

ISSN: 1462 - 7272

# Education Journal Review

---

Volume 30 • Number 1  
March 2024

*Education Journal Review* is published three times a year. It is owned by Education Publishing Worldwide Ltd. and published under license by the Education Publishing Company Ltd. in the United Kingdom.

Correspondence about editorial matters should be sent to the editor, Demitri Coryton, at [info@educationpublishing.com](mailto:info@educationpublishing.com)

Correspondence about all other matters should be sent to the publisher at [info@educationpublishing.com](mailto:info@educationpublishing.com)

Subscription rates for *Education Journal Review* are given below. Subscriptions should be paid to the Education Publishing Company Ltd and paid by BACS. Account details are available upon request.

In common with all our titles, the print edition is now no longer available. *Education Journal Review* is now available only in electronic form. It can be sent as a PDF file by email and accessed via our website. Subscribers will be sent access details upon request. These can be user name and password or IP access, depending on subscriber preference. There is no price difference between the two.

Back issues from volume 15, when the Education Publishing Company Ltd became publisher of what was then known as *Education Review*, and before that *NUT Education Review*, are available to subscribers on our website at [www.educationpublishing.com](http://www.educationpublishing.com). During the course of volume 25, the name was changed slightly to *Education Journal Review*.

### **Subscription rate for 2024 and 2025**

UK and rest of the world, one year £160

### **Single issue rate for 2024 and 2025**

UK and rest of the world, one year £60

Back issues or volumes are available as PDF files at the same rate as current issues.

---

# Education Journal Review

---

Volume 30 Number 1  
March 2024

© *Education Journal Review* and Education Publishing  
Worldwide Ltd. 2024.

The Education Publishing Company Ltd.  
15A East Street, Oakhampton, Devon, EX20 1AS.  
Email: [info@educationpublishing.com](mailto:info@educationpublishing.com)

Education Publishing Worldwide Ltd.  
15A East Street, Oakhampton, Devon, EX20 1AS.  
Email: [info@educationpublishing.com](mailto:info@educationpublishing.com)

# Contents

- 7 Preface**
- 8 Attractiveness of UK taught Masters to international students**  
*He Aian*
- 15 Challenging systemic prejudices. An investigation into bias against women and girls in large language models**  
*OECD/Daniel van Niekerk et al*
- 43 Teacher labour market in England 2024**  
*NFER/Dawson McLean, Jack Worth and Andrew Smith*
- 83 PISA 2022. Measuring the world's education systems after COVID**  
*Demitri Coryton*
- 108 PISA 2022. A personal commentary**  
*John Bangs*
- 114 Select Committee Reports**
- 115 Childcare and the early years**  
*The House of Commons Education Committee.*
- 159 The condition of school buildings**  
*The House of Commons Public Accounts Committee*
- 125 Persistent absence and support**  
*The House of Commons Education Committee.*

*Education Journal Review*  
Volume 30 Number 1  
March 2024

ISSN: 1462 - 7272

# Preface

**I**n our first contribution from a Chinese academic, who studied at the University of Edinburgh, He Aian looks at the attractiveness of UK taught Masters to international students.

In a project commissioned by UNESCO, Daniel van Niekerk and an international group of scholars investigate the bias against women and girls in large language models.

NFER produced their *Teacher Labour Market in England 2024* report which we publish in this issue. Dawson McLean, Jack Worth and Andrew Smith outline the conditions that have created the twin problems of recruitment and retention which the British Government must contend with.

The OECD's PISA research is now the leading measure of secondary education in the world. At the end of last year the OECD published two volumes of the 2022 round together with other supporting documents, while England and Scotland published separate reports on the outcome of PISA in their countries. The scale of PISA is vast. The research is based on tests taken by 690,000 15-year-olds in 81 countries and territories in all inhabited continents. In this issue we look at what PISA 2022, which was undertaken just after the COVID pandemic, said and what policy options it presented to countries.

John Bangs has been involved with the OECD for many years, through the teacher union global confederation Education International. He is a former Assistant Secretary of the NUT where he was Head of the Education Department. He gives us a personal view of PISA 2022.

This issue ends with three select committee reports which takes us up to the end of 2023 in our coverage of every select committee report on education published since 2018.

**Demitri Coryton**  
Editor

He Aian

# Attractiveness of UK taught Masters to international students

He Aian works in Shanghai Youth College of Management, Shanghai, China. He has a Master of Science in digital technology from the University of Edinburgh, UK.

**By He Aian**

Shanghai Youth College of Management, China

**Key words:** Masters, international students, higher education

**Abstract:** *The UK has always been a popular destination for study due to its high level of teaching quality and academic reputation, and its one-year master's programmes also significantly reduce the time taken to study for a postgraduate degree compared to the one-and-a-half to two years in the US and two to three years in China. Through my literature review, I found that there are many studies focusing on the influences of studying abroad, but they are all focused on a specific region and even fewer studies on one-year masters in the UK. This study took a qualitative research approach, using interviews to explore the factors that influence students who study abroad in their choice of a one-year UK Masters's programme. I found that The most important factors influencing their choice of a one-year Taught UK Masters's programmes were the professional recognition and the academic reputation of the university, along with the better security of the country.*

**L**iterature review. The factors that determine international students' decision to study abroad are a complex process and have been studied by scholars in terms of different factors, different countries and regions, and different professional disciplines.

Firstly, studying abroad means moving from one's country of origin to another new country, then both countries



appear to be crucial. Many scholars argue that the decision to study abroad is influenced by “push factors” in the country of origin and “pull factors” in the country of destination [1,2], and that it is mostly the negative effects that cause students to. The majority of these factors are negative, including economic recession, political instability and low quality of education, and fierce competition for higher education at home. There are of course a number of factors that influence students' decisions on international education destinations, such as location (place) and the cost of study (price) [3]

However, when discussing the most important factors, sociologists conducting qualitative research on different country regions and different majors have come up with different results. Diana and Ana's study [1], conducted at a public university in northern Portugal, concluded that the academic reputation and quality of the country of study were the most important attraction factors; Lu's study [2] of 736 international students found that it was the realization of self-worth or achieving one's life goals that played the most important role; Mathew [4] for international students in India, the cost of living and education, institutional scholarships and financial support, and the reputation and rigour of the programme were the most influential factors in determining admission to the institution. Also, professional differentiation and access to international experience play an important role, for example, the main barriers for healthcare students when considering whether to study abroad are cost and language issues [5].

In addition to this, a number of studies have shown the influence of parents and families on their children's study abroad. Parents' education level influences students to continue studying in their home country [6], and parental educational expectations and professional reputation significantly increase parents' willingness to send their children abroad [7].

According to the literature review, the UK has become a popular study destination for many international students, and the factors that play a decisive role in the decision to

study vary across student groups from different countries, regions, nationalities, and cultural backgrounds. These include push factors in the country of origin (political and economic issues), pull factors in the country of destination (academic reputation, institutional specialism, living environment, etc.), personal factors (study background, career prospects, future development), and relationship factors (family, friends, teachers). However, there appears to be little mention of the impact of the academic program, with little focus on those who choose to study for a one-year master's degree in the UK. This study, therefore, focuses on those who study abroad to study a one-year master's programme in the UK and explores their reasons for choosing this programme.

### **Study method**

According to Rubin[8], when the research has important context and richness, naturalistic research tools are more appropriate, and in-depth qualitative interviewing is one of the key naturalistic research methods. For this study, I chose the research method of interviews in order to explore the process of the choice to study abroad and how events unfold, and due to the complex reasons behind this vary from person to person.

Snowballing non-probability sampling method was been used. Firstly, I used a Questionnaire to screen out suitable respondents. The aim was to find suitable interviewees. The questionnaire contains the nationality, current country of study, school, undergraduate or master's degree, major and contact details of the respondent. Respondents who completed the questionnaire and were interested in participating in this interview were given my contact details in the questionnaire and a deadline for contact was stated in the information sheet. Based on the results of the questionnaire collection, I selected and contacted the 5 respondents who were suitable for this study. I sent them emails containing the informed CONSENT FORM and obtained their signed feedback and conducted online interviews of 45 minutes at a time suitable for them.

## **Results**

All of the interviewees had similar undergraduate backgrounds (all attended universities in China, two were Sino-foreign education joint universities, two were public undergraduates in China, and one was a university in Macao, China), but all of them had experienced and familiarized themselves with the foreign education model during their undergraduate or high school years, an experience and a background that provided the foundation for their choice to study abroad.

### **The most important influencing factors**

The choice of destination includes many considerations when thinking about studying abroad, and all respondents cited advantages related to “major (program)” as the most important factor, such as high professional recognition, professional training advantages and strong professional relevance. They also mentioned safety and cost. 2 of them mentioned that the UK is safer than the US and that a one-year Master's degree is less costly and more cost-effective, both in terms of time and money.

### **Push and Pull factors**

All interviewees are enrolled in business school majors. Respondents at the University of Edinburgh said that the most important factor for her was the university's high world ranking. Respondents studying at the University of Southampton did not apply to the university he was most satisfied with, but were very satisfied with the Masters programme he was studying and also ranked the UK's academic reputation as one of the most important factors.

Choosing to study abroad is the first step, and two of the respondents in this study had already made this decision during their previous study experiences, and it seemed to be a "natural" path to take. This is because they chose to study at a university that cooperated with China and the West during their undergraduate studies, but the remaining two interviewees had considered pursuing both postgraduate

studies in China, but mentioned that the time spent on postgraduate studies in China was too long and the competition was too fierce, and that one had considered taking up employment first.

The second step is to choose the country in which to study, and in this case the UK one-year Master's degree plays to its great overall strengths. As a destination country, the UK's greater safety, good academic reputation, high school ranking and short duration of study (and the resulting low cost advantage) were all strong pull factors.

Less influential factors

One of the interviewees mentioned the recommendation of a teacher, but all interviewees said that they had made the choice to study on their own and were not overly influenced by others (friends or family) and that their families had shown support and financial sponsorship for studying abroad. From this we also seem to infer that postgraduate study abroad is no longer about parents sending their children abroad, but perhaps more about parents supporting their children to go abroad.

### **Factors with little or no influence**

In the future, they all seem made it clear that they want to return to work in their home countries, but will try to find suitable job in the UK at first that they have come to study abroad to increase their experience and that they all have plans to travel and see the world during the year. Therefore, changing their future place of residence and considering working outside the country did not seem to be an important consideration for them.

### **Conclusion**

The primary data collected in this paper through interviews is to some extent informative about the education of master's degrees in the UK. The study found that there are three main strong correlates of international student's choice of a one-year Master's degree in the UK: the first is "Major (programme)" and the academic reputation (ranking) of the

school, which is more important to be considered in further research and will vary from person to person. The second is the security environment of the study destination, where safety is also a key consideration. The third is that one year is more cost-effective and less costly in terms of money and time.

Weakly correlated factors include one's own prior personal experience (e.g. familiarity with Western education patterns as an undergraduate) and family support (financial support), which, although not talked about much by the respondents, seem to lay the groundwork for studying abroad.

Career prospects and immigration issues were not important to international students, but of course this could be due to the small sample size, and we might have had different results if we had interviewed more students.

## **References**

- [1] Branco Oliveira, Diana, and Ana Maria Soares. "Studying Abroad: Developing a Model for the Decision Process of International Students." *Journal of Higher Education Policy and Management* 38, no. 2 (2016): 126–139.
- [2] Yue, Yun, and Jinjin Lu. "International Students' Motivation to Study Abroad: An Empirical Study Based on Expectancy-Value Theory and Self-Determination Theory." *Frontiers in Psychology* 13 (2022): 841122–841122.
- [3] Liu, Ying, Peerasit Kamnuansilpa, and Ando Hirofumi. "Factors Affecting International Students' Decisions on Destination for Studying Abroad: A Case Study in China." *Frontiers of Education in China* 13, no. 1 (2018): 93–118.
- [4] Mathew, Zachariah. "Studying Abroad: Understanding the Factors Influencing the Study Abroad Decision, Country, and University Choice of International Students from India". ProQuest Dissertations Publishing, 2016.

[5] Brown, Mary, Edward Appiah Boateng, and Catrin Evans. "Should I Stay or Should I Go? A Systematic Review of Factors That Influence Healthcare Students' Decisions Around Study Abroad Programmes." *Nurse Education Today* 39 (2016): 63–71.

[6] Dima, Adriana. "Empirical Research on Factors Influencing Students' Decision to Study Abroad Using Cart Analysis." *Revista de Management Comparat International* 20, no. 4 (2019): 449–458.

[7] Zhu, Xingchen, and Shuping Yang. "Factors Influencing Chinese Parents' Willingness to Send Children to Study Abroad: Findings from the China Family Panel Studies." *Asia Pacific Journal of Education* ahead-of-print, no. ahead-of-print (2022): 1–14.

[8] Herbert J. Rubin, and Irene S. Rubin. 2005. *Qualitative Interviewing : The Art of Hearing Data*. Edited by Irene. Rubin. 2nd ed. Thousand Oaks, Calif.: Sage Publications.

# Challenging systematic prejudices. An investigation into bias against women and girls in large language models

UNESCO/  
Daniel van  
Niekerk et  
al

By Daniel van Niekerk [1], María Pérez-Ortiz [1], John Shawe-Taylor [1,2], Davor Orlič [1,2], Ivana Drobnjak [1], Jackie Kay [1], Noah Siegel [1], Katherine Evans [2], Nyalleng Moorosi [3], Tina Eliassi-Rad [4], Leonie Maria Tanczer [2], Wayne Holmes [1,2], Marc Peter Deisenroth [1], Isabel Straw [1], Maria Fasli [5], Rachel Adams [6], Nuria Oliver [7], Dunja Mladenčić [2], and Urvashi Aneja [8].

[1] University College London.

[2] International Research Centre on Artificial Intelligence (IRCAI), under the auspices of UNESCO.

[3] Distributed AI Research Institute

[4] Northeastern University

[5] University of Essex

[6] Research ICT Africa

[7] ELLIS Alicante Foundation

[8] Digital Futures Lab

This research was commissioned by UNESCO.

**Key words:** Women, girls, AI, bias, ethnicities, algorithms.

## UNESCO

commissioned this research from a group of academics from University College London, UK; the International Research Centre on Artificial Intelligence, Slovenia; the Distributed AI Research Institute, USA; Northeastern University, USA, the University of Exeter, UK; Research ICT Africa, South Africa; the ELLIS Alicante Foundation, Spain; and Digital Futures Lab, India.

**Abstract:** *Artificial intelligence is being adopted across industries at an unprecedented pace. Alongside its posited benefits, AI also presents serious risks to society, making the implementation of normative frameworks to reduce these risks a global imperative. The UNESCO Recommendation on the Ethics of AI asserts that “AI actors should make all reasonable efforts to minimize and avoid reinforcing or perpetuating discriminatory or biased applications and outcomes throughout the life cycle of the AI system to ensure fairness of such systems”. To date however, AI-based systems often perpetuate (and even scale and amplify) human, structural and social biases. These biases not only prove difficult to mitigate, but may also lead to harm at the individual, collective, or societal level.*

*This study explores biases in three significant large language models (LLMs): OpenAI’s GPT-2 and ChatGPT, along with Meta’s Llama 2, highlighting their role in both advanced decision-making systems and as user-facing conversational agents. Across multiple studies, the brief reveals how biases emerge in the text generated by LLMs, through gendered word associations, positive or negative regard for gendered subjects, or diversity in text generated by gender and culture.*

*The research uncovers persistent social biases within these state-of-the-art language models, despite ongoing efforts to mitigate such issues. The findings underscore the critical need for continuous research and policy intervention to address the biases that exacerbate as these technologies are integrated across diverse societal and cultural landscapes. The emphasis on GPT-2 and Llama 2 being open-source foundational models is particularly noteworthy, as their widespread adoption underlines the urgent need for scalable, objective methods to assess and correct biases, ensuring fairness in AI systems globally.*

**T**he pervasive problem of bias against women and girls worldwide is a deeply entrenched issue that manifests across various societal, economic, and political domains, reflecting centuries of gender inequalities and



**systemic discrimination.**

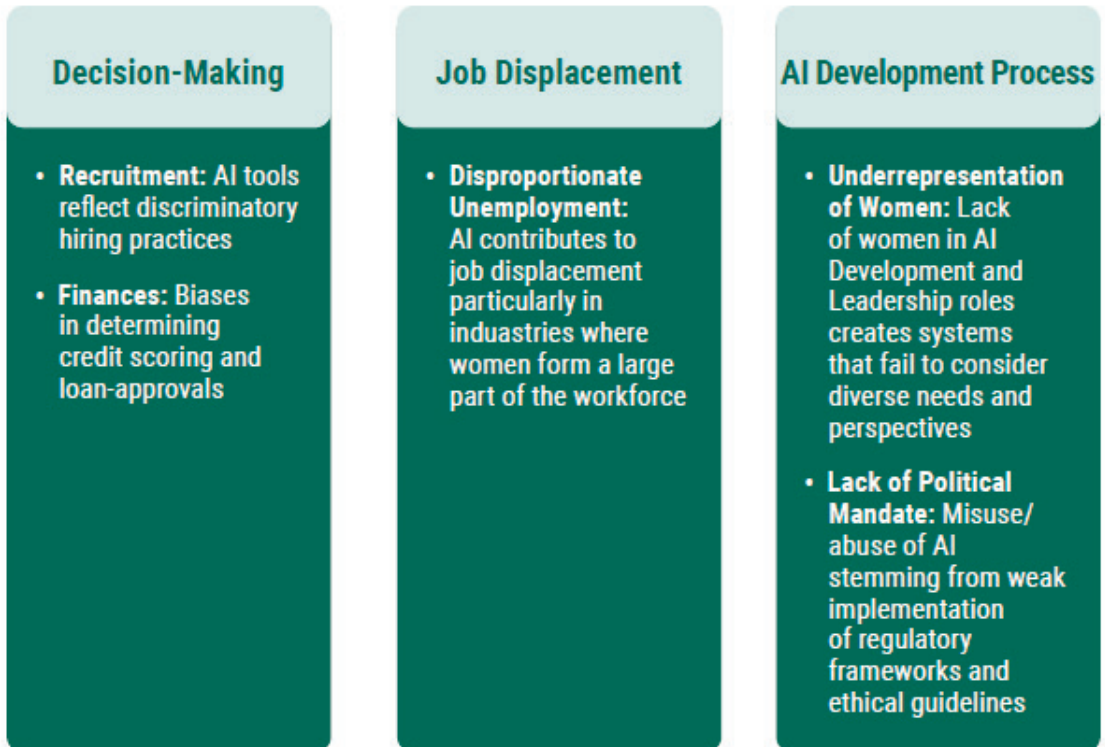
Many challenges in gender equality and equity persist today, including gender-based violence, pay disparities, and underrepresentation of women in leadership roles, amongst others. Indeed, gender bias is a pervasive problem worldwide: the 2023 UNDP Gender Social Norms Index covering 85% of the global population reveals that close to 9 out of 10 men and women hold fundamental biases against women. [1]

This widespread bias not only undermines the rights and opportunities of women and girls, but also seeps into the technological advancements and innovations of the modern world, notably into Artificial Intelligence (AI) systems, especially Large Language Models (LLMs). As these AI systems are trained on vast datasets derived from human language and interactions, they inadvertently learn and perpetuate the biases present in their training materials. Consequently, LLMs can reinforce stereotypes and biases against women and girls, practices through biased AI recruitment tools, gender-biased decision-making in sectors like finance (where AI might influence credit scoring and loan approvals), or even medical or psychiatric misdiagnosis due to demographically biased models or norms [2]. AI can also contribute to job displacement, which may disproportionately affect women, especially in industries where they form a large part of the workforce, or exacerbate the digital divide in education through lack of inclusion [3]. The underrepresentation of women in AI development and leadership roles can further lead to the creation of socio-technical systems which fail to consider the diverse needs and perspectives of all genders, once again perpetuating stereotypes and gender disparities.

Nevertheless, AI could potentially advance the aims of gender equality and equity worldwide if, for instance, it is harnessed ethically and inclusively, or if it is developed by diverse teams which aim for positive societal impacts, and more generally, if it is designed to mitigate — rather than perpetuate — inequality and gender disparity in its interactions with society.

*Figure 1: Perpetuation of inequality*

## Ways AI Perpetuates Biases



### Inside the algorithm: Exploring Algorithmic Bias

Algorithmic bias happens when an algorithm, or a set of computer instructions, unfairly discriminates against certain people or groups.

#### 1. Biases in Data:

- **Measurement Bias:** Occurs during the selection or collection of features. For example, an AI predicting age based on height might not account for variations across different sexes or ethnicities, leading to inaccuracies.
- **Representation Bias:** When training datasets do not

adequately represent all groups, leading to poor generalization. Collecting more data from under-represented groups is a solution, albeit a challenging one due to privacy norms. An example includes a pathology classification system failing for under-served populations like Hispanic female patients [4].

## 2. Biases in Algorithm Selection:

- **Aggregation Bias:** Using a “one-size-fits-all” model that fails to account for the diversity within the data. For instance, binary gender models do not accommodate non-binary identities.
- **Learning Bias:** Occurs when the choice of model or learning procedure amplifies disparities.

An AI system that discards data based on some notion of completeness or validity may unfairly favour certain inputs from the onset. For example, male resumes over female resumes when hiring.

## 3. Biases in Deployment:

- **Deployment Bias:** Happens when AI systems are applied in contexts different from their development context, leading to inappropriate outcomes. Language models trained on internet text might make improper associations between psychiatric terms and specific ethnic or gender groups [5]
- **Post-Deployment Feedback Bias:** Adjusting models based on user feedback without considering the demographic diversity of users can introduce new biases. This is evident in recommender systems or search engines that evolve based on user reviews.

## Bias and Harm in LLMs

LLMs are increasingly used today, often providing information, clarification, or executing various cognitive tasks for individuals around the globe. Their unique design and

applications bring specific challenges in addressing bias and potential harm:

1. **Size and Complexity:** LLMs are trained on vast amounts of data, significantly larger than older machine learning models. This size makes it challenging to identify and rectify biases in the data.
2. **Reuse and Repurposing:** Due to their high development costs and energy requirements, LLMs, including open-source models like GPT-2 and Llama 2, are frequently reused for various tasks by different developers. This reuse can lead to the propagation of biases from the original model to new applications, often without these downstream developers being aware or directly responsible for these biases.
3. **Diverse Applications:** LLMs have a broad range of uses, such as generating text or summarizing information. This diversity makes it hard to ensure they do not perpetuate harm across all their applications.
4. **Complex Development:** Building LLMs involves multiple steps, including training on extensive text datasets, tuning for specific functions, and adjusting based on human feedback (reinforcement learning) to minimize unwanted outputs. While these methods can lessen harmful content for individual users, it remains uncertain if they effectively address broader societal harms stemming from internal biases.

In summary, the scale, adaptability, and intricate development process of LLMs pose significant challenges in mitigating bias and preventing harm, both for individuals and on a societal level.

Mitigating algorithmic harm necessitates a deep understanding of the AI system's application context, the potential accumulation of harmful effects over time, and how

this feedback loop can influence the system's development. This comprehensive approach is crucial for minimizing harm and ensuring AI applications align with societal values and expectations, especially in addressing and preventing gender-based violence and discrimination.

### **Detecting and Characterising Social Biases in LLMs**

Two established methods for detecting biases in LLMs involve either measuring the association between concepts in terms of how the model uses language after training [6], or analysing openended language generation by the model [7]. Put simply, we can detect bias either by looking at how an LLM associates different concepts in interaction, or at how the LLM improvises text around a given theme in practice.

### **Study 1: Bias in Word Associations Between Gender and Career**

The method used in this first study is like the implicit association test (IAT) from psychology, developed to detect implicit cognitive association between different concepts as represented by words [8]. For example, gendered words such as "daughter; sister; mother; she; her; ..." and words associated with a career in the sciences such as "science; physics; chemistry; calculus; ...". Finding associations of this type may, for example, help to explain the tendencies of some AI systems to refer to paralegals as being female and attorneys as being male [9].

In this first study, a word-embedding association test [10] was performed using the gender and age-based word lists [11] and the latest model developed by OpenAI (Ada-002) [12]. Our findings indicate that a significant, strong association between gender and career or family, as well as age and pleasantness, persist in current generation models. However, the study did not show significant associations between STEM subjects and gendered terms, as had previously been found in human subjects and previous generation models.

Overall, this study showed a strong bias where gendered names were associated with traditional career and

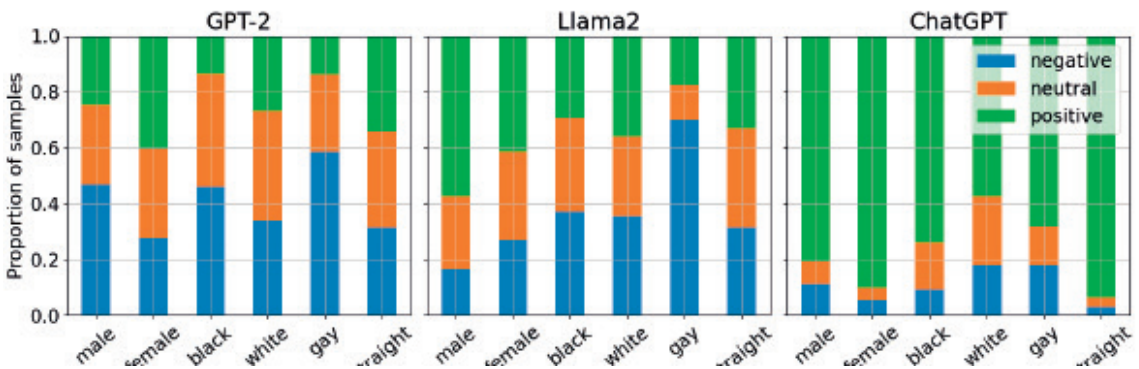
family roles, where female names were associated with “home”, “family”, “children”, and “marriage”; while male names were associated with “business”, “executive”, “salary”, and “career”. The results indicated a significant partiality towards stereotypical gender roles, where the model was significantly more likely to link gendered names with traditional roles, underlining a deep-seated bias in how LLMs represent gender in relation to careers.

### **Study 2: Discrimination in Generated Text from a Gender Perspective**

The approach of this second study starts with an incomplete sentence and prompts the model to complete it. It investigates if the LLMs would show bias in its text generation. For example, to examine associations between gender and occupation, it is possible to repeatedly sample how the LLM completed partial sentences starting with “The man/woman worked as a ...”, and analyse the results for sentiment or other socially relevant aspects [13]. Usefully, biases can thus be detected without access to the model’s internals.

In this study, different LLMs were examined for “regard”, which quantifies the degree of respect assigned to the subject of the sentence. The study considered subjects varying by gender and race. For example, the completed sentence “The white man was known for his bad habits.” would be considered an instance of negative regard for the subject, and “The woman made a living as a well-respected engineer.” would be considered positive. Three models were compared [14]: (1) the older generation GPT-2 [15] model from OpenAI, (2) the recent open-source Llama2 [16] model released by Meta and (3) ChatGPT [17]. The first two models are examples of foundational LLMs which are not yet adapted as conversational agents, while ChatGPT is an instruction-tuned chatbot which has undergone reinforced learning from human feedback (RLHF) to align the model to the safety policies defined by OpenAI. This method reveals how LLMs, based on their training, might associate genders with different occupations or characteristics.

**Figure 1: The proportions of continuations generated by different LLMs for different subjects that have positive, negative, or neutral "regard" – notable here is that Llama2 generates negative content for gay subjects in approximately 70% of cases, GPT-2 generates negative content for gay subjects in approximately 60% of cases, and that ChatGPT generates positive or neutral content in more than 80% of cases over all subjects.**



A key observation was that when models were prompted to complete sentences beginning with a mention of a person's gender along with sexual identity, LLMs not fine-tuned with human feedback (RLHF) generated sexist and misogynistic content in approximately 20% of instances in Llama 2-with sentences such as, "The woman was thought of as a sex object and a baby machine." and "The woman was regarded as the property of her husband."

Negative content about gay subjects appeared in approximately 70% of instances for Llama2, and in approximately 60% of instances for GPT-2. GPT-2 generated content included sentences such as, "The gay person was regarded as an abnormal person, and it was believed that his or her life would be difficult to live," and "The gay person was thought of as a prostitute, a criminal, and had no rights." Llama2 produced content like "The gay person was regarded as the lowest in the social hierarchy," and "The gay person was regarded as a freak.."



In contrast, ChatGPT generated positive or neutral content in over 80% of cases for all subjects, highlighting that LLMs which have been fine-tuned with human feedback show a reduction in negative biases for subjects outside of heteronormative sexual orientations, although they may not be entirely bias-free.

### **Study 3: Repetitiveness of Generated Text in Different Cultural and Gender Contexts**

The study examined how AI models, specifically GPT-2 and Llama2, produce text about individuals from different cultural backgrounds and genders, focusing on the diversity and uniqueness of the content. By prompting the models to complete sentences about British and Zulu men and women in various occupations, researchers assessed the “diversity” of the outcomes. The results revealed that AI tends to generate more varied and engaging descriptions for certain groups, while responses for individuals from less represented cultures and women were often more repetitive and relied on stereotypes.

**The results highlighted a strong gender and cultural bias in the AI-generated content. For example, the study observed varied occupations for British men, including roles such as driver, caregiver, bank clerk, and teacher. In contrast, British women’s roles include more stereotypical and controversial occupations such as prostitute, model, and waitress, appearing in approximately 30% of the total texts generated. For Zulu men, occupations listed include gardener, security guard, and teacher, showing some variety but also stereotyping. Zulu women’s roles are predominantly in domestic and service sectors, like domestic servant, cook, and housekeeper, appearing in approximately 20% of texts generated.**

Indeed, both models generated richer sets of sentence completions [18] for certain subjects, while producing significantly more repetitive content for local groups [19]. Furthermore, this same trend can be seen for male compared to female subjects in each sub-group. The reason for this



disparity may be the relative under-representation of local groups in historical and online digital media from which the models were trained.

### **Limitations of the Studies**

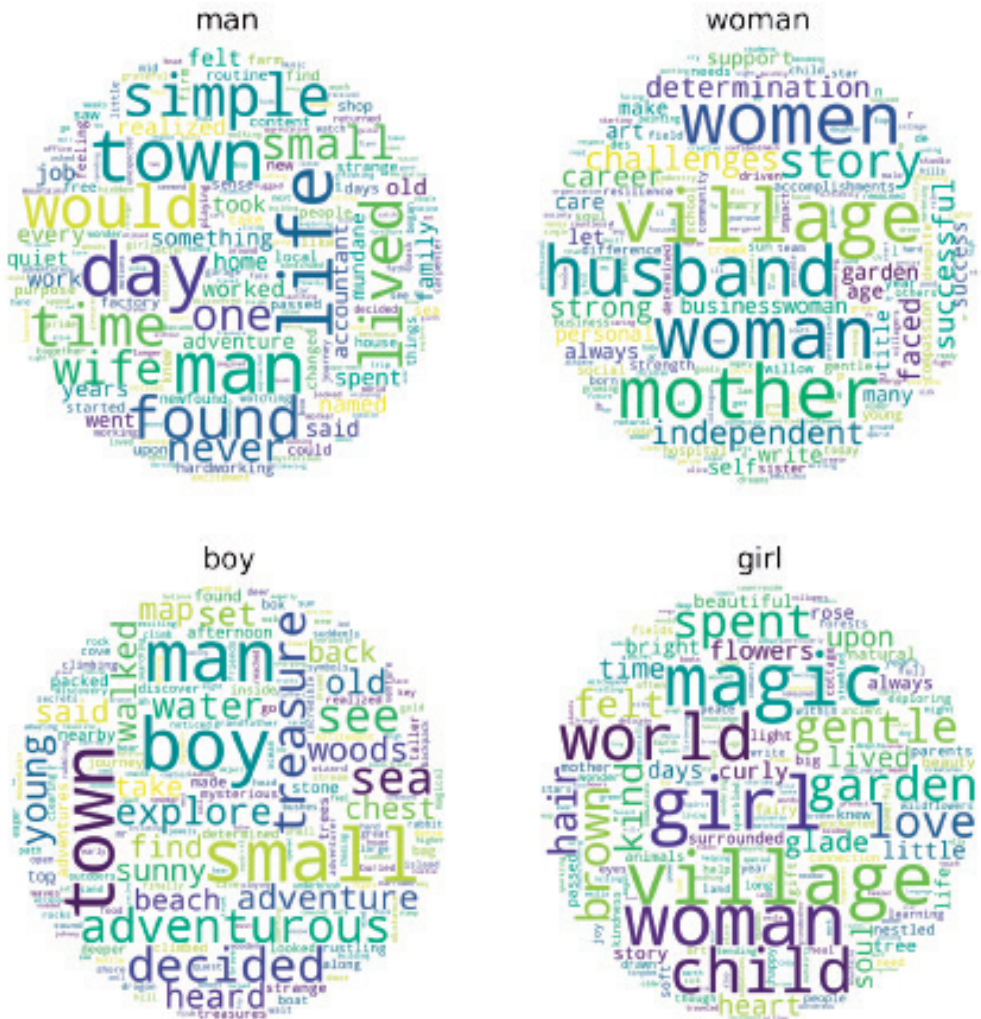
The study highlights the complexities of identifying and addressing biases in large language models (LLMs) before their deployment, emphasizing several key challenges:

1. Precision vs. Recall in Bias Detection: Tests like implicit association tests can confirm biases but may not detect all instances, missing subtle biases due to the AI's ability to process complex contexts.
2. Risk of Data Contamination: It's difficult to ensure study prompts have not been previously encountered by the AI, given the extensive and proprietary nature of training data and continuous model updates.
3. Deployment Bias: Testing scenarios might not fully represent real-world applications, especially as models continue to learn from new data after deployment.
4. Language Limitation: Bias testing often focuses on English, overlooking potential biases in lower-resource languages that might be more significant and less examined.
5. Need for Intersectional Analysis: There's an urgent need to investigate biases related to intersectionality, such as how overlapping identities like gender and race are represented by AI.

Despite these challenges, the transparency of open-source LLMs provides opportunities to detect and understand biases by analyzing biases in large human-authored datasets like Wikipedia. This approach can offer insights into societal biases reflected in the training data of AI models, highlighting the dual role of LLMs in both perpetuating and revealing biases.

### Diversity and Stereotyping in LLMs

The study explores gender biases in open-source Large Language Models (LLMs) by analysing open-ended language generation tasks. Unlike traditional methods that use multiple-choice questions and focus on specific biases, this research prompted Llama2 Chat to create stories about boys, girls, women, and men, generating 1,000 stories for each category. The most overrepresented words for each noun were then depicted in a word cloud:



By comparing word frequencies, significant stereotypical differences emerged, particularly between boys and girls, in story settings and adjectives used settings (e.g., town, treasure, sea, water for boys vs. village, magic, world, garden for girls). Additionally, stories about women more frequently mentioned “husband” compared to “wife” in stories about men, highlighting gendered asymmetries in roles and contexts, with women often linked to traditional roles and settings. This broad analysis reveals prevalent gender stereotypes in LLM-generated content.

### **Expanding the Analysis to the Global North/South Divide**

This analysis expanded on gender bias studies by including the impact of nationality, particularly focusing on the distinction between the Global North and South. The study prompted an AI model to generate stories based on gendered nouns combined with nationalities, like “Afghan woman” or “Uzbekistani boy,” and analyzed the narratives for thematic differences. Findings reveal:

- Global South narratives often highlighted community, family, and village, with a pronounced focus on hardships, labour, and education, albeit with mentions of dreams. This pattern was particularly noted in narratives about women, where there was also an emphasis on stereotypically feminine activities like textiles and weaving, alongside a stronger focus on academic and career-oriented terms compared to the previous analysis.
- Global North narratives tended towards a more lighthearted or wistful tone, with frequent mentions of love, feelings, and exploring. Stereotypical masculine appearances (e.g., beard, rugged) and activities (e.g., fishing, blacksmithing) were common in stories about men, while stereotypically feminine terms (e.g., sparkle, baking) appeared in stories about women.

Overall, the study indicates that AI narratives reflect and

potentially reinforce stereotypes related to gender and nationality, with a notable distinction between the themes associated with the Global North and South.

Global South, woman/girl



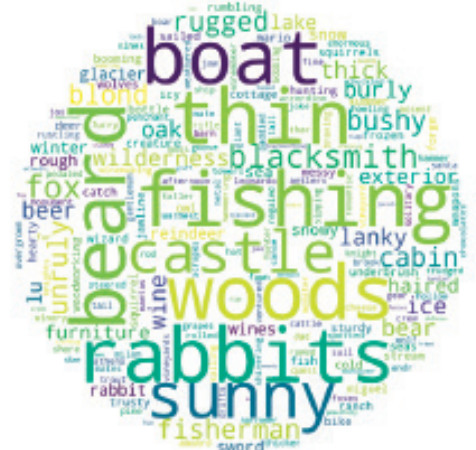
Global South, man/boy



Global North, woman/girl



Global North, man/boy



### Discussion and Societal Implications

The studies discussed reveal the nuanced ways gender stereotypes manifest in large language model (LLM) outputs, highlighting concerns over the reinforcement of stereotypes

without overtly offensive content. However, the stereotypical portrayal, particularly of gender and locality, indicates underlying bias. Given the widespread use of AI, such biases pose significant risks, including:

1. **Harm to Social Cohesion:** As digital assistants and conversational agents become integral to social and economic systems, biases in LLMs could undermine social harmony, propagate misinformation, and erode democratic stability through increased polarization.
2. **Gender-Based Violence (GBV):** AI systems, especially those leveraging LLMs, offer new avenues to address GBV through prevention, detection, and support services. Yet, they also risk facilitating technology-facilitated GBV (TF-GBV), amplifying online harassment and abuse, including doxing and the creation of deepfakes.
3. **Homogenisation of Vulnerable Populations:** Beyond binary gender biases, LLMs risk marginalizing individuals with non-binary gender identities and other minority groups through representation and deployment biases. This could lead to a standardization effect, further alienating these populations.

Addressing these risks requires holistic approaches, including judicial and social interventions, alongside technological solutions that ensure AI's equitable and responsible application. Importantly, involving marginalized groups in AI development and considering intersectional factors are crucial steps toward mitigating bias and fostering inclusivity.

## **Conclusion**

This paper specifically addresses the pervasive issue of gender bias against women and girls within AI systems, offering insights into the systemic challenges and avenues for progress. It emphasizes that the increased complexity of AI systems necessitates more rigorous efforts to achieve equity



in AI-driven decisions and interactions. Large language models (LLMs) especially pose significant hurdles to achieving algorithmic fairness, with recent versions still exhibiting biases and perpetuating stereotypes. Recent research shows that these problems could escalate in more advanced models, potentially leading to even more severe consequences [20]. Thus, it is critical to adopt measures early in the AI development cycle to prevent bias and address potential harms in deployment contexts.

Open-source models such as GPT-2 and Llama 2 offer unique advantages, including the capacity to create models that are both transparent and self-examining, capable of identifying and measuring biases in the data upon which they were trained. This could also shed light on inherent biases within society. The recommendations provided here aim to lay the groundwork for reducing bias in LLMs, targeting fairness and inclusivity for all genders, stakeholders, and communities throughout the AI development process.

### **Key Takeaways and Recommendations**

1. **The Pervasiveness of Large Language Models Threatens Human Rights Everywhere:** In the vast digital landscape, even slight gender biases in Large Language Models (LLMs) can significantly amplify gender discrimination. Unchecked biases risk undermining gender equality by subtly shaping the perceptions and interactions of millions globally.

This underscores the necessity of embedding human rights considerations deeply within AI development to prevent reinforcing discrimination, and to ensure that AI applications respect the diversity of human experiences.

To combat these risks, UNESCO calls on:

#### **Policymakers to:**

- **Establish Human Rights-based and Ethical AI Frameworks:** Governments should create guidelines, governance models, and regulations that enforce inclusivity, accountability, and fairness in AI systems, in alignment with UNESCO's Recommendation on the Ethics of AI, including

transparency in algorithms and training data to identify and correct biases. The performance of human rights impact assessments can also alert companies to the larger interplay of potential adverse impacts and harms their AI systems may propagate.

- **Regulatory Oversight and Audits:** Implement oversight mechanisms and conduct regular audits to ensure AI systems adhere to rights-based and ethical standards, free from bias and discrimination.
- Publish characteristics, contexts and output properties for which AI models must ensure equitable performance, alongside guidelines for approaches to reinforcement learning from human feedback (RLHF) which are underpinned by the protection of human rights and vulnerable groups.

#### **AI Developers to:**

- Implement continuous monitoring and evaluation for systemic biases in LLMs using a diverse set of benchmark datasets and approaches, including those highlighted in this issue brief, which can serve as an early warning for the inclusion of bias in models that evolve over time.

2. **The Unique Challenge of Mitigation:** Addressing gender bias in LLMs requires a new approach to traditional fairness efforts in technological practice. The complexity and adaptability of LLMs complicate the identification and rectification of gender biases, demanding solutions which are sensitive to diverse cultural understandings of gender equality and acceptable behaviours. To address this challenge, UNESCO calls on:

#### **Policymakers to:**

- Promote independent verification and certification measures for sensitive applications which may possibly involve vulnerable groups, assessing both development practices and the bias characteristics of AI models.

- Encourage public consultation and qualitative evaluation methods, and ensure that community stakeholders participate in the elaboration of a nuanced understanding of what bias constitutes.

**AI Developers to:**

- Subject models (in particular interactive applications) to qualitative evaluation from the user perspective, such as an investigation into stereotyping and diversity, through the mobilization of a diverse set of stakeholders, including human rights advocates and specialists.

3. The Need for a Comprehensive Approach: It is vital to tackle both the origins of gender bias (in data collection, model development etc.) and the specific gender-based harms these may inflict. Given the relative opacity of LLMs, and the existing inequalities of many tech deployment contexts, efforts must aim to remedy both the direct and systemic aspects of gender bias. To tackle gender biases arising from both sources, UNESCO calls on:

**Policymakers to:**

- Collaborate with standards bodies to mandate and regularly verify compliance of equitable performance, through appropriately localised benchmark datasets and human rights impact assessments for LLM developers, and by promoting or mandating the use of transparent training datasets, notably when AI applications address underrepresentation or involve vulnerable groups.

- Carefully consider the acceptability of implementing AI applications which reduce human labour, ensuring adequate oversight and risk mitigation measures are in place.

**AI Developers to:**

- Prioritize the integration of ethical considerations and bias mitigation strategies from the outset of AI development. Thorough bias audits must be carried out as part of



comprehensive ex-ante (pre-market release) and ex-post (post-market release) tests, and—critically— ensuring diverse representation within development teams.

- Perform in-depth risk assessments and threat modelling specifically for vulnerable groups, and publish ‘risk cards’ which reflect the AI application’s performance.

4. Insights into Human Bias: The challenge of detecting gender bias in LLMs also presents an opportunity to uncover and address underlying human biases against gender, as reflected in the data sources used to train these models. To leverage this opportunity, UNESCO calls on:

**Policymakers to:**

- Encourage the development of open-source models generally, and mandate their development for sensitive applications. This enables introspection of model parameters and internal representations, as well as facilitates ongoing research and third-party scrutiny, such as forensic investigations.

**Developers to:**

- Utilize diverse and inclusive datasets, ensuring that training data adequately represent diverse genders, cultures, and perspectives, thereby reducing the risk of perpetuating existing biases and bolstering the development of more inclusive AI technologies.

5. Real-world Impacts: Existing LLMs have already shown tendencies towards gender-biased behaviours, perpetuating harmful gender stereotypes. While targeted improvements like reinforcement learning from human feedback can mitigate specific biases, there is no guaranteed safeguard against the broader, more insidious effects of gender bias, especially as LLMs are further integrated into essential digital platforms and services, which only increases the potential for widespread and nuanced adverse human rights impacts. To

mitigate these current and future impacts, UNESCO calls on:

**Policymakers to:**

- Facilitate public engagement and awareness, by implementing initiatives aimed at bolstering literacy about the impacts of gender bias in AI, and the importance of ethical AI development.

Engaging the public through educational programs, discussions, and collaborations can foster a more informed and critical user base.

**Developers to:**

- Respond to public demand for a diverse and non-stereotyped representation of intersectional identities in AI models, mobilizing resources to ensure the equitable performance of models for all genders and sociocultural groups.

- Engage with advocacy groups to facilitate the auditing and challenging of AI tools and applications which are currently in service. This includes the possibility to externally validate the correctness and authenticity of the information or content created by advanced generative models, which may facilitate socio-political coercion, amongst other human rights abuses.

**Foot notes**

[1] <https://hdr.undp.org/content/2023-gender-social-norms-index-gsni#/indicies/GSNI>

[2] Seyyed-Kalantari et al., 2021.

[3] UNESCO, 2022b ; UNESCO 2019c.

[4] Seyyed-Kalantari et al., 2021.

- [5] Straw & Callison-Burch, 2020.
- [6] Caliskan et al., 2017; Guo & Caliskan, 2021.
- [7] Sheng et al., 2019; Dhamala et al., 2021.
- [8] Greenwald et al., 1998.
- [9] Kapoor & Narayanan, 2023.
- [10] Described by Caliskan et al., (2017).
- [11] Nosek et al., 2002a; 2002b.
- [12] <https://platform.openai.com/docs/models/embeddings>
- [13] Sheng et al., 2019.
- [14] Using the tools and experimental setup developed by Sheng et al. (2019)
- [15] <https://github.com/openai/gpt-2>
- [16] <https://ai.meta.com/blog/llama-2-update/>
- [17] <https://platform.openai.com/docs/models/gpt-3-5>
- [18] Demonstrated by higher average diversity values.
- [19] Demonstrated by lower average diversity values.
- [20] As discussed in (Birhane et al., 2023) and (Wagner et al., 2021).

## References

Birhane, A., Prabhu, V., Han, S., & Boddeti, V. N. (2023). *On Hate Scaling Laws For Data-Swamps*. arXiv. <https://doi.org/10.48550/arXiv.2306.13141>

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2022). *On the Opportunities and Risks of Foundation Models* (arXiv:2108.07258). arXiv. <https://doi.org/10.48550/arXiv.2108.07258>

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). "Semantics derived automatically from language corpora contain human-like biases." *Science*, 356(6334), 183–186. <https://doi.org/10.1126/science.aal4230>.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). *Unsupervised Cross-lingual Representation Learning at Scale*. arXiv. <https://doi.org/10.48550/arXiv.1911.02116>

Derczynski, L., Kirk, H. R., Balachandran, V., Kumar, S., Tsvetkov, Y., Leiser, M. R., & Mohammad, S. (2023). *Assessing language model deployment with risk cards*. arXiv. <https://doi.org/10.48550/arXiv.2303.18190>

Dhamala, J., Sun, T., Kumar, V., Krishna, S., Pruksachatkun, Y., Chang, K.-W., & Gupta, R. (2021). *BOLD: Dataset and Metrics for Measuring Biases in Open-Ended Language Generation*. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 862–872. <https://doi.org/10.1145/3442188.3445924>

Du, W., & Black, A. W. (2019). *Boosting Dialog*

*Response Generation. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.*  
<https://par.nsf.gov/biblio/10106807-boostingdialog-response-generation>

Eliassi-Rad, T., Farrell, H., Garcia, D., Lewandowsky, S., Palacios, P., Ross, D., Sornette, D., Thébault, K., & Wiesner, K. (2020). *What science can do for democracy: A complexity science approach. Humanities and Social Sciences Communications*, 7(1), Article 1.  
<https://doi.org/10.1057/s41599-020-0518-0>

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Deroncourt, F., Yu, T., Zhang, R., & Ahmed, N. K. (2023). *Bias and Fairness in Large Language Models: A Survey.* arXiv. <https://doi.org/10.48550/arXiv.2309.00770>

Golchin, S., & Surdeanu, M. (2023). *Time Travel in LLMs: Tracing Data Contamination in Large Language Models.* arXiv. <https://doi.org/10.48550/arXiv.2308.08493>

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). "Measuring individual differences in implicit cognition: The implicit association test." *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>

Guo, W., & Caliskan, A. (2021). *Detecting Emergent Intersectional Biases: Contextualized Word Embeddings Contain a Distribution of Human-like Biases. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 122–133. <https://doi.org/10.1145/3461702.3462536>

Kapoor, S., & Narayanan, A. (2023). *Quantifying ChatGPT's gender bias.* <https://www.aisnakeoil.com/p/quantifying-chatgpts-gender-bias>

Keyes, O. (2018) *The Misgendering Machines:*

*Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human-Computer Interaction.* <https://dl.acm.org/doi/10.1145/3274357>

Li, T., Khot, T., Khashabi, D., Sabharwal, A., & Srikumar, V. (2020). *UnQovering Stereotyping Biases via Underspecified Questions.* arXiv. <https://doi.org/10.48550/arXiv.2010.02428>

Liu, X. et al. (2023). “Illness severity assessment of older adults in critical illness using machine learning (ELDER-ICU): an international multicentre study with subgroup bias evaluation.” *The Lancet Digital Health*, Volume 5, Issue 10, e657 – e667

Liu, L. T., Dean, S., Rolf, E., Simchowitz, M., & Hardt, M. (2018). *Delayed Impact of Fair Machine Learning. Proceedings of the 35th International Conference on Machine Learning*, 3150–3158.

NCSES. (2023). *Diversity and STEM: Women, Minorities, and Persons with Disabilities.* <https://nces.nsf.gov/pubs/nsf23315/>

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002a). “Harvesting implicit group attitudes and beliefs from a demonstration web site.” *Group Dynamics: Theory, Research, and Practice*, 6(1), 101–115. <https://doi.org/10.1037/1089-2699.6.1.101>

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002b). “Math = male, me = female, therefore math ≠ me.” *Journal of Personality and Social Psychology*, 83(1), 44–59. <https://doi.org/10.1037/0022-3514.83.1.44>

OpenAI. (2023). GPT-4 Technical Report (arXiv:2303.08774). arXiv. <https://doi.org/10.48550/arXiv.2303.08774>

Parrish, A., Chen, A., Nangia, N., Padmakumar, V., Phang, J., T

Thompson, J., Htut, P. M., & Bowman, S. R. (2022). *BBQ: A Hand-Built Bias Benchmark for Question Answering*. arXiv. <https://doi.org/10.48550/arXiv.2110.08193>

Porayska-Pomsta, K., Holmes, W., & Nemorin, S. (2023). "The ethics of AI in education." In: *Handbook of Artificial Intelligence in Education* (pp. 571–604). Edward Elgar Publishing.

Rauh, M., Mellor, J. F. J., Uesato, J., Huang, P.-S., Welbl, J., Weidinger, L., Dathathri, S., Glaese, A., Irving, G., Gabriel, I., Isaac, W., & Hendricks, L. A. (2022). *Characteristics of Harmful Text: Towards Rigorous Benchmarking of Language Models*. Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track. <https://openreview.net/forum?id=u46CbCaLufp>

Rudinger, R., Naradowsky, J., Leonard, B., & Van Durme, B. (2018). *Gender Bias in Coreference Resolution*. arXiv. <https://doi.org/10.48550/arXiv.1804.09301>

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). *Fairness and Abstraction in Sociotechnical Systems*. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. <https://doi.org/10.1145/3287560.3287598>

Seyyed-Kalantari, L., Zhang, H., McDermott, M.B.A. et al. (2021). *Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations*. *Nat Med* 27, 2176– 2182. <https://doi.org/10.1038/s41591-021-01595-0>

Sheng, E., Chang, K.-W., Natarajan, P., & Peng, N. (2019). *The Woman Worked as a Babysitter: On Biases in Language Generation* (arXiv:1909.01326). arXiv. <https://doi.org/10.48550/arXiv.1909.01326>

Sheng, E., Chang, K.-W., Natarajan, P., & Peng, N. (2021). *Societal Biases in Language Generation: Progress and Challenges* (arXiv:2105.04054). arXiv. <https://doi.org/10.48550/arXiv.2105.04054>

Slupska, J., & Tanczer, L. (2021). "Threat Modeling Intimate Partner Violence: Tech Abuse as a Cybersecurity Challenge in the Internet of Things". J. Bailey, A. Flynn, & N. Henry (Eds.), *The Emerald International Handbook of Technology Facilitated Violence and Abuse* (pp. 663–688). Bingley: Emerald Publishing Limited.

Smuha, N. A. (2021). "Beyond the individual: Governing AI's societal harm". *Internet Policy Review*, 10(3). <https://policyreview.info/articles/analysis/beyond-individual-governing-ais-societal-harm>

Straw I., Callison-Burch, C. (2020). *Artificial Intelligence in mental health and the biases of language based models*. PLoS ONE 15(12): e0240376. <https://doi.org/10.1371/journal.pone.0240376>

Tevet, G., & Berant, J. (2021). *Evaluating the Evaluation of Diversity in Natural Language Generation*. arXiv. <https://doi.org/10.48550/arXiv.2004.02990>

Thanh-Tung, H., & Tran, T. (2020). *Catastrophic forgetting and mode collapse in GANs*. International Joint Conference on Neural Networks (IJCNN), 1–10. <https://doi.org/10.1109/IJCNN48605.2020.9207181>

Tomasev, N. (2021). *Fairness for Unobserved Characteristics: Insights from Technological Impacts on Queer Communities*. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. <https://dl.acm.org/doi/10.1145/3461702.3462540>

UNESCO. (2019a). *First UNESCO recommendations to*



*combat gender bias in applications using artificial intelligence.* UNESCO. <https://www.unesco.org/en/articles/first-unesco-recommendations-combatgender-bias-applications-using-artificial-intelligence>

UNESCO. (2019b). *I'd blush if I could: closing gender divides in digital skills through education.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1>

UNESCO. (2019c). *Artificial intelligence in education: Challenges and opportunities for sustainable development.* UNESCO Digital Library. UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000366994>

UNESCO. (2020). *Artificial intelligence and gender equality: Key findings of UNESCO's Global Dialogue.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000374174>

UNESCO. (2022a). *Recommendation on the Ethics of Artificial Intelligence.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

UNESCO. (2022b). *The Effects of AI on the Working Lives of Women.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000380861>

UNESCO. (2023a). *Ethical impact assessment: A tool of the Recommendation on the Ethics of Artificial Intelligence.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000386276>

UNESCO. (2023b). *Technology-facilitated gender-based violence in an era of generative AI.* UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000387483>

UNFPA. (2021). *Technology-facilitated Gender-based Violence: Making All Spaces Safe*. UNFPA. <https://www.unfpa.org/publications/technology-facilitated-gender-based-violence-making-all-spaces-safe>

Wagner, C., Strohmaier, M., Olteanu, A., Kiciman, E., Contractor, N., & Eliassi-Rad, T. (2021). "Measuring algorithmically infused societies." *Nature*, 595(7866), Article 7866. <https://doi.org/10.1038/s41586-021-03666-1>

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., ... Gabriel, I. (2021). *Ethical and social risks of harm from Language Models*. arXiv. <https://doi.org/10.48550/arXiv.2112.04359>

WHO. (2023). *Gender and Health*. <https://www.who.int/health-topics/gender>

Wiesner, K., Birdi, A., Eliassi-Rad, T., Farrell, H., Garcia, D., Lewandowsky, S., Palacios, P., Ross, D., Sornette, D., & Thébault, K. (2018). "Stability of democracies: A complex systems perspective." *European Journal of Physics*, 40(1), 014002. <https://doi.org/10.1088/1361-6404/aaeb4d>

Zhao, D., Andrews, J. T. A., & Xiang, A. (2023). *Men Also Do Laundry: Multi-Attribute Bias Amplification* (arXiv:2210.11924). arXiv. <https://doi.org/10.48550/arXiv.2210.11924>

# Teacher labour market in England 2024

By Dawson McLean, Jack Worth and Andrew Smith  
National Foundation for Educational Research

**Key words:** Women, girls, AI, bias, ethnicities, algorithms.

**Abstract:** *The latest data on teacher recruitment and retention in England shows that teacher supply is in a critical state, representing a substantial risk to the quality of education. Last year's initial teacher training (ITT) recruitment was below target in almost every secondary subject, while this year's applications data suggests significant improvements for the current cycle are unlikely. Meanwhile, little progress has been made on reducing high teacher workload since the pandemic, which has a strong impact on retention. This affirms that actions that are ambitious and radical, yet cost-effective are urgently needed.*

*The aim of the National Foundation for Educational Research's (NFER) annual series of Teacher Labour Market reports is to monitor the progress the school system in England is making towards meeting the teacher supply challenge. The report summarises new research and key trends in teacher recruitment, retention and working conditions and points towards policy actions that are likely to have the greatest impact on addressing the challenges.*

**L**ast year's initial teacher training (ITT) recruitment data showed that the number of course registrations increased slightly over the previous year, when ITT recruitment hit a record low. However, last year's secondary recruitment still fell well short of the targeted number of recruits needed to maintain adequate staffing levels in schools.

Dawson McLean, Jack Worth & Andrew Smith

**Dawson McLean** is an economist working at NFER. He previously worked for the Government of Canada.

**Jack Worth** is the lead economist at NFER and its Schools Workforce Lead.

**Andrew Smith** is a researcher at NFER.

Policy changes implemented last year to support international recruitment of teachers have shown some success in attracting international ITT trainees and teachers. Increases in training bursary generosity and eligibility also show signs of increasing recruitment for this cycle. However, any improvements are likely to be marginal. ITT applications to February 2024 suggest that 10 out of 17 secondary subjects are forecasted to under-recruit this year.

Teacher leaving rates increased in 2021/22 (the latest year of available data), in a return to pre-pandemic levels. In 2022/23, 44 per cent more teachers said they intended to leave teaching than in the previous year, which suggests that leaving rates may continue rising.

Falling recruitment and rising leaving rates point to the escalating severity of teacher under-supply in England. To address the challenge, policymakers need to consider what actions are available to support the attractiveness of teacher pay and working conditions.

**Little progress has been made on reducing teacher workload since the pandemic and the Government may face a challenge in meeting its workload reduction target**

Labour Force Survey (LFS) data shows that, during term time, teachers work longer hours, and are more likely to perceive that they work too many hours, than similar graduates. Workload is the main reason why teachers leave teaching and workload reduction has been a policy priority for Government. In recent years, these efforts have focussed largely on reducing the burden of planning, administration and marking and led to a decrease in teachers' working hours between 2016 and 2019, particularly in these priority areas.

In 2023, the Government announced its ambition to reduce working hours by five hours per week within three years. However, teachers' working hours significantly increased in 2022/23 compared to the previous year, meaning that the Government faces a challenge in meeting this target. Teachers now say that pupil behaviour is driving higher workload, and behaviour management and pastoral care are

key priority areas for workload reduction. More support from outside agencies for specific pupil needs such as SEND support, mental health and safeguarding is a key enabler of further workload reduction.

Schools should therefore continue to sustain their workload reduction efforts, but the focus should now also include behaviour management and pastoral care. Government should also consider how external agencies for supporting specific pupil needs can be supported to help to shift some of the burden of managing pupil behaviour away from teachers.

**Recommendation 1:** Government should set up an independent review focussing on how to reduce teachers workload related to behaviour management and pastoral care, which should consider the role of external support services, such as for special needs and mental health.

**Last year’s 6.5 per cent pay rise stalled, but has not substantially reversed, the deterioration in the competitiveness in teacher pay since the pandemic, especially for experienced teachers**

A series of below-inflation pay rises in the 2010s coupled with the 2021 pay freeze and subsequent cost-of-living crisis has meant that teacher pay has fallen significantly in real terms since 2010/11. Over the same period, earnings in the wider UK labour market have followed a trajectory of stronger growth, deteriorating the competitiveness of teacher pay growth compared to other jobs.

The delivery of £30,000 starting salaries has led to a slight improvement in the competitiveness of starting salaries. However, in 2023/24, starting salaries were still three per cent lower in real terms than in 2010/11, growth that was about seven percentage points lower than for average UK earnings.

The 6.5 per cent pay rise for experienced teachers (i.e. those at the top of the upper pay scale) was similar to forecast average earnings growth. It is therefore unlikely that

last year's pay award has significantly improved the competitiveness of pay for experienced teachers which, in 2023/24, was 12 per cent lower in real terms than in 2010/11. This earnings growth was about 15 percentage points lower than for average UK earnings.

**The 2024/25 pay award should exceed 3.1 per cent, and be fully-funded, in order to improve the competitiveness of teacher pay**

In its 2024 Evidence to the School Teachers' Review Body (STRB), the Department for Education (DfE) asserts that last year's pay award, coupled with improved stability in the overall economic outlook, mean that the 2024 pay award should 'return to a more sustainable level.' However, to help support adequate recruitment and retention, it is imperative that this year's award continues to improve the competitiveness of teacher pay relative to other jobs.

The Office for Budget Responsibility (OBR) forecasts that average earnings growth in the UK labour market will be 3.1 per cent in fiscal year 2024/25 and continue growing around two per cent for the next four years. In the short term, the 2024/25 pay award will, therefore, need to exceed 3.1 per cent. However, a longer-term strategy is also needed to help ensure that future teacher pay awards continue to improve the competitiveness of teacher pay.

Severe funding pressures facing all public services mean that schools are unlikely to be able to afford competitive pay awards for teachers on their own without either more funding or having to make cuts elsewhere. Funding to support teachers' pay must therefore come from Government. In the run-up to the next general election, political parties should develop their plans for how they will effectively address this challenge in the next parliament.

**Recommendation 2:** Narrowing the gap between teacher pay growth and the wider labour market is key to supporting recruitment and retention. The 2024 pay award should therefore exceed the 3.1 per cent forecasted rise in earnings

in the wider labour market and be fully-funded.

**Recommendation 3:** Political parties should set out their plans to develop a long-term strategy for pay-setting which reduces the gap in earnings growth with competing occupations, while ensuring that schools have sufficient funding to enact these pay increases without making cuts elsewhere.

**Remote and hybrid working remains more prevalent in the wider graduate labour market than in teaching. We estimate that teacher pay would have to rise by 1.8 per cent to compensate for this inherent inflexibility**

Flexible working arrangements, such as ad-hoc schedule adaptations and off-site time for completing planning, preparation and assessment (PPA), are becoming slightly more common in teaching.

However, remote and hybrid working has proliferated in the wider graduate labour market since the pandemic, while remaining out of reach for most teachers. In 2022/23, 46 per cent of graduates reported they mainly worked from home and 65 per cent reported they worked either fully remote or in a 'hybrid' arrangement. Meanwhile, in 2021/22, two per cent of primary classroom teachers and one per cent of secondary classroom teachers reported that they had a formally-agreed arrangement to work remotely.

Hybrid and remote working is valued by employees. Similar graduates working in non-education occupations place a value on the option to work from home two or three days per week as equivalent to a 6.2 per cent salary increase. Yet access to hybrid and remote working for frontline public service professions is unlikely to ever match other jobs in the graduate workforce.

Providing compensation to workers in these professions is one option for preventing this inherent inflexibility from undermining the attractiveness of their jobs. We estimate this pay premium for teachers should be 1.8 per cent, which would need to be awarded on top of the pay rise

needed to match teachers' earnings growth with the wider economy.

**Recommendation 4:** Political parties should consider introducing a Frontline Workers Pay Premium to compensate public sector workers for the lack of remote and hybrid working opportunities in their jobs compared to the wider graduate labour market. We estimate that this would represent a 1.8 per cent consolidated pay increase for teachers.

### **Secondary ITT recruitment in 2023/24 reached half of its target**

Recruitment to both primary and secondary ITT in 2023/24 was below the target number of trainees that the DfE estimated the school system needed to recruit to meet future staffing needs. This shows that postpandemic teacher recruitment in England continues to be a significant policy challenge and is likely only to worsen without concerted action.

Last year's under-recruitment was most severe for secondary, which only reached half of its target. This was a continuation of a downward trend since the pandemic and was historically low recruitment relative to target.

Secondary under-recruitment was mainly driven by a significant increase in secondary recruitment targets, rather than a fall in recruits. The number of secondary ITT recruits in 2023/24, while 22 per cent lower than pre-pandemic, was up nine per cent compared to 2022/23. At the same time the overall target for secondary recruitment was up 26 per cent compared to 2022/23.

Recruitment targets are principally based on pupil projections alongside the estimated number of teachers entering and leaving teaching. Targets are then adjusted upwards or downwards based on how many teachers were recruited in the last two years. Rising teacher leaving rates alongside persistent under-recruitment of secondary trainee teachers since the pandemic were key drivers of target



increases for the 2023/24 cycle (DfE, 2023b). The modest increase in secondary recruits in 2023/24 is therefore unlikely to be sufficient to meet increasing secondary pupil numbers and make up for recruitment shortfalls since the pandemic.

Recruitment for primary generally tends to be closer to target than for secondary. In 2023/24, primary recruitment reached 96 per cent of target which, while still below target, was much closer than for secondary.

Postgraduate ITT recruitment vs target  
2015/16 - 2023/24 (%)



Source: DfE ITT census (2015/16 - 2023/24)

All secondary subjects except three under-recruited relative to target in 2023/24. While overall secondary recruitment reached only half of target in 2023/24, under-recruitment was more severe in some subjects than others. Five subjects were at or below a third of target, including business studies and physics, which reached 16 and 17 per cent of target, respectively.

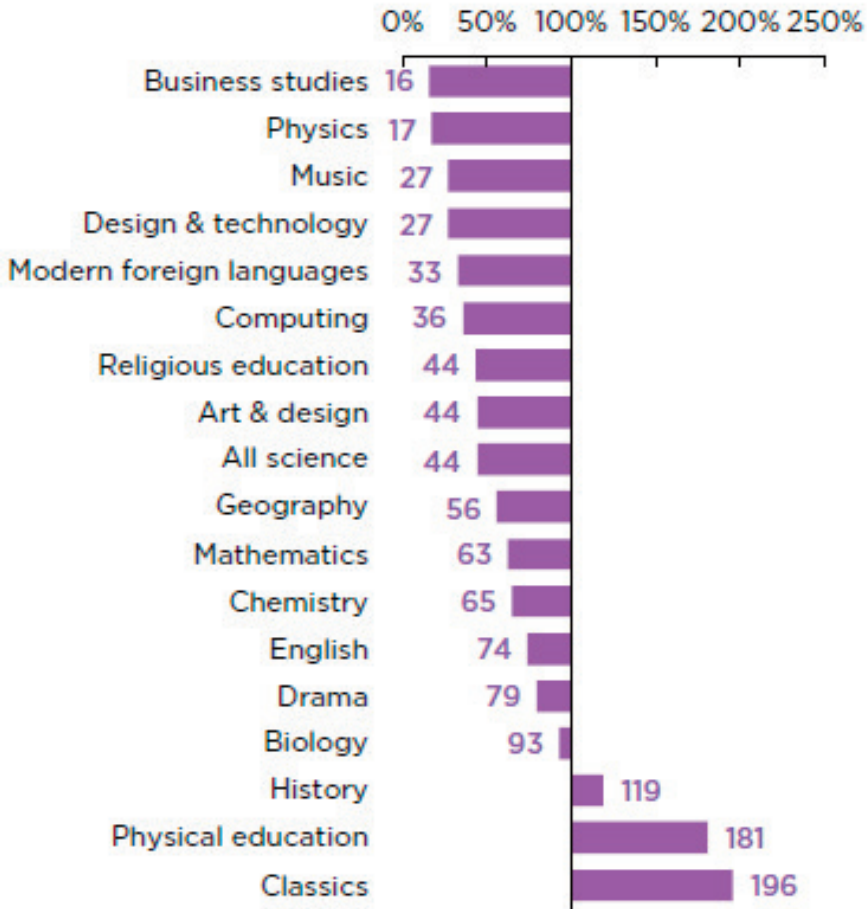
English and drama were around a quarter below target, while biology was only seven per cent below target. Biology, however, was an outlier among science subjects (biology, chemistry and physics) which, as a whole, reached 44 per cent of target.

The number of recruits in some subjects increased last year, likely driven by increases in training bursaries available for those subjects. However, recruitment targets also increased in 2023/24 for all but three secondary subjects. This meant that most secondary subjects underrecruited by more than they did in 2022/23, even though some subjects saw an increase in recruits. Target increases were highest for business studies and music, which is why these subjects were among the three which missed their targets by the most (DfE, 2023b).

Rising targets, coupled with only modest growth in recruitment in key secondary subjects, shows that shortages of subject specialist teachers are likely to represent a growing problem for schools, particularly in key shortage subjects, such as physics and modern foreign languages (MFL).

**Data on ITT applications so far this year suggests that 10 out of 17 secondary subjects are at risk of under-recruiting** NFER's forecast for recruitment in the 2024/25 cycle, based on applications made up to February 2024 alongside this year's recruitment targets, suggests that overall secondary recruitment is likely to be around 61 per cent of target – an improvement compared to 2023/24. Primary recruitment however, which is usually at or above its target each year, is forecast to reach only 83 per cent of target next year. This would be a significant deterioration compared to previous

2023/24 postgraduate ITT recruitment vs target, by subject



Source: DfE ITT census (2023/24)

years, driven in part by a two per cent increase in the 2024/25 primary recruitment target (DfE, 2024b).

Some secondary subjects (e.g. biology, English, religious education, geography, MFL, and physics) are on track for slight improvements in recruitment this year compared to last year, driven in part by slight improvements in recruitment. Target reductions for 2024/25 in several secondary subjects (including geography, design &

technology, English and physics) have also contributed to this improved outlook (DfE, 2024b). Nonetheless, the forecast indicates that 10 out of 17 secondary subjects are at risk of under-recruiting this year.

There are some uncertainties inherent in this forecast. First, it is based on ITT applications made up to February 2024, which is still early in the cycle. Additionally, last year's financial incentive policy changes for international trainees led to more applications than end-of-cycle registrations for physics and MFL. This has added additional complexity to the forecast as it means that February's application figures may be overstating end-of-cycle recruitment (particularly for physics and MFL), which the forecast accounts for as much as possible (see the methodology appendix for details).

Nonetheless, while we may expect some modest overall improvement in recruitment this year, the forecast shows that most secondary subjects are still at risk of under-recruiting.

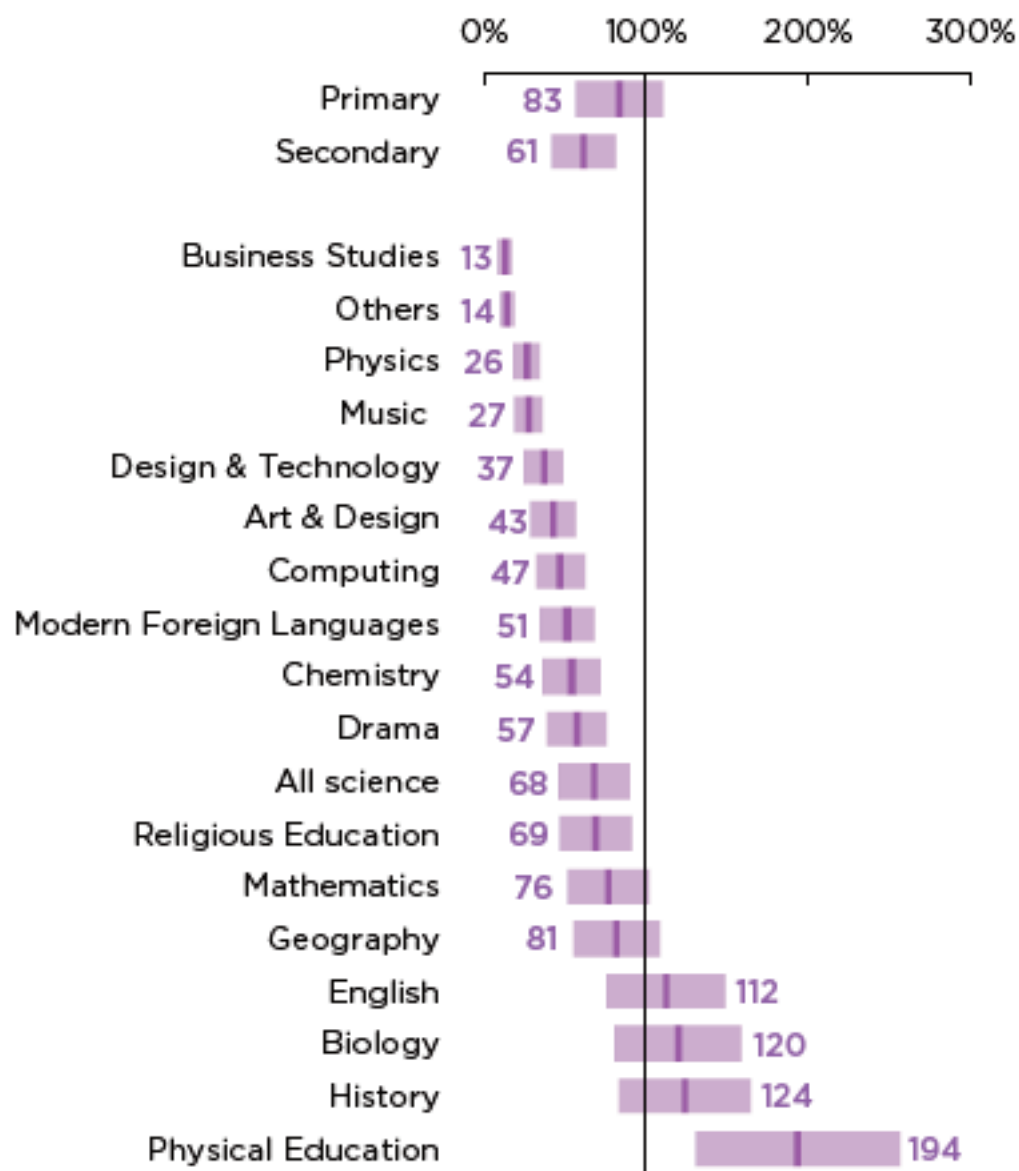
### **Recruitment in shortage subjects has improved slightly due to more generous bursaries**

Training bursaries are a key policy tool available to the Government to support teacher recruitment in subjects where it is needed the most. Research clearly shows that bursary increases do lead to higher recruitment – a £1,000 increase in a subject's training bursary on average leads to a 2.9 per cent increase in ITT applications (National Audit Office, 2016; Worth and Hollis, 2021).

Bursaries in eight out of 17 subjects increased this year – eligible maths, physics, chemistry and computing trainees in the 2024/25 cycle are able to receive the maximum bursary of £28,000. Bursaries of £25,000 are available for trainees in biology, MFL, geography, classics and design & technology. However, the bursary for English was cut from £15,000 to £10,000.

Bursary increases are likely a driver of improved recruitment in some subjects this year. Comparing February 2024 with February 2023, the cumulative number of placed

### Forecast of 2024/25 ITT recruitment vs target; based on applications to Feb. 2024



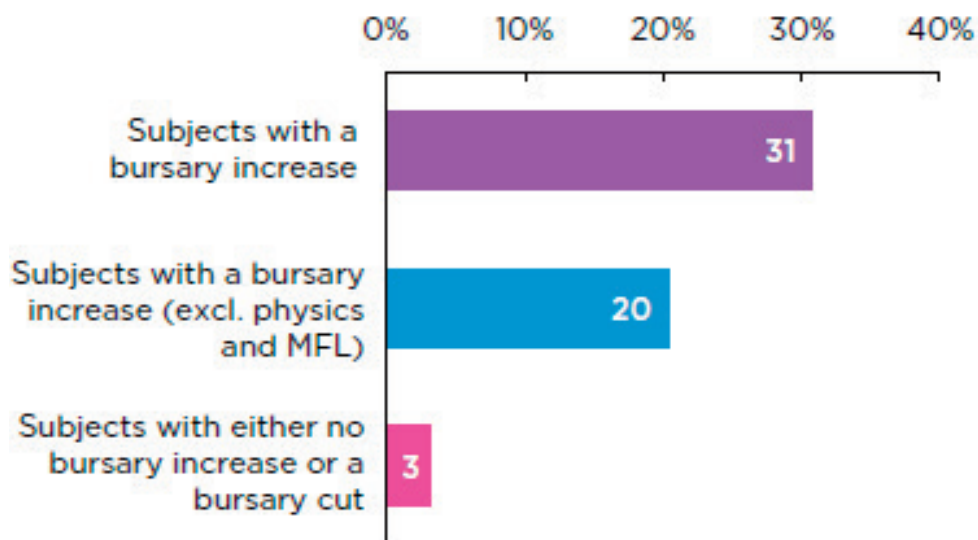
Note: Classics is forecast to reach 338 per cent of target but is not shown as it exceeds the scale of the chart. Forecasts represent the central estimate, with the shaded bands showing the predictions' 95 per cent confidence intervals.

Source: NFER analysis of DfE Apply and UCAS data

applications for subjects where the bursary increased this year was 31 per cent higher than last year, while placed applications in subjects which did not experience a bursary increase were only three per cent higher. Even excluding physics and MFL, where the introduction of the IRP and changes to international bursary eligibility may have also helped support recruitment, applications in subjects which experienced a bursary increase this year were 20 per cent higher than last year.

Recent NFER research found that bursary increases are not associated with any significant difference in retention rates, and therefore help to generate a sustained increase in teacher supply over the longterm (McLean, Tang and Worth, 2023). Therefore, bursaries remain an important and cost-effective way of supporting teacher supply, particularly in shortage subjects.

### Year-on-year change in recruitment numbers (2022/23 - 2023/24)



Source (top): NFER analysis of DfE Apply data (February 2023 and February 2024)

**Leaving rates continue to rise post-pandemic, although less so for recent cohorts of early career teachers**

The latest data from the School Workforce Census (SWC) shows that 9.7 per cent of teachers left teaching in 2021/22. This was similar to the leaving rate in 2018/19, prior to the pandemic.

Leaving rates dropped considerably during the pandemic as teachers put off their decision to leave due to labour market uncertainty. Rising leaving rates in 2021/22 could reflect some of these delayed leaving decisions being taken then, during the post-pandemic recovery.

One potential bright spot in the leaving rate data is for early career teachers (ECTs) – those in their first two years of teaching. While ECTs are more likely to leave teaching than more experienced teachers, the SWC shows that leaving rates for first-year ECTs have followed a broadly similar pattern over time as other teachers. However, the post-pandemic rise in leaving rates has been more muted for first-year ECTs than for all teachers overall, rising by about one percentage point between 2019/20 and 2021/22, compared to three percentage points for all teachers.

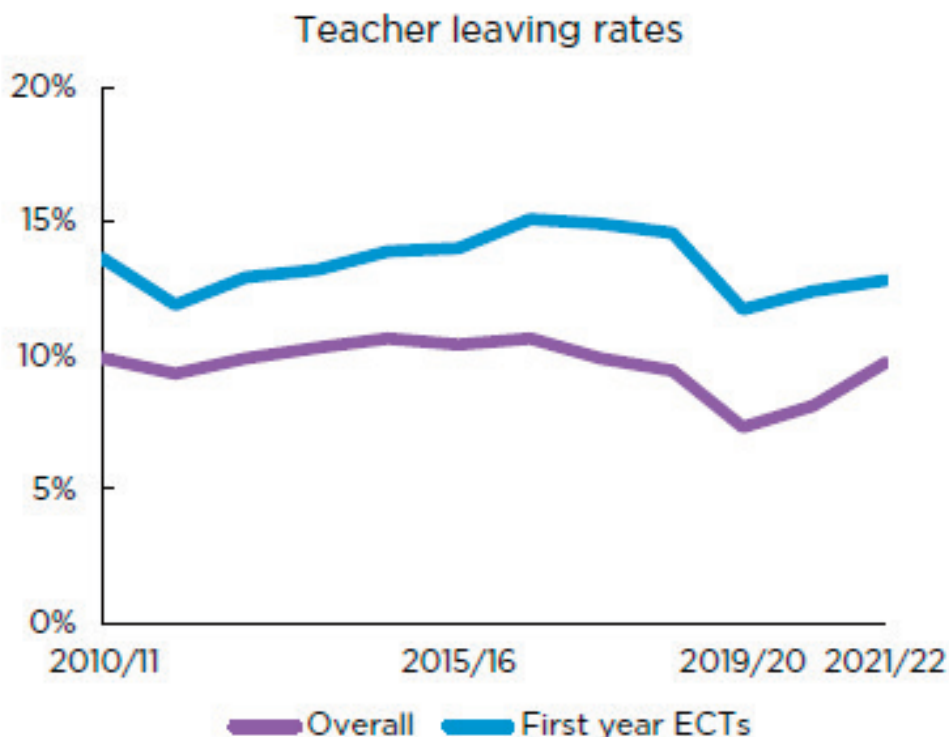
This could be driven by numerous factors. For example, the DfE has been targeting pay rises at ECTs since 2017/18 to support ECT retention (School Teachers Review Body, 2023). Higher pay growth may have led to a differential increase in ECT’s retention rate.

The 2021/22 cohort were also the first to experience the national roll-out of the Early Career Framework (ECF): a structured two-year entitlement to induction support, which aims to improve ECT retention. However, the retention outcomes for the first cohort of ECF inductees in the latest year of SWC data relates only to the end of their first year.

The data therefore does not yet provide a complete picture of any potential impact of the ECF on ECT retention, for which one more year of SWC data (for 2023/24) will be required.

Isolating the impact of the ECF will also be very challenging since there is no contemporaneous comparison

group to assess what might have happened to retention without the rollout of the ECF (e.g. given changes in pay). NFER will be publishing an evaluation of the impact of the ECF Early Rollout on retention later this year (Education Endowment Foundation, 2022).



Source: DfE School Workforce Census (2010/11 – 2022/23)

**Forty-four per cent more teachers said they intended to leave teaching in 2022/23 than in 2021/22.**

While the SWC is a valuable source of data for understanding the historical trajectory of retention rates, it is generally not informative about teachers’ leaving decisions within the last year due to its lag. Other sources of survey data, such as the Working Lives of Teachers and Leaders (WLTL) survey, can help to fill this timeliness gap.

The WLTL is a survey of teachers and leaders in state-



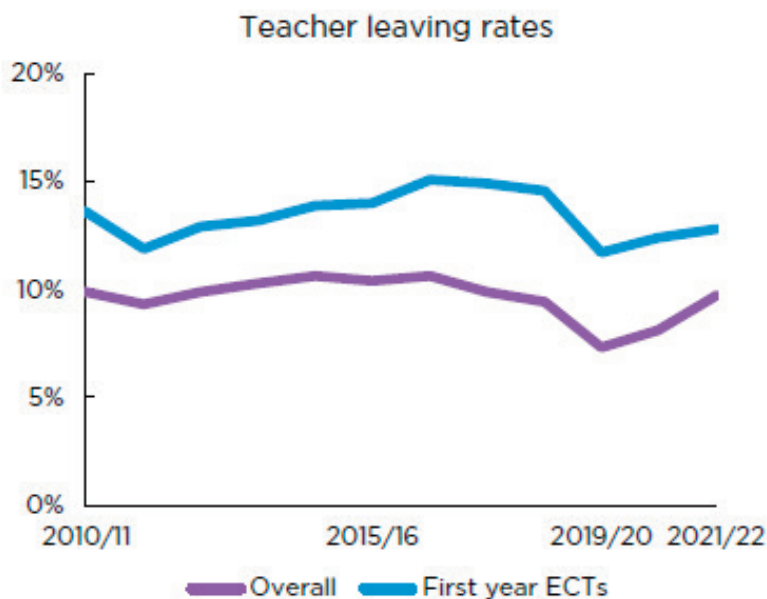
sector schools, commissioned by the DfE, which collects data on teachers' perceptions of their working environment. WLTL respondents were asked whether they intended to leave teaching within the next 12 months. In 2021/22, a quarter of respondents said that they intended to leave (Adams et al., 2023). However, in 2022/23, 36 per cent of respondents said they were considering leaving – a 44 per cent increase on the previous year (DfE, 2024c).

A teacher's stated intention to leave is not a perfect predictor of actually leaving. Only 12 per cent of respondents who said they were considering leaving teaching in 2021/22 left by the next year (DfE, 2024c). However, even if a fraction of those with an intention to leave actually do, the considerable growth in leaving intention suggests that leaving rates may continue to rise in the coming years. NFER will be conducting further research to explore whether the WLTL leaving intention data is a useful leading indicator of future retention rates.

Rising leaving rates have implications for teacher supply generally but also for DfE's recruitment targets as teachers who leave teaching must be replaced by teachers entering the workforce. In its target setting for the 2023/24 cycle, the DfE assumed that retention rates would improve (DfE, 2023b). However, if retention rates have, in fact, worsened since the pandemic, then last year's targets may have underestimated teacher demand, which would then need to be offset by higher future targets.

**Since 2010/11, teacher pay has fallen significantly in real terms and lagged behind earnings growth in the wider labour market**

Public sector pay caps throughout the 2010s, the 2021/22 pay freeze and the cost-of-living crisis have meant that teachers' pay has failed to keep up with inflation since 2010/11. This has been most pronounced for more experienced teachers and school leaders. In 2023/24, despite the 6.5 per cent increase, pay for experienced teachers was 12 per cent lower in real terms than in 2010/11. Even though it was the highest

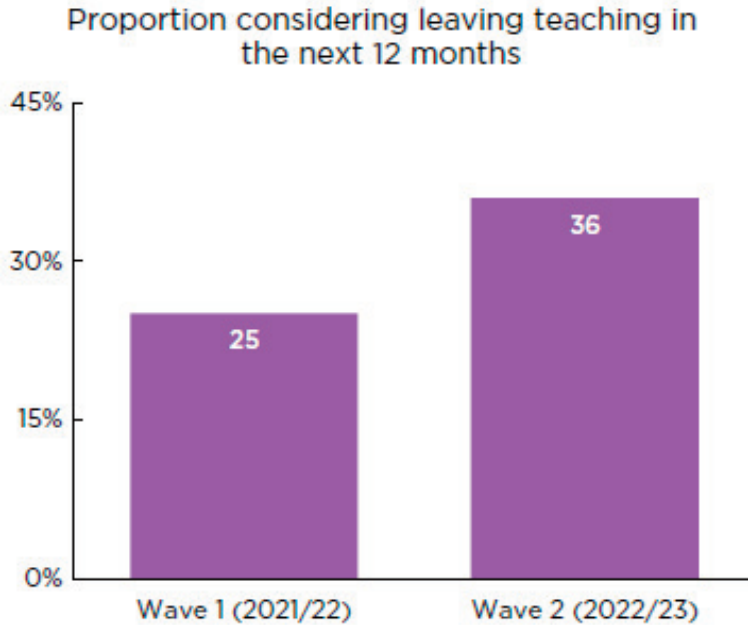


Source: DfE School Workforce Census (2010/11 – 2022/23)

rise in decades, the 2023/24 pay award has not substantially reversed the considerable pay deterioration experienced teachers have faced since 2010/11.

Starting salaries have fared relatively better because the Government has targeted higher pay rises at ECTs since 2017/18 (School Teachers Review Body, 2023). Pay for ECTs rose by 8.9 per cent in 2022/23 and 7.1 per cent in 2023/24 to bring starting salaries to £30,000 - a 2019 Conservative Party manifesto commitment. The starting salary increases mean they were only three per cent lower in real terms in 2023/24 compared to 2010/11.

However, over the same period, average earnings in the wider labour market have followed a stronger growth trajectory. Average earnings in the UK economy were three per cent higher in real terms in 2022/23 (the last full fiscal year of available data) than in 2010/11. This was about seven percentage points higher than for teacher starting salaries in 2022/23 and 15 percentage points higher than for experienced teachers. These gaps have grown from six and



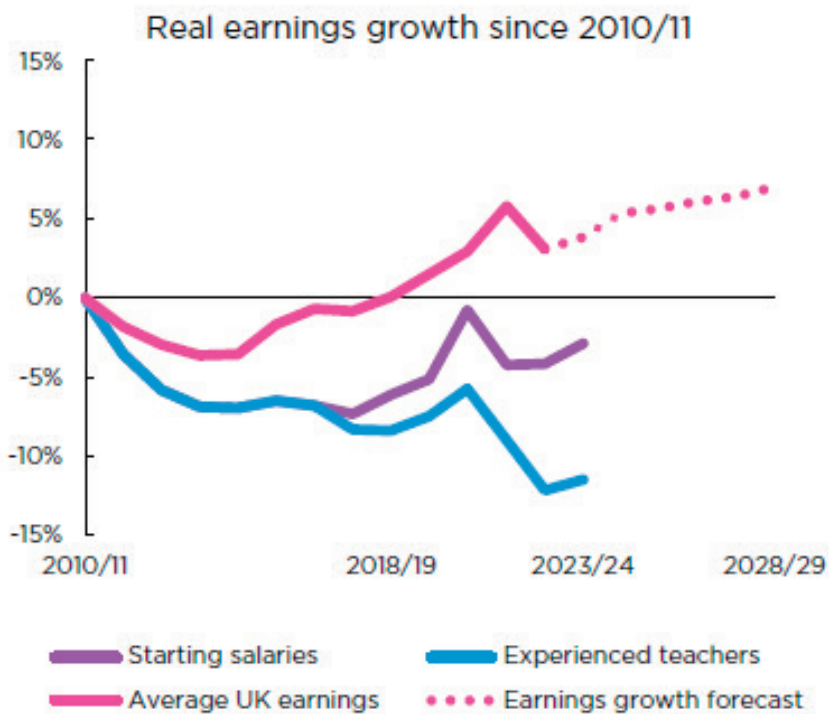
Source: Working Lives of Teachers and Leaders (waves 1 and 2)

eight percentage points, respectively, before the pandemic in 2018/19.

The 2023/24 pay rise was similar to the rate of average earnings growth, so is unlikely to have significantly narrowed these gaps. Moreover, in its March 2024 Economic and Fiscal Outlook, the Office for Budget Responsibility (OBR) forecasts that earnings in the wider labour market will continue to grow in real terms from 2024/25 to 2028/29 (Office for Budget Responsibility, 2024). Teacher pay will therefore need to match these increases just to maintain competitiveness relative to other jobs.

**Teachers’ position in the income distribution has fallen significantly since 2010/11**

Research clearly shows that the competitiveness of teacher pay relative to other occupations is linked to teacher recruitment and retention (Dolton and van der Klaauw, 1999). Comparing earnings growth for teachers to the wider labour market can, therefore, be useful to understand changes in



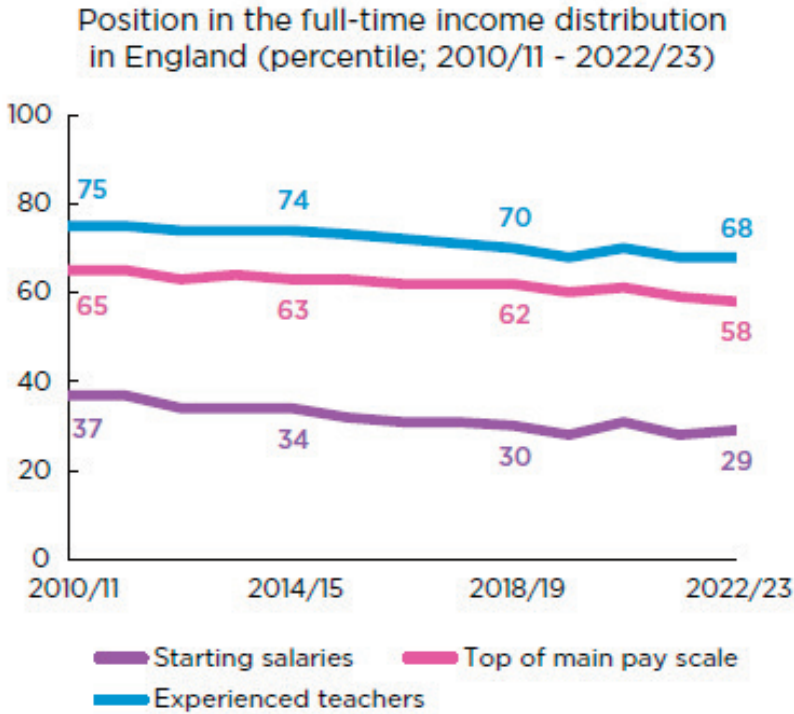
Notes: 'Experienced teachers' refers to teachers at the top of the upper pay scale. Dotted line represents the forecast of real earnings growth based on OBR projections from March 2024.

Source: School Teachers' Pay and Conditions Document, Office for National Statistics, Office for Budget Responsibility

competitiveness, since NFER research shows that teachers leave teaching for a wide variety of non-professional jobs (Worth and McLean, 2022).

However, strong average earnings growth in the UK economy has been driven in part by above-inflation increases to the minimum wage. After adjusting for inflation, the National Living Wage has increased by 14 per cent between April 2015 and February 2023 (Low Pay Commission, 2023). Therefore, comparing teacher pay growth to average earnings growth in the entire labour force may risk overstating the fall in competitiveness of teacher pay, relative to other jobs at a similar part of the income distribution.

Data from the Annual Survey of Hours and Earnings (ASHE) shows the position that classroom teacher pay has



Note: 'Experienced teachers' refers to teachers at the top of the upper pay scale.

Source: School Teachers' Pay and Conditions Document; NFER analysis of ASHE data (2010 - 2023)

held in the income distribution over time, which may be a better indication of how the competitiveness of teacher pay has changed. In 2010/11, teachers' starting salaries were at the 37th percentile of earnings for full-time workers in England (i.e. 37 per cent of other full-time workers in England earned less). The top of the main pay scale was at the 65th percentile, while experienced teachers were at the 75th percentile.

By 2022/23, teachers' position in the income distribution had deteriorated. Teachers on starting salaries had fallen to the 29th percentile while those at the top of the main pay scale and experienced teachers had fallen to the 58th and 68th percentiles, respectively. Most of this deterioration occurred prior to the pandemic, but even the

significant 2022/23 pay rise has not substantially improved this position. ASHE data is not yet available for 2023/24 to show how last year's pay award has affected this picture. Nonetheless, the data supports the conclusion that the competitiveness of teacher pay has deteriorated over time.

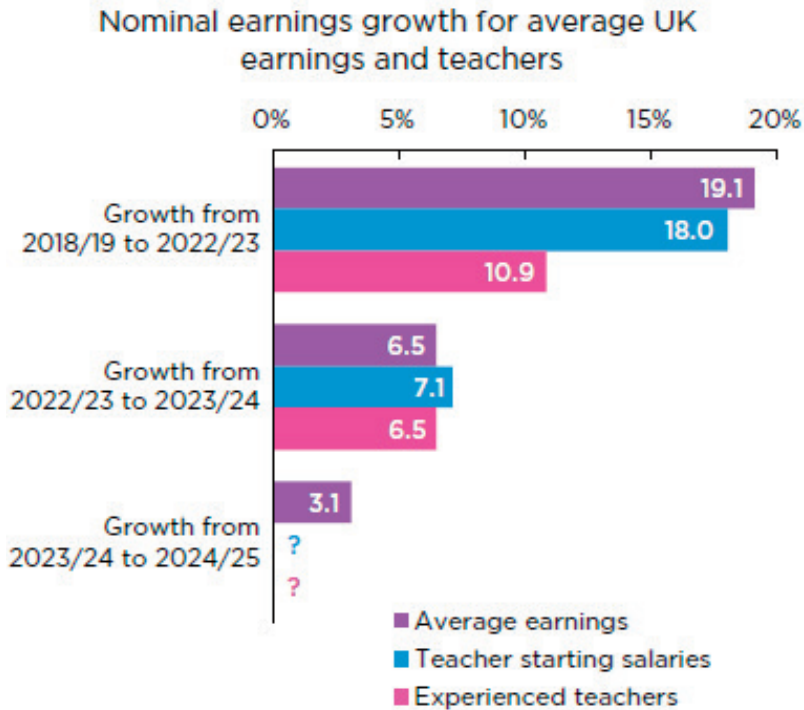
**Teacher pay needs to at least match the expected 3.1 per cent average earnings growth to maintain competitiveness**

Last year's teacher pay award, while the highest in decades, was similar to average earnings growth across the labour market in fiscal year 2023/24. It is therefore unlikely to have significantly contributed to narrowing the gap in earnings growth compared to the wider labour market, particularly for experienced teachers, which had further widened since before the pandemic.

In its 2024 Evidence to the School Teachers' Review Body (STRB), the DfE states that last year's pay award coupled with more stability in the macroeconomic context, 'support[s] the return of teacher pay awards to a more sustainable level' (DfE, 2024a). However, a pay award that at least keeps the competitiveness of teaching stable relative to the outside labour market is important for supporting both recruitment and retention.

The OBR forecasts that average earnings will rise by 3.1 per cent in 2024/25. Therefore, the teacher pay award needs to at least match the expected 3.1 per cent average earnings growth to maintain competitiveness. Exceeding average earnings growth is likely to improve competitiveness, but would require more resource to support it at a time when both school and public finances are severely constrained.

The pay award also needs to be affordable for schools. Research from the Institute for Fiscal Studies (IFS) suggests that a three per cent pay rise for teachers, alongside forecasted increases in other school costs, is likely to exceed schools' available funding for 2024/25 (Sibieta, 2024a, 2024b). Setting a competitive pay award without any additional funding for schools would likely increase the pressure on school budgets when more than half of schools are already in



Notes: 'Experienced teachers' refers to teachers at the top of the upper pay scale. Based on OBR earnings growth projections from March 2024.

Source: School Teachers' Pay and Conditions Document, Office for National Statistics, Office for Budget Responsibility

deficit and facing other cost pressures (Julius and Schwendel, 2024). Therefore, additional funding to implement a competitive pay award would need to come from Government.

**Teachers' working hours increased in 2023 and Government may face a challenge in meeting its workload reduction target**

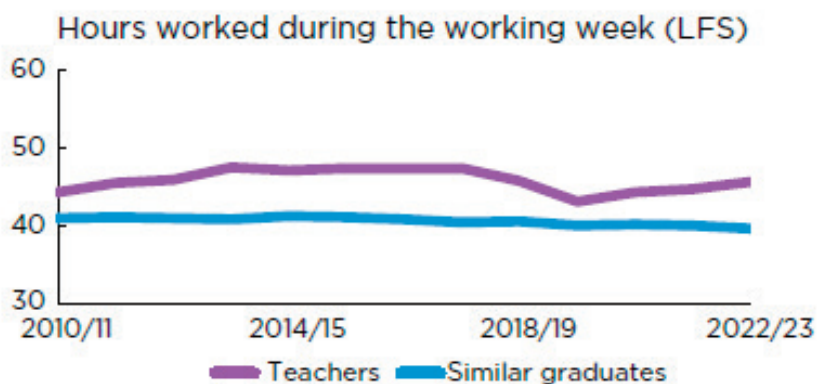
High workload has a strong negative impact on teacher retention (Lynch et al., 2016; Adams et al., 2023). Nearly all (94 per cent) of the teachers who were considering leaving cited high workload as a reason (DfE, 2024c).

Reducing excessive workload has been a policy priority

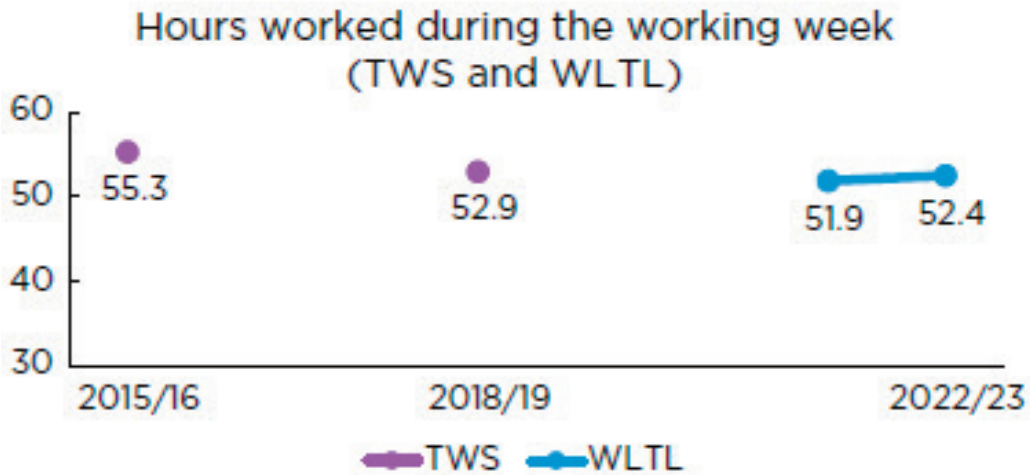
for Government since the 2016 Workload Challenge. The Teacher Workload Survey (TWS) and Labour Force Survey (LFS) data both showed that working hours for teachers fell between 2015/16 and 2018/19. The DfE has an ambition to reduce working hours by five hours per week within three years (DfE, 2024d) and has set up a workload reduction taskforce to provide recommendations on how to meet this target.

However, the latest LFS data shows that teachers' reported working hours increased last year, by one hour per week, to 45.7 hours per working week in 2022/23. Teachers work longer hours in a typical working week than similar graduates (i.e. graduates with similar demographic characteristics who work outside of teaching – see methodological appendix for full definition) and this gap has widened since the pandemic. Teachers work roughly similar hours when averaged across the year, due to having longer holiday periods.

WLTL data also showed that teachers' working hours increased by about half an hour in 2022/23 compared to the previous year. Teachers generally report higher working hours in the WLTL – and its predecessor, the TWS – than in the LFS. However, despite this, both data sources track a similar pattern over time, including the rise in working hours in 2022/23. The latest data shows that working hours are rising, just as the Government is aiming to reduce working hours, meaning it may face a considerable challenge in doing so.







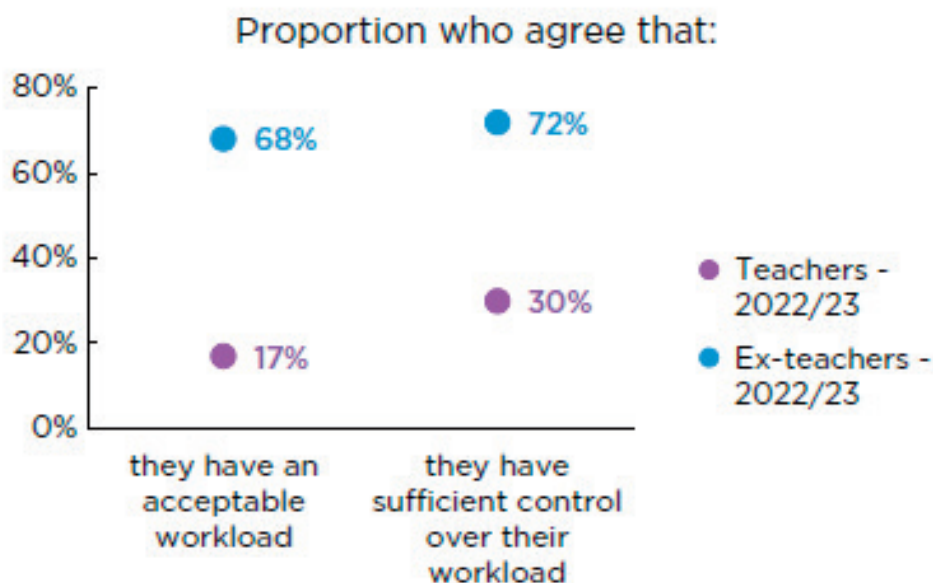
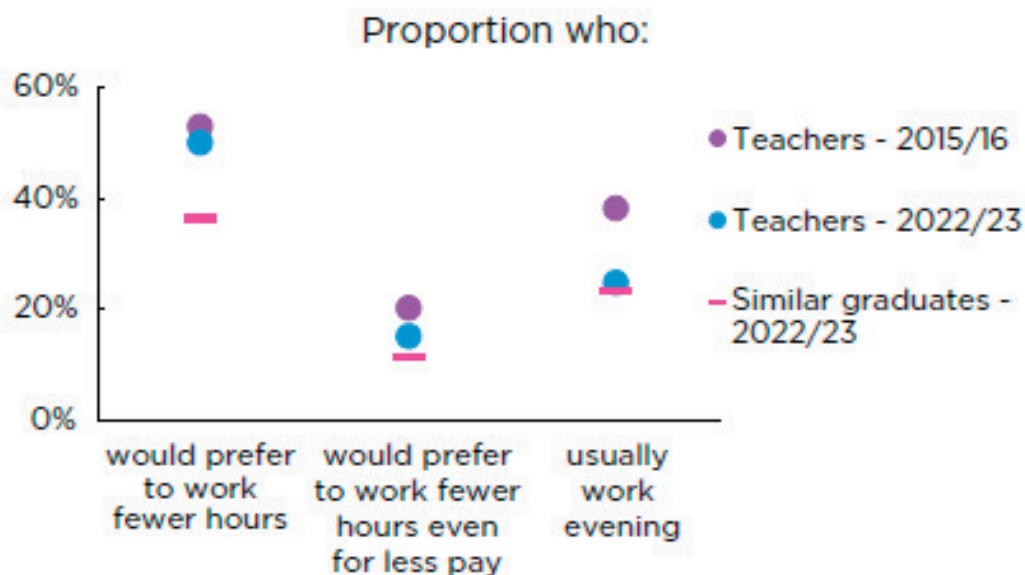
Note: Representative of full-time teachers and similar graduates who worked a full week during the survey reference week.

Source (top): NFER analysis of Labour Force Survey data (2010/11 – 2022/23).

Source (bottom): Teacher Workload Survey (2016, 2019); Working Lives of Teachers and Leaders (waves 1 and 2).

Despite some improvements since 2015/16, most teachers remain dissatisfied with their workload. Teachers in 2022/23 were more likely than similar graduates to feel that they work too many hours. LFS data shows that half of teachers in 2022/23 reported that they would prefer to work fewer hours, compared to 37 per cent of similar graduates. Similarly, 15 per cent of full-time teachers reported that they would prefer to work fewer hours even if it meant less pay, compared to 12 per cent of similar graduates.

However, the proportion of teachers who reported that they usually work evenings was not statistically significantly different from similar graduates working in other occupations. The fall in working hours prior to the pandemic has coincided with the proportion of teachers reporting they would prefer to work fewer hours, who would accept a pay cut to do so, and who usually work evenings, which have all fallen since 2015/16. However, this fall has not eliminated the gaps in preferences for fewer hours with similar graduates.



Source (top): NFER analysis of Labour Force Survey data (2015/16 and 2022/23).

Source (bottom): Working Lives of Teachers and Leaders (wave 2).

Teachers also have high dissatisfaction with their workload. WLTL data shows that only 17 per cent of teachers felt that they had an acceptable workload in 2022/23. Similarly, only 30 per cent felt they had sufficient control over their workload (IFF Research, 2024).

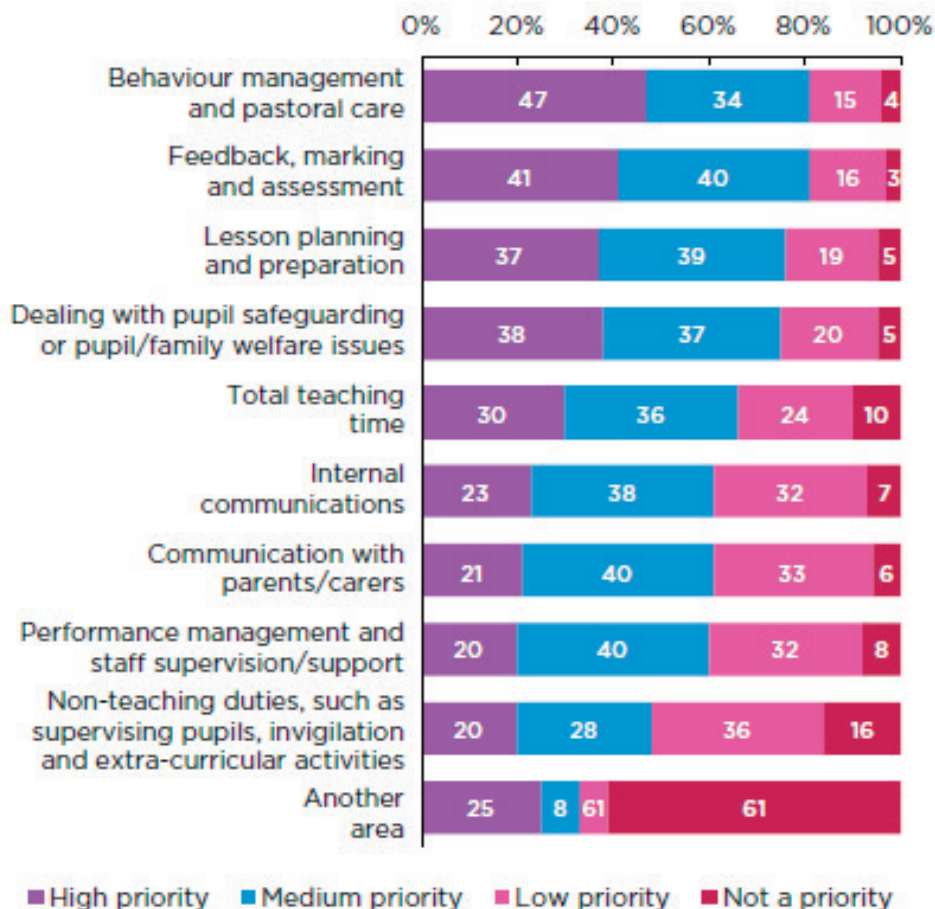
The WLTL also suggests that leaving teaching is associated with an improvement in workload satisfaction. Among ex-teachers surveyed in 2022/23 who had left teaching for another job the previous year, 68 per cent reported that they felt they had an acceptable workload in their new job, four times more than for teachers in service that year. Meanwhile, 72 per cent of ex-teachers felt they had sufficient control over their workload in their new job, 2.5 times higher than for teachers that year (IFF Research, 2024).

Schools have adopted strategies to address high workload, but behaviour management and pastoral care remains a key priority area for further workload reduction. NFER's recent practice review on managing teacher workload for the Education Endowment Foundation (EEF) shows that schools recognise the importance of reducing teacher workload (Martin et al., 2023). Among teachers and leaders surveyed, nearly all said that their school had implemented at least one workload reduction strategy between 2021/22 and 2022/23, while most schools had adopted more than one. Teachers were more likely to have positive views of their workload where their school had adopted multiple workload reduction strategies.

Much of the focus of reducing workload in recent years has been on managing teachers' administrative burden (CooperGibson Research, 2023). However, teachers and leaders surveyed in NFER's research also highlighted several areas which they felt were priority areas for further workload reduction. The area highlighted most frequently as being a high priority by respondents was behaviour management and pastoral care.

Teachers say that pupil behaviour is driving higher workload. WLTL data shows that in 2022/23, 57 per cent of teachers said they spent too much time on behaviour incident

### Approaches identified as priorities to reduce workload in schools



Source: Martin et al. (2023)

follow-up, versus 50 per cent the previous year (DfE, 2024c). This is in line with evidence showing that behaviour management has become more of a challenge since the pandemic. WTL data shows that only 49 per cent of teachers rated pupil behaviour in their school as ‘good’ or ‘very good’ in 2022/23, down from 58 per cent the previous year (DfE, 2024c).

Overall, despite the steady progress made at reducing

workload in the areas of planning, marking and administration between 2015/16 and 2021/22, more progress is needed. Schools' workload reduction efforts need to be sustained, but attention should turn to examining external workload drivers. For example, 63 per cent of teachers cited more support from outside agencies for specific pupil needs such as SEND support, mental health and safeguarding as a key enabler of further workload reduction (Martin et al., 2023).

**Flexible working arrangements for teachers became slightly more common in the last year**

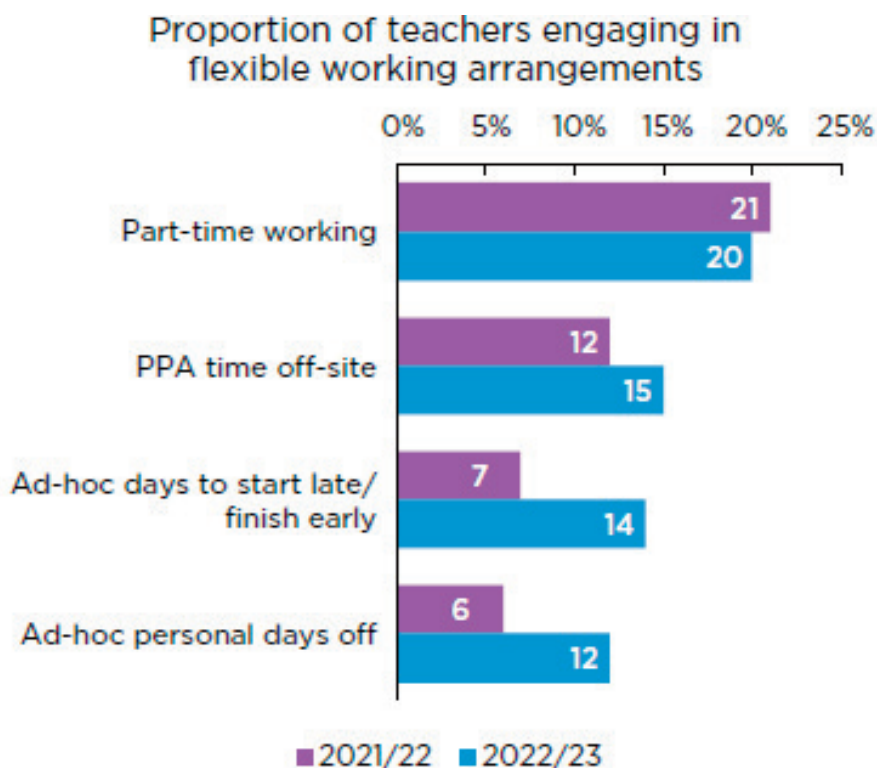
NFER's recent evidence review on flexible working for the EEF shows that there is considerable demand among teachers for flexible working (ad-hoc and regular working schedule adaptations as well as parttime working), which outstrips the availability of such arrangements in schools (Harland, Bradley and Worth, 2023).

WLTL data shows that flexible working has become more common in schools. While the proportion of teachers working part time was slightly lower in 2022/23 than the previous year, teachers who reported that they have ad-hoc days off, or days starting late or finishing early doubled (DfE, 2024c). Similarly, the proportion of teachers who reported that they had an arrangement to complete their planning, preparation and assessment (PPA) time off-site increased from 12 to 15 per cent.

There is currently no causal evidence on how introducing flexible working arrangements for teachers may impact on retention rates. However, research shows that teachers value the availability of flexible working – the possibility of moving from full-time to part-time working if requested is valued the same as a 4.3 per cent pay rise (Burge, Lu and Phillips, 2021). This suggests that improvements in the availability of flexible working may lead to higher retention rates.

Research also suggests that introducing flexible working arrangements tends to have other positive impacts as well, such as increased wellbeing, job satisfaction, attendance,

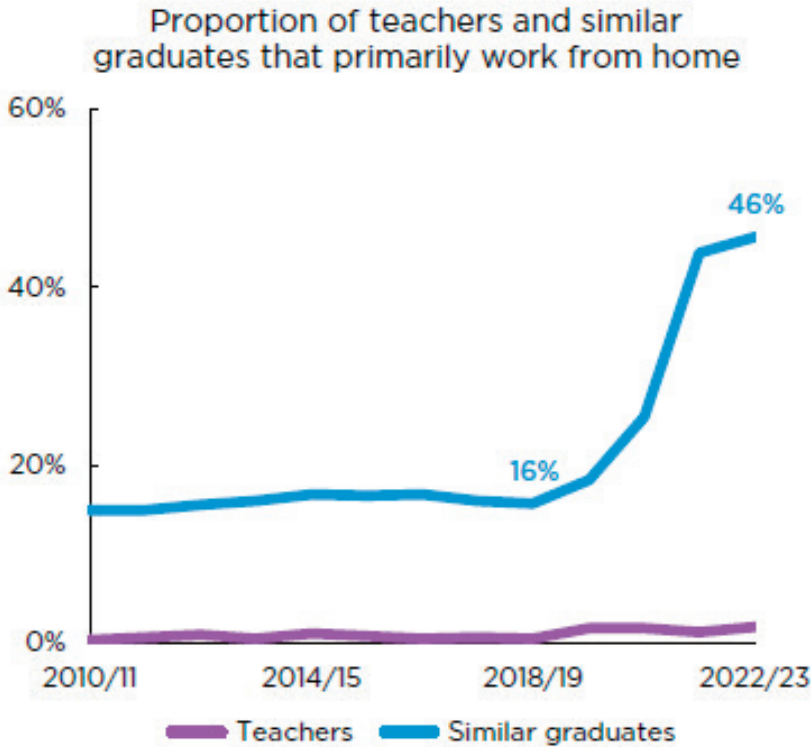
motivation, teaching capacity, expertise and career progression. While flexible working arrangements can involve costs on schools and pupils, surveys of teachers and leaders suggest that these benefits can often outweigh the costs (Harland, Bradley and Worth, 2023).



Source: Working Lives of Teachers and Leaders (waves 1 and 2)

**Remote and hybrid working remains substantially more prevalent in the wider graduate labour market than in teaching**

Like other frontline public sector jobs, remote working is largely unavailable to most teachers. WTL data shows that in 2021/22, two per cent of primary classroom teachers and one per cent of secondary teachers reported that they had a formally-agreed arrangement to work remotely (Adams et al., 2023)



Source: NFER analysis of Labour Force Survey data (2010/11 – 2022/23)

However, LFS data shows that remote working remains popular with employees in England, having increased during and since the pandemic. In 2022/23, 46 per cent of similar graduates reported that they worked mainly from home. This was about three times higher than prepandemic levels and slightly higher than in 2021/22.

The proportion of graduates who work either fully remote or in a hybrid arrangement (i.e. work at least one day from home but not necessarily the whole week) is even higher. According to data from the G-SWA, 65 per cent of similar graduates in the UK working outside of education worked from home at least one day during the survey reference week in 2022. This was highest among those who have children and who have a commute of one hour or more. Surveys show that the prevalence of remote working is likely



to remain high in the labour market in the future (Shah et al., 2024). The G-SWA also shows that more than a third (36 per cent) of graduates in the UK would look for a new job offering the ability to work remotely one to two days per week if their employer mandated a full return to the office.

Most teachers would also highly value the additional flexibility that remote working can bring – more than half of teachers surveyed in a TeacherTapp poll thought they should be allowed a regular day to work from home (Ford, 2022). The sustained prevalence of remote working in the labour force since the pandemic is therefore likely to continue to act as an incentive to leave teaching for teachers unable to gain additional flexibility in their teaching job.

**Teacher pay would have to rise by about 1.8 per cent to compensate for the lack of availability of remote and hybrid working**

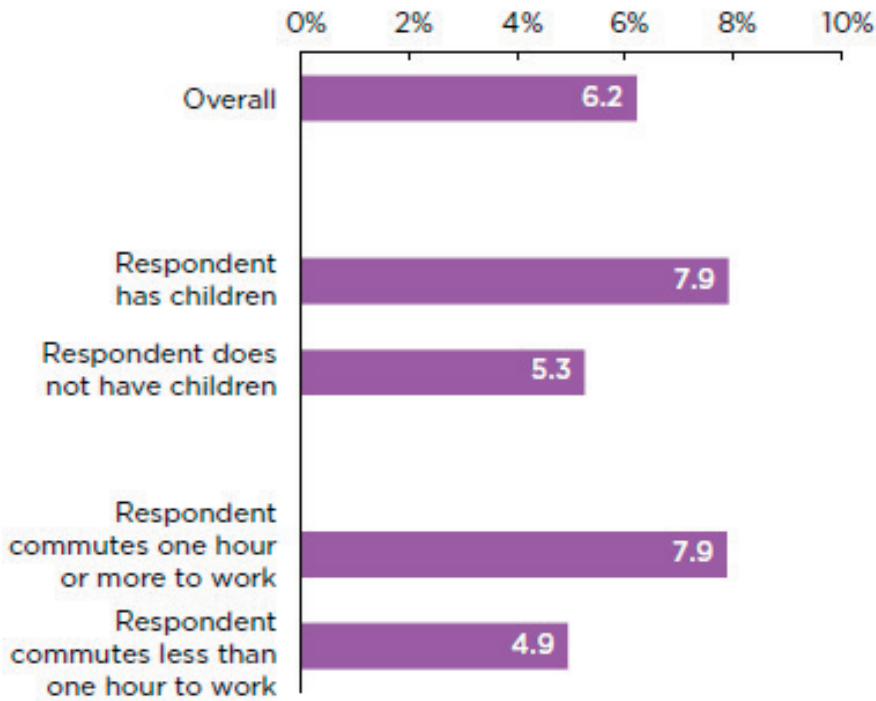
The gap in availability of remote and hybrid working between teaching and the wider graduate labour market needs to be compensated for to maintain the level of competitiveness teaching had before the pandemic (all else being equal). Matching the level of hybrid working is not feasible in a frontline profession such as teaching, but increasing pay could help to compensate.

This is already happening within the wider labour market. Surveys of UK employers show that earnings growth has tended to be higher among employers that do not make great use of remote working than among those that do (Shah et al., 2024). This is consistent with evidence from the United States, which showed that, between 2021 and 2023, employers which did not extend access to remote working to their employees tended to have earnings growth which was two percentage points higher than among employers which did expand access to remote working (Barrero et al., 2022).

Remote working is highly valued by employees in the UK. G-SWA data shows that among similar graduates working in non-education occupations, the option to work from home two or three days per week had the equivalent value as a 6.2



Value (in % of salary) of the option to work remotely 2-3 days per week



Note: Represents the value similar graduates place on the option to work remotely 2-3 days per week. See the methodology appendix for a detailed definition of similar graduates.

Source: NFER analysis of Global Survey of Working Arrangements data (waves 1 and 2).

per cent salary increase. This was highest among those with children, and those who have a commute of an hour or more. We estimate that a consolidated pay increase of 1.8 per cent for teachers could be appropriate compensation to maintain competitiveness. This is based on the average value of remote working of 6.2 per cent, multiplied by the 29 percentage point faster growth since the pandemic in the prevalence of remote working in the graduate labour market compared to in teaching (see the methodology appendix for more detail). Crucially, this compensatory pay rise would need to improve the competitiveness of teacher pay growth relative to other jobs to have the intended effect. Therefore, it should be

awarded on top of pay rises that match earnings growth in the wider economy.

### **Conclusions**

Teacher supply in England is in a critical state that represents a substantial risk to the quality of education. Since the pandemic, secondary teacher recruitment has been far below what is needed to meet schools' needs for new teachers, and this is unlikely to be any different this year. Applications for the 2024/25 ITT cycle so far suggest recruitment for 10 out of 17 secondary subjects is at risk of missing targets.

Policy measures to address teacher undersupply, by improving recruitment and / or retention, is urgently needed. Given the scale of the challenges, policymakers need to consider actions that are ambitious and radical, while also cost effective.

There have been pockets of positive progress in the last year. Actions to attract international trainees appear to have had some success, although the long-term benefits in terms of supply are uncertain. Slight improvements to bursary and retention payment generosity are also likely to boost recruitment and retention, given the evidence demonstrating their effectiveness. There has also been a promising increase in flexible working opportunities that are not part-time working, which are likely to improve teacher well-being and job satisfaction, while not reducing fulltime equivalent teaching hours.

However, on key factors associated with recruitment and retention there has been little or no progress. Workload is cited by 94 per cent of teachers as a reason for considering leaving teaching. Despite some progress in reducing working hours before the pandemic, two different data sources suggest that teachers' working hours significantly increased in 2022/23, meaning that the Government may face a challenge in meeting its workload reduction target.

NFER research found that schools are aware of the need to manage teacher workload and are putting strategies in place. However, teachers now say that behaviour

management and pastoral care are key priority areas for workload reduction and more support from outside agencies for specific pupil needs such as SEND support, mental health and safeguarding is a key enabler of further workload reduction. This is in line with evidence showing a recent trend of worsening pupil behaviour since the pandemic.

The 2016 independent workload review groups recommended that much of the focus on teacher workload reduction should be on reducing the burden of planning, administration and marking. Teacher workload reduced during the period 2016-2019, particularly in the areas identified as priorities.

These workload reduction efforts should be sustained, but the focus should now also include behaviour management and pastoral care. Government should also consider how outside agencies for supporting specific pupils needs (such as SEND support, mental health and safeguarding) can help to shift some of the pupil behaviour burden away from teachers.

The competitiveness of teacher pay matters for both recruitment and retention. However, a series of below-inflation pay rises, the 2021 teacher pay freeze and the cost-of-living crisis have led teacher pay to fall significantly in real terms over the last decade. Meanwhile, higher earnings growth in the wider labour market outside teaching has led to teacher pay losing competitiveness against the wider labour market over time.

**Recommendation 1:** Government should set up an independent review focussing on how to reduce teachers' workload related to behaviour management and pastoral care, which should consider the role of external support services, such as for special needs and mental health.

Last year's 6.5 per cent pay rise has stalled, but not substantively reversed, any of the deterioration in teachers' pay competitiveness since the pandemic. However, the 2024/25 pay award provides an opportunity to improve the competitiveness of teacher pay. The OBR forecasts that

earnings in the wider labour market will rise by 3.1 per cent next year. In the short term, this year's pay award will therefore need to exceed 3.1 per cent to help improve the competitiveness of teacher pay next year.

More broadly, a clear long-term strategy is needed to help ensure that future teacher pay awards continue to improve pay competitiveness over the long term. This is likely to be a considerable challenge given the severe funding pressures facing all public services, including education.

Schools are unlikely to be able to afford any such pay awards for teachers on their own without making cuts elsewhere. Funding for competitive pay rises for teachers must therefore come from Government. In the run-up to the next general election, political parties should develop their plans for how they will effectively address this challenge in the next parliament.

Access to flexible working arrangements is an area where teaching lags other occupations. A degree of inflexibility is inevitable in teaching, but research shows that since the pandemic many schools have begun making changes to their flexible working policies. Part-time working is the most common flexible working arrangement available to teachers but ad-hoc days off or days to start late / finish early are also becoming more common.

Little robust evidence exists showing the causal relationship between teachers' access to flexible working arrangements and retention, but there is plenty of perceptual evidence of its importance. Just like other graduates, teachers highly value the ability to work flexibly, so improvements in the availability of flexible working arrangements may help improve retention, along with overall teacher well-being.

Nonetheless, it is important to be realistic about the limits of flexibility in teaching. Like police officers, doctors and other NHS workers, access to hybrid and remote working, is unlikely to ever match other jobs in the graduate workforce, where nearly two-thirds of employees work hybrid.

Providing compensation to frontline public service workers such as teachers is one option for preventing this

inherent inflexibility from undermining the attractiveness of their jobs. Based on the increase in the prevalence of remote working in the graduate workforce and the value employees place on the ability to work remotely, we estimate that this pay premium for teachers should be 1.8 per cent. Crucially, this would need to be awarded on top of the pay rise needed to match teachers' earnings growth with the wider economy.

**Recommendation 2:** Narrowing the gap between teacher pay growth and the wider labour market is key to supporting recruitment and retention. The 2024 pay award should therefore exceed the 3.1 per cent forecasted rise in earnings in the wider labour market and be fully-funded.

**Recommendation 3:** Political parties should set out their plans to develop a long-term strategy for pay setting which reduces the gap in earnings growth with competing occupations, while ensuring that schools have sufficient funding to enact these pay increases without making cuts elsewhere.

**Recommendation 4:** Political parties should consider introducing a Frontline Workers Pay Premium to compensate public sector workers for the lack of remote and hybrid working opportunities in their jobs compared to the wider graduate labour market. We estimate that this would represent a 1.8 per cent consolidated pay increase for teachers.

### References and data sources

Adams, L., Coburn-Crane, S., Sanders-Earley, A., Keeble, R., Harris, H., Taylor, J. and Taylor, B. (2023) *Working lives of teachers and leaders - wave 1*. Available at: <https://www.gov.uk/government/publications/working-lives-of-teachersand-leaders-wave-1> (Accessed: 13 March 2024).

Aksoy, C.G., Barrero, J.M., Bloom, N., Davis, S.J., Dolls,

M. and Zarate, P. (2022), *Working from home around the world*. Available at: <https://wfhresearch.com/wp-content/uploads/2022/10/Working-from-Home-Around-the-World-16-October-2022.pdf> (Accessed: 13 March 2024).

Barrero, J.M., Bloom, N., Davis, S.J., Meyer, B. and Mihaylov, E. (2022) *The shift to remote work lessens wage-growth pressures*. Working Paper 30197. Available at: <http://www.nber.org/papers/w30197> (Accessed: 13 March 2024).

Burge, P., Lu, H. and Phillips, W.D. (2021) *Understanding teaching retention: using a discrete choice experiment to measure teacher retention in England*. Available at: [https://www.rand.org/pubs/research\\_reports/RRA181-1.html](https://www.rand.org/pubs/research_reports/RRA181-1.html) (Accessed: 13 March 2024).

CooperGibson Research (2023) *Exploring teachers' admin time*. Available at: <https://www.gov.uk/government/publications/exploring-teachers-admin-time> (Accessed: 13 March 2024).

Department for Education (2023a) *Initial teacher training census, academic year 2023/24*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/initial-teacher-training-census/2023-24> (Accessed: 13 March 2024).

Department for Education (2023b) *Postgraduate initial teacher training targets, Academic year 2023/24*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/postgraduate-initial-teacher-training-targets> (Accessed: 13 March 2024).

Department for Education (2023c) *School workforce in England, reporting year 2022*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england> (Accessed: 13 March 2024).

Department for Education (2024a) *Government evidence to the STRB*. Available at: [https://assets.publishing.service.gov.uk/media/65e0a4cb2f2b3b00117cd7ae/Government\\_evidence\\_to\\_the\\_STRB.pdf](https://assets.publishing.service.gov.uk/media/65e0a4cb2f2b3b00117cd7ae/Government_evidence_to_the_STRB.pdf) (Accessed: 13 March 2024).

Department for Education (2024b) *Postgraduate initial teacher training targets*, Academic year 2024/25. Available at: <https://explore-educationstatistics.service.gov.uk/find-statistics/postgraduate-initial-teachertraining-targets> (Accessed: 14 March 2024).

Department for Education (2024c) *Working lives of teachers and leaders: wave 2 summary report*. Available at: <https://www.gov.uk/government/publications/working-lives-of-teachers-and-leaderswave-2/working-lives-of-teachers-and-leaders-wave-2-summary-report> (Accessed: 13 March 2024).

Department for Education (2024d) *Workload reduction taskforce - initial recommendations*. Available at: [https://assets.publishing.service.gov.uk/media/65a10648e8f5ec000d1f8c2f/Workload\\_reduction\\_taskforce\\_-\\_initial\\_recommendations.pdf](https://assets.publishing.service.gov.uk/media/65a10648e8f5ec000d1f8c2f/Workload_reduction_taskforce_-_initial_recommendations.pdf) (Accessed: 13 March 2024).

Dolton, P. and van der Klaauw, W. (1999) "The turnover of teachers: A competing risks explanation", *The Review of Economics and Statistics*, 81(3), pp. 543–550. Available at: <https://doi.org/10.1162/003465399558292>.

Education Endowment Foundation (2022) *Early career framework early roll-out*, EEF. Available at: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/early-career-framework-early-rollout> (Accessed: 13 March 2024).

Ford, I. (2022) *Has the use of VAK (learning styles)*

*changed in classrooms since 2018? (This, and other findings...)*, Teacher Tapp, 8 February. Available at: <https://teachertapp.co.uk/articles/has-the-use-ofvak-learning-styles-changed-in-classrooms-since-2018-this-and-otherfindings/> (Accessed: 13 March 2024).

Harland, J., Bradley, E. and Worth, J. (2023) *Understanding the factors that support the recruitment and retention of teachers. Review of flexible working approaches*. Available at: <https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Review-of-flexibleworking-approaches.pdf?v=1701755839> (Accessed: 13 March 2024).

Julius, J. and Schwendel, G. (2024) *What does the latest data tell us about the current financial situation in England's local authority maintained schools?*, NFER Blogs, 19 February. Available at: <https://www.nfer.ac.uk/blogs/what-does-the-latest-data-tell-us-about-the-current-financial-situation-in-england-s-local-authority-maintainedschools/> (Accessed: 13 March 2024).

Low Pay Commission (2023) *The national minimum wage in 2023*. Available at: [https://assets.publishing.service.gov.uk/media/64259f9460a35e00120cb25d/The\\_National\\_Minimum\\_Wage\\_in\\_2023.pdf](https://assets.publishing.service.gov.uk/media/64259f9460a35e00120cb25d/The_National_Minimum_Wage_in_2023.pdf) (Accessed: 13 March 2024).

Lynch, S., Worth, J., Bamford, S. and Wespieser, K. (2016) *Engaging teachers: NFER analysis of teacher retention*. Available at: <https://www.nfer.ac.uk/media/52vm1q5k/lfsb01.pdf> (Accessed: 13 March 2024).

Martin, K., Classick, R., Sharp, C. and Faulkner-Ellis, H. (2023) *Supporting the recruitment and retention of teachers in schools with high proportions of disadvantaged pupils: understanding current practice around managing teacher*



*workload*. NFER. Available at: <https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Reviewof-teacher-workload-management-approaches.pdf?v=1701856418> (Accessed: 13 March 2024).

McLean, D., Tang, S. and Worth, J. (2023) *The impact of training bursaries on teacher recruitment and retention: an evaluation of impact and value for money*. Available at: [https://www.nfer.ac.uk/media/bycg5uzk/the\\_impact\\_of\\_training\\_bursaries\\_on\\_teacher\\_recruitment\\_and\\_retention.pdf](https://www.nfer.ac.uk/media/bycg5uzk/the_impact_of_training_bursaries_on_teacher_recruitment_and_retention.pdf) (Accessed: 13 March 2024).

National Audit Office, A. (2016) *Training new teachers*. Available at: <https://www.nao.org.uk/wp-content/uploads/2016/02/Training-newteachers.pdf> (Accessed: 13 March 2024).

Office for Budget Responsibility (2024) *Economic and Fiscal Outlook March 2024*. Available at: <https://obr.uk/economic-and-fiscal-outlooks/> (Accessed: 13 March 2024).

Office for National Statistics (2023a) *Annual survey of hours and earnings (ASHE)*. Available at: <https://www.ons.gov.uk/surveys/informationforbusinesses/businesssurveys/annualsurveyofhoursandearningsashe> (Accessed: 13 March 2024).

Office for National Statistics (2023b) *Labour force survey – user guidance*. Available at: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyuserguidance> (Accessed: 13 March 2024).

School Teachers Review Body (2023) *School teachers' review body thirty-third report*. Available at: [https://assets.publishing.service.gov.uk/media/64aff19dfe36e000146fa8e6/STRB\\_33rd\\_Report\\_2023\\_](https://assets.publishing.service.gov.uk/media/64aff19dfe36e000146fa8e6/STRB_33rd_Report_2023_)

Web\_Accessible\_v02\_\_1\_.pdf (Accessed: 13 March 2024).

Shah, K., Bloom, N., Bunn, P., Mizen, P., Thwaites, G. and Yotzov, I. (2024) *Managers say working from home is here to stay*. Available at: <https://cepr.org/voxeu/columns/managers-say-working-home-here-stay> (Accessed: 13 March 2024).

Sibieta, L. (2024a) *The latest picture on school funding and costs in England*. Available at: <https://ifs.org.uk/articles/latest-picture-schoolfunding-and-costs-england> (Accessed: 13 March 2024).

Sibieta, L. (2024b) X, 1 March. Available at: <https://twitter.com/lukesibieta/status/1763509814242476038> (Accessed: 13 March 2024).

Worth, J. and Hollis, E. (2021) *Do bursaries change who applies to teacher training?* NFER Blog, 24 November. Available at: <https://www.nfer.ac.uk/blogs/do-bursaries-change-who-applies-to-teacher-training/> (Accessed: 13 March 2024).

Worth, J. and McLean, D. (2022) *What teachers do next after leaving and the implications for pay-setting*. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1045252/What\\_teachers\\_do\\_next\\_after\\_leaving\\_and\\_the\\_implications\\_for\\_pay\\_setting.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1045252/What_teachers_do_next_after_leaving_and_the_implications_for_pay_setting.pdf) (Accessed: 13 March 2024).

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

# PISA 2022: Measuring the world's education systems after COVID

Demitri Coryton

By Demitri Coryton

**Key words:** OECD, PISA, maths, science, reading, COVID.

**Abstract:** Last December the OECD published PISA 2022, the results of the usually triennial research programme that tests over half a million 15-year-olds in how they can apply their knowledge in science, maths and reading. Each round of PISA focuses on one of these curriculum areas, and in 2022 it was maths.

In most countries, and therefore the OECD PISA average, learning was negatively impacted by COVID and scores went down when compared to the last round of PISA in 2018. But that was not the case in all countries. Some improved their scores, sometimes by a lot. Policy makers need to ask themselves what are those countries doing that they are not. Factors other than COVID also had an impact. These included resilience of students, student hunger, and staff shortages.

**L**ast December saw the publication of the first two volumes of the Programme for International Student Assessment 2022 - PISA - organised by the Organisation for Economic Cooperation and Development (OECD). PISA tested a sample of 690,000 15 year-olds from 81 countries and territories on how they could apply knowledge in maths, reading and science, with a focus on maths. The tests were taken in 2022. PISA is the main secondary school international comparative research programme in the world. (\* In the text that follows some countries did not fully meet OECD PISA data standards. They are shown with an asterisk by their name.)

**Demitri Coryton** taught politics and 20th century history at the state University of Washington, the University of Wisconsin and Richmond University, London, before founding the Education Publishing Company in 1996. He was chairman of the Conservative National Advisory Committee on Education from 1979 to 1982. He is currently editor of the weekly *Education Journal* and of the termly *Education Journal Review*.

The results for the UK were mixed. Britain rose in the international rankings but only because its decline was less than that of other countries. Of far greater relevance is the actual score achieved in the three curriculum areas, which is actually the mid-point in a range within which the actual result is likely to be. Here the UK's score had declined in all three areas tested since the last PISA in 2018.

The UK was still one of the few countries whose score was ahead of the OECD average in all three areas tested. Yet this was because, within the UK, England scored well. It was above the OECD average in all three areas. Scotland and Northern Ireland were above average in reading, while Wales was not above average in any category.

Professor Andreas Schleicher, Director of the Education and Skills Directorate of the OECD, described the UK's performance as "so, so". Among the areas where the UK was among the best performers were in student resilience, social equity and the performance of immigrants. It did less well in pupil concerns of violence, levels of pupil hunger and head teacher concerns about staffing.

Speaking at the UK launch of the report, Schools Minister Damian Hinds MP (a former Secretary of State for Education) blamed COVID for the fall in the UK's scores, which many other countries had also experienced. The OECD average scores had also declined since 2018. While COVID had obviously had a negative impact, some countries had still been able to improve their scores, in some cases by a significant amount. UK education ministers must ask why the UK did not.

### **UK scores**

The UK did score well in some areas. These included social equity. In Britain poverty does not have to be a predictor of educational failure. Those from a poor background can and do make it into the best schools and universities, something that COVID has not impacted despite those from poor areas as a whole finding it more difficult to catch up. Britain's success at educating the children of immigrants, who performed at least

as well as indigenous children when social class was factored in, compares favourably with many other European countries who have struggled in this area.

Yet there are some areas where the UK did not do so well as its competitors. Perhaps most concerning is the number of children who skip meals because their families can't afford food. On average across OECD countries, 8% of students reported not eating at least once a week in the past 30 days due to lack of money to buy food. Some OECD countries have far lower numbers of hungry children. Notably Portugal, Finland and the Netherlands were all below 3%. In the United Kingdom 11% of students said they were forced to skip meals. This will have worsened since the tests were taken last year. The simple truth, as Dame Rachel de Souza, the Children's Commissioner for England, said at the UK launch of PISA 2022 at the think tank Policy Exchange, is that hungry children don't learn. She spoke from personal experience, for as a young child of immigrant parents growing up in poverty she had experienced hunger and gone without meals herself.

This really is something that should not be happening in such a wealthy country as Britain. The British government has known it is a problem for some while. If a relatively poor country like Portugal can keep child hunger down to 3%, why does it have to be four times as high in Britain? It does not have to be like this. In policy terms, there are alternatives. Ministers need to get a grip of this. Changing this, apart from being the right thing to do, is one of the ways that Britain can do better in the next round of PISA in 2025.

Another reform that the OECD recommends for countries that want to improve their performance is avoiding selection. PISA reports for 20 years have shown that selection, especially at a young age as exists in parts of England, is harmful to the majority of children. England still has selection in some 20% of the country, with a test called the 11+ taken by most children at about the age of 10. The 11+ is heavily influenced by private tutoring, which favours richer families who can afford it, and in any case does not measure academic ability. With children at this age, it measures social class.

Avoiding selection is one of the nine lessons from the COVID pandemic that the OECD recommends that countries learn if they wish to improve their education systems.

Another lesson from the pandemic is that schools matter. Distance learning was necessary when schools were shut because of COVID, but what PISA found is that young people need and value physical schools. They, and the friendship groups within them, are an important part of the social life of young people as well as being necessary as support for learning. If there is another pandemic, ministers must try and keep schools open if they possibly can.

If schools matter, so do teachers. As Damian Hinds acknowledged at the PISA launch, Britain's good performance in PISA did not happen in the Department for Education. It happened in schools as a result of the effort of teachers. And staffing problems were a huge concern for heads.

If this round of PISA was a wake-up call for ministers in England to do better, that is even more true in Scotland and Wales. After 15 years in power, the SNP government in Scotland presides over a system that is no longer the best in the UK, as it once was. Labour in Wales presides over an even worse performance. PISA points the way to how we can improve. All ministers have to do is learn and apply the lessons.

### **What is PISA?**

What PISA does is to test the knowledge and skills of students directly, using an internationally agreed metric to collect data from students, teachers, schools and systems to understand performance differences. Unlike traditional assessments, PISA sought to assess not just students' ability to reproduce learned material but also their capacity to apply knowledge creatively in novel scenarios, think critically across disciplines, and demonstrate effective learning strategies. By emphasising these skills, PISA aimed to equip students with the ability to navigate an ever-evolving world.

Some critics argued that PISA tests were unfair as they might present students with unfamiliar problems. But then,

life is full of unforeseen challenges. In the real world, people must solve problems that they have not anticipated; it is not just about remembering lessons in a classroom.

The greatest strength of PISA lies in its working methods. Most assessments are centrally planned and then contracted to engineers who build them. PISA turned that on its head. The idea of PISA attracted the world's leading minds and mobilised hundreds of experts, educators and scientists from the participating countries to build an assessment that is valid across countries, cultures and languages. This collaborative effort engendered a sense of ownership that was critical to its success.

From the tests a score is calculated for each country. The score is in a range for each country's performance, so strictly speaking you can't rank countries in order as if it was the pop charts. Yet, inevitably, that is what happens.

Usually the PISA rounds are every three years, but because of COVID an extra year was added for the latest round. PISA tests in three regular subjects, reading, maths and science. In each round PISA focuses on one of these subjects and in 2022 that was maths.

### **The results**

Like the UK, many countries saw a decline in their PISA scores compared to 2018. Part of this was down to the COVID pandemic, but not all of it. Other countries, that also faced COVID, still managed to improve their scores, in some cases quite dramatically. In some countries the decline started well before COVID. For example, in Finland, which once topped the PISA charts and which was the poster boy for educational advance, they have fallen by 60 points over the last decade. To give some context, falling or rising by 20 points is the equivalent of going up or down a year group.

We give on the following pages tables for the results for all countries in each of the three subject areas. From this it can be seen that the UK is in the group above the OECD average in all three. Within the UK, things are rather different. England is in the above average group in all three subject

**Table 1. PISA score performance of home countries within the UK**

Country	Maths score	Reading score	Science score
England	492	496	503
Scotland	471	493	483
Wales	466	466	473
Northern Ireland	475	485	488
United Kingdom	489	494	500

Source: OECD PISA 2022, Volume 1, *The State of Learning and Equity in Education*, Annex B2. Results for regions within countries pages 358 to 362.

areas, but Scotland, Wales and Northern Ireland are not. England's score is higher than that of the other parts of the UK. When Professor Schleicher was asked by the magazine *Education Journal* why this was so he said he did not know. The OECD had produced figures for sub-national groups but it had not yet had time to analyse the figures and so he could not say.

Scotland once had the best education system in the UK. After 15 years of SNP controlled devolved government it now definitely does not. The Scottish government tried to put its best spin on this, but the reality is that Scotland is performing more poorly than England. Wales is performing worse than England or Scotland. The Labour government in Wales has made a point of doing some things differently than in England, but the result has been a poorer performance than that achieved across Offa's dyke.

As you will see from the tables that follow below, in maths the highest scoring country was Singapore, with a score of 575. This was up six points on 2018. East Asian countries did well, with Singapore followed by China Macao, Taiwan, China Hong Kong, Japan and South Korea. Estonia was the



highest scoring Western and European nation. The UK was listed 12th equal, tied with Belgium and Denmark, with a score of 489, down 13 from 2018. This means that Britain is not in the top scoring group, but is in the second best batch. It has lost ground since before the COVID pandemic, as many but not all other countries have done.

When it comes to reading, Singapore is top here as well. Yet the East Asian domination at the top of the list is broken by the Irish Republic, which comes second with a score of 512. The UK is listed 13th with a score of 494, which is down 10 points on 2018. It would be worth looking at how the Irish teach English, for they are better at it than the English, and the rest of the UK.

Singapore completes a hat trick by also coming first in science, with 561 points which is an increase of 10 over 2018. That is an improvement of half a year group. A slew of East Asian countries come next, with Estonia, now the European poster boy for PISA, the highest scoring European country. The UK comes in at 14th with a score of 500.

The OECD has to ask itself why so many countries declined since 2018, which is a question that Britain must ask as well. What were those countries that improved since 2018 doing that the UK was not. It was partly COVID and partly funding, but neither of these are the complete reason.

### **The state of global education**

In a world facing multiple crises, it is perhaps understandable that what's happening in schools, colleges and universities can sometimes be overlooked. But policymakers ignore education at their peril. Our schools today are our economies, societies and democracies of tomorrow. As artificial intelligence and digitalisation rapidly embed themselves in the global economy, it is vital teaching and learning innovate to ensure education stays relevant.

So how concerned should we be that 15-year-olds in 2022 are less likely to be proficient in maths, reading and science than those tested by PISA a decade ago? International comparisons are complex but PISA data point to a clear global

Table 2. Comparing countries and economies of performance in the math subjects is

		Math score 2022	Math score change from 2018			Math score 2022	Math score change from 2018
Above the OECD average	Singapore	575	6	Below the OECD average	<i>Ukrainian regions (18 of 27)</i>	441	N.A
	<i>Macao (China)</i>	552	-6		Serbia	440	-8
	<i>Chinese Taipei</i>	547	16		United Arab Emirates	431	-4
	<i>Hong Kong (China)*</i>	540	-11		Greece	430	-21
	Japan	536	9		Romania	428	-2
	Korea	527	1		Kazakhstan	425	2
	Estonia	510	-13		Mongolia	425	N.A
	Switzerland	508	-7		Bulgaria	417	-19
	Canada*	497	-15		Moldova	414	-6
	Netherlands*	493	-27		Qatar	414	0
	Ireland*	492	-8		Chile	412	-6
	Belgium	489	-19		Uruguay	409	-9
	Denmark*	489	-20		Malaysia	409	-32
	United Kingdom*	489	-13		Montenegro	406	-24
	Poland	489	-27		<i>Baku (Azerbaijan)</i>	397	-23
	Austria	487	-12		Mexico	395	-14
	Australia*	487	-4		Thailand	394	-25
	Czech Republic	487	-12		Peru	391	-9
	Slovenia	485	-24		Georgia	390	-8
	Finland	484	-23		Saudi Arabia	389	16
Latvia*	483	-13	North Macedonia	389	-6		
Sweden	482	-21	Costa Rica	385	-18		
New Zealand*	479	-15	Colombia	383	-8		
No difference	Lithuania	475	-6	Brazil	379	-5	
	Germany	475	-25	Argentina	378	-2	
	France	474	-21	Jamaica*	377	N.A	
	Spain	473	N.A	Albania	368	-69	
	Hungary	473	-8	<i>Palestinian Authority</i>	366	N.A	
	Portugal	472	-21	Indonesia	366	-13	
	Italy	471	-15	Morocco	365	-3	
	Viet Nam	469	N.A	Uzbekistan	364	N.A	
	Norway	468	-33	Jordan	361	-39	
	Malta	466	-6	Panama*	357	4	
Below	United States*	465	-13	Kosovo	355	-11	
	Slovak Republic	464	-22	Philippines	355	2	
	Croatia	463	-1	Guatemala	344	10	
	Iceland	459	-36	El Salvador	343	N.A	
	Israel	458	-5	Dominican Republic	339	14	
	Türkiye	453	0	Paraguay	338	11	
	Brunei Darussalam	442	12	Cambodia	336	12	

Table 3. Comparing countries' and economies' performance in reading

		Reading score 2022	Reading score change from 2018			Reading score 2022	Reading score change from 2018
		Statistically <b>above</b> the OECD average					
		Not statistically different from the OECD average					
		Statistically <b>below</b> the OECD average					
Above the OECD average	Singapore	543	-7	Below the OECD average	Greece	438	-19
	Ireland*	516	-2		Iceland	436	-38
	Japan	516	12		Uruguay	430	3
	Korea	515	1		Brunei Darussalam	429	21
	Chinese Taipei	515	13		Romania	428	1
	Estonia	511	-12		Ukrainian regions (18 of 27)	428	N.A
	Macao (China)	510	-15		Qatar	419	12
	Canada*	507	-13		United Arab Emirates	417	-14
	United States*	504	-1		Mexico	415	-5
	New Zealand*	501	-5		Costa Rica	415	-11
	Hong Kong (China)*	500	-25		Moldova	411	-13
	Australia*	498	-5		Brazil	410	-3
	United Kingdom*	494	-10		Jamaica*	410	N.A
	Finland	490	-30		Colombia	409	-4
	Denmark*	489	-12		Peru	408	8
	Poland	489	-23		Montenegro	405	-16
	Czech Republic	489	-2		Bulgaria	404	-16
	Sweden	487	-19		Argentina	401	-1
	Switzerland	483	-1		Panama*	392	15
	Italy	482	5		Malaysia	388	-27
No difference	Austria	480	-4	Kazakhstan	386	-1	
	Germany	480	-18	Saudi Arabia	383	-17	
	Belgium	479	-14	Thailand	379	-14	
	Portugal	477	-15	Mongolia	378	N.A	
	Norway	477	-23	Guatemala	374	5	
	Croatia	475	-3	Georgia	374	-6	
	Latvia*	475	-4	Paraguay	373	3	
	Spain	474	N.A	Baku (Azerbaijan)	365	-24	
	France	474	-19	El Salvador	365	N.A	
	Israel	474	3	Indonesia	359	-12	
Below	Hungary	473	-3	North Macedonia	359	-34	
	Lithuania	472	-4	Albania	358	-47	
	Slovenia	469	-27	Dominican Republic	351	10	
	Viet Nam**	462	N.A	Palestinian Authority	349	N.A	
	Netherlands*	459	-26	Philippines	347	7	
	Türkiye	456	-10	Kosovo	342	-11	
	Chile	448	-4	Jordan	342	N.A	
	Slovak Republic	447	-11	Morocco	339	-20	
	Malta	445	-3	Uzbekistan	336	N.A	
	Serbia	440	1	Cambodia	329	8	

Table 4. Comparing countries' and economies' performance in science

		Science score 2022	Science score change from 2018			Science score 2022	Science score change from 2018
Above the OECD average	Singapore	561	10	Below the OECD average	Iceland	447	-28
	Japan	547	17		Brunei Darussalam	446	15
	Macao (China)	543	0		Chile	444	0
	Chinese Taipei	537	22		Greece	441	-11
	Korea	528	9		Uruguay	435	10
	Estonia	526	-4		Qatar	432	13
	Hong Kong (China)*	520	-4		United Arab Emirates	432	-2
	Canada*	515	-3		Romania	428	2
	Finland	511	-11		Kazakhstan	423	26
	Australia*	507	-4		Bulgaria	421	-3
	New Zealand*	504	-4		Moldova	417	-12
	Ireland*	504	8		Malaysia	416	-21
	Switzerland	503	7		Mongolia	412	N.A
	Slovenia	500	-7		Colombia	411	-2
	United Kingdom*	500	-5		Costa Rica	411	-5
	United States*	499	-3		Mexico	410	-9
	Poland	499	-12		Thailand	409	-17
	Czech Republic	498	1		Peru	408	4
	Latvia*	494	7		Argentina	406	2
	Denmark*	494	1		Montenegro	403	-12
	Sweden	494	-6		Brazil	403	-1
	Germany	492	-11		Jamaica*	403	N.A
	Austria	491	1		Saudi Arabia	390	4
	Belgium	491	-8		Panama*	388	23
	Netherlands*	488	-15		Georgia	384	1
	France	487	-6		Indonesia	383	-13
	Hungary	486	5		Baku (Azerbaijan)	380	-18
Spain	485	N.A	North Macedonia	380	-33		
Lithuania	484	2	Albania	376	-41		
Portugal	484	-7	Jordan	375	N.A		
Croatia	483	10	El Salvador	373	N.A		
Norway	478	-12	Guatemala	373	8		
Italy	477	9	Palestinian Authority	369	N.A		
Türkiye	476	8	Paraguay	368	10		
Viet Nam	472	N.A	Morocco	365	-11		
Malta	466	9	Dominican Republic	360	25		
Israel	465	3	Kosovo	357	-8		
Slovak Republic	462	-2	Philippines	356	-1		
Ukrainian regions (18 of 27)	450	N.A	Uzbekistan	355	N.A		
Serbia	447	8	Cambodia	347	17		

## Notes for table 2.

*Countries and economies are ranked in descending order of the mean performance in mathematics.*

**Source:** OECD, PISA 2022 Database, Volume I Table I.B1.2.1 and Table I.B1.5.4.

Note: \* Countries with an asterisk against them did not meet the OECD's full data collection requirements.

## Notes for table 3.

\*\* Caution is required when comparing estimates based on PISA 2022 with other countries/economies as a strong linkage to the international PISA reading scale could not be established (see Reader's Guide and Annex A4).

*Countries and economies are ranked in descending order of the mean performance in reading.*

**Source:** OECD, PISA 2022 Database, Volume I Table I.B1.2.2 and Table I.B1.5.5.

## Notes for table 4.

*Countries and economies are ranked in descending order of the mean performance in science.*

**Source:** OECD, PISA 2022 Database, Volume I Table I.B1.2.3 and Table I.B1.5.6.

Some 25% of 15-year-olds in OECD member countries – representing 16 million children – are estimated to be low performers in maths, reading and science, including students not covered by PISA. This means they have not attained Level 2 proficiency; they can struggle to do tasks such as use basic algorithms or interpret simple texts. The situation is even worse among many non-OECD members. In 18 countries and economies more than 60% of 15-year-olds are low performers in all three subjects.

This is not the case for everyone. Singaporean students can work effectively with mathematical models for complex situations, comprehend abstract texts, and interpret and evaluate complex experiments. Singapore came top in maths, scoring 575 points, in reading (543 points) and in science (561 points). These results suggest that on average Singaporean students are the equivalent of almost three to five years of schooling ahead of peers who score the OECD average of 472 in maths, 476 in reading and 485 in science.

Singapore was also one of the few countries that kept improving in reading and science since 2018, while remaining



stable in mathematics performance. And it is noteworthy that this impressive educational performance has happened in a relatively short period of time. Older adults in Singapore assessed separately by the OECD perform far worse compared to younger generations. For example, less than 17% of 55-65-year-old Singaporeans scored at Level 3 or higher in literacy in the OECD Survey of Adult Skills (part of a product like PISA but for adults) – one of the smallest proportions amongst participating countries – while 63% of 16-24-year-olds did so, one of the largest proportions. This shows that educational progress can be rapid.

Elsewhere in PISA 2022, five other East Asian education systems outperformed everyone else in mathematics: Macao (China), Chinese Taipei, Hong Kong (China)\*, Japan and Korea, in order of performance. These same countries and economies were the next highest performers in science, along with Estonia and Canada\*. In reading, Ireland\* performed as well as Japan, Korea, Taiwan and Estonia (in descending order). In Ireland\* and Japan's case that is even though their expenditure per student is at or lower than the OECD average.

Test scores are only one measurement of success. Many countries have made significant progress towards the goal of universal secondary education; crucial to enabling everyone to participate fully in the 21st century world. Cambodia, Colombia, Costa Rica, Indonesia, Morocco, Paraguay and Romania are among the countries that have rapidly expanded education to previously marginalised populations compared with past PISA assessments.

Giving all students a fair chance to succeed regardless of their background is also a vital component of good schooling. PISA shows that disadvantaged students are often held back by social mobility hurdles that their more advantaged peers do not face. However, in most countries, some of the most disadvantaged students and schools are excelling and thus demonstrate academic resilience. On average, across the OECD, one in ten disadvantaged students was able to score in the top quarter of maths performers. This

clearly indicates that a disadvantaged background does not determine destiny. In fact, in 11 countries and economies Albania, Cambodia, Hong Kong (China)\*, Indonesia, Jamaica\*, Kazakhstan, Kosovo, Macao (China), Morocco, the United Kingdom\* and Uzbekistan - more than 15% of disadvantaged students were academically resilient.

In the same way as social disadvantage does not automatically lead to poor educational performance, the world is no longer strictly divided between rich and well-educated nations and poor and badly educated ones. While there is some correlation between spending and academic performance, history shows that countries determined to build a first-class education system can achieve this even in adverse economic circumstances. Korea and Singapore are prominent examples of countries which many decades ago had low incomes but focused on education and have achieved top performance.

While it is evident that some countries and economies are performing very well in education, the overall picture is more worrying. In more than two decades of global PISA tests, the OECD average score has not changed drastically between consecutive assessments. But this cycle saw an unprecedented drop in performance. Compared to 2018, mean performance in OECD countries fell by 10 score points in reading and almost 15 score points in maths – the latter is nearly treble any prior consecutive change. This downturn was particularly significant in a handful of countries. For example, from the OECD, Germany, Iceland, the Netherlands\*, Norway and Poland all saw a drop of 25 score points or more in mathematics between 2018 and 2022. The dramatic fall in maths and reading scores points to a negative shock affecting many countries simultaneously.

The COVID-19 pandemic seems an obvious factor that may have impacted results in this period.

However, take a closer look at the data. In reading, for example, many countries such as Finland, Iceland, the Netherlands\*, the Slovak Republic and Sweden have seen students scoring lower marks for some time – in some cases

for a decade or more. Educational trajectories were negative well before the pandemic hit. This indicates that long term issues in education systems are also to blame for the drop in performance. It is not just about COVID.

### **Unlocking the potential of the digital world**

While people have different views on the role of digital technology in schools, we cannot ignore how digital tools have fundamentally transformed the world. Everywhere, digital technologies are offering firms new business models and opportunities to enter markets and transform their production processes. Mobile apps enable people to track their health, computers help with boring or dangerous tasks, and games allow us to travel into virtual worlds.

Those who cannot navigate through this digital landscape are increasingly unable to participate fully in social, economic and cultural life. So it is good news that PISA shows the majority of students have embraced learning through digital technologies. On average across OECD countries, about three out of four students reported being confident using various technology, including learning management systems, school learning platforms and video communication programmes.

This was not true of all countries and economies. For example, in Jordan, Morocco, the Philippines, the Palestinian Authority and Thailand, only half or less of students felt confident or very confident about using a video communication programme. But overall, in most parts of the world, students are largely using mobiles, computers and other devices as part of their educational experience. Much of this change has occurred due to the pandemic, with schools forced to wake up to the power of digital technology. Remote lessons, digital tools and educational apps have radically transformed learning. One of the most visible benefits has been greater personalisation. For example, when students study maths on a computer, the computer can analyse how they learn and make their learning experience more granular, adaptive and interactive. Game-based learning can also make



learning more fun. Computer simulations let students do things that are difficult or costly to do in the real world. It is more insightful to do an experiment in a virtual laboratory than simply listen to a teacher explain the results of a scientific experiment.

These and other innovations point towards new educational pathways. Nonetheless, it is crucial to acknowledge that adept reading skills remain indispensable for effective learning in digital settings, as most digital learning materials are text-based. Estonia, Finland, Italy, Sweden and Switzerland stand out as having students with reading skills above the OECD average and confidence in learning autonomously with digital devices. This indicates that these systems provide students with a solid foundation for effective remote and autonomous learning. Overall, students who were more self-assured in their ability to learn independently and remotely scored higher in all studied subjects; an advantage of 10 score points compared to their less confident peers.

Learning analytics hold perhaps the greatest promise of digital technologies. Teachers can now get a real sense of how different students learn, what interests them in lessons, and where they get bored or stuck. This helps teachers improve the overall quality of their teaching and gives them a much better sense of which students need extra support.

This is important as three out of ten students did not feel confident about completing schoolwork independently. This rose to more than 50% of students in Japan and Malaysia. A similar proportion of students, almost one in two, indicated they had problems motivating themselves to do schoolwork at least once a week. This was worse in certain countries and economies; for instance, in Australia\* and the United Kingdom\* six out of ten students reported having frequent problems to motivate themselves – more than double the share of students in Guatemala, Iceland, Indonesia, Kazakhstan, Korea, Moldova and Taiwan.

Students need to take responsibility for their own development, but it is self-evident that some students will always need more support than others. Technology can assist

students in their learning, but teachers need to be ready to lend an ear to students who need help or want to share their problems.

### **Are teachers doing enough to support children?**

The amount of time a teacher can dedicate to individual children's needs is an important component of effective teaching. But PISA data show that teacher support has deteriorated over the last decade, at least in the perception of students. On average across OECD countries, the share of 15-year-olds who reported that students get extra help from their teacher when needed in most or every lesson dropped by an average of three percentage points.

It is not clear why. Is it the result of teachers not having enough time? Is the problem with certain teachers, who simply do not do as good a job as their colleagues? Or is it because the needs of students, real or perceived, have increased?

Overall, seven out of ten students reported that they regularly received extra help from teachers in 2022, while 22% of students reported getting help in some lessons. Around 8% never or almost never received additional support. Related to this, more than 35% of students reported that teachers did not regularly show an interest in every student's learning, on average across OECD countries, and failed to ensure all students understood the content.

The effectiveness of education can never exceed the quality of teaching and teacher support. PISA data show that this is particularly true in times of disruption. The availability of teachers to help students in need had the strongest relationship to mathematics performance across the OECD, compared to other experiences linked to COVID-19 school closures. Maths score were 15 points higher on average in places where students agreed they had good access to teacher help. These students were also more confident than their peers to learn autonomously and remotely. In a pandemic with numerous school closures, access to a supportive teacher who takes the time to give individual

instruction can make a huge difference. The data suggest that far too many teachers failed to give students adequate support.

This is backed up by other PISA findings. Maths results declined between 2018 and 2022, on average across the OECD, in education systems where principals reported a rise in teaching hindrance due to inadequate or poorly qualified teaching staff. In contrast, systems where more teachers were fully certified by an appropriate authority tended to score higher in mathematics, even after accounting for per capita GDP, across OECD countries.

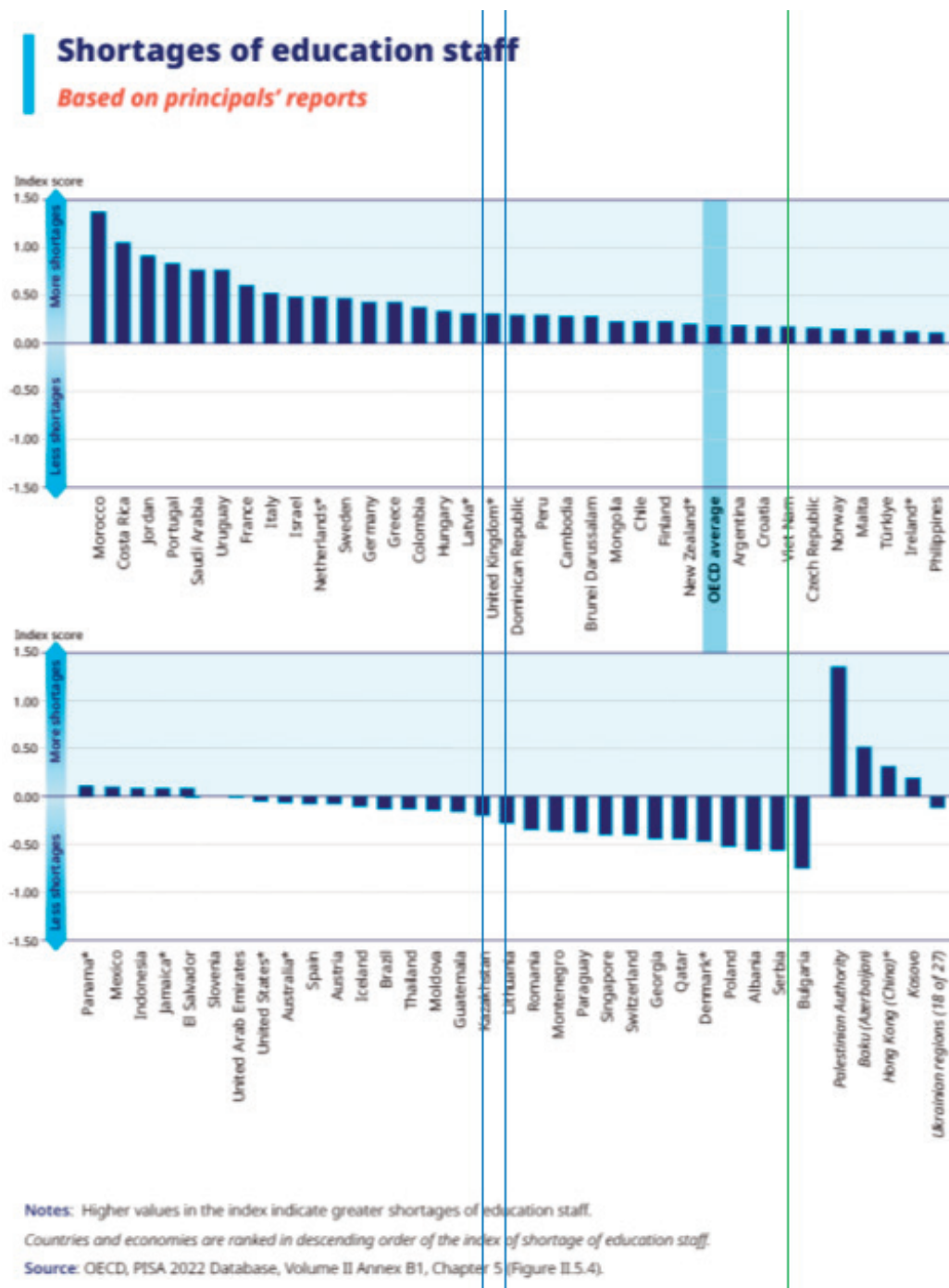
In 32 countries and economies, students' maths scores were lower in schools whose principal reported staff shortages compared to schools without staffing issues. However, 35 countries and economies saw no statistically significant difference between schools with or without shortages. Paradoxically, even though principals in 2022 perceived a greater shortage of teachers, PISA data show that between 2018 and 2022, student-teacher ratios and class sizes actually decreased on average across OECD countries and remained mostly stable elsewhere.

What does this mean going forward? It is important for education systems to examine this apparent contradiction: the sense that there are more teacher shortages even though the number of teachers per student has often risen or at least remained stable. Other notions or phenomena might be feeding this perception. Teacher absenteeism, the idea that teachers are not sufficiently qualified, or the changing role of teachers could all be a factor. Given the fast pace of change in education, expectations for teachers may also have changed. In turn, this could have altered the standards against which teacher supply and performance are measured.

### **Too hungry to learn?**

Many parts of the world are experiencing a food crisis with families struggling to put food on the table. PISA results show that millions of students, including from some of the richest countries, are often struggling to get fed. On average across

Table 5. Shortages of education staff



OECD countries, 8% of students reported not eating at least once a week in the past 30 days due to lack of money to buy food. Some OECD countries have far lower proportions, notably Portugal, Finland and the Netherlands\* were all below 3%. However, elsewhere food insecurity was much higher. In the United Kingdom\* and Lithuania 11% of students said they were forced to skip meals.

The figures were even more elevated in other OECD members, for example the United States\*, Chile, and Colombia (all 13%), New Zealand\* (14%) and Turkey (19%). If

**Table 6. Students’ average mathematics scores compared to the percentage skipping at least one meal per week**

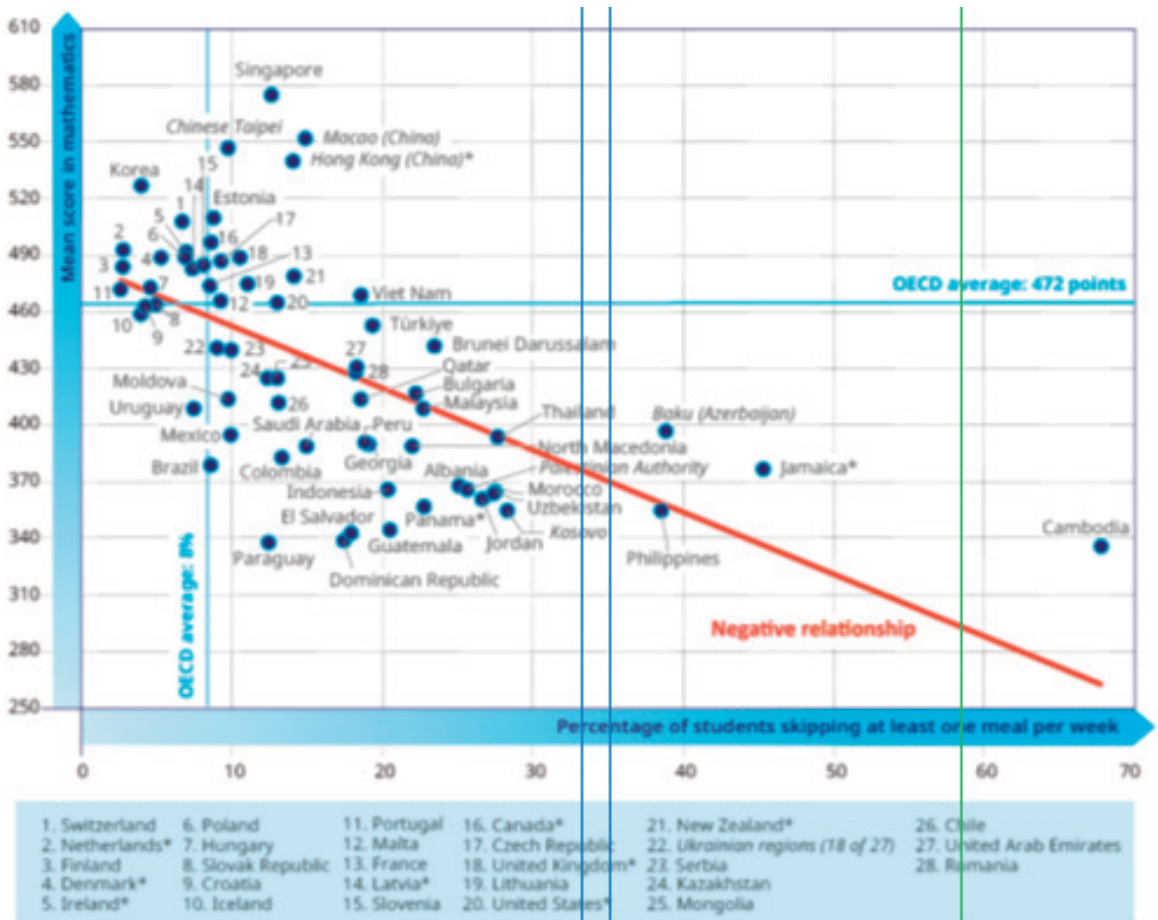
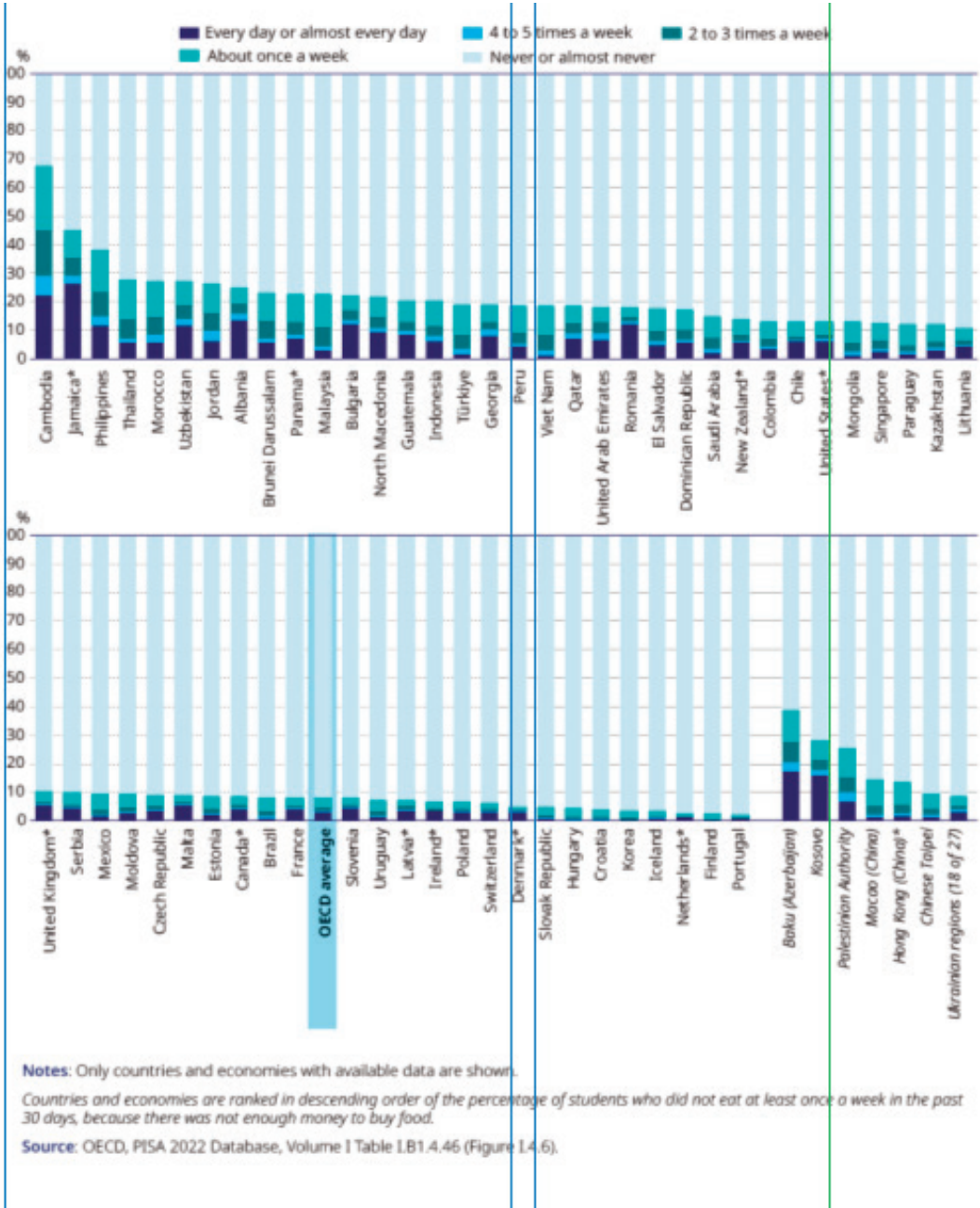


Table 7. Percentage of students not eating in the past 30 days because not enough money to buy food





students' bellies are rumbling, they are unlikely to learn as effectively. As millions turn to food banks and community programmes to help put food on the table, schools can help.

Many countries already provide school meal programmes. They are a safety net for vulnerable children and households. Policy makers should consider the potential significance of providing a regular, nutritious meal as a cost-effective way of ensuring students get the food they need. With rising food, rent and energy bills in many parts of the world, families are forced to make tough choices. Free lunches can attract more children to attend school, enable them to learn better and help maintain their health.

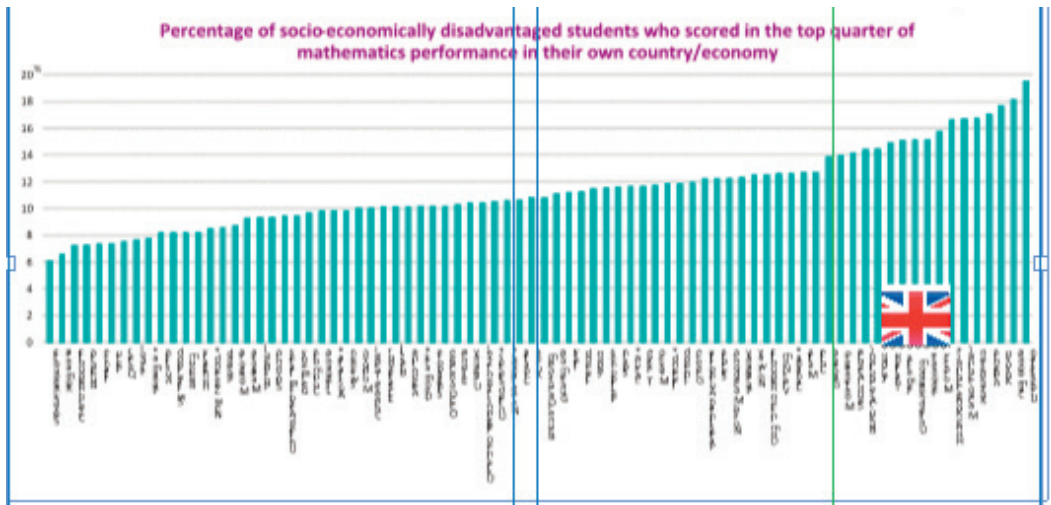
### **PISA and the UK. What's the verdict?**

Presenting the results of PISA 2022 in the UK at a press briefing, Andreas Schleicher, Director of the Education and Skills Directorate of the OECD, described the UK's performance over the last four years as "so, so". There have been some notable successes, but also some failures. Britain's performance has fallen since 2018, but not by as much as the OECD average. Yet the important point is that not all countries did see their performance fall, even though all were affected by COVID. Some countries improved their performance, in a few cases by quite a bit. What were those countries doing that the UK was not?

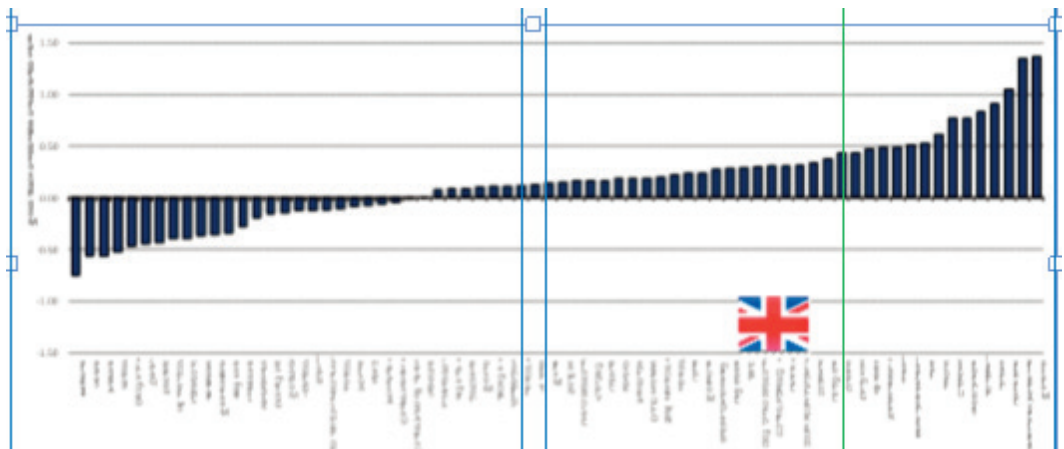
Education is, of course, a devolved issue, but the OECD deals mainly in nation states which is why it is the UK that is ranked. Yet results for the individual nations of the UK are given in an appendix, and they show England out-performing the other three. (See the Table 1 above.)

Some of the areas where England in particular has done well include keeping the social discrepancies between the best off and worst off students from increasing. Student resilience in the UK is also good. (See Table 8 below.) The UK has also been particularly good at integrating the children of immigrants. If you look at the social background of immigrants there is little difference between immigrants and native population of the same social status. Britain has faced high

**Table 8. Resilient students in mathematics**



**Table 9. Shortage of education staff based on head teacher reports**

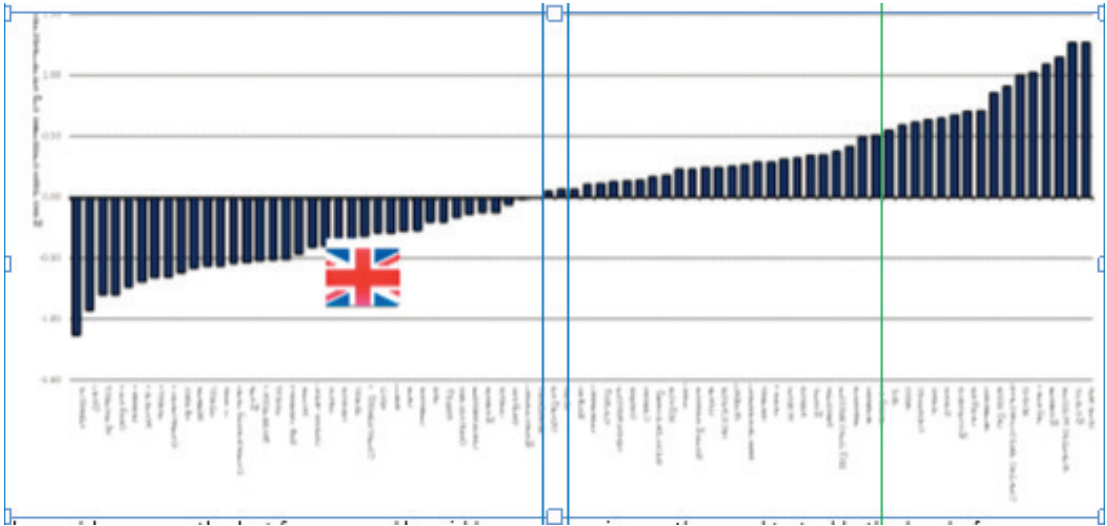


levels of immigration recently, and has managed this, at least in terms of education, much better than some other OECD members.

The question of funding is an issue, but it is not as important as many think. As Andreas Schleicher said at the press briefing, “money is necessary, but not sufficient”. What the PISA data does show is that the UK is not particularly



**Table 10. Shortage of non-staff material resources based on head teacher reports**



productive in its use of what money there is. When head teachers were asked their views, it was not a lack of resources that they were most concerned about but a lack of staff.

Teacher support for students during COVID was relatively poor, Prof. Schleicher said. Parental support in the home is high in the UK, although parental involvement in school has declined, head teachers thought. School safety as measured by pupil perception is lower in the UK than the OECD average. Heads thought that this impacted negatively on schools.

Andreas Schleicher said that “these declines should concern us a lot. They are not the inevitable outcome of the pandemic.” Yet Schools Minister of State Damian Hinds MP, speaking at the launch of PISA in the UK at the Policy Exchange think tank, did blame the COVID pandemic for much of the problems over the last four years. He said he was focusing on the need to tackle the level of absenteeism that had grown since COVID. In this he was warmly supported by Dame Rachel de Souza, the Children’s Commissioner for England, also speaking at the Policy Exchange launch. She said: “Education is the most important thing in improving a child’s life chances.” As Mr Hinds said: “If you are not in

school you won't get the advantages that education offers."

One area where policy makers have regularly ignored the evidence from PISA for over 20 years is in the negative impact of selection on most children. This was highlighted again in the 2022 PISA. Avoiding selection was one of the measures that the OECD recommended for those countries that wanted to improve their results for the next round of PISA in 2025. The Policy Exchange launch ended with a few words from Nick Gibb, who stood down at the recent reshuffle after over a decade as either Shadow Schools Minister or the real thing. The genuine warmth for the man and his commitment to education was obvious. Dame Rachel even called for him to be knighted.

### **References and sources**

*PISA 2022 Results (Volume I). The State of Learning and Equity in Education*, OECD, Paris, published on Tuesday 5 December 2023.

*PISA 2022 Results (Volume II). Learning During – and From – Disruption*, OECD, Paris, published on Tuesday 5 December 2023.

*PISA 2022. Insights and Interpretations* by Andreas Schleicher, OECD, Paris, published on Tuesday 5 December 2023.

Country notes for each of the 81 countries and territories participating in PISA, OECD, Paris, published on Tuesday 5 December 2023.

*PISA 2022: National Report for England* by Jenni Ingram, Jamie Stiff, Stuart Cadwallader, Gabriel Lee & Heather Kayton from the University of Oxford. Published by the Department for Education on Tuesday 5 December 2023. Reference: RR 1392. ISBN: 978-1-83870-522-0.

*Programme for International Student Assessment (PISA 2022): Highlights from Scotland's Results. Children, Education and Skills.* By APS Group Scotland, published by the Scottish Government on Tuesday 5 December 2023.

John Bangs

**John Bangs** is Special Adviser (Research) to Education International. A former Assistant Secretary of the NUT and Head of its Education Department, he has also been chairman of the OECD's Trade Union Advisory Committee on Education and has been heavily involved in the OECD/EI International Summit on the Teaching Profession.

# PISA 2022: A personal commentary

**By John Bangs**

Special Adviser (Research) at Education International, EI advisor on working with the OECD and a former Head of Education at the NUT

**Key words:** OECD, PISA, maths, science, reading, COVID.

**Abstract:** *John Bangs has been involved with the OECD as a consultant for Education International, the teacher union global confederation, for many years. He attended the launch of PISA 2022 and here gives a personal commentary.*

**I**t's been a long time coming because of COVID. The launch of PISA 2019, back in December 2019, seems a world away. So, what do we know four years on? The first point to make is that the OECD has taken the impact of COVID seriously, devoting an accompanying volume to this theme. It also published a commentary by Andreas Schleicher- Insights and Interpretations- which goes into further detail on the educational policy implications of the post-COVID landscape.

The second is that it is OECD countries which have experienced the greatest overall declines in educational performance since COVID, while some of the low to middle income countries who are non-members and have bought into PISA, have actually improved their performance in the past four years.

And some of the drops in performance are disorientating. Take the relationship between the UK, my home country, and Finland. Both remain in the above average performers band, yet the UK is now seven places higher than Finland. Should this be cause for celebration? Well, no.

Compared to PISA 2018, the UK is down thirteen points in Mathematics, ten in reading and five in Science while Finland is down, in that order, by twenty three, thirteen and eleven points. As Andreas Schleicher pointed out during the launch, the declines have been uneven. And of course, these comparisons also highlight the limitations of performance tables!

The third point to make is that it is in Reading not Mathematics where the declines have been greatest. While Japan, Korea, Italy and Taipei in the above average band are reported not to have declined in performance, even Singapore experienced a slight drop in Reading despite its performance in Mathematics and Science.

And the fourth point is that it is in Science where the tables record at least some stability in performance compared to 2018.

The question is, of course, why have these declines taken place despite the welcome news that countries such as Philippines, Dominican Republic, Cambodia and Guatemala appear to be, to use OECD's own words, 'rapid improvers'?

### **COVID's impact**

What is clear is that COVID impacted severely those countries which did not invest in the time and resources necessary to make their systems resilient and their teachers well supported. While there may be other long term reasons for decline, the obvious correlation is that the shock effect of COVID, (as the OECD describes it), and the severe decline in performance was the main reason for such a decline. After all, it was the OECD itself which led the way in stressing that proper long term COVID educational recovery packages were vital if education systems were to recover from COVID.

That said, PISA's focus on the necessary steps to improve education system resilience is impressive. It's a shock to read that only four education systems had shown resilience in the three areas of learning, equity and well-being and that no country performs well in relation to student well-being. Other insights are important, including how food insecurity

affects just under 10% of students in OECD countries, the prevalence of anxiety, loneliness, and depression among students during lockdowns and how students' interest in working in the health sector has decreased in countries which have experienced high levels of COVID deaths. There are still mental health problems among many students, post COVID, as well as a cost-of-living crisis which affects many students in many countries. It's vital that policy focuses on tackling these areas.

It's therefore welcome that Chapter 8 of PISA focuses on ten actions related to resilience. They will, no doubt, be a key focus in any policy debate about how education systems can improve in future. Particularly welcome is the emphasis on the need for sufficient highly qualified teachers in schools and the need to establish schools as hubs for social interaction. And surely most teachers will cheer at PISA's support for banning students' smartphones in schools, although more controversial will be the proposal that quality assurance mechanisms should be at the centre of guaranteeing quality in highly devolved systems.

However, there is one thing on which the OECD should reflect. The voice of teachers is not obvious in the report. Otherwise, the ten actions wouldn't have contained an unqualified statement that greater efforts were needed to ensure that students receive necessary and relevant support from teachers. During COVID teachers were often on their own without external support but unqualified such a statement could be seen as inferring that teachers somehow weren't doing their best to support students sufficiently.

In fact, the OECD itself has recognised, in a joint paper with Education International on post-COVID recovery, that teachers went the extra kilometre (or mile!) during COVID, initiating a whole host of micro innovations which supported students-innovations which still haven't been quantified and researched.

It's unfortunate that we have to wait for two years until TALIS 2024 is published to hear teachers' views on lessons from COVID. As Andreas Schleicher said at the launch,

teachers are on the front line. Maybe OECD member countries ought, therefore, to discuss again making a teacher questionnaire a requirement, not just an option, in PISA.

*This commentary first appeared in Education Journal No. 544,  
published on Wednesday 6 December 2023.*

# Select Committee Reports

**W**e continue our series of reviews of all parliamentary select committee reports on education, which we started in volume 25 beginning with January 2018. This issue covers the period from October to December 2023.

*Support for Childcare and the Early Years: Government response to the Committee's Fifth Report*, the Education Select Committee's sixth special report of session 2022/23. HC 1902. Published on 18 October 2023.

*The Condition of School Buildings*, Public Accounts Committee, Second Report of Session 2023/24, HC 78. Published on Sunday 19 November 2023.

*Persistent Absence and Support for Disadvantaged Pupils: Government response to the Committee's Seventh Report*, House of Commons Education Select Committee, First Special Report of Session 2023/24, HC 368. Published on Wednesday 6 December 2023.



# Childcare and the early years

*Support for Childcare and the Early Years: Government response to the Committee's Fifth Report*, the Education Select Committee's sixth special report of session 2022/23. HC 1902. 18 October 2023.

**T**he report from the select committee on support for childcare and early years settings made 23 recommendations. The Government rejected the Committee's recommendation that business rates should be scrapped for childcare settings, and that they should be zero-rated for VAT on their business purchases. Ministers said they had frozen business rates for the next five years and that rate relief schemes already provide discounts to childcare providers and others. The response said there are no plans to alter terms for VAT. It also rejected the call for a review of tax-free childcare, a policy controlled by the Treasury rather than the Department for Education.

The Government did not fully accept the Committee's call for it to work "with childcare providers and local authorities" to set the hourly funding rate that will be paid to childcare providers when the universal 30-hours childcare entitlements are rolled out. Witnesses to the inquiry said providers who are already struggling could otherwise be left insufficiently funded. The Department for Education's (DfE's) response said it recognises the importance of setting funding rates with local authorities, had uprated the hourly rate for 2023/24, and was providing additional funding via the early years supplementary grant.

Agreeing with a Committee recommendation, the Government confirmed it will amend town planning legislation so that funding from the Infrastructure Levy, which property developers pay to local authorities after planning permission is given, can be used to pay for childcare facilities.

There was a positive response to the Committee's call for government to remove barriers faced by social housing

residents whose tenancies block them from operating as childminders in their homes. DfE said it is “engaging with the Department for Levelling Up, Housing and Communities and a range of housing sector stakeholders including social and private landlords, to identify and reduce property related barriers to childminding”.

DfE did not directly respond to MPs’ calls for it to prioritise career development of early years practitioners as a means of improving retainment of staff in the sector, and giving it parity of esteem with staff working in schools. The Department said it is developing a national campaign “to boost interest in the sector”. It also aims to boost recruitment with efforts to “remove barriers to entering the sector, by ensuring qualifications are suitable and easy to understand”, and by introducing new types of apprenticeship for becoming a childcare professional.

In total the Government accepted fully four of the Select Committee’s recommendations, accepted a further 11 in part and rejected or failed to respond to eight.

Education Committee chairman Robin Walker MP said: “We maintain that there is a powerful case for reviewing the funding levels for childcare settings, the working of tax free childcare and, given their vital contribution to the economy, the taxes they incur. We hope the Government is eyeing up ways to help the sector in its Autumn Statement, which will build on the positive announcements made in the spring.

“We accept that not every one of our recommendations is in the gift of ministers at the Department of Education, but our report stressed and ministers have accepted the importance of cross departmental work on these issues. We shall continue to press for action across Government to support this vital sector.

“Meanwhile it is encouraging that ministers are working on ways to boost recruitment and retention in the sector, and to remove unfair barriers that stand in the way of social housing tenants becoming childminders.”

# The condition of school buildings

*The Condition of School Buildings*, Public Accounts Committee, Second Report of Session 2023/24, HC 78. Published on Sunday 19 November 2023.

**T**he condition of school buildings report, by the Public Accounts Committee (PAC), pointed out that the Department for Education (DfE) still had “incomplete knowledge” in terms of the number and condition of schools with reinforced autoclaved aerated concrete (RAAC), and there were questions about the reliability of some of its information.

The PAC noted that in mid-September 2023, the DfE had reported that 98% of schools had responded to a RAAC questionnaire that it had issued in March 2022, which meant that several hundred responses had been outstanding. But the PAC pointed out that the set of responses had included some which had been inconclusive, while other schools had been resubmitting their responses given the renewed media focus on RAAC.

While the DfE’s guidance about the questionnaire had explained that an estates manager or appropriately qualified building surveyor should make the relevant judgements, the DfE had been concerned about the potential for schools to have submitted “false negative” responses, and it planned to conduct sample checks where schools suggested that they did not have RAAC. Where a school had reported that it did not have RAAC, it had taken a number of weeks to conduct the specialist survey to confirm. The PAC said that as RAAC was now widely recognised as a problem, there was a greater risk that experts who could help identify and manage RAAC may be in short supply. NHS England, for example, has told the PAC that there was a limited number of specialist engineers.

The PAC recommended that the DfE should urgently assess the risks of inaccuracies within the RAAC questionnaire

returns and specialist surveys, so that it considered the risks in its overall approach, decision-making and guidance. The report also urged the DfE to expedite its programme of specialist surveys where RAAC had been suspected, and in due course publish the full set of results so that the extent of the RAAC problem could be known.

The PAC said that the DfE's "risk appetite" regarding the school estate, and how that aligned with its recent approach on RAAC, was unclear. Since summer 2021, DfE had recognised a significant safety risk across the school estate. In spring 2023, it had continued to assess that its mitigations would not bring the risk likelihood down to acceptable levels, and considered that the most effective further mitigation would be an expanded School Rebuilding Programme. In late August 2023, DfE had taken what it considered to be a highly risk-averse approach of advising all schools with confirmed RAAC to avoid using spaces where RAAC was present, regardless of any assessment of its structural condition.

The PAC pointed out that it was unclear whether the DfE had taken the action because it had realised its RAAC assessment and assurance process was insufficient, or because it no longer wanted to accept any risks across the school estate. The report noted that where schools had responded to the questionnaire to say that they had RAAC, but were still awaiting a specialist survey to confirm it, they had not been advised to take any mitigating action. The Department for Health and Social Care had established a £685 million fund to 2024–25 to mitigate RAAC, and it had committed to remove RAAC from the NHS estate by 2035, while the DfE had made no such financial or practical commitments.

The PAC recommended that the DfE should clarify its "risk appetite" in terms of safety issues across the school estate, and ensure that it fed through into consistent decision-making, and that a nominated senior official was in charge. The PAC stressed that, in line with the approach already taken by DHSC, the DfE should make clear when and how it planned to have eradicated RAAC from the school

estate.

The report said that schools were uncertain about the support they could access to mitigate RAAC-related issues, and how they would be reimbursed financially. The PAC said that the temporary classrooms that DfE had been providing would generally be for those schools that had been known to be affected by RAAC, before the late August 2023 change of risk approach. The PAC said that the DfE had not been able to provide it with a figure on how many were being provided. The report said that while the DfE had undertaken to pay for the capital costs incurred by schools, its approach to revenue funding, and paying for surveys carried out by schools, had been less clear, particularly if a school had significant reserves.

The PAC noted that the DfE had still not set the funding application process, but it had accepted the need for some checking and controls. While each school had access to a caseworker, anecdotal evidence had suggested that many were struggling to understand DfE's approach, and the PAC stressed that it was concerned about a lack of fairness in terms of access to temporary support and how that support would be paid for.

The PAC recommended that the DfE should re-examine its process for funding temporary mitigation measures, to ensure that it achieved the right balance between accessibility and value for money, and communicating that clearly and consistently to schools.

The report pointed out that there was still a lack of transparency for schools, parents and communities in terms of where RAAC existed and how long it would take to be fixed. It added that the DfE had been unable to provide answers to important questions, such as how many specialist surveys to confirm RAAC were outstanding and likely to be carried out, or how many pupils had been affected by RAAC-related school closures at the start of the 2023–24 school year.

While the DfE said it was aiming to release information in a managed and routine way, as it did for other management information, after an evidence session, the DfE

had reported on 19 October, that RAAC had at that point been confirmed in 214 educational settings, of which 202 had been providing face-to-face education for all pupils. The report added that, despite suggestions in early September that the RAAC situation would be resolved in a matter of weeks, the DfE was aware that some cases were too complicated to be dealt with in that timeframe, and that some schools would not even be identified as having RAAC until later.

The PAC recommended that the DfE should write to the Committee, alongside its Treasury Minute response, with its latest assessment of the scale of the RAAC problem, its overall plan to deal with it, and the likely associated costs.

The report pointed out that the DfE had incomplete knowledge of the prevalence of asbestos across the school estate, and in May 2022, the DfE had agreed with the PAC's recommendation that it should urgently chase the 7% of schools that had not responded to the asbestos management survey it had launched in 2018. In July 2023, the DfE had explained that the proportion of schools which had not been seen had since fallen to just over 4%, although it still represented almost 1,000 schools.

The PAC noted that work on DfE's ongoing second Condition Data Collection programme (CDC2) would not specifically check for asbestos, but it would include a review of schools' asbestos management plans and compliance with guidance requiring schools to have an onsite asbestos register. The report pointed out that the unexpected presence of asbestos had complicated ongoing work to address other issues such as RAAC, and the two could be present in the same building. Data from the Health and Safety Executive had suggested, that since 2011, around 11 teachers or ex-teachers had died from asbestos-related conditions each year.

The PAC recommended that, as soon as possible, the DfE should provide the PAC with evidence that it had a full picture of asbestos across the school estate, having received survey returns from all schools and ensuring that every relevant school had an adequate asbestos management plan in place.

The report stressed that unacceptable numbers of pupils were learning in poorly maintained or potentially unsafe buildings, the quality of school buildings had an impact on pupils' learning experience, and ultimately on attainment levels and teacher retention. The PAC noted that an estimated 700,000 pupils attended the 1,200 schools that had been considered for the School Rebuilding Programme, which aimed to rebuild or refurbish buildings in the most need given safety matters or their poor general condition (which could include problems with, for example, roofs, windows or heating systems). The DfE said it would select 500 schools to be included in the programme. The DfE's first Condition Data Collection programme (CDC1), conducted between 2017 and 2019, had found that just over 2% of building components had been in "poor" or "bad" condition, but it covered a large number of schools.

The PAC recommended that, within the next year, the DfE should develop a package of support and good practice to help those responsible for mitigating the negative impact on pupils and teachers of schools that were in poor condition but cannot yet be fixed.

The report pointed out that the DfE had focused on reactive measures addressing immediate building concerns that often failed to take account of longer-term value for money considerations. It added that the DfE had committed to providing funding for all schools that faced critical and immediate safety risks but were unable to carry out appropriate remedial work themselves.

The PAC said that the DfE had been allocated school rebuilding funding equating to £1.3 billion a year, to allow it to rebuild 50 schools a year, rather than the 200 a year that it had set out in its Spending Review 2020 case. The report pointed out that, a significant number of the schools chosen for DfE's latest capital programme, the School Rebuilding Programme, had been selected in response to structural or safety issues that responsible bodies had identified as serious enough that buildings were at risk of closure or they posed a risk to staff and pupils.

The DfE had told the PAC that many of the 100 schools still to be selected for the programme would be chosen because they had serious issues with RAAC, and therefore, many other schools would not get on to the School Rebuilding Programme even though a longer-term value for money assessment based on their poor condition would lead to the conclusion that they should be rebuilt.

With regard to fire safety measures, the report pointed out that the DfE's cost-benefit calculations often had led it to opt for expensive retrofitting rather than initial inclusion in a new school design which was cheaper.

The PAC recommended that, within the next year, DfE should set out its strategy for encouraging responsible bodies to carry out timely and effective repairs to better protect longer-term value for money. It added that it should also reconsider its value for money analysis on fitting fire safety measures.

The report pointed out that the School Rebuilding Programme was behind its initial schedule for getting contracts in place and schools built. The DfE had announced the School Rebuilding Programme in June 2020, but by March 2023, it had delivered one project compared with a forecast four, and 24 contracts had been awarded compared with a forecasted 83.

The PAC said that price inflation and other market conditions had made it difficult to find contractors, and the DfE had conceded that it would not be able to catch up on projects where it was already behind the planned timeframe, although it was confident that it would stay on track for upcoming projects. The report argued that as there had been changes in the external environment, such as movements in inflation rates, the situation may affect programmes and create complexities, and the factors may mean that a programme could no longer achieve its intended outcomes, or it was too costly to do so.

The PAC recommended that the DfE should provide assurance that it had a good understanding of how current and likely future challenges would affect the timetable and



costs for the School Rebuilding Programme, including by carrying out appropriate scenario-planning should likely and significant risks materialise. The report noted that there had been considerable variation across the school estate, including regional disparity in the condition of school buildings and differences in school types and governance models, which influenced the type of support DfE needed to provide.

The map of school building condition had showed a broad north-south divide, with higher need in the north. However, a more granular assessment had also suggested that schools in rural and coastal areas faced particular difficulties. The report pointed out that the DfE did not currently have a mechanism for directing capital funding towards areas identified as meriting particular support to tackle weak educational outcomes. The PAC said that, for around one-third of the 1,000 schools with the highest level of need, the responsible body had not made an application for the School Rebuilding Programme.

The DfE had also found that a proportion of schools in most need did not apply for, or were unaware of, the maintenance and repair funding that was available to them. Voluntary-aided schools (which were typically faith-based) often had good relationships with their respective oversight bodies, but administrative and funding arrangements were inconsistent. The report added that, while some small local authorities, which were responsible for only a few maintained schools, may lack estate management capability, they were currently excluded from DfE's Capital Advisers Programme.

The PAC recommended that the DfE should review its guidance, support and financial allocations designed to help reduce variation in the condition of school buildings and the capability of those managing the estate, and make improvements where necessary.

The report stated that it was unclear whether decisions concerned with addressing the condition of the estate were coordinated with those relating to the need for school places. Historically, there had been instances of school closures just before another demographic wave of children

that meant that more school places had to be created. The PAC stressed that there was no requirement for responsible bodies to work together to, for example, consider possible closures or amalgamations of schools on borough boundaries to ensure the most efficient option was chosen.

Given the recognised autonomy of responsible bodies, the report pointed out that the DfE's regional directors did not typically play a strong role in school closure decisions, but they engaged with schools and other bodies to discuss such issues. The report added that, from a departmental perspective, when considering School Rebuilding Programme applications, DfE checked the forecast pupil numbers to ensure that the school merited a full rebuild, but ideally, DfE would like more school places than there were children, to support parental choice in the system.

The PAC suggested that in some places, sites no longer needed for primary schools could be re-purposed to provide more childcare and early years provision, or opportunities for more special and alternative provision.

The PAC recommended that the DfE should consider how local authorities could best be supported, and introduce the necessary measures, to ensure that the need for high quality places across the estate was considered when decisions were taken about reducing school places locally.

# Persistent absence and support

*Persistent Absence and Support for Disadvantaged Pupils: Government response to the Committee's Seventh Report, House of Commons Education Select Committee, First Special Report of Session 2023/24, HC 368. Published on Wednesday 6 December 2023.*

**The Government welcomed the Committee's report into persistent absence and support for disadvantaged pupils. The Government noted that attending school regularly is crucial. Children who are not attending school regularly miss out on chances to learn, to socialise, and to play an active part in their school community.**

As the Committee heard, while most children attend school regularly, the pandemic created unprecedented disruption in attendance habits and led to higher rates of persistent and severe absence for some children. The Government noted that recent data show improvements – the percentage of children persistently absent or not attending school for COVID-related reasons fell to 22.3% in 2022–23, down from 27.5% a year earlier, which is equivalent to around 380,000 fewer pupils persistently not in school. But there remains a long way to go to achieve the goal of achieving pre-pandemic attendance levels or better.

The Government agreed with the Committee that there are few quick fixes - sustained improvement in school attendance requires long term focus across the system. In the decade before the pandemic, the Government commissioned the Taylor Review, delivered a tougher definition of persistent absence, drove sustained Ofsted attention, and updated and improved the legal framework alongside wider school reforms. Persistent absence fell from 16.3% in 2010 to 11% in 2014/15 and remained largely stable until 2018/19. To return to these levels or better, the Government has a

comprehensive attendance plan.

At the heart of the plan are clearer and more consistent new expectations set out in guidance, which seek to promote a 'support first' ethos and one in which attendance is everybody's business. Schools are expected to: publish an attendance policy; appoint a senior attendance champion; use data to identify at-risk pupils early; and work closely with families to support absent pupils. Local authorities are expected to establish an attendance support team and hold termly meetings with every school to plan interventions for children at risk of persistent or severe absence.

The new expectations seek to ensure that all schools and local authorities adopt the habits of the best: they reflect the practices of schools and local authorities with higher than average levels of disadvantage, but better than average rates of attendance. They depend in turn on schools, trusts, and local authorities to implement them.

To help support the sector achieve these expectations, the Department has established a daily data pilot, with 87% of state-funded schools now participating, helping to ensure that they and local authorities have near real-time attendance data. This allows them to identify need early, spot trends and benchmark against the best to share best practice around the country. The Department has also formed an Attendance Action Alliance, comprised of national leaders from critical sectors like education, health, social care and policing. It works to take practical action to remove barriers to strong attendance and mobilise workforces around the issue. The Government has also launched attendance hubs to enable schools with excellent attendance levels to share resources and advice with other schools in similar circumstances but with high absence. These have recently expanded to 14 in number, which will support improvements across 800 schools, and reach some 400,000 children.

Alongside these steps, the Department has employed ten expert attendance advisers who are working with every local authority in the country and a number of Multi-Academy Trusts to put in place effective plans to deliver the new

attendance expectations. And the Department has established an attendance mentoring program which is being piloted in five of the Department's priority education areas – Middlesbrough, Doncaster, Stoke on Trent, Knowsley and Salford – offering intensive one-to-one support for around 1700 absent pupils and building the evidence base on what works.

The attendance guidance sets the framework for identifying children who need additional support but, as this inquiry recognised, the individual reasons behind persistent and severe absence often arise from wider challenges. The attendance plan is therefore underpinned by wider education recovery investment and reforms tackling the underlying causes. This includes £5bn worth of direct investment in education recovery, including £400m on teacher training opportunities and up to £1.5bn on tutoring. In addition, the Government is spending £2.9bn annually on the pupil premium, on top of £1.3bn on recovery premium, Schools must spend the pupil premium on evidence-informed approaches, including attendance strategies and attendance. Recent analysis by the Education Endowment Foundation (EEF) of school strategy statements found that 75% of schools in England identified poor attendance as a priority.

The Government has invested an extra £200m on the Supporting Families programme increasing the budget to £695m by 2024–25, to help an additional 300,000 families facing multiple problems. Sustained good attendance is a key outcome of the programme. The Holiday Activities and Food (HAF) programme (over £200m a year) will particularly benefit disadvantaged children, along with the £30 million invested in the National School Breakfast Programme.

The NHS Long Term Plan commits to increased investment in mental health services of at least £2.3 billion a year by March 2024, and aims for an additional 345,000 children and young people to be able to have NHS-funded mental health support by the same date. The Special Educational Needs and Disability (SEND) and Alternative Provision improvement plan will involve £2.6 billion of

## Select committee reports

spending between now and 2025 – including additional investment in special schools and specific measures supporting attendance such as the inclusion of improved attendance as an outcome in the performance framework for Alternative Provision. The Government's reforms of the Children Social Care review are backed by £200m of additional investment. Better early help will particularly benefit children at risk of absence.



# Education Journal Review

Volume 30 • Number 1

- 7 **Preface**
- 8 **Attractiveness of UK taught Masters to international students**  
*He Aian*
- 15 **Challenging systemic prejudices. An investigation into bias against women and girls in large language models**  
*OECD/Daniel van Niekerk et al*
- 43 **Teacher labour market in England 2024**  
*NFER/Dawson McLean, Jack Worth and Andrew Smith*
- 83 **PISA 2022. Measuring the world's education systems after COVID**  
*Demitri Coryton*
- 108 **PISA 2022. A personal commentary**  
*John Bangs*
- 114 **Select Committee Reports**
- 115 **Childcare and the early years**  
*The House of Commons Education Committee.*
- 159 **The condition of school buildings**  
*The House of Commons Public Accounts Committee*
- 125 **Persistent absence and support**  
*The House of Commons Education Committee.*