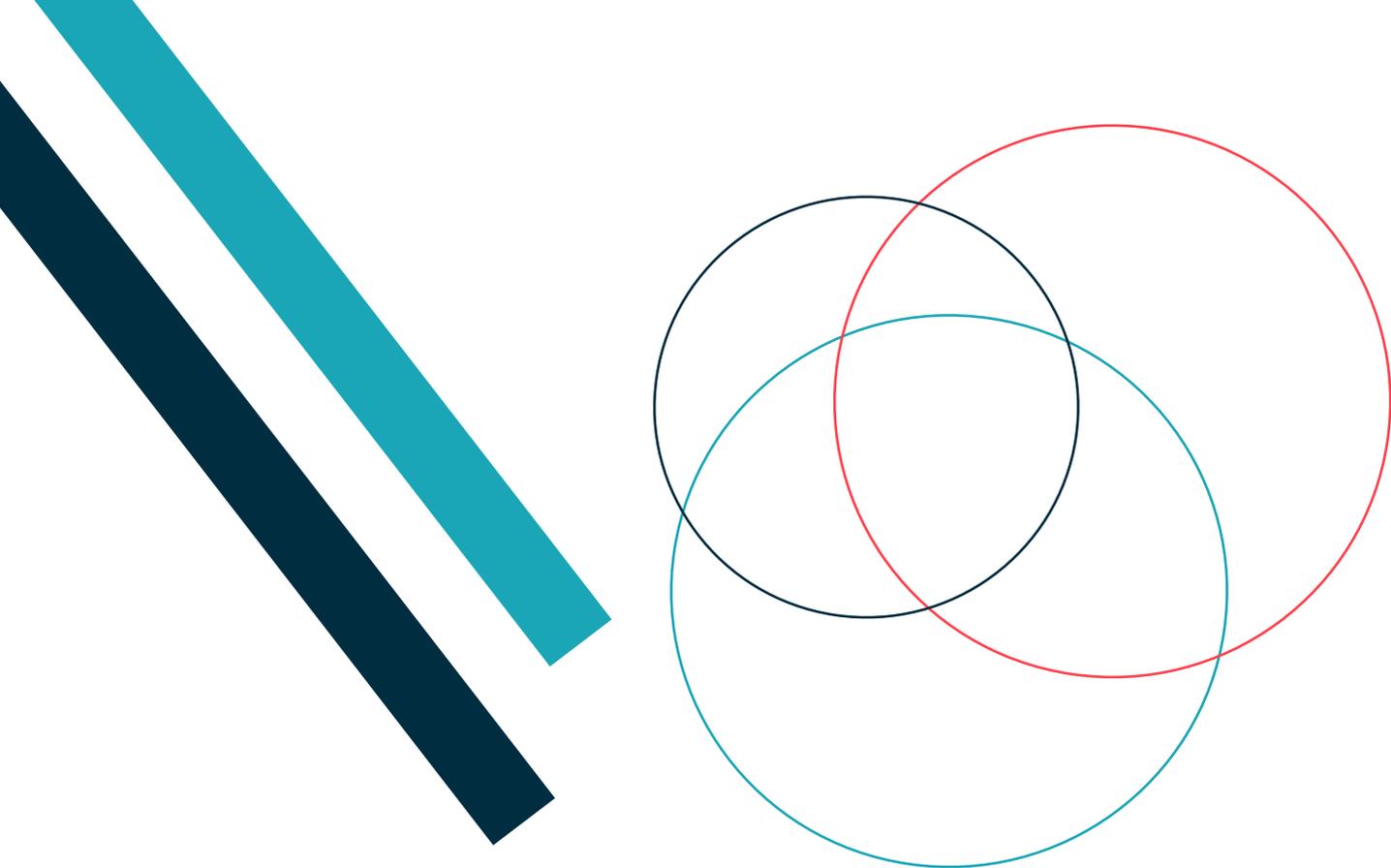
References

Education Endowment Foundation (2016) *A Marked Improvement?: A review of the evidence on written marking.* Accessed at: https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/written-marking

Hattie, J. (2008) *Visible Learning, Taylor and Francis: Australia*

Kuepper-Tetzel, C. E., & Gardner, P. L. (2021) *Effects of temporary mark withholding on academic performance.* Psychology Learning & Teaching

Shute, V. (2007) *Focus on Formative Feedback, Review of Educational Research* 78(1)

Smith, E. & Gorard, S. (2005) *"They don't give us our marks": the role of formative feedback in student progress*, Assessment in Education: Principles, Policy and Practice, 12 (1) 7-20

UCL (2019) *Verbal Feedback Project*

Willingham, D. T. (2021) *Why don't students like school?* Jossey Bass: London

Wiliam, D. (2016) *The Secret of Effective Feedback,* Looking at Student Work, 73(7) 10-15

An investigation into handling low-level disruption in a year 8 Latin classroom

Emma Desmond

Abstract

My research aims to develop a guide to managing low-level disruption in a Year 8 Latin classroom and to explore the impact of environmental factors that contribute to this behaviour. My environment for enquiry is an independent selective boys school (University College School), where Latin is compulsory in Year 7 and 8, but becomes an optional subject from Year 9. I have examined the impact of gender, subject choice and high attainment on low-level disruption.

My research followed an action research approach, in which I interviewed three students from my Year 8 Latin class that volunteered to speak to me about student focus. I also held two staff focus groups, each with five members of staff across a selection of subjects, to gain an understanding of their experience in the management of low-level disruption in the classroom. After I had gained an insight into student and staff perception of low-level disruption, I produced my artefact, "A guide to managing low-level disruption for teachers." I implemented this guide in my Year 8 Latin lessons and used both self-reflection and observation notes from a colleague to determine how useful it seemed.

I found that the guide is a useful tool and combines a number of key principles that help enable effective classroom management, with a particular emphasis placed on the importance of knowing the students. I found that the students responded well to consistent use of the artefact and were especially motivated by reward systems. I would recommend the use of the guide for teachers, but recognise the ongoing challenge of implementation. My overarching conclusion is that the consistent implementation of the guide might be initially difficult for teachers, but perseverance may lead to improved learning habits among students and will simplify behaviour management.

Introduction and Context

The purpose of my research is to explore the environmental factors that contribute to low-level disruption in Year 8 students of Latin and to test practical strategies for teachers to use to manage this kind of disruption in the classroom. This is an important area of research as low-level disruption can prevent students from learning and it is essential that teachers can gain the control needed to protect students' learning and development.

Research Questions

1) To what extent do the environmental factors of gender, high attainment and subject choice affect low-level disruption in students?

2) How can low-level disruption be managed in this environment to ensure that learning outcomes of other students are not affected?

Behaviour: Definition

It is important to provide a clear definition of the nature of low-level disruption that I refer to throughout my research. Tennant (2004 p53) describes low-level disruption as "off-task behaviour which is not concerned with the teacher's intention for that lesson." From my experience, a whole array of behaviours fall into this category, including but not limited to, students chatting, making silly noises, fiddling with equipment, rocking on their chairs and calling out answers. I agree with Tennant that low-level disruption is likely to be interpreted differently by each individual teacher and therefore, it is best to define it as behaviour that is not violent or abusive, but recognise that as Lewis states, it can equally "interfere with the rights of other pupils" (Lewis 2009 p40) and consequently should not be tolerated by the teacher. Arguably, one of the biggest challenges of managing low-level disruption is that whilst "many teachers regard behavioural problems of students as a massive disruptive factor in their teaching" (Schwab, Eckstein & Reusser 2019 p228) the temptation as a teacher, is often to ignore behaviour deemed to be comparatively minor, in order not to draw even greater attention to it in the classroom. Feldman (2001, p138) states that "we erroneously tend to ignore some of the low-level acts of classroom incivility in the hope that they will go away" which in turn can cause the students to repeat and exacerbate their actions. Therefore, exploring ways to manage low-level disruption should be a priority to both reduce the impact on other students' learning outcomes and improve the well-being of teachers.

Methods and Results

In order to explore my research questions, I took an action research approach to analyse the behaviour of students in one of my Latin classes, produced a guide to managing the low-level disruption that is currently present in my classroom and then implemented the guide in my teaching to make a tangible change. I then evaluated whether the implementation of the guide had a positive impact on the behaviour of students in the classroom and thus, alleviated the disruption to others.

As I am fortunately placed in the midst of the issues reflected by my research project on a daily basis, as the teacher in the classroom surrounded by low-level disruption, I used my situation to my advantage. I wanted to improve my own practice in managing this behaviour to limit the effects on the learning outcomes of all students in the class. I felt it would be worthwhile to assess the behaviour of Year 8 students in the classroom, but I also needed to consider the limitations to my study, based on the changes to normality because of the unprecedented circumstances of the Covid-19 pandemic.

I took a qualitative approach to collect rich data through a variety of interviews. I conducted semi-structured one-to-one interviews with three Year 8 students, after I sought informed consent from all those involved for their participation in my research. I analysed the opinions, feelings, emotions and experiences of the individual students, by exploring that which they chose to share with me. I also carried out two focus group interviews with members of the teaching staff, in addition to my initial one-to-one interviews of students. I planned to use the qualitative data from the one-to-one interviews of the students and the focus group of the staff to develop and create my artefact; a guide to managing low-level disruption in the classroom. After I had produced my artefact, I implemented its use in my classroom practice and carried out a series of lessons,

two of which were observed by a member of my department. I analysed notes made by my observer, alongside my self-reflections to determine whether the use of the artefact helped to reduce low-level disruption and consequently improved the learning outcomes of the students in the class.

Artefact – A Guide to Managing Low-Level Disruption in the Classroom for Teachers

1) Know your students
- Target students with individual needs (SEN)
- Identify and engage the "leaders" of low-level disruption
- Show you care about the students and greet them as they enter
- Carefully arrange seating plans to diffuse disruptive groups

2) Apply classroom rules consistently
- Control the space as the teacher and move around the classroom
- Students should not enter before the teacher
- Respect for teacher and peers must be enforced
- Starter should be on board for students to complete upon entry

3) Address disruption
- Teacher should explain what specific behaviour needs to stop and why
- Teacher should email student and CC in the Form Tutor/Deme Warden consistently to follow up on disruptive behaviour
- Teacher should keep behind disruptive students to have a one-to-one chat at the end of the lesson to provide a choice for their future behaviour
- Teacher should deflect off-topic questions by asking students to hold onto them/write them down and come back to them later

4) Make use of rewards and praise
- Write names on the board when students do something well
- Engage disruptive students with questions they can answer successfully and then praise them
- Plan for low-level rewards that build up to school-wide rewards
- Use a box of extension tasks that implement the reward system
- Email student with Form Tutor/Deme Warden to consolidate praise

5) Engage students with the lesson
- Differentiate for lower/higher attaining students
- Use directed questioning for students
- Include variety of tasks that can be teacher-led or independent work
- Reduce time to complete each activity
- Incorporate competition into lessons
- Teacher should use tone of voice and body language/gestures

Discussion and recommendations

I found that it is challenging to determine the impact of gender in my school environment because I am not able to contextualise boys' behaviour in a mixed learning space, or whether girls would behave similarly or differently in a single-sex environment. I have noticed in my teaching career that girls and boys present their low-level disruption differently, but I do not think I can deduce that disruptive behaviour is due to gender. However, I did find it interesting that the

20

teachers seemed more conscious of the role that gender could play in the classroom, than the students did.

I established that high attaining students do not necessarily have good study skills and that there is still a spectrum of attainment that could affect student behaviour. This explains why there are plenty of instances of low-level disruption in high attaining schools and I found that it can be most effectively managed through teachers providing differentiation of materials, individualised support and engaging lessons.

I deduced that students are more likely to display positive behaviour in lessons if they have chosen to continue their study of the subject. I also found that positive behaviour can have a domino effect in the same way that low-level disruption can and that engaging some students can result in more students following suit. It is important for teachers to apply consistency in asserting rules and consolidating praise to gain the respect of the students.

I personally found that the artefact was an effective tool to manage low-level disruption when used consistently in lessons and encompasses the key areas needed to promote positive behaviour from students. I found that the artefact can be just as effective for students with SEN, as it is for those without. I recognise that the challenge for teachers is juggling the consistent delivery of the different elements of the artefact, in addition to the delivery of the lesson content and upholding the other requirements of teaching.

I feel that all the elements of the artefact are important, but that the most effective management techniques stem from knowing the individual students and personalising support, conversations and tasks, alongside making use of rewards and praise, particularly by incorporating a low-level reward system. I would like to disseminate my artefact across the departments at my school and encourage all of them to develop their own low-level

reward system. I believe that the students will respond best to these rewards, if they see them being used across the school with minor tweaks depending on the subject. It is a way for teachers to show appreciation and recognition of the efforts of individuals, even if they are not the most high achieving students. Teachers might also need to refine the system over time, so that it does not become less effective as the students become more comfortable with it.

I believe that this research and artefact will transform my practice by allowing me to manage low-level disruption better and create a positive, supportive and consistent learning environment for both students and staff. I believe it has the power to combat the most significant behaviour issue present in my school, if it is adapted appropriately and used consistently by all members of staff across the school.

References

Feldmann, L.J. (2001). Classroom civility is another of our instructor responsibilities. *College Teaching,* 49(4), 137–140.

Lewis, Ramon. (2009). *Understanding Pupil Behaviour.* Routledge: London.

Schwab, Susanne, Eckstein, Boris, and Reusser, Kurt. (2019). "Predictors of Non-compliant Classroom Behaviour of Secondary School Students. Identifying the Influence of Sex, Learning Problems, Behaviour Problems, Social Behaviour, Peer Relations and Student–teacher Relations." *Journal of Research in Special Educational Needs* 19(3), 220-31.
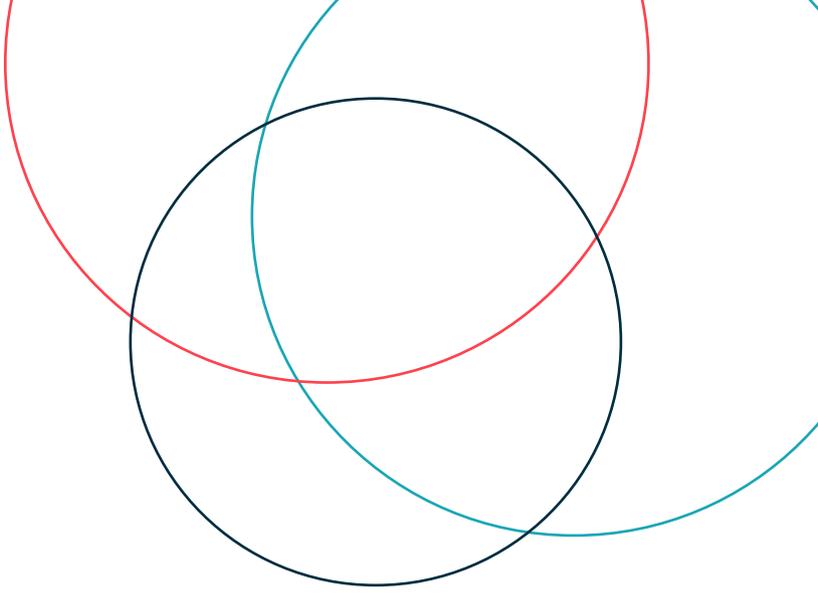
An article exploring the impact of single-sex environments on the retention of girls studying A-level Physics

Amy Holland

Abstract

This article explores two research studies carried out into the impact of single-sex classrooms on girls studying physics in comparison to co-educational settings. Both research papers look to conclude whether single-sex physics classes have a positive impact on young girls studying the subject, with particular focus on their engagement, retention and, self-concept of their ability. The aim of this literature study was to determine whether the creation of a single-sex environment to study physics within a co-educational sixth form setting could encourage the retention of girls at A-level physics and increase the number of girls pursuing physics beyond A-level.

Background

Females continue to be under-represented within scientific fields which is evident in both their time at school and pursuit of careers within the field. In 2019, 76% of jobs in STEM subjects were held by males (STEM Women, 2021) and this is a problem that the Institute of Physics has been addressing since 2011. Data from 2018 shows that 68% of schools with girls in England send fewer than two girls to A-level physics whilst 44% of schools send no girls at all (Institute of Physics, 2018). Data within the report demonstrates that girls perform just as well as boys in physics which suggests that there are other reasons that girls are choosing to move away from subjects with a scientific focus.

Subjects such as physics have been viewed traditionally as having a strong masculine attribution and these gender-science stereotypes can influence a young person's aspirations to pursue a subject (Makarova, et al., 2019). Findings have suggested that the use of mentors and role models in addition to training in unconscious bias would have a positive impact on the gender imbalance in STEM subjects (Institute of Physics, 2017; Jamieson, 2018). It has also been observed that the uptake from single-sex classrooms is notably different than from co-educational maintained schools and the Institute of Physics (2018) found that single-sex independent schools send almost four times as many girls to A-level as mixed-sex schools (Institute of Physics, 2018).

Summary of the papers

1. Abraham, J. & Barker, K. (2020) Motivation and Engagement with Physics: a Comparative Study of Females in Single-Sex and Co-educational Classrooms. Research in Science Education (Australasian Science Education Research Association), 50(6), pp. 2227-2242

Background and Research Aims

Within this paper, Abraham and Barker (2020) explore whether a single-sex environment for females can have a positive impact on attitudes and engagement in physics. This paper formed part of a larger project that examined the intentions of students studying physics in their first year of senior high school. Their research reports findings from 90 female physics students across eight single-sex and co-educational schools in New South Wales, Australia. As this research has been carried out in Australia, the first year of senior high school is equivalent to the first year of A-level studies in the United Kingdom. Similar to the A-level structure in England, students have the option to opt out of physics after the first year of senior secondary physics.

This study aimed to identify whether there are any differences for female students studying physics at single-sex and co-educational establishments, with a focus on comparing the students' "motivation, engagement, and sustained enrolment plans" (Abraham & Barker, 2020, p. 2227). Abraham and Barker acknowledged a gap in the research by focusing their study on specific physics topics being studied rather than an assumption that the motivation and engagement for a student will be consistent over the course of study.

Design and Data Collection

There were four data collection points during the study which corresponded with the completion of four physics modules in the curriculum. Participants in the study were asked to consider six areas at each of the data collection points. These focused on the level of interest, performance perceptions, sex-stereotypes attitudes, sustained engagement, intention to continue studying the subject and, the utility of the module. Data was collected using the Physics Motivation Questionnaire (Abraham & Barker, 2014) which allowed them to measure 'attitude' in a scientifically validated manner and analyse the data statistically. In addition to this data collection, they initially asked the students to complete a survey to explore their backgrounds to allow for a comparative study.

Main Findings and Conclusions

Abraham and Barker's findings revealed some significant differences across the data collection points. Single-sex (female) cohorts demonstrated slightly higher values for performance perceptions in addition to the interest and utility values when studying the modules. Gender role beliefs were consistent across the modules and after the year, motivation levels were consistent for both cohorts with no notable difference between the two groups for discontinuing the subject. Overall, Abraham and Barker's findings provided evidence of similarities between the two school structures and students were found to be highly motivated in both environments. Although there was some difference in the data collected, it did not provide sufficient evidence that single-sex classes lead to improved motivation.

2. Kessels, U. & Hannover, B. (2008) When Being a Girl Matters Less: Accessibility of Gender-related Self-knowledge in Single-sex and Coeducational Classes and Its Impact on Students' Physics-related Self-concept of Ability. British Journal of Educational Psychology, 78(2), pp. 273-289.

Background and Research Aims

Kessels and Hannover (2008) explore whether there are positive effects on girls' self-concept of ability in physics in single-sex education and test whether this is due to the lower access of "gender-related self-knowledge" (Kessels & Hannover, 2008). Given that self-concept of ability can be an important component of overall achievement outcomes in a subject, they aimed to test the extent to which single-sex settings could promote a positive self-belief of ability for girls studying physics (Kessels & Hannover, 2008). Secondly, Kessels and Hannover wanted to understand why this style of classroom could have such an impact on female students. Kessels and Hannover's study was carried out in German schools which provided the opportunity to randomly assign students at the co-educational schools into single-sex and co-educational classrooms. In other countries, this would have been difficult to implement however due to German laws, they were able to temporarily track the students by gender.

Design and Data Collection

Kessels and Hannover used a sample size of 401 students, with an average age of 14 years, of which 210 were girls from four comprehensive co-educational schools in Berlin, Germany. The students were randomly assigned to a single-sex or a co-educational class for the entire year of education and no other changes were made. At the end of the year, students' opinion of their ability was explored, and students were asked to complete a 5-point Likert scale questionnaire with statements about their personal belief of their ability. Within a smaller sample of 134 students, the students were assessed on their association with gender-stereotyped adjectives. This was to explore whether the educational setting had an impact on their gender-related self-knowledge and in turn affected how they deemed their ability. To measure the accessibility of gender-related self-knowledge, students were asked to judge whether a particular adjective described them. These words were displayed on a computer programme and included typical stereotypes associated with the genders. In addition to the choice of word, latency data was used to determine the timing of a decision in milliseconds.

Main Findings and Conclusions

Kessels and Hannover's investigation into the self-concept of ability in physics determined that girls had a lower self-concept of physics ability in mixed-sex classes compared to single-sex whilst boys reported a high self-concept of physics ability irrespective of the teaching environment. Their research into gender-stereotyped language association demonstrated that boys endorsed more masculine adjectives in co-educational classrooms than in all-boys settings, although classroom setting made no impact on their association with feminine adjectives. Data showed an indication that girls chose more feminine traits more frequently in co-educational settings, however, Kessels and Hannover were unable to conclude a significant

impact of the educational environment on the gender-stereotyped adjectives. Using the latency data, they did, however, observe that girls in mixed-sex classes judged feminine traits more quickly while girls in single-sex settings tended to judge both the feminine and masculine adjectives equally.

Combining their research studies, Kessels and Hannover concluded that the more masculine traits girls associated with, the higher their self-concept of ability. This was supported by the latency data ; the more quickly girls responded to the feminine traits, the lower their self-concept of ability. Overall, Kessels and Hannover confirmed their assumption that girls report a higher self-concept of physics ability in single-sex classes, whilst boys are unaffected by the classroom setting. Their findings propose that the reduced accessibility of gender-bias language in single-sex classrooms has a notable impact on a girl's self-concept of ability.

Reflection

I hoped to use these studies to determine whether the creation of single-sex classes in a co-educational sixth form setting could encourage the retention of girls at A-level Physics and increase the number of girls pursuing physics beyond A-level. Kessels and Hannover's research has allowed me to realise how attitudes surrounding gender stereotypes can be impacted by the context of the classroom. Although arranging this separate classroom setting is not feasible, it has led to the organisation of a weekly discussion group for the female students studying physics at UCS. I look forward to determining whether this collaborative group leads to more positive self-concepts of ability in physics and an increase in retention of the subject. Although this will not be a measure of the impact of single-sex classes, it will provide a support network for the girls outside of the more male-dominated classroom environment. The research I have explored within this assignment has reinforced my passion for promoting the subject and reflecting on this topic has given me further desire to develop strategies that support and guide young women studying the subject.
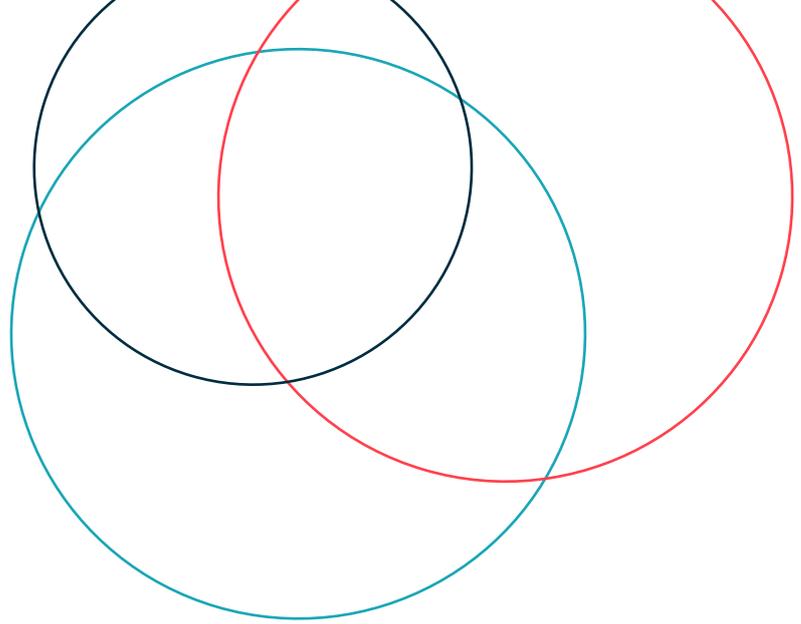
References

Abraham, J. & Barker, K., (2014) Sustaining young people's enrollment intentions in relation to physics: Development and validation of a tool. *The Australian Journal of Education & Developmental Psychology,* Volume 14, pp. 93-116.

Abraham, J. & Barker, K., (2020) Motivation and Engagement with Physics: a Comparative Study of Females in Single-Sex and Co-educational Classrooms. *Research in Science Education (Australasian Science Education Research Association),* 50(6), pp. 2227-2242.

Institute of Physics, (2012) *It's Different for Girls: The influence of schools*, London: s.n.

Institute of Physics, (2017) *Improving Gender Imbalance*, London: s.n.

Institute of Physics, (2018) *Why not physics? A snapshot of girls' uptake at A level,* London: s.n.

Jamieson, V., (2018) Women in physics: Why there's a problem and how we can solve it https://www.newscientist.com/article/mg24032031-900-women-in-physics-why-theres-a-problem-and-how-we-can-solve-it/#ixzz6rWxGBW1X. *New Scientist,* 10 November, p. Issue 3203.

Kessels, U. & Hannover, B. (2008) When Being a Girl Matters Less: Accessibility of Gender-related Self-knowledge in Single-sex and Coeducational Classes and Its Impact on Students' Physics-related Self-concept of Ability. *British Journal of Educational Psychology,* 78(2), pp. 273-289.

STEM Women, (2021) *Women in STEM | Percentages of Women in STEM Statistics.* [Online] Available at: https://www.stemwomen.co.uk/blog/2021/01/women-in-stem-percentages-of-women-in-stem-statistics[Accessed 11 April 2021].

To what extent does Shen's self-discovery vocabulary learning method impact learners' ability to recognise and produce two character words in written Mandarin?

Maria Pia Maggioni

. . . .

Abstract

In the academic year 2020/2021 I carried out an action research project with my year nine Mandarin students as part of my participation in the Subject Expert Micro-Programme[1] organised by the UCL Institute of Education Confucius Institute for Schools.

This research was carried out by four classroom teachers, based in four different secondary schools across England. The research topic chosen derived from observations that the teacher researchers made in their daily practices. Despite having a diverse teaching background, all teacher researchers found the learning of Chinese words (retention, recognition and production) one of the most challenging aspects of Chinese teaching pedagogy. The research was thus carried out to explore how beginner Mandarin learners could learn Chinese vocabulary more efficiently.

1 A continuing professional development course developed by UCL IOE Confucius Institute for Schools for qualified school teachers of Mandarin.

Introduction and context:

The study was conducted by four teacher researchers who were teaching at four different secondary schools (11-18 years old) across England. The teacher researchers worked collaboratively to design and carry out the study uniformly, with unavoidable slight discrepancies due to the nature of classroom research.

The study consisted of three main parts: (1) Intervention, (2) Vocabulary baseline and post-tests, and (3) Student survey. The whole process took between three to five weeks.

The research focused on the effective learning of words 词, not single characters 字. Taking reference from Shen and Xu's (2015) self-discovery method, three intervention worksheets were developed. Each worksheet contained and evolved around three two-character words extracted from the textbook Jinbu One Chapter Five: Food and Drink.

Set 1　水果　红茶　鸡蛋

Set 2　饺子　牛奶　炒饭

Set 3　草莓　羊肉　葡萄

The worksheets were aimed to guide student engagement in active learning, such as discovering the elements of characters that formed the words and the relationship between the elements or characters of a word.

For example, students were required to find the pinyin or individual meaning of a character or word that they had never learned in school lessons before.

1.2 Use « handwriting » function of Pleco on your phone and look for the pinyin of each of the characters. Put the pinyin with the correct tone in the table below.

Pinyin:		
	红	茶

For compound characters (or 'phonetic-semantic character'), students were asked to find out the semantic component (or 'radical') and were encouraged to make connections with the phonetic element.

2.4 Guess why 火 [fire] is in the character 炒. Explain your guess in English in the box below:

2.5 Guess why 饣 [eat] is in the character 饭. Explain your guess in English in the box below:

For pictographic characters, students were shown the ancient/traditional form and encouraged to make connections with its modern/simplified form.

3.4 Look at the traditional form of 水 below. Explain in English in the box how you will memorise this character in modern form.

Traditional Form (金文)	Modern Form	
水	水	

3.5 Look at the traditional form of 果 below. Explain in English in the box how you will memorise this character in modern form.

Traditional Form (金文)	Modern Form	
果	果	

3.6 Use pinyin to look up what 水果 means exactly on Pleco and explain why it means that in English in the box below:

Before and after each intervention, a recognition test and a production test were given to students that contained the same words in the worksheet, as illustrated below:

| Baseline recognition test Baseline production test | → | Intervention worksheet | → | Post recognition test Post production test |
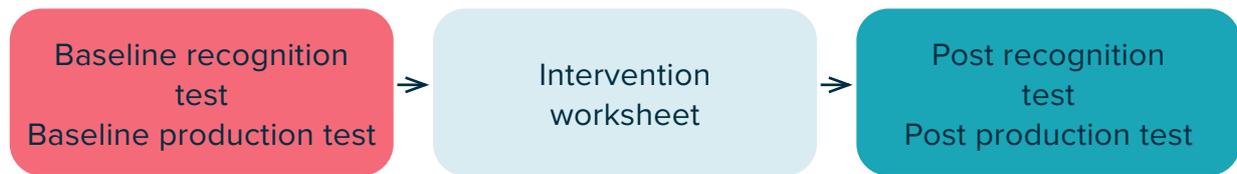
The same cycle ran three times in accordance with the three different sets of vocabulary and their corresponding worksheets.

1. Recognition tests: Students were shown the words (in Chinese characters only) and asked to write down their English meanings. One mark was awarded for the correct English meaning for each word.

2. Production tests: Students were shown the English meaning of the words and asked to write down their Chinese characters. Teacher researchers marked the production tests based on the same marking schemes (as displayed to the right).

3. Students were also given the same questionnaire at the very beginning of the data collection process and right after the last set of vocabulary post-tests. It was aimed at finding out how confident students felt about learning and reproducing Chinese characters before and after the intervention period.

Students were asked to rate each of the following statements on a scale of 1 (strongly disagree) to 5 (strongly agree):

1. I am able to remember the meaning of the words in characters that I have just been taught.
2. I am able to remember the pronunciation of the words that I have just been taught.

Mark Scheme for Production Accuracy Score

3 marks: for each accurate character or character with minor errors such as a malformed stroke in a character, or errors of proportion/balance/alignment between the components within a character

2 marks: for each character with 1-2 extra or missing stroke

1 mark: for each character with 3-4 extra or missing strokes but the general shape of which still allows native speakers to recognize

0 marks: for each character with more than 4 strokes or not recognisable by native speakers

3. I am able to write the words in characters that I have just been taught by copying.
4. I am able to write the words in characters that I have just been taught from memory.
5. I believe that I have the ability to learn words in characters well.
6. I think learning words in characters is easy.

Results:

Finding 1

There was a consistent increment of scores across all schools in the recognition tests after the intervention, as shown in Figure 1 below.
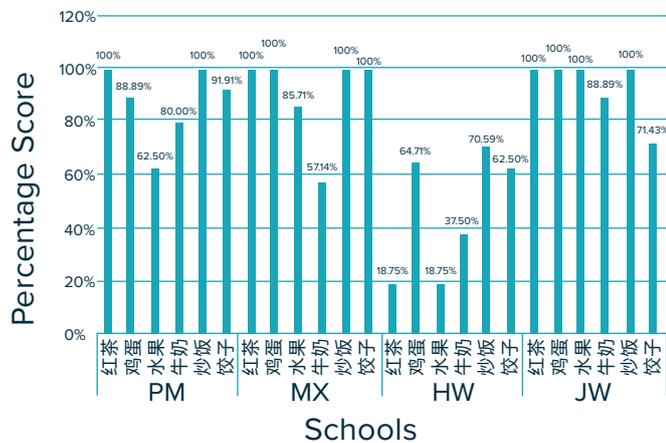

Figure 1. Recognition test: Change of accuracy in % for each of the 6 target words after the intervention.

Below is a breakdown of students' average score increase from each school, showing improvement in the recognition test after the intervention:

School	Score increase in Post-Recognition Test
PM	87.22%
MX	90.48%
HW	45.47%
JW	93.39%

In general, students performed better in some 2-character words than others:

Words in Recognition Tests	Score increase across the schools
红茶	100% increase in 3 schools, 19% in 1 school
炒饭	100% increase in 3 schools, 71% in 1 school
鸡蛋	88% increase in all schools
饺子	81% increase in all schools
水果	67% increase in all schools
牛奶	66% increase in all schools

Finding 2

Students in all schools showed improvement in their ability to re-write learned words in Chinese characters after the intervention.
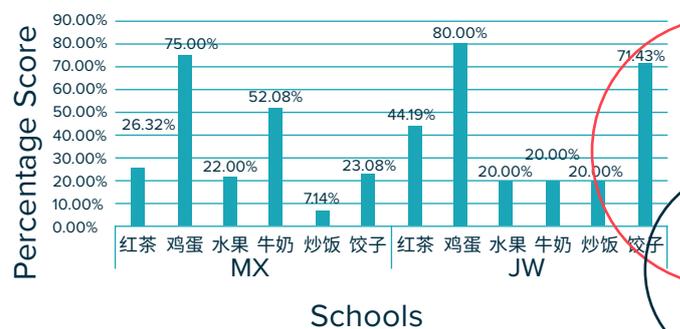

Figure 2. Production test: Change of accuracy in % for each of the 6 target words after the intervention.

There were some errors in data collection for the production tests, hence only data from two schools were included in this analysis.

The average score increase for each school is below:

School	Score increase in Post-Production Test
MX	34.27%
JW	42.60%

Below is a breakdown of students' score increase for each character in the production test after the intervention:

Words in Production Tests	Average scores increase for all schools
鸡蛋	57.50%
饺子	47.26%
牛奶	36.04%
红茶	35.26%
水果	21.00%
炒饭	13.57%

Finding 3

A correlation test has been run between pre-survey and pre- recognition and production tests (Cronbach's alpha: 0.826). Generally, a Cronbach's alpha of 0.70 and above is considered good for social science research, a maximum alpha value of 0.90 is recommended. An alpha value of 0.826 in this study suggested that the survey questions have relatively high internal consistency, indicating that the questions are highly reliable in measuring what the survey is meant to measure (Dennick & Tavakol, 2011).

Based on the relatively high Cronbach's alpha scores mentioned above, correlation tests between pre-questionnaire and pre-tests (both recognition and production tests) were run, which revealed some interesting results as illustrated in Figure 3 and Figure 4.

Correlations		Q1PRE	Q2PRE	Q3PRE	Q4PRE	Q5PRE	Q6PRE	Recognition Baseline
Q1PRE	Pearson Correlation	1	.355**	.490**	.463**	.568**	.459**	.465**
	Sig. (2-tailed)		.009	.000	.000	.000	.001	.000
	N	53	53	53	53	53	53	53
Q2PRE	Pearson Correlation	.355**	1	.441**	.269	.620**	.322*	.156
	Sig. (2-tailed)	.009		.001	.051	.000	.019	.266
	N	53	53	53	53	53	53	53
Q3PRE	Pearson Correlation	.490**	.441**	1	.406*	.332*	.347*	.178
	Sig. (2-tailed)	.000	.001		.003	.015	.011	.203
	N	53	53	53	53	53	53	53
Q4PRE	Pearson Correlation	.463**	.268	.406**	1	.423**	.588**	.408**
	Sig. (2-tailed)	.000	.051	.003		.002	.000	.002
	N	53	53	53	53	53	53	53
Q5PRE	Pearson Correlation	.568**	.620**	.332*	.423**	1	.609**	.192
	Sig. (2-tailed)	.000	.000	.015	.002		.000	.168
	N	53	53	53	53	53	53	53
Q6PRE	Pearson Correlation	.459**	.322*	.347*	.588**	.609**	1	.283*
	Sig. (2-tailed)	.001	.019	.011	.000	.000		.040
	N	53	53	53	53	53	53	53
Recognition Baseline	Pearson Correlation	.465**	.156	.178	.408**	.192	.283*	1
	Sig. (2-tailed)	.000	.266	.203	.002	.168	.040	
	N	53	53	53	53	53	53	53

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3. Correlation between each Pre-Survey question and Pre-Recognition Test.

Figure 3 indicated that the following questions in the Pre-Survey showed relatively high correlation with the Pre-Recognition Test, meaning that the questions are highly related to the score in the test.

- Survey Question 1: I am able to remember the meaning of the words in characters that I have just been taught
- Survey Question 4: I am able to write the words in characters that I have just been taught from memory
- Survey Question 6: I think learning words in characters is easy

Correlations		Q1PRE	Q2PRE	Q3PRE	Q4PRE	Q5PRE	Q6PRE	Recognition Baseline
Q1PRE	Pearson Correlation	1	.567**	.460*	.531**	.767**	.654**	.487*
	Sig. (2-tailed)		.003	.018	.005	.000	.000	.012
	N	26	26	26	26	26	26	26
Q2PRE	Pearson Correlation	.567**	1	.436*	.447*	.539**	.494*	.710**
	Sig. (2-tailed)	.003		.026	.022	.005	.010	.000
	N	26	26	26	26	26	26	26
Q3PRE	Pearson Correlation	.460*	.436*	1	.392*	.156	.121	.329
	Sig. (2-tailed)	.018	.026		.048	.447	.555	.101
	N	26	26	26	26	26	26	26
Q4PRE	Pearson Correlation	.531**	.447*	.392*	1	.465*	.416*	.416**
	Sig. (2-tailed)	.005	.005	.048		.017	.034	.101
	N	26	26	26	26	26	26	26
Q5PRE	Pearson Correlation	.676**	.539**	.156	.465*	1	.830**	.458*
	Sig. (2-tailed)	.000	.005	.447	.017		.000	.019
	N	26	26	26	26	26	26	26
Q6PRE	Pearson Correlation	.654**	.494*	.121	.416*	.830**	1	.457*
	Sig. (2-tailed)	.000	.010	.555	.034	.000		.019
	N	26	26	26	26	26	26	26
Recognition Baseline	Pearson Correlation	.487**	.710**	.329	.416*	.458*	.475*	1
	Sig. (2-tailed)	.012	.000	.101	.035	.019	.019	
	N	26	26	26	26	26	26	26

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4. Correlation between each Pre-Survey question and Pre-Production Test.

Figure 4 indicated that the following questions in the Pre-Survey showed relatively high correlation with the Pre-Production Test, meaning that the questions are highly related to the score in the test.

- Survey Question 1: I am able to remember the meaning of the words in characters that I have just been taught
- Survey Question 2: I am able to remember the pronunciation of the words that I have just been taught
- Survey Question 4: I am able to write the words in characters that I have just been taught from memory
- Survey Question 5: I believe that I have the ability to learn words in characters well
- Survey Question 6: I think learning words in characters is easy

Pedagogical implications:

We found that the intervention worksheet is a very good way of teaching characters. It has enabled students to discover the different meanings for the same character, they can also find out more about the knowledge of the characters eg. the traditional forms and stories behind them.

We found that using the intervention worksheets was a bit time consuming. To resolve this, we suggest changing the layout of the worksheets and limiting them to one page each, making it "seem" even more welcoming and accessible to all students. We hope this will also help overcome the drawbacks shown in the survey. We would happily use the worksheets again especially as a flipped learning resource.

From informal discussions with students, we learned that some students may not feel confident about the new learning format (intervention worksheet). We suggest giving students more time to digest and get used to the learning method to see if students can benefit more from it.

Further research:

Areas that could be developed further are:

- How to move this research from word-level to sentence-level;

- How to move this research from word-level to sentence-level;

- Understanding further the findings about the correlation between pronunciation and character learning;

- The effect of using the intervention worksheet as flip learning for vocabulary;

- The relationship between independent character learning and students' motivation or confidence.

References

Grenfella, M., Harris, V. (2015) *Memorisation strategies and the adolescent learner of Mandarin Chinese as a foreign language,* Linguistics and Education, 31, 1-13

Higes, G. I., Taylor, H. A., Thomas, A. K. (2018) *Study techniques differentially influence the delayed judgment of learning accuracy of adolescent children and college-aged adults,* Metacognition and Learning, 13(5)

Shen, H., Xu, W. (2015) Active Learning: *Qualitative Inquiries Into Vocabulary Instruction in Chinese L2 Classrooms,* Foreign Language Annals, 48(1)

Wang, V., Farmer, L. (2008) *Adult Teaching Methods in China and Bloom's Taxonomy,* International Journal for the Scholarship of Teaching and Learning, 2(2)

Yuan, A. (2020) *Chinese Language Teaching Methodology and Technology*, Vol. 3, Iss. 1, Art. 6

How does low stakes or no stakes testing in every lesson affect student outcomes?

Katie Matthews
Emma Taylor
Andrew Quirke

. . . .

Abstract

Research from the field of cognitive science (see Chandler and Sweller, 1991; Bjork et al, 2013; Dunlosky, 2013; Brown et al, 2015) has highlighted a number of different strategies that can be used to improve the quality of learning. One such example is that of 'retrieval practice' or recreating something you've learned in the past from your memory, and thinking about it right now (Weinstein and Sumeracki, 2018). In other words, being given time to forget content taught in lessons, before being asked to remember or 'retrieve' it.

The aim of this project was to assess how using forms of retrieval practice regularly in lessons affects students' overall performance across Key Stage 3 and 4. Therefore, low stakes or no stakes testing[2] was introduced into every lesson taught at UCS by Emma Taylor (Year 8) and Katie Matthews (Year 10) across the Autumn term and into the Spring term of 2019. Students were given an extra exercise book in which they were expected to carry out the tests during lessons.

Methodology

This project was run as an action research project whereby an intervention was tested to assess its overall effectiveness on student outcomes. As such, low stakes or no stakes testing was introduced into every lesson taught at UCS by Emma Taylor (Year 8) and Katie Matthews (Year 10) across the Autumn term and into the Spring term of 2019. Students were given an extra exercise book in which they were expected to carry out the tests during lessons. These tests could take place at any point during the lesson, and consisted of short answer quiz questions, multiple choice tests, brain dumps and other similar retrieval exercises designed to test recall.

The intervention was evaluated in the following three ways:
1. A focus group was conducted with selected Shell students at the beginning and end of the study
2. A survey was issued to all Remove students participating in the study in order to ascertain their views on the effectiveness of the strategy.
3. Performance data for the Remove was collected and analysed from the beginning and end of the study

2 Low stakes or no stakes tests are defined as follows:
Low stakes: A record of the student's test result is taken by the student but not necessarily by the teacher. Could be peer marked
No stakes: No record of the test result is taken by the student or teacher

Results and analysis:

From the survey and focus groups:

When asked if they found retrieval practice beneficial, themes that were common amongst students' answers were:

- Being able to see clearly where they needed to improve
-
- Being able to keep track of their own progress and see where they were building knowledge by getting questions right in future quizzes
-
- Being able to see how they could use the strategy beyond the class tests
-
- Feeling more positive before subsequent higher stakes tests

When pupils were asked if they found the process of retrieval practice useful they were overwhelmingly positive (see figure 1)

Would you like to do this activity in other subjects?

37 responses
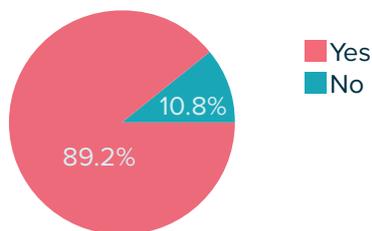
Figure 1. Pie chart showing the extent to which students would like to see retrieval exercises replicated in lessons.

"it helped to secure that knowledge that was on the edge of being forgotten. Because of the questions, I found that my recall in practice tests was better without previous revision."

"I could use it for a range of other subjects, and I think this type of practice could apply to all of them."

"I still feel worried about upcoming tests but I feel that it may have helped to prepare me and may reduce the stress."

The vast majority of students surveyed felt it helped them to improve their academic outcomes in Chemistry (see Figure 2)

Has it help you do better in chemistry?

37 responses
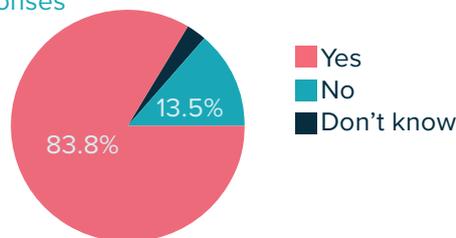
Figure 2. Pie chart showing the extent to which students felt the retrieval exercises had enabled success in chemistry

"yes because my test results are significantly higher than last year and I am able to remember things that I had previously forgotten"

From the performance data:

Effect size has been popularised as a means of showing the comparative effect of an intervention on student academic attainment by Hattie (2009).

Here, we look at the effect size of the intervention on a group relative to the same group's performance in a previous exam. This minimises the possible influence of different teachers which our preliminary analysis demonstrated to be a significant

influence for teachers teaching the same topic. Table 1 shows the key statistics for the group that received the intervention and the calculated effect size. Figure 1. shows a visual representation of their results demonstrating the improvement in both the median result (thick horizontal line in the box).

This indicates that the intervention not only had a positive impact on the whole group but a very significant impact on the less academically successful students. For comparison, this intervention has a similar effect size as effective feedback, student learning strategies and peer tutoring, as highlighted by Hattie (2009).

	Pooled	Periodic Table (post-intervention)	Structure and bonding (pre-intervention)
Mean (%)	88	93	83
Median (%)	91	96	84
Std (%)	12	7	15
Number	78	39	39
Effect Size	0.79		

Table 1. Effect size and key statistics for a test group on an end-of-unit test before receiving the intervention and the same group after receiving the intervention on a different end-of-unit test in Chemistry.

Conclusion:

Overall our findings show a positive effect on student outcomes.

Students were able to articulate how the intervention has supported their learning and they could identify evidence of personal academic improvement across the term. Furthermore, our evidence shows that the intervention has helped to improve their confidence in a test setting. Unintended outcomes include students utilising their retrieval practice books as revision tools. Statistical analysis showed an intra-group effect size of 0.79, and although there are confounding variables to consider, this is still extremely significant in an educational setting. Since carrying out this particular project, the findings have been disseminated to the UCS teaching community and low stakes/ no stakes testing is now a regular feature of lessons at the school.
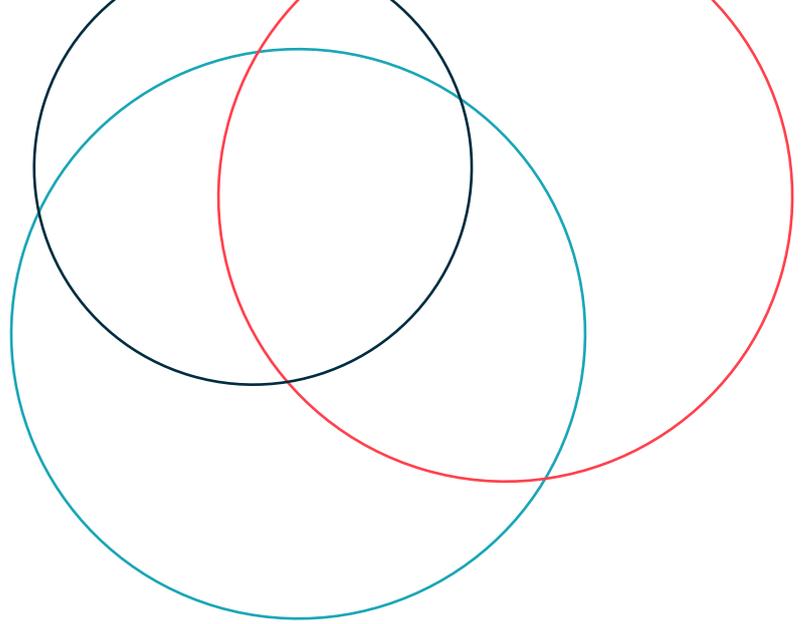
References

Bjork, R., Dunlosky, J., and Kornell, N. (2013). Self regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology,* 64, 417-444

Brown, P., Roediger, H., and McDaniel, M. (2015). *Make it stick: The science of successful learning.* Boston: Harvard Press

Chandler, P., and Sweller, J. (1991). Cognitive Load Theory and the format of instruction. *Cognition and Instruction* 8(4): 293- 332

Dunlosky, J. (2013). Strengthening the student toolbox: Study strategies to boost learning. *American Educator* 37(3): 12-21

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement.* London: Routledge

Weinstein, Y., and Sumeracki, M. (2018). *Understanding How We Learn*. New York: Routledge
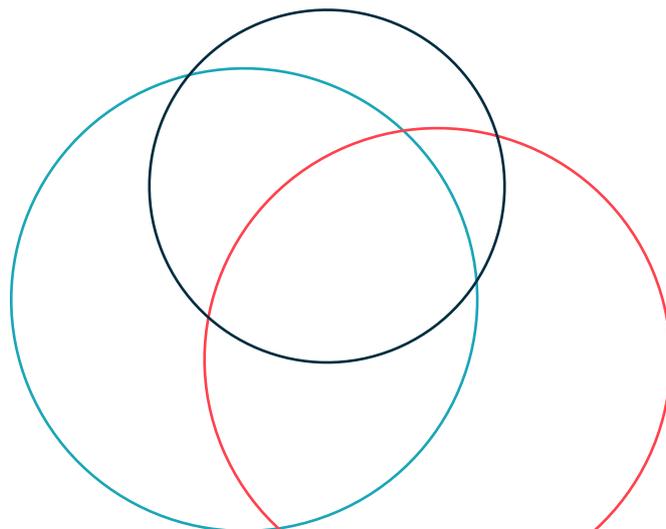
Extended Essay
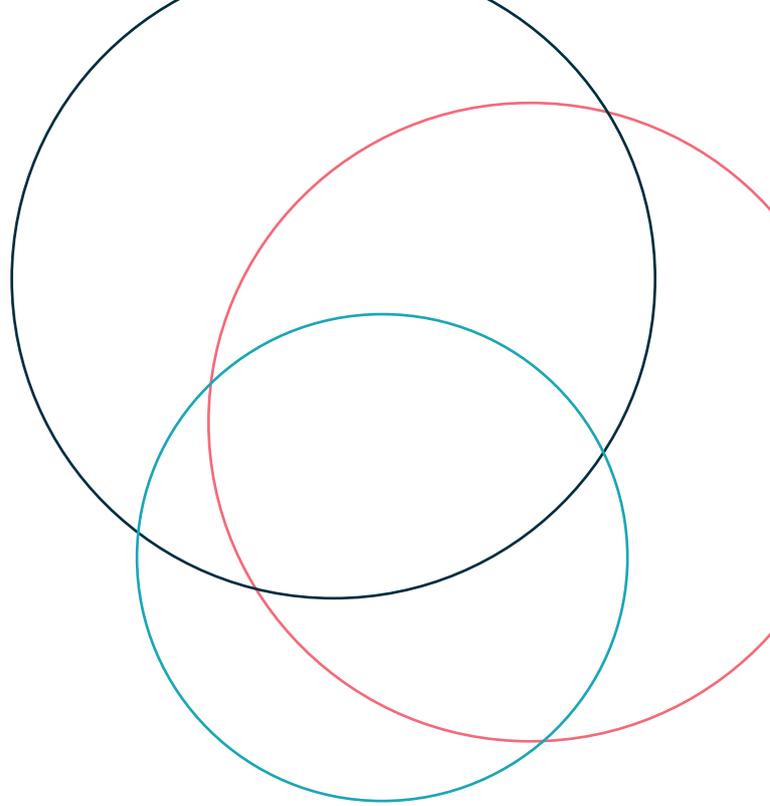
A 'powerful instrument for policy reform' or 'impoverishing our classrooms'?

Analysing the influence of PISA on global education policy

Abby Caplin

Why focus on PISA?

When the first OECD Programme for International Student Assessment (PISA) test took place in 2000, 32 countries and 200,000 students took part; as of the last survey in 2018, there were 79 countries and economies and 710,000 students participating, with 88 countries scheduled to take part in PISA 2021 (now PISA 2022). Some estimate that by 2030, over 80% of countries across the world will be taking part in PISA. By sheer numbers alone, it is undeniable that PISA is presently the most influential global education test, particularly in relation to its impact on global education policy.

Angel Gurria, OECD Secretary General, wrote in the foreword to the 2015 PISA results:

Over the past decade...PISA has become the world's premier yardstick for evaluating the quality, equity and efficiency of school systems.

(OECD, 2015)

This notion of PISA as the 'premier yardstick' for evaluating national education systems is not a sentiment that many scholars, policymakers or stakeholders would agree with. Many argue that PISA has in fact had a detrimental effect on global education since its inception. With this in mind, taking into account the origins of the PISA test, the exponential increase in participation over the last 20 years and the escalating role of the OECD, this essay will discuss both the beneficial and detrimental effects of PISA on global education reform, and what PISA's role, if any, should be in shaping education policy development.

The advantages of PISA for policy reform

Proponents of the triennial PISA survey often espouse one of the key advantages of PISA as being the way that it highlights examples of 'best practice'. These proponents, namely the OECD and governments who utilise

PISA data to inform policy, emphasise the significance of PISA in enabling countries to identify features of, and learn from, the top-performing nations and/or economies and then apply these features in their own contexts. The OECD suggests that PISA results are intended to be used for; diagnostic information at the country level, comparisons over time within each country and comparisons with other countries (Taut and Palacios, 2016). In this way, PISA becomes what certain academics have described as 'a tool for mimetic isomorphism', which enables countries to look at each other's education systems over time and distinguish which practices to borrow or lend (Di Maggio and Powell, 1983; Lockheed and Wagemaker, 2013, cited by Wiseman and Waluyo in Volante, 2018). A number of academics highlight that:

> PISA results have traditionally been promoted as an instrumental way to help governments "borrow" policies from effective education systems, often referred to as 'reference societies' in the hope of emulating their high achieving success withing their own national education context. (Morgan and Volante, 2016; Sellar and Lingard, 2013, cited by Volante, 2018)

When looking at the results of the testing, countries may look at league-toppers, alongside interrogating those who have slipped in the rankings, to ascertain the causes of their 'success' or 'failure'. Countries may then utilise this data to identify examples of 'best practice' to help inform their own education policies.

Furthermore, there are additional potential risks for the 'policy borrowers', some of which lie with the testing itself. Dan Murphy highlights that:

> PISA is one of the largest non-experimental research exercises the world has ever seen...In non-experimental research, 'causation should not be inferred from correlation'. (Gorard, cited in Murphy, 2014)

This suggests that, whilst the resulting conclusions may not be false, countries cannot necessarily or accurately deduce which particular educational policies are effective from a high-ranking PISA score, which limits their ability to apply 'best practice' effectively. Murphy goes on to note the importance of secondary analysis when looking at qualified data such as that of PISA testing in order to fully legitimise potential policy decision-making (Murphy, 2014). This approach is evident in the case of Germany who successfully utilised PISA data to reform their teacher training programme following their PISA shock in 2000. In the development of the resulting SINUS programme, addressing low-quality teaching in science and mathematics, German policymakers drew on PISA data to inform their programming. However, crucially, they did not draw on PISA data alone, but rather focused on additional findings from research on teaching and learning, particularly in relation to studies conducted after TIMSS 1997, recognising the necessity for secondary data to validate the PISA findings.

Another issue with attempting to apply 'best practice' is the risk of creating global education hegemony, linking back to the previously mentioned idea of PISA as a tool for 'mimetic isomorphism'. On the surface, this may not seem problematic if this hegemony leads to more effective systems and better transnational educational outcomes. Schleicher notes that a key advantage of PISA is that it provides policymakers with, 'helpful tools to improve
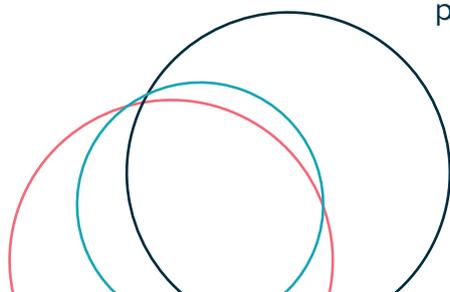
the quality, equity and efficiency in education, by revealing some common characteristics of students, schools and education systems that do well' (Schleicher, 2007). However, encouragement for nations to strive for a normative set of 'common characteristics' does not account for the social, cultural, religious or financial individualities of the high-performing nations or those seeking to borrow their policies, whereby certain educational approaches may not be suitable or effective in different countries. Financially, another area which is not considered, even in terms of PISA rankings themselves, is GDP and the proportion of national expenditure on education, a level of funding that the 'policy borrowing' nation may not have similar access to. This may therefore make the attempt to effectively implement similar policies impossible.

Kerstin Martens notes:

> Since rating and ranking activities by the OECD appear to be based on objective criteria…It puts states under pressure to import and apply models for education which seem to have worked better in other countries instead of continuing on their own path. (Martens, 2012)

Therefore, whilst application of best practice has its advantages, in many respects it reduces national autonomy by underestimating the value of individual nations providing quality education that is suitable for their own contexts and disregards the suitability of different policies in different national contexts.

An additional benefit of PISA data with regards to policy reform is that it has the capacity to provide robust, verified international comparison points for countries looking to make reforms, alongside the ability to then track any resulting changes in educational outcomes over time. While standards-based

reform has traditionally relied on national and regional large-scale assessment results for its data, many national governments have turned or are turning to PISA to inform their curricular and assessment reforms (Volante, 2019). Some of this can be attributed to the way in which the OECD infrastructure makes PISA data accessible and digestible through its open-access resource documents and reports such as Pisa in Focus, Teaching in Focus and Education Indicators in Focus, which provide vital comparative data, and as previously mentioned, is often formulated in response to core educational questions. This encourages governments or organisations to utilise this data when considering education reform, helping to make PISA results a more practically valuable resource for policymakers than data from other international tests such as TIMSS or PIRLS (Volante, 2016), notwithstanding that secondary data must be taken into account to ensure robust policy development.

Moreover, PISA data can establish a strong argument for certain educational reforms, by providing diagnostic information at the country level about the efficacy of their respective education systems. In many ways, PISA acts as a non-binding international agreement on educational goals that systems worldwide should strive to achieve, so when countries do not achieve expected rankings, this often provides the impetus for them to make direct reforms to address these issues. As with the earlier example of Germany, PISA data often enables countries to home in specifically on weaknesses in their education systems and then put in place policies to address them. The same can be said of the United States which, in the same vein as Germany, utilised poor PISA performance to address teaching quality through initiatives such as No Child Left Behind and the Every Student Succeeds Act (ESSA), and Brazil, which utilised PISA data to improve students' proficiency in mathematics following
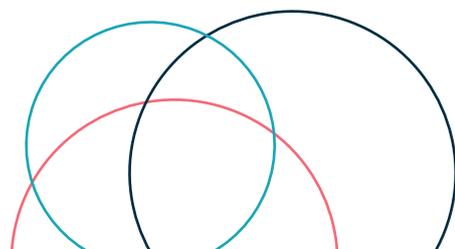
poor PISA 2003 results. What must be noted here, is that modifying systems in line with PISA scores which, whilst robust, has fundamental issues with its content and data reporting, can lead to inadequate or ineffectual reform. In some respects, the OECD has acknowledged this issue by stating that one of its key objectives moving forward is to, 'strengthen the policy relevance and analytical power of PISA, including establishing best practice for linking PISA with national assessments' (OECD, n.d. p8. cited by Volante and Fazio, 2018) This suggests a recognition that for PISA to have stronger 'policy relevance', the OECD must demonstrate greater coherence with individual nations' national assessments to provide the data robustness that may be lacking in PISA alone, and which may currently hinder countries from relying on PISA as an indicator of specific areas which require reform.

PISA and its drawbacks for policy reform

Since PISA's inception, a significant segment of the academic community has contested the role of the OECD and PISA in promoting large-scale educational reforms. Utilising the upsurge in media interest in PISA, this discontent was publicly shared through in a 2014 open letter to Andreas Schleicher, published in The Guardian newspaper, signed by over 80 academics calling on Schleicher to address some of their fundamental concerns with what they described as the 'testing juggernaut' of PISA.

Most importantly they expressed their concerns that PISA results have:

> Begun to deeply influence educational practices in many countries. As a result of Pisa, countries are overhauling their education systems in the hopes of improving their rankings.

> Lack of progress on Pisa has led to declarations of crisis and 'Pisa shock' in many countries, followed by calls for resignations, and far-reaching reforms according to Pisa precepts. (Andrews et al. in The Guardian, 2014)

Whilst the concerns that these academics express are genuine and highlight far-reaching issues brought about through PISA, I would argue that their response is too partisan. The letter does not acknowledge any of the benefits of PISA and is lacking in concrete data or examples to justify their perspective. Instead their point of view is presented as more of an emotive reaction to what they perceive to be the somewhat immoral nature of PISA and the influence of the OECD, a perspective that is refuted in the OECD's response, although again without much tangible data and ignoring some of their fundamental queries. Nevertheless, the negative aspects of PISA that are highlighted, and the potential issues this generates with regards to policy reform in education, are significant and warrant further analysis.

One issue the letter brings to light is the way that PISA rankings can cause nations to adopt short-term policy strategies designed to help them quickly climb the PISA rankings. This notion contradicts accepted research that substantial, meaningful educational change can take decades to come to fruition. That is not to say that short-term fixes cannot in themselves generate meaningful change, but rather that they may deter nations from making necessary, more extensive changes for fear that these will take too long to generate tangible results, particularly those which may improve their PISA rankings. In terms of policy reform, PISA may also detract from focus on harder to measure or intangible educational objectives, particularly ones that are not assessed by PISA due to their less evident coherence with economic competency. This includes aspects such as the
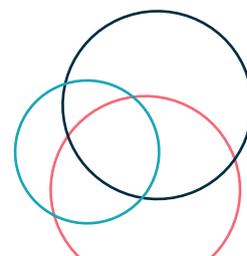
artistic, moral, physical and social development of students, tacitly suggesting that there is only one way of looking at education and what makes it effective. Ultimately, PISA may be narrowing the long-term view of the purpose of education, making it test- and data-centric rather than focusing on the holistic development and overall wellbeing of students.

William Smith notes that:

> Educational assessment through testing has become the most legitimate indicator of educational quality, performance and 'output' in schools around the world. (Smith, 2016)

This emphasises the notion that testing is a measure of potential 'output', with schools relying on this mechanism to prove their competence. However, it does not address the impact that this may be having on students' wellbeing. There is evidence to suggest continuous cycles of global testing adversely affects students and their learning in the classroom, causing more teaching towards the test, more scripted 'vendor' made lessons, less professional autonomy for teachers and increased stress in schools, all of which also significantly endangers the wellbeing of students and teachers (Andrews et al., 2014).

The role of the OECD in PISA also engenders a number of potential risks when utilising PISA data for policy reform. Linking back to the OECD's fundamental role as an economic organisation, PISA is naturally biased to the economic role of schools in the creation of human capital, versus how to prepare students for the social and personal aspects of adult life such as moral action, self-development and civic responsibility. This again leads to the possibility of governments, perhaps inadvertently, introducing policies which serve economic ends but don't take a holistic

view of required improvements to national education systems. Furthermore, by influencing the policies nations then prioritise for reform, PISA also takes the role of governing away from the nation state, by involving multiple actors and scales in the policy production and implementation process in respective countries (Rizvi and Lingard, 2010), eroding national autonomy. International organisations such as the OECD, through PISA, are then able to utilise their soft power to compel nation states to adopt particular policies (Bieber and Martens, 2011; Meye and Benavot, 2013; Mahon and McBride, 2008, cited in Volante, 2016), where countries feel obliged to comply in order to maintain their global standing as part of the PISA brand. Scholars have suggested that these policies are directly associated with the rise of neoliberal forms of governance and 'new managerialism' which involves these international organisations steering countries policies at a distance (Lingard et al., 2016), further eroding the autonomy of nation states.

Moreover, there are a range of cultural and linguistic issues with the test questions, namely the seeming impossibility of achieving linguistic and cultural equivalence across countries given the idiosyncratic nature of languages and the diverse array of participating countries. Representatives from countries participating in PISA 2000 were asked to comment on aspects such as item suitability, topic familiarity and cultural concerns, suggesting that the OECD did attempt to seek cultural equivalence at the time of PISA's launch. However, the organisations responsible for this were Belgium, Canada, Finland, France, Germany, Japan, the Netherlands, the United Kingdom and the United States, a selection heavily skewed towards developed Western states, increasing the possibility of cultural biases. Currently, PISA test items and questionnaires are developed under the responsibility of the PISA Governing Board, which includes

representatives from all member and partner nations (other than Brazil). However, complete cultural and linguistic parity is never going to be possible, which could disadvantage certain nations, particularly where the written questions are based on topics which have less cultural relevance. Research suggests that students respond better in exam situations when they have some context for what they are being asked about (Bracey, 2005). This lack of cultural context may make certain countries' results an inaccurate representation of their educational standing. This lack of cultural equivalence can also influence test response rates in different countries. For example, in France, there was a high rate of non-response in PISA 2000, which a French ministry official suggested was related to the fact that, culturally, French students often prefer not responding when they are unsure of an answer, rather than write something incorrect. Therefore, whilst PISA data can't account for the nuances of each participating nation or economy, not accounting for cultural differences may hinder the ability of the data to provide an accurate and representative picture of a country's education system, thereby making policy decisions based on this information alone, a risky approach for governments and policymakers to take.

This feeds into the fact that PISA does not, partially due to practicality, account for the curricula or learning style of each participating country, either through its formatting or test protocols. For example, some students taking the test may be unaccustomed to taking tests in either of the required constructed response or multiple choice formats, or in a set time-frame, creating a bias towards nations where students learn and are assessed using those frameworks. This relates to the way in which PISA doesn't link its questions into any individual country's curricula, but rather utilises questions devised by PISA researchers. The

OECD asserts that they test skills which are essential for students to succeed in modern society rather than assessing actual curriculum knowledge (OECD, 2014a). However, the ability of students to effectively translate questions on themes such as financial literacy into real-life skills, remains to be assessed. This reinforces the suggestion that PISA is creating its own curriculum and assessing a set of skills based on what the OECD believes students should know, and which are predominantly focused on projected future economic utility, rather than what they are actually taught in school. This reaffirms the risks associated with utilising PISA data to inform policy change, alongside the impossibility of comparing countries without potential margins of error being taken into consideration, which they aren't in PISA rankings. This raises further questions over ILSAs in general, which in many respects have become more of a public competition leading to a high likelihood of misguided policy borrowing in an attempt to avoid national and international shaming, particularly by sensationalist media, for poor performance.

Looking to the future

Looking to the future, the PISA test, whilst firmly embedded in the global education sphere, continues to be highly contested and controversial with regards to its influence over education policy. The OECD has argued, and many concur, that PISA has 'created valuable opportunities for transnational policy collaboration and should be credited with promoting high and more equitable learning outcomes across various student populations' (Schleicher, 2009). However, given its inherent flaws, it is not enough to rely on PISA alone. To avoid a myopic approach, policymaking must be a dynamic process where PISA data is just one aspect of the decision-making process. Whilst PISA can undoubtedly highlight key areas for reform, as has been seen predominantly in the case of Germany, what
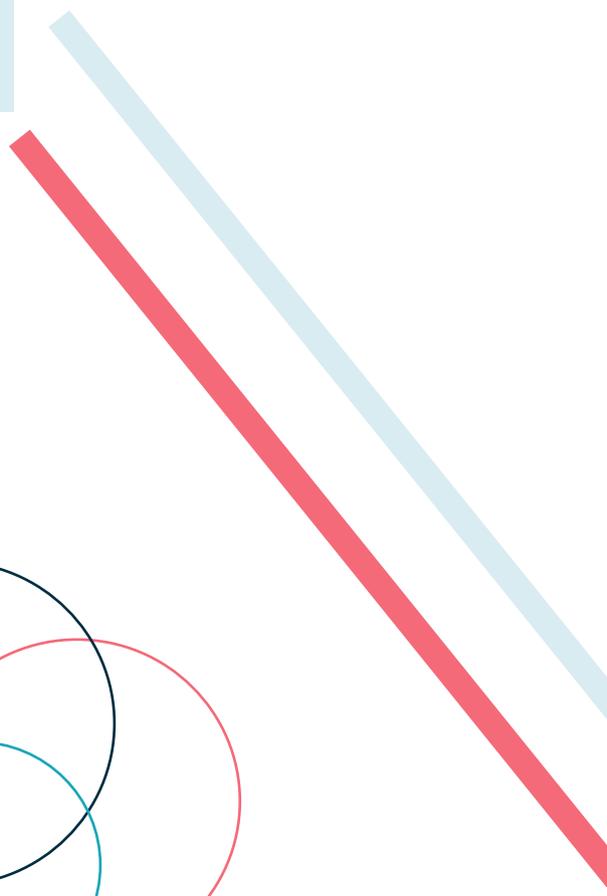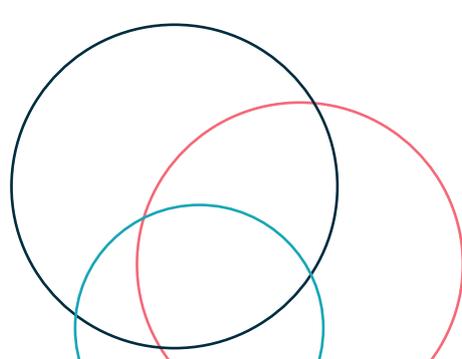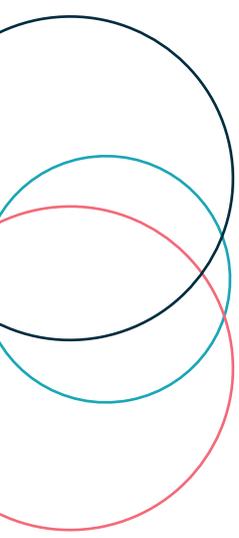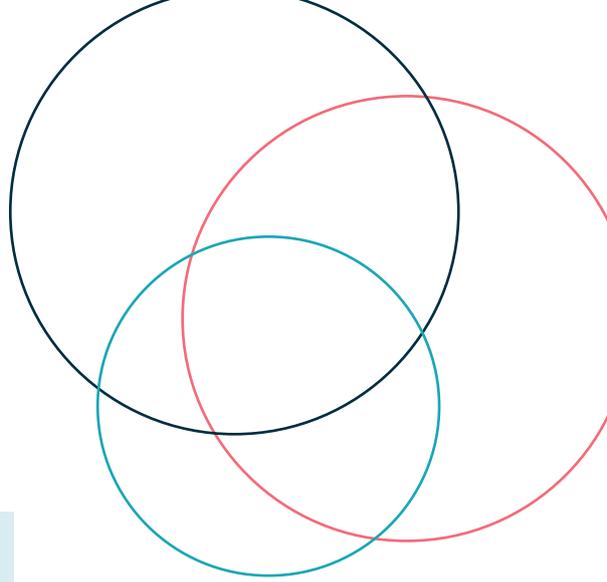
will continue to be the most significant factor for policymakers is not PISA data, but their nation's particular individualities. This includes factors such as geopolitical and financial status, alongside unique cultural, religious and historical contexts, as prime explanatory factors for their policy decision-making. Policymakers must also continue to examine the wider educational landscape, including the use of secondary data and their own national assessments, to enable the formulation and implementation of relevant and effective policy change. What remains to be seen is whether the OECD's continual lack of transparency, combined with PISA's evident methodological shortcomings, result in a future global shift away from PISA and the OECD as barometers for national educational performance. Or, will the OECD's ongoing quest for educational dominance - and the appetite for competition and global standing that has been generated since PISA 2000 - continue into the future, regardless of PISA's flaws and inadequacies as a tool for influencing policy reform.
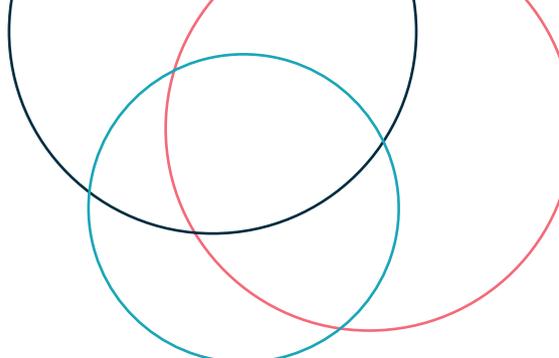
References

Andrews, P., et al. (2014). *OECD and PISA tests are damaging education world wide - academics.* The Guardian. Viewed 15 January 2021. <www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics.

Baird, J-A, Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, TL. And Daugherty, R. (2011). *Policy effects of PISA.* Pearson UK. <http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/10/Policy-Effects-of-PISA-OUCEA.pdf>

Bonnet, G. (2002). *Reflections in a Critical Eye: On the pitfalls of international assessment*, Assessment in Education: Principles, Policy & Practice, 9:3, p.387-399.

Bracey, G. (2005). *Put out over PISA.* Phi Delta Kappan Magazine. Viewed 16 February 2021. <www.journals.sagepub.com/doi/10.1177/003172170508601018>.

Breakspear, S. (2012). *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance.* OECD. Paris, France. Viewed on 7 February 2021. <http://www.oecd-ilibrary.org/education/the-policy-impact-of-pisa_5k9fdfqffr28>.

Dobbins, M. and Martens, K. (2012). *Towards an education approach à la finlandaise?* French education policy after PISA. Journal of Education Policy. 27 (1), p.23-43.

EUISS. (2012). *ESPAS report 'Global Trends 2030 Citizens in an interconnected and polycentric world.* EUISS. Viewed 28 December 2020. <www.iss.europa.eu/content/espas-report-%E2%80%98global-trends-2030-citizens-interconnected-and-polycentric-world%E2%80%99>.

Gorgen, K. and Hopfenbeck, T. M. (2017). *The Politics of Pisa: The media, public and policy responses in Norway and England.* European Journal of Research, Development and Policy. Volume 52, Issue 2, p.192-205.

Hanushek, E.A., and Woessmann, L. (2012). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation.* Journal of Economic Growth, 17 (4), p.267-321.

Lingard, B et al. (2016). *Globalising educational accountabilities*. New York, NY: Routledge.

Lipset, S. M. (1958). *Some Social Requisites of Democracy: Economic Development and Political Legitimacy.* The American Political Science Review, Vol. 53, No. 1 (March 1959). Viewed 2 February 2021. <www.jstor.org/stable/1951731>.

Morgan, C. and Volante, L. (2016). *A review of the OECD's international education surveys: Governance, human capital discourses and policy debates.* Policy Futures in Education, 14(6), p.775-792.

Murphy, D. (2014). *Issues with PISA's use of its data in the context of international education policy convergence.* Policy Futures in Education, 12(7), p.898.

OECD. (1960). *The Convention on the Organisation for Economic Co-operation and Development.* OECD. Viewed 21 February 2021. <www.oecd.org/general/conventionon theorganisationforeconomicco-operationanddevelopment>.

OECD. (2014a). *About PISA.* Paris, France. Viewed 11 February 2021. <www.oecd.org/ pisa/aboutpisa/PISA-trifold-brochure-2014>.

OECD. (2014b). *Response of OECD to points raised in Heinz-Dieter Meyer and Katie Zahedi 'Open Letter'.* Policy Futures in Education. 12(7), 878-879.

OECD. (2000). *PISA 2000 Technical Report.* OECD. Viewed 6 January. <www.oecd.org/ pisa/data/33688233.pdf>.

OECD. (2005). *Education at Glance* 2005: OECD Briefing Note for United States. OECD. Paris, France.

OECD. (2015). P*ISA 2015 Results in Focus:* OECD. Viewed 8 January. <www.oecd.org/ pisa/pisa-2015-results-in-focus.pdf>.

Piro, J. (2019). *The Primacy of Pisa.* Maryland: Lexington Books.

Pons, X. (2016). *Tracing the French Policy PISA debate: A policy configuration approach.* European Educational Research Journal. Vol. 15 (5), p.580-597.

Rizvi, F. and Lingard, B. (2010). *Globalizing education policy*. New York, NY: Routledge.

Schleicher, A. (2007). *Can competencies assessed by PISA be considered the fundamental school knowledge 15-year-olds should possess?* Journal of Educational Change 8: p.349–357.

Schleicher, A. (2009). *International benchmarking as a lever for policy reform*. In Hargreaves, A. and Fullan, M. (Eds.), Change wars. Bloomington: Solution Tree. p.97-115.

Schnepf, S.V. and Volante, L. (2018). *PISA and the Future of Global Educational Governance.* p.217-216 in (Ed.) Volante, L. (2018). The PISA Effect on Global Educational Governance. New York: Routledge.

Smith, W.C. (2016). *The global testing culture: Shaping education policy, perceptions and practice.* Oxford, UK: Emerald Group.

Steiner-Khamsi, G. and Waldow, F. (2018). *PISA for scandalisation, PISA for projection: the use of international large-scale assessments in education policy making – an introduction, Globalisation,* Societies and Education, 16:5, p.557-565.

Strauss, V. and Zhao, Y. (2019). *Expert: How PISA created an illusion of education quality and marketed it to the world.* Washing Post. Viewed 16 February 2021. <www.washingtonpost.com/education/2019/12/03/expert-how-pisa-created-an-illusion-education-quality-marketed-it-world.

Taut, S. and Palacios, D. (2016). *Intended and Unintended interpretations and uses of PISA results: A consequential validity perspective.* RELIEVE, 22(1), art. M8. Viewed 6 February 2021. <www.researchgate.net/publication/305076641>.

Volante, L. (2016). *The intersection of international achievement testing and educational policy: Global perspectives on large-scale reform.* New York, NY: Routledge.

Volante, L. (2018). *The PISA Effect on Global Educational Governance.* New York: Routledge.

Volante, L. and Ritzen, J. (2016). *The European Union, education governance and international education surveys*. Policy Futures in Education 2016, Vol. 14(7) p.988–1004. Viewed 11 January 2021. <www. journals.sagepub.com/doi/pdf/10.1177/1478210316652009>
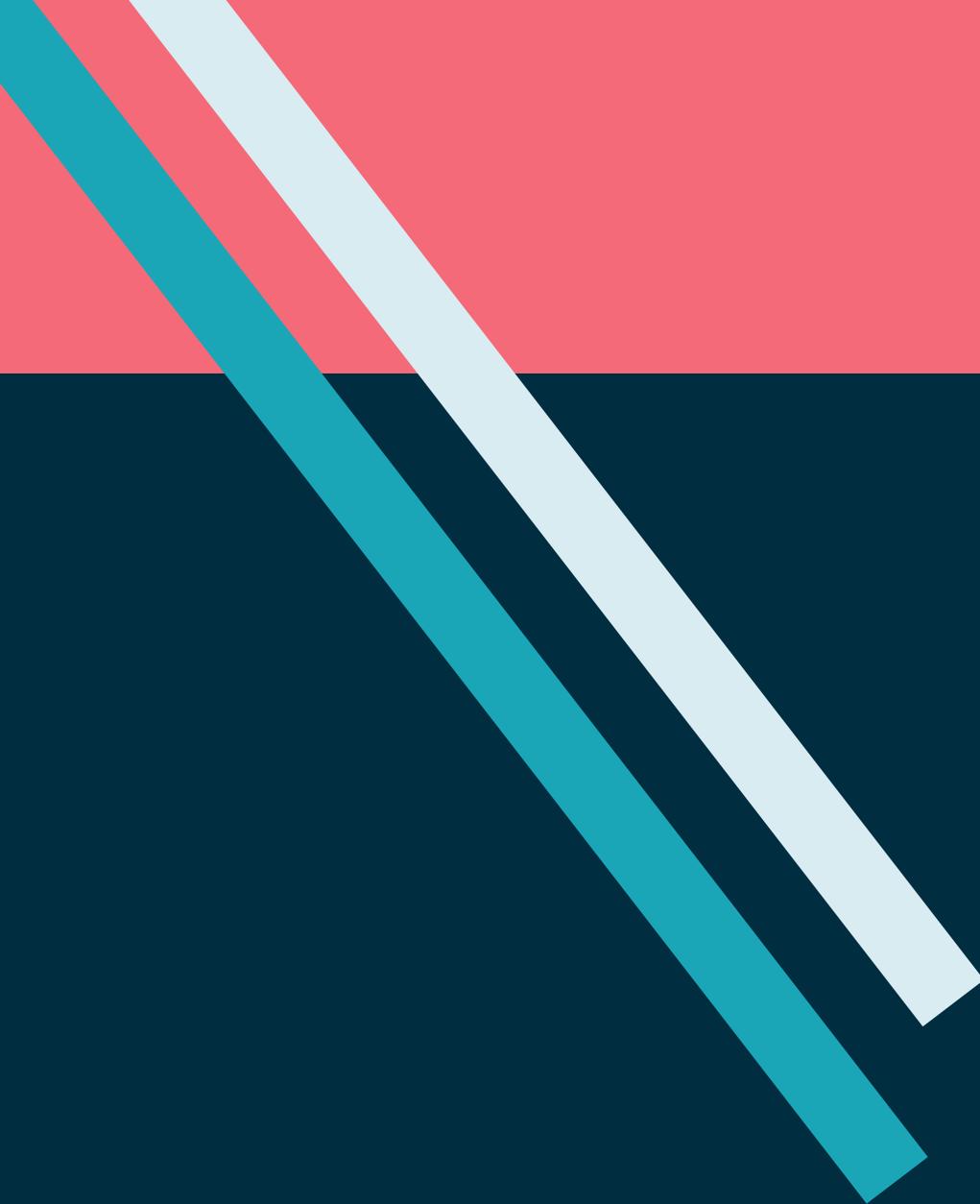
Book Reviews
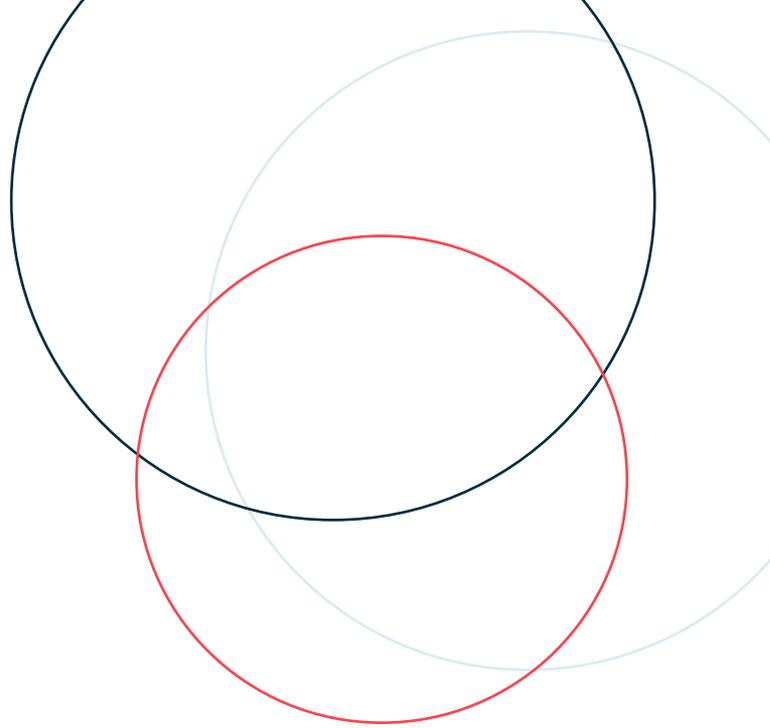
Reading, Writing and Racism: Disrupting Whiteness in Teacher Education and in the Classroom (2021) by Bree Picower

Sophie Bennett
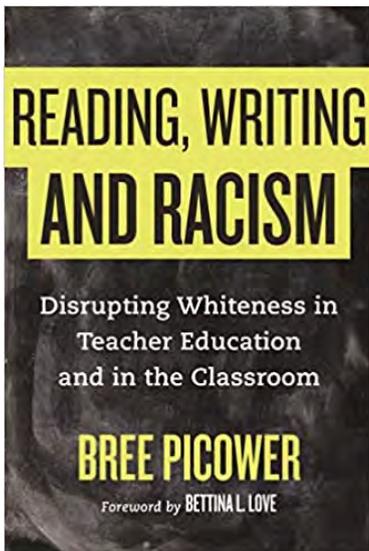
Throughout this book, Picower makes a compelling case for the urgent reflection and action of teachers to 'disrupt whiteness' in their classrooms and more broadly in the curriculum.

What is most impressive about the book is Picower's clear confrontation and then explanation of the issue of racism in education. The opening chapter does this really clearly by outlining historical and structural explanations of racism and Whiteness. This is underpinned by Picower's warning that currently "the curriculum functions to maintain dominant power structures" and emphasis of the power teachers have in exacerbating this by asserting that "education has the potential to function as a tool that reproduces inequality".

Of particular interest is Picower's sharp and useful illustration of 'Whiteness'; she defines the term as an ideology and 'way of being' to maintain the broader system of White Supremacy, usually through the conscious and unconscious actions of White people. This is quite carefully explained as being well hidden and doesn't seek to blame or shame White educators. Instead, Picower suggests that their beliefs on race are so deeply ingrained that they will likely be unaware of them or how complicit they are in maintaining them.

The early chapters of the book examine the structural issues of inequality through what feels like a sociological lens before moving to examine viral incidences of what trended online as #curriculumsowhite in recent years. Picower frames each as underpinned by specific 'curricular tools of whiteness' and says she chose to take this approach because these "singular examples reflect the entire body of the school curricula".

The book contains a number of these but below are two that stood out most to me:

- The No One Is To Blame Tool: this is when teachers frame atrocities committed by mostly White people to BIPOC (Black, Indigenous and people of colour) as stories with no victims or perpetrators. Picower uses the example of a textbook that referred

to enslaved African people as 'workers' or 'immigrants' by 'the Atlantic Slave Trade'.

- The Not That Bad Tool: this is when teachers downplay the violence of past oppressions by sanitizing the picture they present their students with. The example given is of a textbook published in 2018 that, in reference to slavery in the US, says "there were many kind and generous owners" and "many slaves may not have been terribly unhappy".

This anecdotal approach followed by analysis of the 'tools' behind each shocking example of racism allows one to reflect on their own language, resources and practice in the classroom.

What was also really useful were Picower's short 'reframes' that she uses to encourage White teachers to reframe their Whiteness. Those on 'internalised reframes' were especially insightful and were each followed with guidance in the book. They included:

- Recognising the systemic benefits of being White.
- Realising the biases that come with being White.
- Owning a White racial identity.

Additionally, throughout the book, Picower addresses a number of interesting and challenging tensions in the broader field of racial justice, especially those concerning her own Whiteness. She mentions the discomfort she has with the term 'allyship' for it enforcing a power dynamic in racial justice in which White people are working 'on behalf of' BIPOC, the slope towards 'savourisim' and the criticisms made of White anti-racist activists making money and careers from their work. Picower shares with the reader her 'principles' that she has formed to guide her action as a White person engaged in racial justice work. This aspect of the book is also hugely enlightening and useful, not just for teachers but for anyone interested in furthering their understanding of the role of whiteness in the classroom.
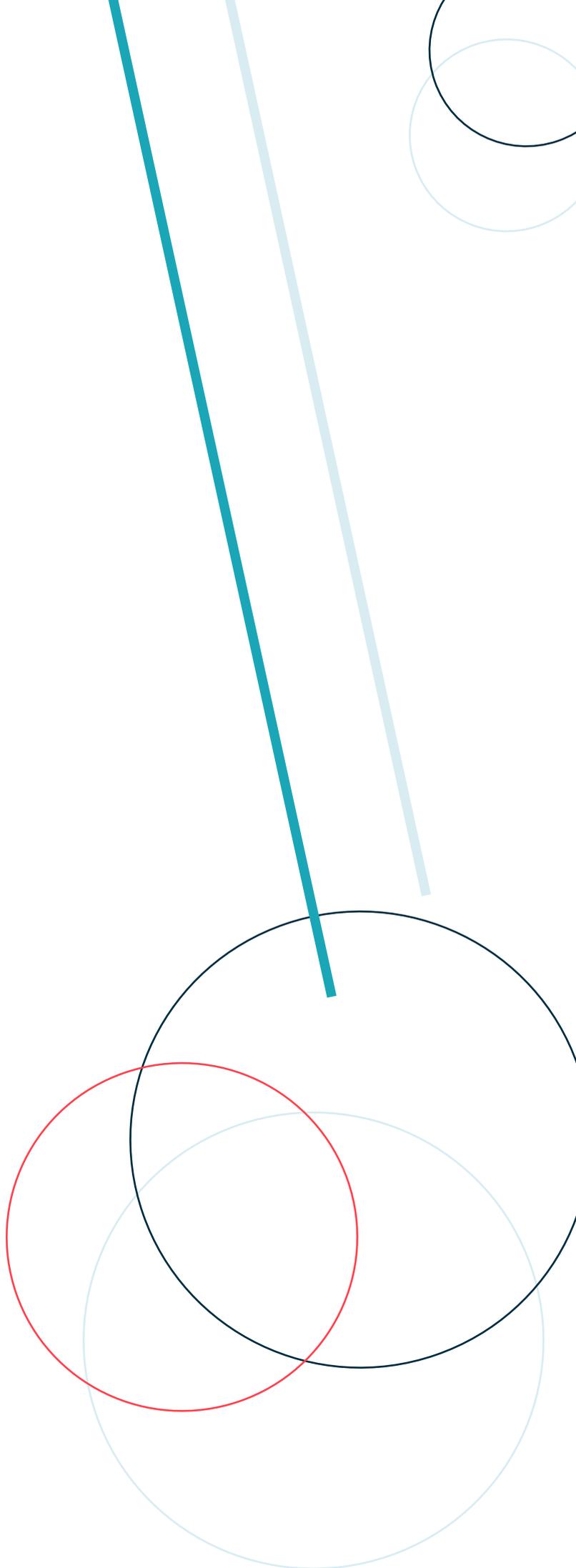
Retrieval Practice: Research and Resources for every classroom (2019) by Kate Jones

Charlotte Hawes
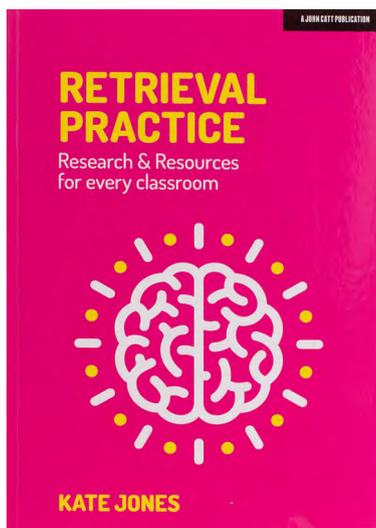
Retrieval Practice is a well written book, with a clear rationale for the use of research based practice in teaching, and plenty of classroom ready ideas. Kate Jones is a teacher, and she writes for teachers, with clarity and authority; she knows her stuff, and demonstrates the value of regular low stakes recall activities on improving pupil performance. She also presents a compelling case as to why teachers could and should engage more with education based research, whilst also recognising the time pressures we all face. As such, I found this book useful in two ways; both developing more RP strategies, and feeling motivated and able to engage with more education research.

Jones highlights that teachers need to be open to learning and keeping up with research, and at times it is challenging - I felt very out of touch when I read about the remarkable resources pupils are finding and generating online, including students, so called 'Study Tubers' who stream their study sessions to support others. But rather than guilt inducing it was inspiring; I have quoted on several occasions, when presenting my pupils with yet another shiny new Retrieval Practice technique, that 'Memory is the residue of thought!' (Willingham, pg 138).

The techniques she suggests are easy to use and embed - I am huge fan of her mantra 'Low Effort High Impact', and can see how this may resonate with many classroom teachers. One idea I particularly like is 'Cops and Robbers', where pupils are given limited time to recall all they can on a topic, then 'steal' ideas from other pupils. As an example, pupils may be given three minutes on 'Filter Theory evaluation points', where they write only in the left hand side of the page, and after time is up read what their neighbour has written and add that to their list on the right hand side, possibly in a different colour and so highlighting what they need to review. It's simple and quick, and because it's pupil led it allows plenty of time to walk around the room, checking to see who's struggling. Pupils enjoy the collaboration and are aware of topics of weakness without the pressure of a grade or score. In a recent pupil survey in Psychology most pupils were very positive about RP both in terms of the activities (they enjoyed them) and the outcomes (they led to better recall).

I can see that finding the time to read education literature may seem unrealistic to the busy teacher, but if you're going to pick one, make it this. When opening comments come from Tom Sherrington... 'a wonderfully written book about a vital subject', then expectations are high, and I was not disappointed. The author draws significantly on cognitive psychology research and makes it accessible. The core ideas can be summarised in a few short statements, but that's not to say I didn't find the whole thing a worthwhile read, in substantiating the concepts and providing so many ideas and examples. For me the key idea was to ensure that retrieval is happening for every pupil at the start of lessons; not just the one with the hand up. Like many colleagues I come up against this issue over and over, and this challenged me to embed the change, and force a habit change, and, in all honesty, whip out the mini whiteboards a lot more often.

The Boy Question (2021) by Mark Roberts

Laura McGill
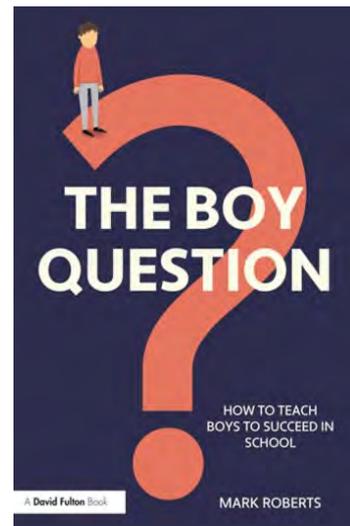
For many decades, schools have grappled with the problem that boys lack motivation to succeed, have poor attitudes to learning, lower literacy levels and a reluctance to read for pleasure or write at length. Mark Roberts' *The Boy Question* - a follow-up to his smash hit *Boys Don't Try* - offers workable solutions to one of the biggest problems teachers face – the fact that boys are not achieving what girls do the classroom, in both primary and secondary schools, and in the private and state sector. It answers nine key questions about how teachers and schools can best tackle boys' academic underperformance and offers interesting insights into why, as well as outlining some of the negative attitudes and preconceptions which may be holding boys back. By providing evidence and research based insights, as well as his own personal experience of topics that range from motivation to misbehaviour, to role models and writing, Roberts then goes on to answer the question, "What can we do about it?" The text, which is due to be published in 2022, even reflects on the devastating effect the current pandemic and lock downs have had on boys' academic performance. According to a report from the National Foundation for Educational Research, a survey of teachers 'found that 21 percent say that boys have fallen further behind normal expectations than girls,' while in stark contrast 'only 1 percent felt that girls had fallen behind.' Older boys in particular seemed to have slipped back during the pandemic, as 'almost two fifths of secondary teachers reported that boys were more behind their usual learning levels than girls.'

I think *The Boy Question* is incredibly useful for anyone who works in a school, particularly classroom teachers and school leaders, and also parents to boys. Each chapter is research based and full of practical and applicable strategies for the classroom. I found the two classroom analogies very helpful, and the case studies were compelling, particularly around how to encourage creative writing and getting boys reading more. Although I found many of the points raised in *The Boy Question* were both interesting and pertinent, I felt it was not entirely relevant to our teaching practice at UCS as we do not have a significant or obvious attainment gap between boys and girls. Also, UCS just has girls in the Transitus and Sixth Form and only sets students for maths, while most of the research and case studies in *The Boy Question* focuses on large, mixed comprehensive schools, with high numbers of disadvantaged students from white working-class backgrounds in boy-heavy bottom sets. One key point that I have taken away from reading this text is that it reinforced my belief that no hands up or 'cold-call' questioning is one of the most valuable strategies in our teaching toolkits to ensure universal participation for all of our students – boys and girls, alike.

Boys Don't Try? Rethinking masculinity in schools (2019) by Matt Pinkett and Mark Roberts

Jay Thomson
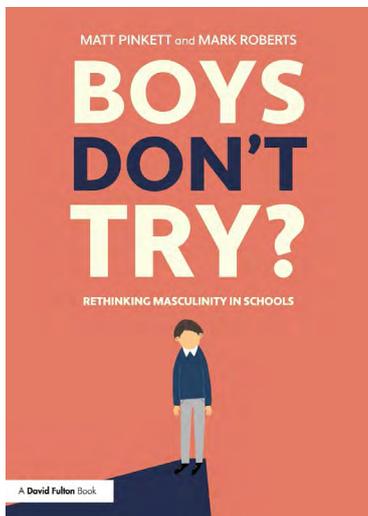
In his 1979 song *Boys Keep Swinging*, David Bowie sang that 'Nothing stands in your way when you're a boy'. Well, in 2022 it's perhaps discussions of 'toxic masculinity' in schools and in society at large, or 'white working class' boys caught in a seemingly eternal educational trough, or male youths in the deathly throes of shank-happy nocturnal activity, or the fact that the boys who become men are growing up psychologically brittle and more likely to be victims of their own suicide. Quite a lot seems to be standing in the way of boys now.

This book is both timely and bleak. The authors, both boys at one time, are seasoned educators and present a powerful picture of the ongoing crises with boys and education. But this book isn't just an extended cavil against boys finding their way in education: it also offers solutions. For instance, chapter 6 advocates more education on the thorny topic of pornography and its malign influence on inexperienced young men. Would you learn to drive by watching F1? It's an unreal world that relentlessly shows unreal and unachievable expectations. In chapter 4 Matt Pinkett reminds us of the 'disturbing parallels' between now and the 16th century (p.62). Even then there was internalised and externalised violence by young men, drug taking, depression (then politely called 'melancholy'), erratic and murderous behaviour, and finally, teenage suicide. He is, of course, talking about the well-known figure of Romeo. Does this example demonstrate that this is a fact of being a young man, a kind of unalterable transhistorical truism?

It is a complicated picture and the book uses masses of evidence and studies from various Anglophone countries. The authors discuss the 'dominant strand of masculinity that sees school work and high achievement as effeminate and uncool' (p.49). Might this link to a form of misogyny where anything vaguely female is seen as inferior? Is it due to the fact that the media is simply saturated by sports? At times the authors display a kind of binary thinking: 'Unfortunately, boys who show proficiency on the football pitch, and dream of the riches of the Premier League, are even less likely to value academic qualifications' (p.49). The authors could possibly have explored why these two things are perceived as mutually exclusive. Is the problem one

of low class expectations and capitalism, a place where goals are only seen in relation to stitched polyurethane? The vast majority of their examples come from the state sector; is the picture different in the independent sector and why would it be different? And why are the working class seen as one homogenous mass anyway?

This book is an effective blend of theory and practical solutions. We are told in the paragraph 'Winning ways to engage boys' that boys enjoy lessons when they have the opportunity to 'get up out of their seats and move around', partake in role play and debates, and 'surprising events'. Research shows that boys responded very well to dissecting squid during Biology, then drawing things with the squid's ink and finally 'turning these cephalopods into calamari' (p.9). Is the answer then a pedagogical fusion? Science and art mixed with practical, nutritious results? Is there something in the male mind that needs to see tangible results?

This is a meticulously well-researched and exhaustive book. It is serious and slightly funny in turns which makes it more palatable. However, at times the flurry of questions choke the answers. But the questions posed here are urgent and hugely relevant and it may well be in an enquiry beset with so many variables, there will always be more questions than answers.
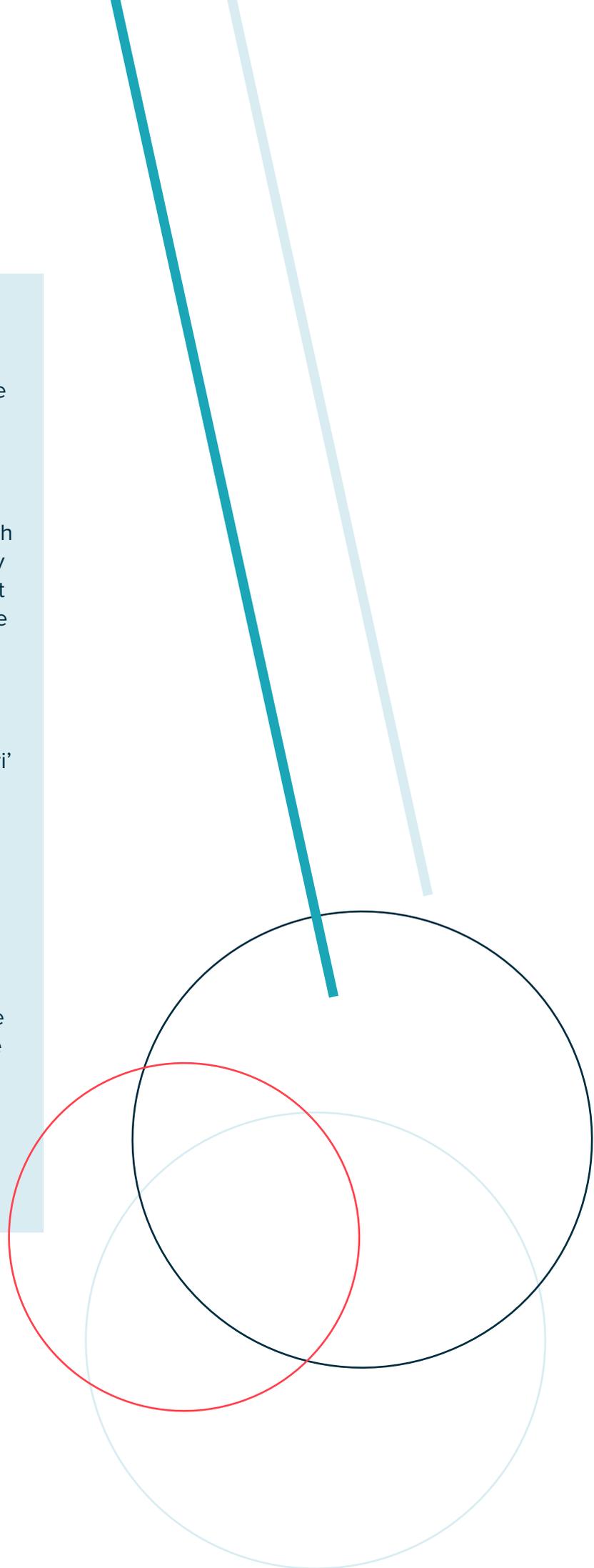
Motivated Teaching (2020) by Peps Mccrea

Kimberley Ward
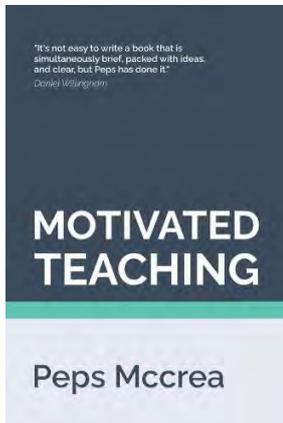
Peps Mccrea does a fantastic job of condensing a plethora of research into an ultra-concise, easy to read manual on motivation in teaching. Every teacher should read this. He scaffolds his ideas around five core drivers of motivation for learning: securing success, running routines, nudging norms, building belonging and boosting buy in. Peps describes and explains each of these drivers and then gives some practical tips for applying them in the classroom. There is so much to this very short book – in terms of insights per number of words, you are getting a pretty good deal. Accessible to all and very easy to dip in and out of, with a 'notes and further reading' list at the end of each chapter giving the option to delve further into the research surrounding his insights on motivation.

The word 'motivation' is thrown around all too easily in teaching, and is often used to describe a trait pertaining to a particular student, i.e. 'student x is unmotivated' or 'student y is a highly motivated student'. Peps challenges this idea and invites the reader to think of motivation as a context specific drive to complete an individual task, and most importantly, something that can be influenced by economic, social and metacognitive factors.

I feel a sense of irony as I sit writing this book review, having let the deadline slip, struggling to find the motivation to get my thoughts down. I find myself thinking about the five core drivers of motivation that Peps Mccrae describes so eloquently in this book. How might they have helped me become more motivated towards my goal? I have no previous experience of writing book reviews for teachers, so have never given myself the opportunity to 'secure success' but can't help but wonder if I might have been more motivated had I previously experienced success in writing for such an audience. Peps suggests that the biggest factor that influences student motivation in a specific context is their success in that domain. If a student expects they are likely to be successful – because of previous success in that same domain – then they are likely to put in more effort because they can expect a pay-off.

My allocation of attention and therefore cost of writing this review was quite high because – well – the last time I managed to read a whole book in a few weeks was – let's just say before I had kids. Reading all 128 pages of this very short, very concise book was a feat in itself for a working mother of two. Peps would have suggested I got myself into a routine of reading ten pages a night and making notes as I went but of course I left it to the last minute and read it all in one go so I failed on the 'running routines' driver. This is all to do with cost/benefit analysis: securing success increases the benefit to the student and running routines reduces the cost because if we put good routines in place, we generally make it easier for students to participate.

My husband is an avid reader and seeing him read often encourages me to dip into one of the many books on my bedside table but unfortunately he was away on business on the lead up to this deadline – so alone on the reading front, I also failed to 'nudge norms'. This is the idea that as humans, we are constantly looking to those around us for signals regarding what we should be doing – something that can and should be used to our advantage in the classroom! We tend to feel safe going with the majority and so if a teacher publicises when the majority is doing something good, this will encourage the de-motivated to follow their peers' example.

As for 'building belonging' and 'boosting buy in', I will let you read the book yourselves and I promise you that the benefit of reading this book far outweighs the cost of time it takes to read it!