Education Endowment Foundation

Evaluating Teacher recruitment strategies using fully randomised paired conjoint experiments

Further appendices

October 2025

Rebecca Allen, Iain Ford, John Jerrim, Loic Menzies, Sam Sims, Burak Sonmez and Karen Wespieser

Teacher tapp

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, nurseries and colleges to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were re-endowed with an additional £137m, allowing us to continue our work until at least 2032.

For more information about the EEF or this report please contact:

Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents of Appendix

Appendix A – Literature review

This literature review examines existing research on teacher job preferences, focusing on the factors that influence recruitment and retention in England's education system. It explores the impact of salary, workload, flexibility, and working conditions on teachers' career decisions, drawing on both quantitative and qualitative studies. Additionally, it considers the role of school leadership, policy, and job attributes in shaping teacher preferences, with particular attention to the challenges faced by schools in disadvantaged areas. Given the ongoing teacher shortages, particularly in secondary education and high-need subjects, understanding these preferences is crucial for designing effective policies to attract and retain teachers. The review synthesises findings from national and international studies, including survey data, discrete choice experiments, and meta-analyses, to provide a comprehensive overview of the key determinants of teacher job choices.

Teacher preferences for different job attributes

England faces significant challenges in teacher recruitment and retention, especially in secondary education shortage subjects amid rising pupil numbers (Allen et al., 2023). Latest data from the School Workforce Census (SWC) shows that 9.7% of teachers left teaching in 2021/22 (Department for Education, 2024a). However, larger numbers of teachers are thinking of leaving the profession according to two surveys. Wave 2 of the Working Lives of Teachers and Leaders reports that 36% of teachers are considering leaving teaching in the next 12 months (Department for Education, 2024b). A Teacher Tapp survey shows only 63% of teachers think it is likely that they will still be a teacher in three years' time (Allen et al., 2024). Initial teacher training (ITT) recruitment continues to fall below what is needed to ensure sufficient staffing levels in schools (McLean et al., 2024). Three-quarters of senior secondary school teachers said they have experienced difficulties with at least one appointment they were involved in during 2024 (Allen et al., 2024).

Salary

Salary is a key factor in teacher retention, and the real-terms decline in teacher pay since 2010/11 has likely contributed to ongoing shortages in England (Worth et al., 2018). Several recent studies have documented the effectiveness of financial incentives, particularly bursaries, in recruiting and retaining teachers in high-shortage subject areas. McLean et al. (2023) find that increases in bursaries are positively associated with both teacher recruitment and retention, thereby supporting long-term teacher supply. Their analysis indicates that "*a £10,000 increase in a subject's bursary was associated with an 18 per cent increase in the number of trainees in that subject*" (p. 7), though with substantial variation across different cohorts. This aligns with findings from Benhenda and Sims (2022), who examined targeted salary uplifts for early career maths and physics teachers. They found that teachers eligible for these payments were 23% less likely to leave state-funded schools during the years they received them. Similar patterns have been observed in the United States, where early career salary incentives have been linked to improved retention rates in shortage subjects (Feng and Sass, 2018; Bueno and Sass, 2018). However, a recent systematic review concluded that while financial incentives show promise in attracting teachers to the profession and to specific regions, subjects, or hard-to-staff schools, the evidence remains inconclusive (See et al., 2020a).

Survey evidence on teacher pay presents a more mixed picture. Data from the TALIS 2018 survey suggest that around half of teachers in England were dissatisfied with their pay. However, compared to teachers in other countries, responses from those working in England were relatively favourable (Jerrim and Sims, 2019). More recent surveys indicate increasing dissatisfaction. In the 2023 *Working Lives of Teachers and Leaders* survey, almost three-quarters of teachers reported being dissatisfied with their salary, with a similar proportion expressing concerns about their long-term pay prospects compared to alternative career paths (Department for Education, 2024b). Evidence from Teacher Tapp (2023a) suggests that when given a choice between a 10% pay rise and a 10% reduction in workload, 57% of teachers preferred the pay increase, while 43% prioritised the workload reduction. Notably, even when the salary increase was reduced to 3%, 39% of teachers still opted for the financial incentive.

Studies using conjoint analysis that are similar to this study provide further insights into teacher job preferences. Burge et al. (2021) find that teachers exhibit loss aversion (i.e. greater sensitivity to salary reductions than to equivalent gains). For

instance, while a 1% increase in final pension was valued as equivalent to a 0.55% increase in annual salary, a 1% reduction in pension required an additional 1.67% salary increase to compensate. Similar studies in the US have examined teacher preferences for pay structures. Johnston (2024) finds that more effective teachers are more likely to favour pay structures incorporating performance-related elements, while Fuchsman et al. (2020) report that teachers only become significantly concerned about the type of pension scheme as they gain more experience.

Workload

Over half of teachers in England report having an unmanageable workload (Jerrim and Sims, 2019), which is also cited as a major reason for leaving the profession (Perryman and Calvert, 2020). In a conjoint analysis of job preferences, Burge et al. (2021) found that reducing workload was highly valued by teachers. Their study estimated that each 1-percentage-point increase in workload would require a 2.72% increase in annual pay to compensate, whereas a 1% reduction in workload was valued as equivalent to a 0.77% increase in annual pay.

However, longitudinal quantitative studies directly linking workload concerns to teacher retention are relatively rare. A notable exception is Sims and Jerrim (2020), who used data from the TALIS 2018 study for England, linked to administrative records, to examine how various working conditions (including views on workload) were associated with teachers leaving their school or the profession by the following academic year. Contrary to conventional wisdom, they report that "workload does not show a robust relationship with either turnover or job satisfaction" (p. 7). However, other studies have found at least a correlation between teachers' perceptions of their workload and their intentions to remain in their current school. For instance, using cross-sectional data, Jerrim (2024) found that each standard deviation increase in teachers' satisfaction with their workload was associated with a 0.07 standard deviation increase in their likelihood of staying at their school. Teacher Tapp (2023b) also reports an association between teacher wellbeing and reported working hours.

Flexibility

Many teachers highly value flexible or part-time work (Worth, 2023; Sharp et al., 2019), but this is challenging to implement. However, with other competing jobs now offering greater opportunities to work from home since the COVID-19 pandemic, flexible working has become a key barrier to recruiting and retaining teachers. Recent research by the NFER indicates that while flexible working arrangement amongst teachers may be on the rise, less than one-in-six report being able to take planning, preparation and assessment (PPA) time off-site, are able to start late or finish early, or are able to take ad-hoc personal days off (McLean et al., 2024). They thus note how the demand for such flexible working arrangements currently outstrips supply, with surveys suggesting that it is highly valued by employees. This is supported by evidence from the conjoint analysis conducted by Burge et al. (2021), with the ability to move from full-to-part-time (if requested) being worth the equivalent of around a 4% pay rise. That said, only a minority of teachers would prefer a teaching job with much shorter working hours but with more conventional holiday arrangements, and very few would be willing to trade two days of school holiday for an additional one day to be taken at their time of choosing (Allen et al., 2023).

Qualitative research conducted in England has, however, warned that "*Increasing opportunities for flexible working may have a role in helping to retain teachers in the profession, but offering such opportunities without addressing fundamental issues around teacher workload is unlikely to have a significant impact*" (CooperGibson Research, 2018, p.7). And clearly there are major organisational hurdles in implementing flexible working without causing issues for students and colleagues, which we return to below.

Professional development opportunities

While a large-scale meta-analysis of various studies finds a limited relationship between professional development and teacher retention (Gundlach et al., 2024), discrete choice experiments suggest that teachers place some value on professional development as part of their job package.

One study in Costa Rica found that increased peer and pedagogical support had only minor marginal effects—between 1% and 3%—on teacher preferences. However, when combined with higher pay, these non-monetary incentives were more effective in encouraging teachers to move to high-need contexts (Lentini et al., 2024). In the US, Lovison and Hyunjung Mo

(2024) found that teachers were less attracted to an offer of one hour per week of instructional coaching compared to other job attributes, such as the availability of in-school counsellors and nurses. However, they concluded that instructional coaching represents good value for money, as the perceived benefit to teachers exceeds its cost.

In England, Burge et al. (2021) found that teachers were only willing to trade a very small salary increase (0.43%) for an additional professional development day per year. Their findings also indicate that teachers preferred professional development delivered within their school and among colleagues, rather than external training chosen and attended independently. Headteachers and leading practitioners placed a particularly high value on professional development compared to classroom teachers, a preference consistent with survey findings reported by Allen et al. (2024). These findings also highlight that teachers strongly prefer to have autonomy over their professional development choices.

Policy and leadership

There are of course a multitude of other factors associated with the recruitment and retention of teachers, many of which have been investigated in the literature. For instance, when asking teachers why they are considering leaving the teaching profession over the next year in a survey, Adams et al. (2023) found that 76% of teachers said government policy changes were a factor - with only concerns over workload (92%) being a more pressing factor. Issues surrounding accountability pressures (69% of teachers) and managing parental complaints (35%) were also key issues named.

Nguyen et al. (2023) reviewed the evidence base on school leadership, culture, climate and structure for teacher retention. They highlight three interrelated leadership approaches and their associated practices to support teacher retention: (i) prioritising teacher development; (ii) building relational trust; and (iii) improving working conditions. The review also underscores four prominent characteristics of school culture, climate, and structure that promote collegiality, positive school discipline, intellectual stimulation, and equity in workload arrangements and support distribution.

In a meta-analysis of 185 studies by Gundlach et al. (2024), found that while workload and time pressure were only weakly associated with intentions to leave the profession, role conflict showed a much stronger association. In the same meta-analysis, perceived threats to safety had a strong association with their decision to stay in a school. Overall resourcing levels, professional development, administrative support, staff-student ratios and pedagogy had little consistent effect. Overall, teachers reported measures of school leadership, culture and collegiality were associated with intentions to stay in a school. We return to some of these factors below.

Recent research on collegiality has suggested that each standard deviation increase in a school staff member's relationship with their colleagues is associated with around a 0.15 standard deviation increase in their organisational commitment (Jerrim and Sims, 2024). This is consistent with international evidence, which has indicated that interpersonal relationships within school settings are a key driver of staff outcomes (Campoli and Conrad-Popova, 2017; Kurtz and St Maurice, 2018).

Other factors

Several additional factors influence teacher job preferences beyond salary, workload, job flexibility, leadership and professional development opportunities.

Pupil behaviour is a significant concern for many teachers in the UK. A majority express a preference for their school to take firmer action on student misbehaviour (Teacher Tapp, 2023a), and more than half report considering pupil behaviour when deciding where to work (Menzies et al., 2015). Behavioural climate appears to be a particularly strong driver of early-career teacher movement between schools and exits from the profession. Analysis of US data found that a one standard deviation increase in first-year teachers' perceptions of poor behaviour was associated with a nearly sixteen-fold increase in their likelihood of leaving the profession, as well as a more than twofold increase in their likelihood of switching schools (Kukla-Acevedo, 2009). Burge et al. (2021, p. 41) further highlight the strong influence of behaviour on teachers' job decisions, estimating that teachers would require a 40.6% increase in salary to compensate for moving from a school where *"poor behaviour is rarely a problem"* to one where "*poor behaviour from a few students significantly disrupts most lessons*". Viano et al. (2021) find that teachers place high value on consistent enforcement of discipline, sometimes even exceeding the value placed on salary increases or class size reductions.

While smaller class sizes are often assumed to be a key job attraction factor, studies in the US have suggested they are relatively low priority compared to other aspects of school support. Lovison and Hyunjung Mo (2024) and Viano et al. (2021) find that teachers are more likely to prefer jobs in schools offering reduced class sizes, but this was among the least valued job attributes, ranking below access to school nurses, counsellors, and other support staff. Johnston (2021) estimates that a reduction of one pupil per class is valued as equivalent to a $595 salary increase. Similarly, a study of primary teachers in California found that working conditions—including facilities, administrative support, and class size—were more influential in shaping workplace preferences than student characteristics such as socioeconomic background, ethnicity, or achievement levels (Horng, 2009).

Finally, childcare benefits are particularly valued by teachers with children. Lovison and Hyunjung Mo (2024) find that childcare subsidies act as a near-perfect substitute for a salary increase of equivalent value. Given that childcare costs in the UK are among the highest internationally (Statham et al., 2022), such benefits may be especially attractive in this context.

Headteacher preferences regarding employment practices and teacher job candidates

The literature on how headteachers approach teacher recruitment, the employment conditions they prefer to offer, and the types of candidates they favour remains relatively limited, with most studies being descriptive or correlational (e.g., Lynch et al., 2024; Robinson & Gray, 2019). However, several strands of research provide insight into headteachers' hiring practices and preferences, including recruitment strategies, employment incentives, and the attributes they value in teachers.

Recruitment strategies and employment incentives

Engel & Curran (2016) examined how 31 Chicago principals approached recruitment, screening, and selection. They identified common strategies such as early hiring, appointing substitute or student teachers, and leveraging personal networks. However, strategic recruitment was relatively rare and appeared more common in higher-performing schools—though findings were limited by the study's small sample. Similarly, Jabbar (2018) explored recruitment in New Orleans charter schools, where greater flexibility allowed for more varied hiring strategies. While many school leaders supported salary enhancements as an incentive, financial constraints led them to offer alternative benefits such as performance-based bonuses or extra pay for additional duties.

Chong et al. (2024) took a different approach, analysing 908 job adverts from schools in disadvantaged areas of England (Education Investment Areas). This study provided indirect evidence of headteachers' preferences by identifying the most frequently advertised job attributes, including professional development (79.5%), healthcare benefits (30.6%), flexible working (14.1%), workload reduction initiatives (12.9%), behaviour policies (5.3%), class size reductions (5.1%), and childcare support (2.6%). While these findings highlight a range of strategies used to attract teachers, they do not indicate the relative importance of each policy, or the trade-offs involved.

Headteacher preferences for teacher characteristics

We know of only two discrete choice experiments (DCEs) that have explored headteacher preferences for teacher attributes. Giersch & Dong (2018) surveyed 467 North Carolina principals, finding a preference for candidates with advanced degrees and those from highly selective institutions. A master's degree or prestigious undergraduate background increased the likelihood of selection by over 7 percentage points. Experience was also valued, but only up to five years—beyond which additional years (up to 15) did not further increase preference. Principals favoured candidates demonstrating innovative instruction, alignment with the school's mission, and engagement with current research. These findings align with broader psychological research on 'fit,' which suggests that teachers who align with a school's culture and working style are more likely to succeed (Perrone & Meyers, 2023).

A second DCE by Johnson et al. (2024) surveyed 170 Christian school leaders across multiple countries. It found that leaders strongly preferred candidates with extensive teaching or leadership experience and a strong academic background. They also valued candidates who had attended Christian colleges or previously worked in similar school settings, illustrating how hiring preferences can be shaped by the school's specific ethos and educational philosophy. This highlights the importance of studying headteacher preferences within specific national and institutional contexts—such as the UK, where no equivalent DCE has yet been conducted.

Stated preferences and broader trends

Several studies have explored headteachers' self-reported hiring preferences, though these findings may be influenced by social desirability bias or post-hoc rationalisation (Engel, 2012; Davis et al., 2021). A review by Davis et al. (2021) identified a diverse set of criteria considered by headteachers, including appearance, academic credentials, demographic factors (e.g. race and gender), and relational qualities (e.g. rapport with students and colleagues). Engel (2012) found that, compared to leaders in higher-achieving schools, those in lower-achieving schools prioritised classroom management skills and a teacher's ability to improve student test scores—patterns that were also reflected in Giersch & Dong's (2018) findings.

Other research has documented variability in headteacher preferences. The Preferred Teacher Applicant Characteristics Survey (PTACS), developed by Davis et al. (2021), revealed significant heterogeneity in hiring preferences, with female headteachers placing greater emphasis on candidates' personal characteristics. Ingle et al. (2021) further demonstrated that urban school leaders valued personal qualities and professional qualifications above student outcomes or applicant demographics. Meanwhile, qualitative studies reinforce the idea that different school contexts lead to different hiring priorities (Ingle et al., 2011).

A recent review by Perrone & Meyers (2023) summarised key trends in headteacher hiring preferences, concluding that principals generally favour teachers who are fully certified through traditional training routes and have some classroom experience. However, less experienced teachers were sometimes considered more adaptable. Enthusiasm and passion for teaching were also highly valued, as they were seen as foundational qualities that could support professional growth.

Willingness to work in schools serving disadvantaged communities

Schools serving disadvantaged communities in England face greater challenges in recruiting and retaining high-quality teachers than those in more affluent areas. This disparity in access to experienced and qualified teachers likely contributes to the persistent attainment gap between pupils from higher- and lower-income families. A study of the School Workforce Census by Allen and Sims (2018) found that in primary schools with the highest concentration of Free School Meals (FSM) pupils, 4% of teachers are unqualified, compared to 2% in the most affluent schools. In secondary schools, the figures rise to 9% and 5%, respectively. Additionally, schools with more affluent intakes have 12% of teachers with over ten years of experience, whereas the most disadvantaged schools have just 7–8%. This "expertise gap" is particularly pronounced in STEM (science, technology, engineering and maths) subjects. For example, at Key Stage 4, the proportion of teachers without a relevant academic degree is 10 percentage points higher in Maths, 14 in Chemistry, and 22 in Physics in schools serving the most deprived areas.

A substantial body of evidence, both from England and internationally, indicates that schools serving disadvantaged communities face significantly greater challenges in recruiting and retaining teachers compared to those in more affluent areas. Allen and Sims (2018) found that schools in more deprived areas experience higher levels of teacher turnover, with secondary teachers in the highest deprivation quintile being 70% more likely to leave than those in wealthier schools. Recent research by the NFER (Worth & Faulkner-Ellis, 2022) further supports this, with senior leaders in high-FSM schools rating recruitment challenges at an average of 6 on an 8-point scale (where 1 = no challenge and 8 = a great extent), compared to just 1.5 in the most affluent schools. Additionally, Sibieta (2020) found that disadvantaged schools struggle more to fill vacant posts and report 50% more staff sick leave per year on average. Allen and McInerney (2019) revealed that 85% of teachers in schools serving disadvantaged communities believe recruitment difficulties negatively impact the quality of education, in stark contrast to just 18% in independent secondary schools. More recent research by Allen et al.

(2024) highlights that schools in deprived areas frequently extend application deadlines, experience higher candidate withdrawal rates before contract signing, and report greater mid-year disruptions due to teacher absences or resignations. These schools also rely more heavily on non-specialist teachers in GCSE classes. The 2023 Science Teaching Survey by the Royal Society of Chemistry (2023) shows that 48% in high-FSM schools reported understaffing issues compared to 32% in lower-FSM schools, while 42% cited high science staff turnover compared to just 25% in more affluent settings. Jabbar and Holme (2025) describe how this teacher turnover depletes organisational social capital by disrupting networks and relationships between teachers working in teams, weakening shared meanings and goals in schools, and impeding teachers' ability to collectively engage in problem-solving and learning. Taken together, these findings highlight a persistent and systemic disadvantage faced by schools in deprived communities, affecting both teacher stability and student outcomes.

Teachers across all school types recognise that working in disadvantaged schools is more demanding and requires greater skill; however, Allen et al. (2024) found that while workload differences between school types are minimal, the classroom experience varies significantly. Teachers in disadvantaged schools report more frequent behavioural problems, with disruptive behaviour more likely to halt learning compared to more affluent schools. These challenges contribute to lower job satisfaction and higher stress levels, with teachers in deprived schools experiencing more frequent burnout and stronger feelings of frustration, particularly in secondary settings. The Royal Society of Chemistry (2023) further highlights the difficulties faced by science teachers in high-FSM schools, particularly regarding student literacy and numeracy. In these schools, 75% of science teachers identified literacy as a significant challenge, compared to 50% in lower-FSM schools, while 69% reported difficulties with numeracy, compared to 48% in more affluent settings. Poor attendance further exacerbates these issues, with 73% of science teachers in high-FSM schools citing it as a problem, compared to 47% in lower-FSM schools. These findings illustrate the compounded difficulties teachers face in disadvantaged schools, where classroom disruptions, academic barriers, and lower student attendance create an environment that is both more demanding and more stressful than in more affluent settings.

The TALIS 2018 survey in England found that 65% of primary and 61% of secondary teachers strongly agreed that they entered teaching to contribute to society (Jerrim and Sims, 2019). Around 85% also agreed or strongly agreed that they joined the profession to benefit socially disadvantaged students. Taken at face value, these responses suggest that many teachers would actively seek opportunities to work in disadvantaged schools.

However, responses to such questions may suffer from social desirability bias—where individuals answer in ways they believe will be viewed favourably rather than truthfully. Research shows people tend to underreport undesirable behaviours (e.g., drug use) and overreport desirable actions (e.g., charitable giving, voting) (Krumpal, 2013). In teacher surveys, this bias may lead to an overstatement of willingness to work in challenging schools.

Teacher Tapp (2024a) examined teachers' preferences for working in high-FSM schools using the Unmatched Count Technique (UCT) to elicit more candid responses. Rather than answering directly, teachers indicated how many statements they agreed with from a list, some including a sensitive item like, *"I would prefer not to work in a high-FSM school."* Results suggested greater reluctance than direct questioning revealed. While 25% openly admitted a preference against working in high-FSM schools, UCT estimates indicated this was likely an understatement. Similarly, when asked directly, 21% of teachers felt they would be less effective in such schools, but this figure rose to 32% when assessed indirectly. These findings highlight the impact of social desirability bias in teacher surveys and suggest hesitations about working in disadvantaged schools are more widespread than self-reports imply.

A small experimental study in the US showed that all teachers rated job advertisements from schools with lower standardised test scores more poorly, and that this was particularly true for experienced teachers relative to inexperienced teachers (Winter and Melloy, 2005). This is consistent with the perspective that new teachers may have altruistic intentions to work in more challenging environments but that they develop an understanding of the personal cost of doing so over time.

The evidence on how to encourage teachers to work in disadvantaged schools remains limited, with few high-quality experimental studies. However, quasi-experimental policy interventions in the United States, France, and Brazil have

provided insights into the salary increases needed to reduce turnover in high-poverty schools (e.g., Camelo & Ponczek, 2021; Clotfelter et al., 2008; Prost, 2013).

Of particular relevance to our project is a recent study by Lentini et al (2024). They use a conjoint survey experiment with 400 teachers to study the design of incentives to attract teachers to low-performing remote schools in Costa Rica. Results show that up to one third of teachers could be induced to move to remote schools based on a combination of increased salary and other in-kind (housing and transport) incentives. Viano et al (2021) also use conjoint analysis to study 'what teachers want' when it comes to attracting them to work in low-performing schools. They surveyed 811 teachers working in Tennessee schools and found that teachers placed the greatest weight on pupil discipline, salary, leadership support and safety.

A systematic review by See et al. (2020a; 2020b) examined international research on attracting and retaining teachers in hard-to-staff areas, finding that financial incentives are the only approach with strong evidence for attracting teachers to challenging schools, though they are not effective in retaining them. Long-term retention requires a supportive and conducive working environment, while other interventions such as mentoring, additional support, or professional development lack robust evidence of effectiveness, primarily due to weaknesses in the existing research.

Appendix B - RQ1: Focus groups / cognitive testing of job attributes

This appendix describes how focus groups were used to develop the wording of the job attributes for the paired conjoint survey questions. A desk review of literature preceded the focus groups and was used to draw up some prospective wording for the job attributes. Following the focus groups, further refinement of the job attributes took place on the Teacher Tapp survey app, as described in Appendix B.

Campoamor et al. (2024) report that despite pre-testing being an essential stage in designing discrete choice experiments (DCEs), "*fewer than one-fifth of DCE studies report including pretesting in their development*" (2024, p.109, p116). Recent DCEs exploring teacher choices have followed in this vein. For example, Burge et al (2021) report modifications to their study design based on piloting, but not pre-testing. Similarly, Horng (2009) reports having conducted consultation and piloting in advance of her conjoint study of teacher preferences, but not detailed pretesting. She therefore acknowledges that one of the limitations on her findings was that "*different respondents may have interpreted the terminology of the survey differently - for example, 'administrative support' is likely to mean different things to different teachers*" (2009; p705). This is one of the risks our pre-testing process was designed to mitigate.

More broadly, pretesting is a valuable step in increasing validity, reliability and relevance; reducing bias, burden and error; contributing to hypothesis generation; and supporting interpretation (Campoamor 2024). We therefore conducted four focus groups before piloting our survey to aid the development of the research instruments and to aid subsequent interpretation.

Recruitment to focus groups

Participants were recruited via an advert that ran on Teacher Tapp for two days in early August 2024. The advert targeted classroom teachers and middle leaders in state schools, excluding head teachers, senior leaders, and those working in independent schools. As a result, 7,433 potential participants were invited to respond.

A total of 77 individuals responded, providing details on their years of experience in the profession, gender, whether their school was urban or rural, and whether they had parenting responsibilities. Responses from primary school teachers were lower, so all 21 primary teachers who responded were invited to participate. Meanwhile, 26 of the 56 secondary school respondents were randomly selected, with the sample reviewed to ensure a balanced mix of characteristics. Selected individuals were then invited and asked to provide consent to participate.

Due to the particularly large number of participants in the final focus group, the decision was made to split the group. As a result, two breakout rooms ran simultaneously.

Table A1: Number of participants in pilot testing focus groups for RQ1

Session		N Attendees
1 (Primary)		4
2 (Secondary)		6
3 (Primary)		2
4 (Secondary)	Breakout A	5
	Breakout B	4

Methodology

Pretesting was intended to:

- Refine attributes and statements
- Improve clarity and participant satisfaction
- Identify any gaps or sources of bias
- Highlight any unanticipated problems

Rather than the more common approach of interviews, online focus groups were used to allow a larger group of teachers to participate within a limited timeframe. Campbell (2009, p.2) notes that while focus groups are sometimes used for cognitive interviews in survey questionnaire design, they are generally more effective for "identifying overall themes and issues… rather than specific question problems." This is often due to group dynamics, including the influence of social desirability bias, which was observed during our focus groups. These limitations are discussed below. However, we judged that the next stage of questionnaire design—extensive pilot testing—would adequately mitigate this drawback.

The focus group script was designed to explore four "domains of assessment" (Campoamor et al., 2024):

- **Content**: Assessing relevance and comprehensiveness - i.e. whether the chosen attributes accurately captured the areas teachers consider when selecting roles.
- **Comprehension**: Evaluating whether participants could "*envision the proposed scenario and decision conte*xt" (p111).
- **Presentation**.
- **Elicitation**: including exploring trade-offs and underlying heuristics.

Cognitive testing is a commonly used technique for pretesting, as it helps to understand respondents' thought processes when answering questions and the reasoning behind their responses. The focus groups we conducted did not constitute a full cognitive pretest, as the survey had not yet been fully drafted. However, we incorporated several techniques commonly used in cognitive interviewing, including:

- **Think-aloud and verbal probing**: The interviewer observes participants as they make their choices and either asks them to verbalise their thoughts or follows up with targeted questions after each response (Patel-Syed et al., 2024), though the two approaches can be used in combination (Boeije and Willis, 2013).
- **Debriefing**: This involves gathering feedback from participants after completing a task, to explore their thought processes, identify any difficulties encountered, and refine the survey instruments accordingly.
- **Codesign**: This refers to the collaborative development of survey materials, where participants contribute insights and suggestions to improve the wording, structure, and relevance of survey items.

Two rounds of focus groups were conducted, with the instruments refined between rounds, following discussions within the research team. Focus groups followed a semi-structured protocol using a visual stimulus (such as lists of attributes or example job adverts), along with prompts related to credibility, interpretation, and understanding. Where issues were identified, participants were asked to suggest alternatives, which were then reviewed by the research team.

The two rounds of focus groups had slightly different foci. **Round one** (which included separate sessions for primary and secondary school teachers) focused on cognitively pretesting the proposed attributes and statements, particularly in terms of content and comprehension (see above). We prioritised 'soft attributes'—such as school culture and ethos—over 'hard attributes' like pay levels, since it is particularly challenging to develop statements for the former whereas more concrete attributes could be tested as part of the pilot.

Round two (again involving two sessions) provided an opportunity to review statements that had been substantially revised. However, the primary focus was on the choice task. Participants were presented with example job adverts that combined four attributes, and discussions centred on the clarity of presentation and participants' decision-making processes—specifically, how they made their choices, the trade-offs involved, and the information or heuristics they relied upon.

At the start of each focus group, the facilitator explained that the task was a pretest, outlining what this involved and clarifying that responses would be used to improve the study. Throughout the discussion, participants were reminded of the real-world contexts in which they would be making their choices. For example:

> *"I'm going to show you a pair of adverts just as you'd see them in the app. It will appear like a phone screen, and I'd like you to read them and then post in the chat whether you would choose School A or B. Then, we'll talk through how you made that decision, and I'll ask you some questions about it—if that's okay—as well as how you found the task. We're trying to make this as realistic as possible. I know it's not quite the same as doing it on your phone, but try to picture yourself doing this at 3:30 [the time at which the Teacher Tapp app pushes out questions]. Think about how long you would realistically spend on it when doing it for real."*

As noted below, participants explicitly acknowledged that they were making choices differently during the discussion compared to how they normally would. However, a key advantage of recruiting regular Teacher Tapp users was their familiarity with the platform. Prior experience allowed them to make explicit, informed comparisons between the focus group setting and the typical decision-making process when using the app.

Participants were then asked to introduce themselves and describe what they look for when choosing a job. This served both as an icebreaker and as an initial opportunity to explore the relevance and comprehensiveness of the attributes included in our planned study.

Detailed notes were taken throughout the focus groups, and discussions were recorded and transcribed, except for one of the sub-groups in the final focus groups, which was not recorded.

The analysis began with an initial attribute-by-attribute and thematic write-up based on written notes. This focused on immediate actions, such as refining attributes for the second session and identifying any further changes needed before piloting. Unresolved questions requiring testing during the pilot phase were also noted.

A more detailed analysis followed, involving a thorough review of transcripts to refine the thematic analysis and identify key excerpts that exemplified emerging themes. Particular attention was given to:

- Relevance and representation
- Credibility
- The choice process
- Comprehensibility, clarity, and enjoyability
- Ecological validity

An additional goal of this second stage of analysis was to identify insights that could inform the subsequent interpretation of quantitative results.

Findings

Evaluation of attributes

The first part of each focus group aimed to determine whether the selected attributes were relevant and adequately represented valued attributes of a job.

The most commonly referenced theme among participants was school ethos, and this encompassed a range of considerations, including:

- Professional culture: Collegiality, as well as engagement with professional development and research.
- Leadership style and structure: For example, flat hierarchies, 'open' leadership, and staff autonomy.
- Student-centred culture: Emphasis on holistic education and extra-curricular activities.
- Well-being and workplace happiness: Often linked to low staff turnover.
- Links to the local community.

Participants explained that when seeking a job, ethos and school composition were often judged based on the 'feel' of the school during a visit. This ties into questions of ecological validity, which are explored in greater depth below.

Other important themes included the school and community demographics, working conditions (such as workload and flexibility), and student behaviour.

While these themes largely aligned with the attributes selected for the DCE, some possible omissions were identified:

- Leadership and management structure: In practical terms, whether the organisational hierarchy was flat or more rigid.
- Structural considerations: Whether the school was part of a Multi-Academy Trust (MAT), a Local Authority (LA) school, or a church school.
- Uncaptured aspects of ethos: Approaches to extra-curricular provision, community outreach, and what the school prides itself on.
- School composition and demographics.

Additionally, secondary school teachers raised concerns about timetabling, specifically which subjects and classes they would be expected to teach—though this was also relevant for one primary school subject specialist. For example, RE teachers wanted to know whether they would also be required to teach other humanities subjects, while others sought clarity on whether they would be teaching A-Level, GCSE, or Key Stage 3 classes.

Evaluation of draft statements

Numerous refinements were made to the draft statements in response to participants' feedback.

Applicability to different participants' circumstances

For example, references to 'time off for childcare' were broadened to 'family commitments,' as participants without dependents felt they should have equivalent allowances for other caring responsibilities, such as supporting a partner or elderly relative.

Red flags and green flags

Certain terms were unexpectedly interpreted negatively, while others proved particularly positive and salient. For instance, references to 'private healthcare' were seen as either ideologically problematic or as a sign that a school was misallocating resources:

> *"Given how budgets are at the moment, I'd be a bit horrified if I saw this in an advert, if I'm really honest… It would make me wonder what the children weren't getting."*
>
> Primary teacher

Meanwhile, one participant suggested that overly generous onsite childcare could signal an expectation of unreasonable workload. Similarly, statements about 'reduced marking expectations' were seen as potentially indicative of rigid, top-down policies that lacked evidence. References to dedicated staff for pupil behaviour management were also interpreted as a possible indicator that behaviour was a problem:

> *"It looks nice to potentially know that there is a specialist role, shall we say—however, it might then make you question, 'what is the behaviour like on the ground at that school if they need to have a particular member of staff with that as their focus?'"*
>
> Primary teacher

Other statements related to behaviour were perceived as suggesting either a disciplinary school culture or an approach where responsibility was not shared among staff, undermining the much-valued sense of collegiality.

Conversely, some terms emerged as 'green flags', disproportionately valued by participants. These included 'protected' PPA time and 'self-directed' professional development, which were incorporated more explicitly into revised statements.

Identifying 'red' and 'green' flags was particularly important given our study design, as participants would not have the opportunity to investigate concerns arising from the adverts, in the way they might in real-world scenarios. One participant noted that, in practice, a potential red flag would serve as a starting point for discussion during an interview or visit:

> *"It also kind of provides you with the springboard to a question when you go and visit. So, 'what kind of backup are we talking about? Is that backup to do with behaviour in class? Is it to do with argumentative parents? …I think the fact that it's worded that way gives you the potential to be guided in the way that you would ask a question if you were at interview… It might raise a red flag, but it also gives you ideas for what to ask to determine whether that is a red flag or whether it's a green flag, so to speak."*
>
> Secondary teacher

Credibility

Many participants doubted the credibility of certain statements, questioning whether they would realistically translate into practice. As noted below, this could skew the choice process by masking desirable but non-credible job attributes. The degree of scepticism or trust varied between participants and was an important source of heterogeneity, explored further below.

We directly compared the credibility of three potential sources of evidence that might enhance the perceived reliability of intangible attributes (such as collegiality and pupil behaviour):

- Pupil and parent surveys
- Staff surveys
- Ofsted reports

Of these, only pupil and parent surveys were widely considered credible and were therefore prioritised in revised statements:

> *"I would be more curious about what the parents say because that is going to be very truthful."*
>
> Primary teacher

We also explored whether a school's professional development spend per teacher was a better indicator of its commitment to Continuing Professional Development (CPD) than time allocation. Participants strongly favoured time allocation, given the variability in programme costs:

> *"It's very difficult because you can go on a course that's incredibly expensive and it would wipe out the entire budget for the year, or you could be involved in a research project and meeting with colleagues from other schools."*
>
> Primary teacher

More was not always seen as better—excessive CPD pledges were viewed with suspicion. One participant suggested that overly frequent CPD might be more likely to be cancelled and could lead to additional cover requirements, making a middle-ground option more appealing. However, this effect may be less pronounced in an experimental setting, where decisions are made more rapidly and instinctively.

As noted earlier regarding 'green flags', certain phrases or promises made statements more credible, including 'protected' PPA time, 'no-cover' expectations, and concrete measures of staff turnover.

Heterogeneity

There was significant variation in participants' preferences, with some teachers ranking different versions of the same attribute in unexpected ways. For example, a statement intended as the 'gold standard' (e.g., *"The school is committed to cutting workload, including through additional protected PPA time for marking"*) was sometimes ranked below a more basic version (e.g., *"The school seeks to support all staff with manageable workloads"*).

Participants often directly contradicted each other, reflecting different approaches to interpreting statements and making decisions. Some exhibited 'high-trust' decision-making, valuing general statements of intent, whereas others were more cynical, becoming suspicious if a statement seemed too good to be true or lacked specificity.

> *"It's really, really prescriptive. So I feel like I wouldn't be disadvantaged... So that feels really good."*
>
> Secondary teacher

Some of this heterogeneity related to personal circumstances or subject area. For example:

- Marking expectations differed between Maths and English teachers, shaping their views on marking policies
- Childcare provisions were viewed differently depending on participants' children's ages.
- Prescriptive CPD approaches were seen as more appropriate for early career teachers, who were also less likely to be sceptical of Ofsted reports or CPD quality due to lack of experience
- Additional PPA time could be misinterpreted by early career teachers as part of their existing timetable reductions.

A recurring difference emerged between senior leaders and classroom teachers. Senior leaders were more aware of the practical challenges of implementing generous policies, such as flexible time off for family commitments. This underscored the need for subgroup analysis in the main experiment and validated the decision to exclude headteachers and deputy heads from the final sample.

Ambiguity and Specificity

Some broad statements were misinterpreted. For example, *'supportive school community'* was assumed by some participants to refer to parents raising money for the school. Minor wording adjustments were made to clarify intended meanings, such as changing this to *'collaborative, friendly staff'*.

In some cases, think-aloud responses revealed that statements meant to represent different levels of the same attribute were actually being interpreted as entirely separate attributes. This was particularly evident with statements about school environment and ethos, likely due to the multiple dimensions involved.

For example, some original statements combined elements of inclusive leadership and collegiality (Sims, 2021), rather than treating them as distinct attributes:

> *"For me as an individual, that sense of community is quite an important part. So I quite like the first one. However, being in a school where there is a low turnover and the leadership actually has a level of inclusivity is just as important. So it'd be a bit of a tricky one."*
>
> Primary teacher

To address this, alternative statements were restructured to focus on a single attribute at a time.

Choice-task

The second round of focus groups focused on understanding participants' decision-making processes when presented with adverts containing multiple statements representing different attributes.

Choice process

Participants approached their choices in different ways. Some read both adverts quickly before going back to examine each in greater depth. Others read the two separately and selected the one that appealed to them most overall. For some, the decision was made quickly, based on gut reaction—either to the overall impression of the school (see below) or to a particularly appealing or salient detail, such as salary. Others took a more systematic approach, comparing each advert bullet point by bullet point.

The fact that statements related to the same attribute were placed side by side (as would typically be the case in a tabulated DCE) may have encouraged this more analytical strategy.

In both focus groups, some participants initially made a choice but later wanted to switch after reading the adverts more carefully or after responding to a second pair of adverts. One teacher wished there had been a back button to revise their first response, while another noted that if they had spent more time thinking, their response would have better reflected their true preferences. Several participants suggested they might have taken more care if an opening blurb had explicitly encouraged them to read carefully or had provided more information about the study's purpose and the number of tasks involved.

Participants often made holistic judgements about what kind of school an advert described, interpreting statements as symbolic of the school's ethos. For example, one participant found an advert appealing because it seemed carefully put together and focused on attracting the right teacher, which they felt reflected well on the school, regardless of the actual content of the statements. Similarly, references to staff surveys were viewed as evidence of a consultative leadership approach, which was appealing in itself, separate from the survey results:

> *"The fact that the school is actually doing those parent and pupil surveys shows their interaction... their commitment to the whole school community."*
>
> Primary teacher

Statements also interacted with each other, with combinations of statements—such as flexible working, generous PPA time, and comprehensive childcare—being interpreted as an indication that the school genuinely prioritised work-life balance. Analysis should therefore examine whether statements have cumulative or combinatorial effects.

Several times, participants assumed certain schools were private schools and rejected them on that basis—though different participants identified different schools as being private. To counter this, the introductory blurb was adjusted to clarify that all schools in the study are state schools and within reasonable commuting distance.

Because participants often evaluated adverts holistically or symbolically, they did not always respond to different quantities of a benefit as expected. Rather than valuing twice as much of X as twice as good, they tended to think in terms of "*this school has shown it's interested in X*" rather than focusing on the precise amount.

"I'm very reactive with Teacher Tapp... when it pops up at 3:30 on my way home. I just happily click. I don't really think about it, to be honest. I'd certainly look at one day per half term and think, 'actually, that sounds like more hours than the other two, and I'll go for that one.'"

Primary teacher

A contrasting approach was also evident, where some participants skimmed over points they found irrelevant to their circumstances, while others fixated on a single bullet point that was particularly salient to them—such as a statement about health and wellbeing or childcare, depending on their personal situation.

As noted earlier, statements that were not seen as credible were often skipped. In some cases, 'mid-level' statements proved more popular due to their perceived credibility. For example, one participant even dismissed a concrete offer of a 15% pay rise because they did not believe it would actually happen. However, this tendency was less pronounced when comparing full job adverts than when comparing side-by-side attribute statements in the first part of the focus group. Since many participants in the experiment would likely be making quick choices.

When comparing salary increases to other benefits, participants highlighted two factors that could affect trade-offs in unexpected ways:

1. Small pay increases (e.g., 5%) were often considered insignificant, as they would largely be absorbed by tax.
2. At least one participant weighed pay increases against specific benefits, such as assessing whether a given salary increase was worth more or less than free dentistry.

Both trade-off decisions are likely to be influenced by a teacher's baseline salary, which determines their marginal tax rate and the real-world impact of a 10% pay rise, highlighting the need for subgroup analysis taking into respondents' baseline pay.

Comprehensibility, clarity and enjoyability

Participants reported that the statements were clear and that the task was straightforward. They found the use of bullet points helpful and said that aligning statements about the same attributes across the two adverts made comparisons easier.

The example adverts used in the focus groups included four statements per advert, which participants found manageable and we decided to use this number across all experiments.

> *"It would be really intuitive as to what you're looking at and what you need to pick through. So from that side, yeah, it reads really nicely."*
>
> Secondary teacher

Teachers found the task both enjoyable and thought-provoking. They particularly liked the forced-choice element, with some even saying it helped them reflect on their own job priorities. As a result, they did not anticipate skipping these questions when they appeared on Teacher Tapp and believed the engaging nature of the task would encourage careful reading. However, it is important to note that focus group participants were likely to be highly engaged users, which may not be representative of all respondents.

> Interviewer: *Why would you be motivated to do it?*
>
> Primary teacher 1: *I guess it's quite interesting. I quite like those sorts of questions where you have to make a forced decision on things—and then it's always interesting looking the next day to see the results... to see which way people went on them as well.*
>
> Primary teacher 2: *Yeah, and it also makes me think... personally, 'what am I going to be looking for in the next job I apply for?'*

Participants said they would enjoy looking back at other people's responses once the results were published. However, they also noted that because this task was more time-consuming, they would expect the next questions on the platform to be shorter to balance the overall experience.

One readability issue was highlighted: in compound statements, participants were less likely to pay attention to the second part of a statement following a comma. This issue became more pronounced when statements were combined into a full advert, suggesting that avoiding compound sentences may improve clarity.

Ecological validity

Participants described a range of typical Teacher Tapp usage patterns. Some said they were not usually in a rush when completing questions—one described logging into the app as a "brain break" at the end of the school day, while others completed the daily survey on their journey home or after arriving home. However, several noted that if their usual routine was disrupted—such as by a staff meeting—they might have to answer in a more rushed manner, engaging with questions differently by skimming or speeding through responses. All of these factors might impact on the extent to which survey responses reflect real-world decisions.

When describing their typical job-hunting approach, participants were clear that they do not apply for or accept a role based purely on an advert. Instead, a job advert serves as a starting point to spark their interest. They would then visit the school to get a feel for it, particularly its ethos—which, as noted above, is a key factor in their decision-making.

> *"When I'm looking for jobs from a teaching perspective, I want to go and visit before I even consider applying."*
>
> Secondary teacher

> *"The demographic—you kind of pick up if you go and feel what the school prides itself on. To me, when I walk into a school... if I can see the children's work everywhere, that's a really big thing. Whereas if I see all the awards for*

> *'Investors in People' and 'Arts Mark'—yeah, they're great—but that's celebrating what the staff can do more than what the children have done."*
>
> Primary teacher

Job interviews also play a key role in teachers' decision-making, as they provide an opportunity to ask questions about what matters to them and probe concerns raised by the job advert or school visit. In this sense, an interview is not just an opportunity for the school to assess the teacher—it is equally an opportunity for the teacher to assess the school.

Because of this, experiment participants' choices between adverts should not be assumed to directly translate into real-world choices. Ideally, a follow-up DCE based on descriptions of school visits or interviews would provide further insights.

Some participants reported that, even at the initial job search stage, they do not rely solely on adverts. Instead, they cross-reference multiple sources of information. For example, one teacher explained that if they noticed a school had repeatedly advertised the same roles, they would interpret this as a sign of retention problems or an undesirable work environment. Others checked school websites and social media to gain a sense of the school's ethos, particularly looking for community engagement and enrichment activities.

> *"My search process tends to be to go on the school website to find out quite a lot of the more... background information—like the clubs, the other things that go on that you don't necessarily see in the job advert. And I suppose that, for me, is equally as important as whatever they're claiming in the advert. I think it's slightly more realistic."*
>
> Primary teacher

Circumstances also shape how teachers respond to job adverts. One participant noted that in some areas, job opportunities are scarce, meaning teachers do not always have the luxury of choosing between multiple roles. Another explained that her responses in the experiment were influenced by the fact that she had recently been job-hunting, and that she would respond very differently if she were not actively seeking a new role.

Participants suggested that survey respondents should be asked whether they are currently looking for a job to account for this variability. They also noted that responses might differ depending on the time of year, with late spring and early summer coinciding with the main job-hunting season.

Given these considerations, experimental findings should not be assumed to map directly onto real-world hiring decisions. Alternative experiment designs—such as incorporating school visits—should be considered as potential extensions to the study.

Limitations

The use of an online focus group format for a cognitive pretest of a discrete choice experiment was unusual and presented certain challenges. In particular, it made the application of the 'think-aloud' technique more complex, as participants were influenced by group discussion rather than working through decisions independently.

Additionally, the discussion format encouraged participants to reflect on statements and decision-making processes in more depth than they would when completing the tasks in a real-world setting. This heightened level of scrutiny may not accurately reflect how respondents would typically engage with the task, where choices are made more intuitively and under time constraints.

The task used in the focus groups also differed from the planned experiment in that it included fewer bullet points per advert. This may have influenced participants' ability to process and compare information, and the implications of this difference should be considered when interpreting the findings.

A noticeable contrast was observed between Parts 1 and 2 of the focus group discussions. During Part 1, when participants were comparing individual statements, there was considerable cynicism and close scrutiny of the wording. However, in Part 2, when participants assessed full job adverts, they tended to make quicker, less critical decisions, engaging more

holistically with the information. This shift suggests that the experimental context may elicit different cognitive processes compared to a tabulated or attribute-by-attribute format.

The group size also had a substantial impact on the discussion dynamics. One session, which included only two participants, resulted in highly convergent discussion, with comments echoing each other, indicating potential priming effects and social desirability bias. In contrast, the largest group was more challenging to facilitate, as it was difficult to explore each response in depth. Based on these observations, a group size of five to six participants appeared to offer the best balance, allowing for diverse viewpoints while maintaining depth of discussion. This insight informed the decision to split a subsequent session into two smaller groups.

Finally, it is important to acknowledge that participants had volunteered their time to take part in the focus groups during their holidays, suggesting that they were likely among the most engaged users of the platform—potentially 'superfans' of Teacher Tapp. This self-selection bias may mean their approach to using the app may differ from other users. Findings should be interpreted with this in mind.

Appendix C - RQ1: Pilot testing of job attributes

This appendix outlines the process of developing job attributes beyond the initial desk research and focus groups. While the focus groups generated a broader range of possible wordings for several job attributes, the final wording was selected based on individual survey questions.

Professional development

We considered several distinct options for how professional development could be presented within job advertisements. As existing literature rarely identifies the nature of professional development as a key factor in teacher recruitment and retention, our approach was to select the most appealing option.

To determine this, we surveyed 390 teachers on 8th October 2024, asking them to select the professional development offer they found most appealing (percentage of respondents selecting each option shown in square brackets):

- School CPD includes mentoring and six coaching sessions per year [19%]
- School CPD includes six days per year for courses, chosen by you from a list of approved providers [45%]
- School CPD includes six days per year for self-directed CPD [25%]
- School CPD includes six days per year of training in-school with colleagues [11%]

These responses indicate that, if the goal is to maximise appeal, job attributes should prioritise external courses chosen by the teacher.

Workload

For the workload attribute, the focus groups raised concerns that a 35-hour working week might be infeasible. However, as this is established policy in Scottish schools, we sought to determine whether it was seen as credible and appealing among teachers in the Teacher Tapp panel. To explore this, we presented a paired conjoint question in which one option included the 35-hour week:

> *If two schools that were identical in every other way had the following job adverts, which would you prefer?*
>
> *A school offering:*
>
> - *Same salary as your current job*
> - *Parent and pupil surveys show that pupil behaviour is consistently good*
> - *35-hour work week guarantee with support for planning, marking, data, and central handling of non-core tasks*
> - *Collaborative and friendly staff community with protected co-planning time*
>
> *A school offering:*
>
> - *Same salary as your current job*
> - *Pupil behaviour is consistently good*
> - *Commitment to cut workload through manageable expectations*
> - *Collaborative and friendly staff community with low turnover*

Of the 223 teachers surveyed, 70% selected the 35-hour week option, indicating that it was both appealing and credible. We also asked teachers an open-ended question for feedback on their responses. Only one teacher explicitly commented that the 35-hour week seemed infeasible, while many noted that the workload attribute was the dominant factor in their decision.

Next, we surveyed 1,605 teachers to determine whether we should use the phrase "*no mandatory written marking or data entry*" in the job attributes. Based on the responses (percentage selecting each option shown in square brackets), we decided not to use this phrasing.

> *Imagine you are looking at teaching job adverts. Which of the following statements about workload at a prospective school would be most appealing to you?*
>
> - *Commitment to support staff with manageable workloads [32%]*
> - *Commitment to cut workload through manageable expectations [41%]*
> - *No mandatory written marking or data entry [28%]*

Leadership

For the final experiment, we required a unidimensional scale of leadership support. However, during the focus groups, it was unclear whether teachers preferred autonomy or proactive support from leadership. To investigate this, we presented a paired conjoint question, with one option emphasising teacher autonomy and the other proactive support:

> *If two schools that were identical in every other way had the following job adverts, which would you prefer?*
>
> *A school offering:*
>
> - *Same salary as your current job*
> - *Minimum 20% non-contact (PPA) time*
> - *Non-contact (PPA) time distributed throughout week and to be taken on school site*
> - *Leaders promote teacher autonomy*
>
> *A school offering:*
>
> - *Same salary as your current job*
> - *Minimum 20% non-contact (PPA) time*
> - *Option to block non-contact (PPA) time to facilitate late starts and/or early finishes*
> - *Leaders provide proactive support for teachers in their daily work*

Of the 211 teachers surveyed, 79% selected the proactive support option, suggesting it was both appealing and credible.

We also asked teachers an open-ended question for feedback. While a few teachers expressed concerns that "proactive support" could imply micro-management, a larger number explained why they found it more appealing than teacher autonomy.

Behaviour

The focus groups and other feedback highlighted concerns about the language used to describe behaviour management environments. To explore this further, we surveyed 1,618 teachers to determine whether references to parent and pupil survey results should be included in job descriptions. The results clearly indicated that teachers do not find these statements appealing, aligning with feedback from some focus group participants.

> *Imagine you are looking at teaching job adverts. Which of the following statements about workload at a prospective school would be most appealing to you?*
>
> - *High standards of behaviour based on clear policies and consistent expectations [55%]*
> - *Parent and pupil surveys show that pupil behaviour is consistently excellent [19%]*
> - *Pastoral leaders work directly with parents and students to manage behaviour [26%]*

Appendix D - RQ1: Main results supplementary tables

Table A2: Experiment 1 main estimates

	Level	Marginal means			AMCE		
		Est.	S.E.	P-val	Est.	S.E.	P-val
Salary level	Same as current job	0.326	0.004	0	0	NA	NA
	5% higher than current job	0.506	0.004	0.143	0.18	0.006	0
	10% higher than current job	0.669	0.004	0	0.343	0.006	0
Flexible working	Requests considered within statutory frameworks	0.423	0.004	0	0	NA	NA
	Accommodated wherever practical	0.523	0.004	0	0.102	0.006	0
	Commitment to meet all requests	0.552	0.004	0	0.132	0.006	0
Healthcare	Staff health and wellbeing is supported	0.380	0.004	0	0	NA	NA
	Time off for appointments and a wellbeing helpline	0.454	0.004	0	0.071	0.006	0
	Free dentistry, physio, counselling, time off for appointments and a wellbeing helpline	0.666	0.004	0	0.284	0.006	0
Collegiality	Friendly staff community	0.433	0.003	0	0	NA	NA
	Collaborative and friendly staff community with protected co-planning time	0.566	0.003	0	0.133	0.005	0

$N_{profiles}$ = 35100; $N_{individuals}$ = 5850

Table A3: Experiment 2 main estimates

	Level	Marginal means			AMCE		
		Est.	S.E.	P-val	Est.	S.E.	P-val
Salary level	Same as current job	0.300	0.004	0	0	NA	NA
	5% higher than current job	0.496	0.004	0.288	0.197	0.006	0
	10% higher than current job	0.701	0.004	0	0.402	0.006	0
Personal flexibility	School endeavours to support staff with family and personal commitments	0.473	0.003	0	0	NA	NA
	Flexibility, within reason, to meet all important family and personal commitments	0.527	0.003	0	0.058	0.005	0
Childcare	Good childcare available close to the school	0.450	0.004	0	0	NA	NA
	20% off onsite nursery and after-school provision from 0 to 13 years	0.488	0.004	0.003	0.038	0.006	0
	40% off onsite nursery and after-school provision from 0 to 13 years	0.562	0.004	0	0.111	0.007	0
CPD	Professional development is valued and supported	0.454	0.004	0	0	NA	NA
	3 days per year for courses, chosen by you from a list of approved providers	0.491	0.004	0.016	0.035	0.007	0
	6 days per year for courses, chosen by you from a list of approved providers	0.555	0.004	0	0.101	0.007	0

$N_{profiles}$ = 34926; $N_{individuals}$ = 5821

Table A4: Experiment 3 main estimates

	Level	Marginal means			AMCE		
		Est.	S.E.	P-val	Est.	S.E.	P-val
Salary level	Same as current job	0.393	0.004	0	0	NA	NA
	5% higher than current job	0.499	0.004	0.729	0.105	0.006	0
	10% higher than current job	0.609	0.004	0	0.218	0.006	0
PPA time	Minimum 10% for all staff	0.291	0.004	0	0	NA	NA
	Minimum 20% for all staff	0.535	0.004	0	0.244	0.006	0
	Minimum 30% for all staff	0.673	0.004	0	0.384	0.006	0
PPA location	PPA time distributed throughout week and to be taken on school site	0.443	0.003	0	0	NA	NA
	Option to block PPA time to facilitate late starts and/or early finishes	0.558	0.003	0	0.113	0.005	0
Workload	Commitment to support staff with manageable workloads	0.429	0.004	0	0	NA	NA
	Commitment to cut workload through manageable expectations	0.471	0.004	0	0.044	0.006	0
	35-hour week guaranteed, with support for planning, marking, and data, plus central management of non-core activities	0.601	0.004	0	0.174	0.006	0

$N_{profiles}$ = 34800; $N_{individuals}$ = 5800

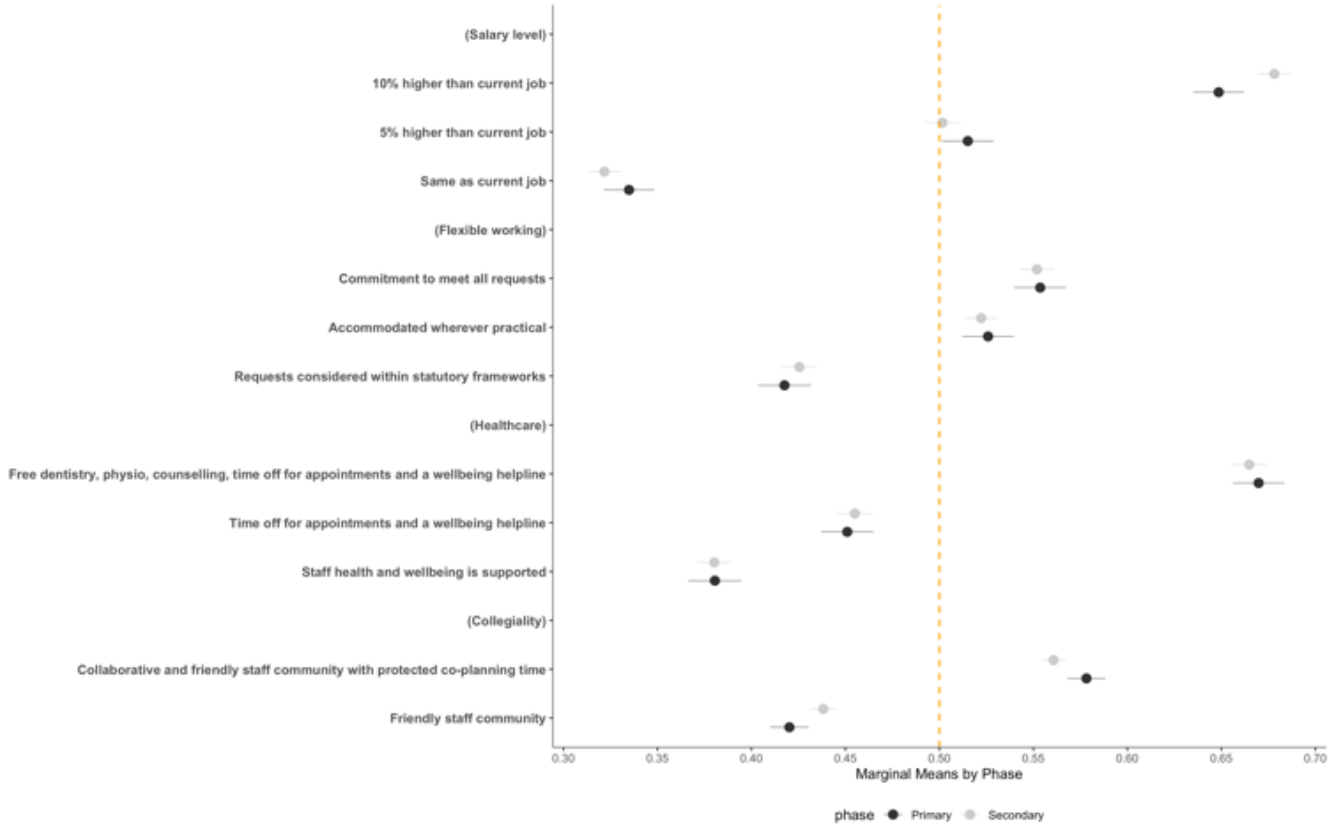
Table A5: Experiment 4 main estimates

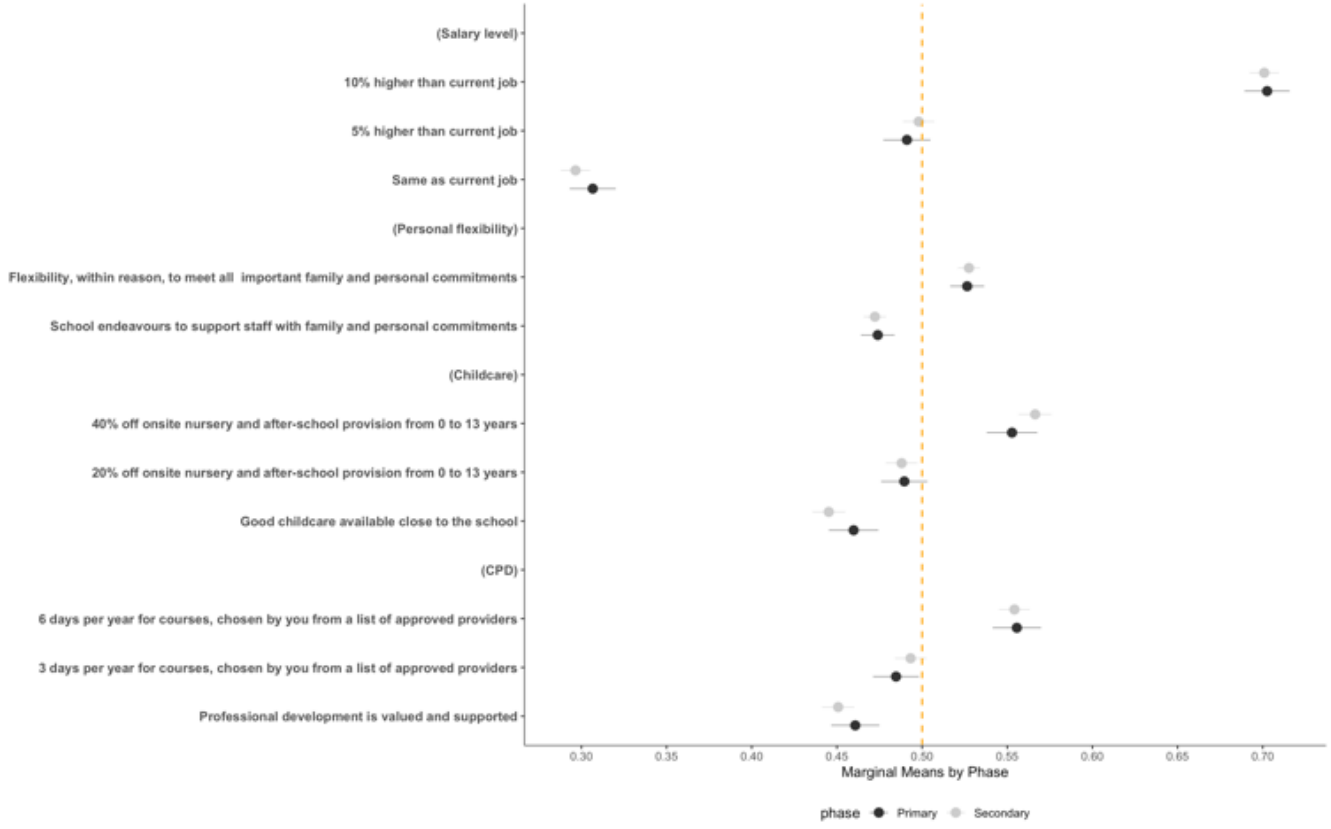
	Level	Marginal means			AMCE		
		Est.	S.E.	P-val	Est.	S.E.	P-val
Salary level	Same as current job	0.332	0.004	0	0	NA	NA
	5% higher than current job	0.498	0.004	0.686	0.167	0.006	0
	10% higher than current job	0.672	0.004	0	0.339	0.006	0
Leadership	Supportive school leadership team	0.480	0.003	0	0	NA	NA
	Leaders and staff set priorities together, with proactive support as needed	0.520	0.003	0	0.039	0.005	0
Class size	No more than 30 students	0.323	0.004	0	0	NA	NA
	No more than 26 students	0.531	0.004	0	0.208	0.006	0
	No more than 22 students	0.647	0.004	0	0.324	0.006	0
Behaviour	Pupil behaviour is consistently good	0.459	0.003	0	0	NA	NA
	High standards of behaviour based on clear policies and consistent expectations	0.541	0.003	0	0.081	0.005	0

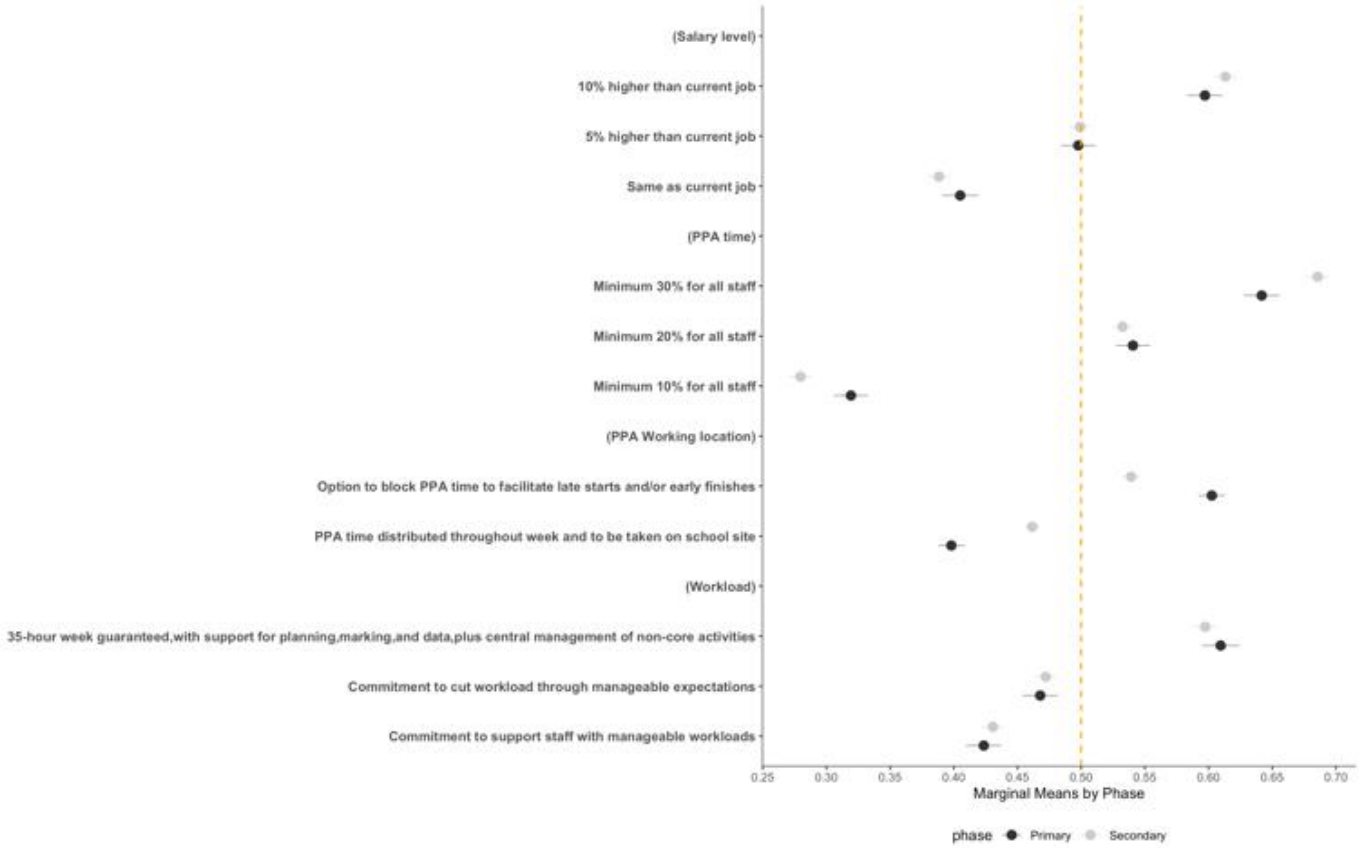
$N_{profiles}$ = 34734; $N_{individuals}$ = 5789

By demographic sub-groups

Figure A1: Experiments 1-4, marginal means by phase






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Phase
phase Primary Secondary

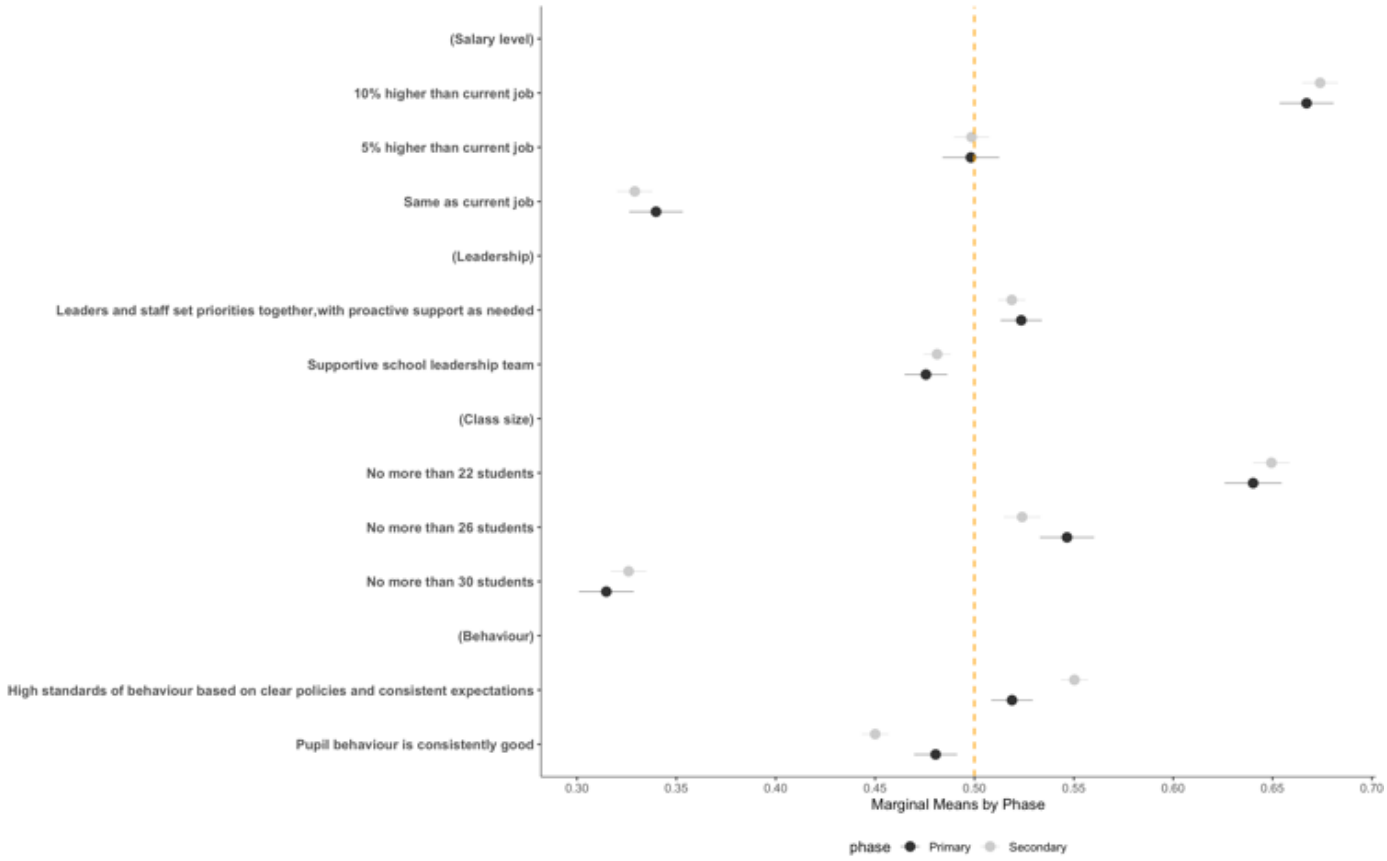
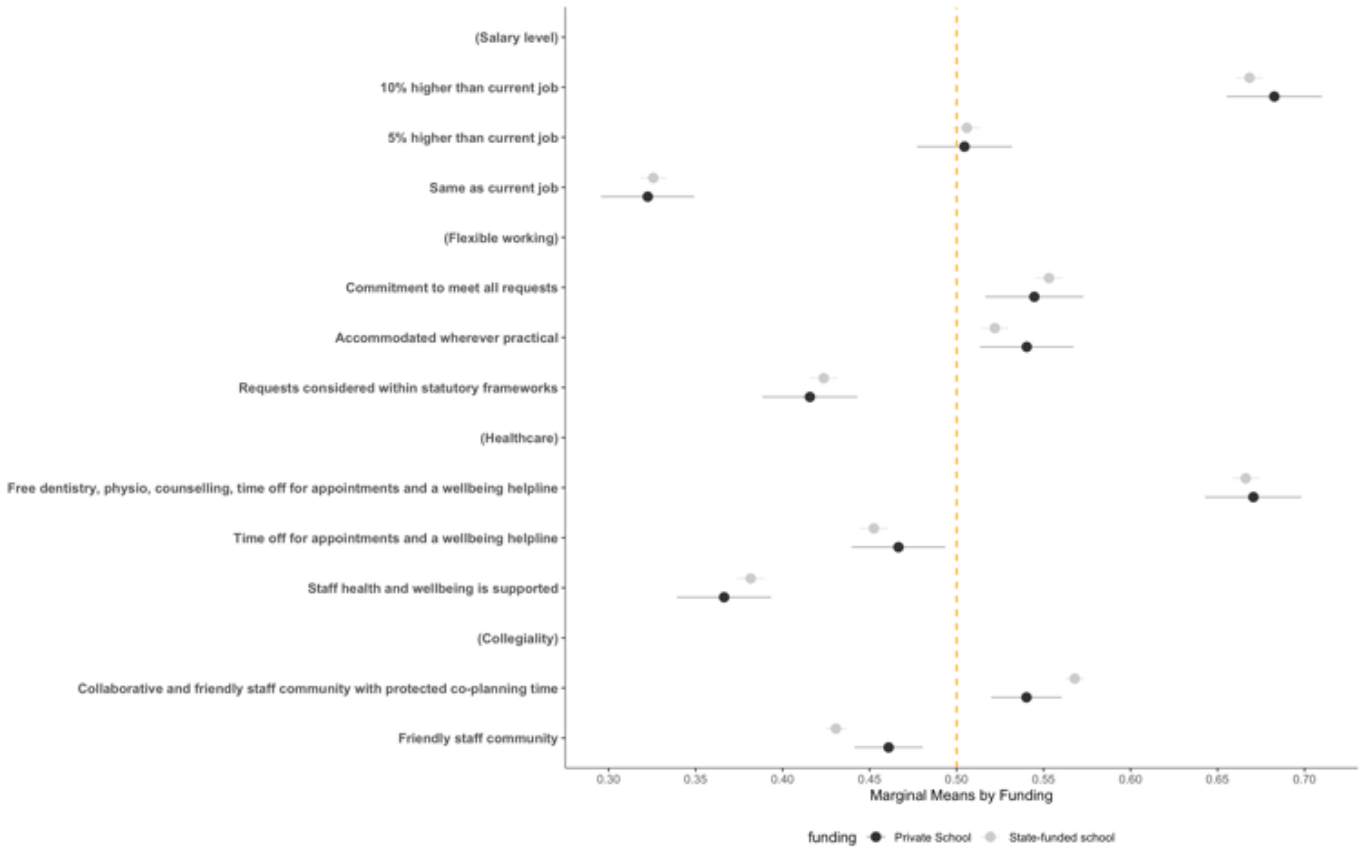
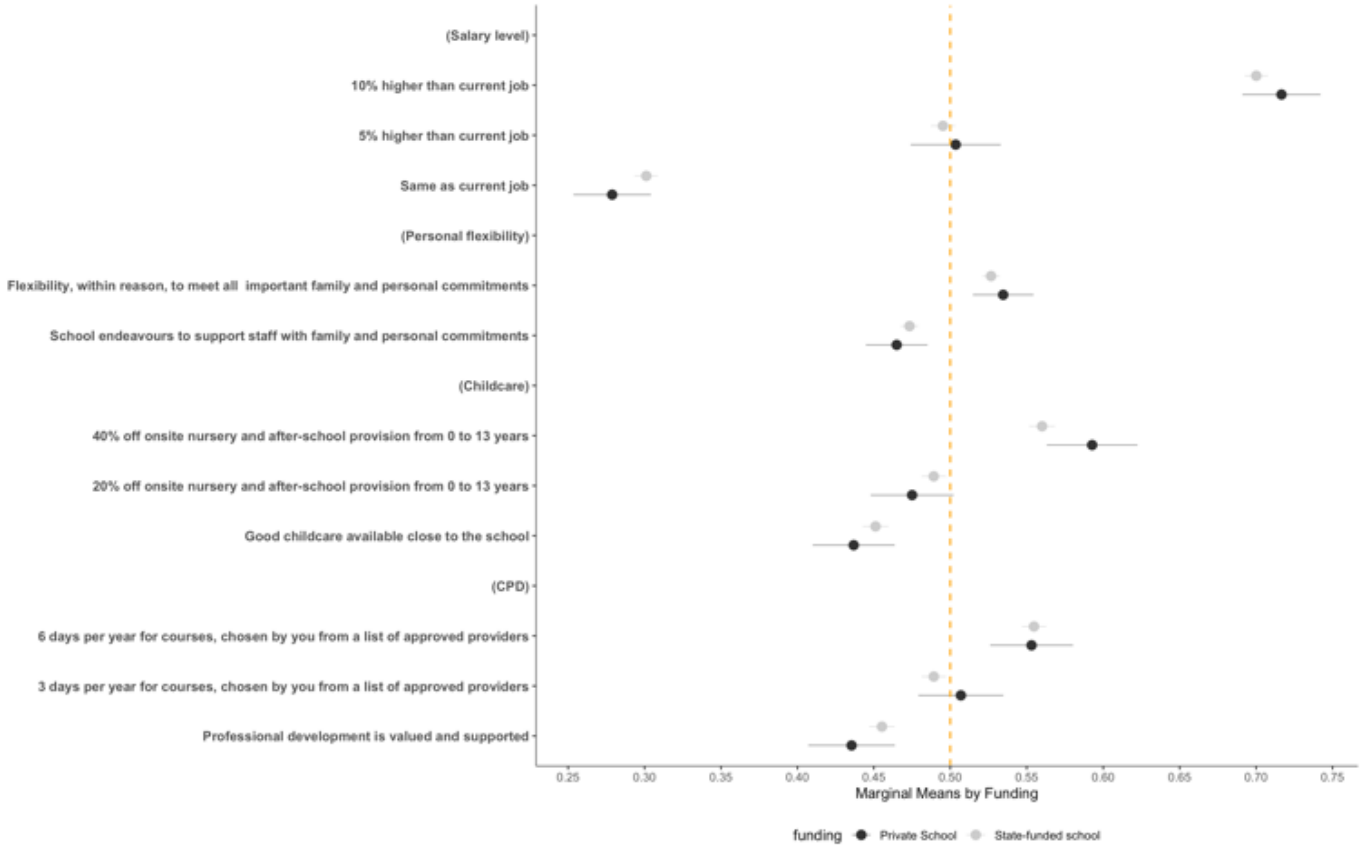
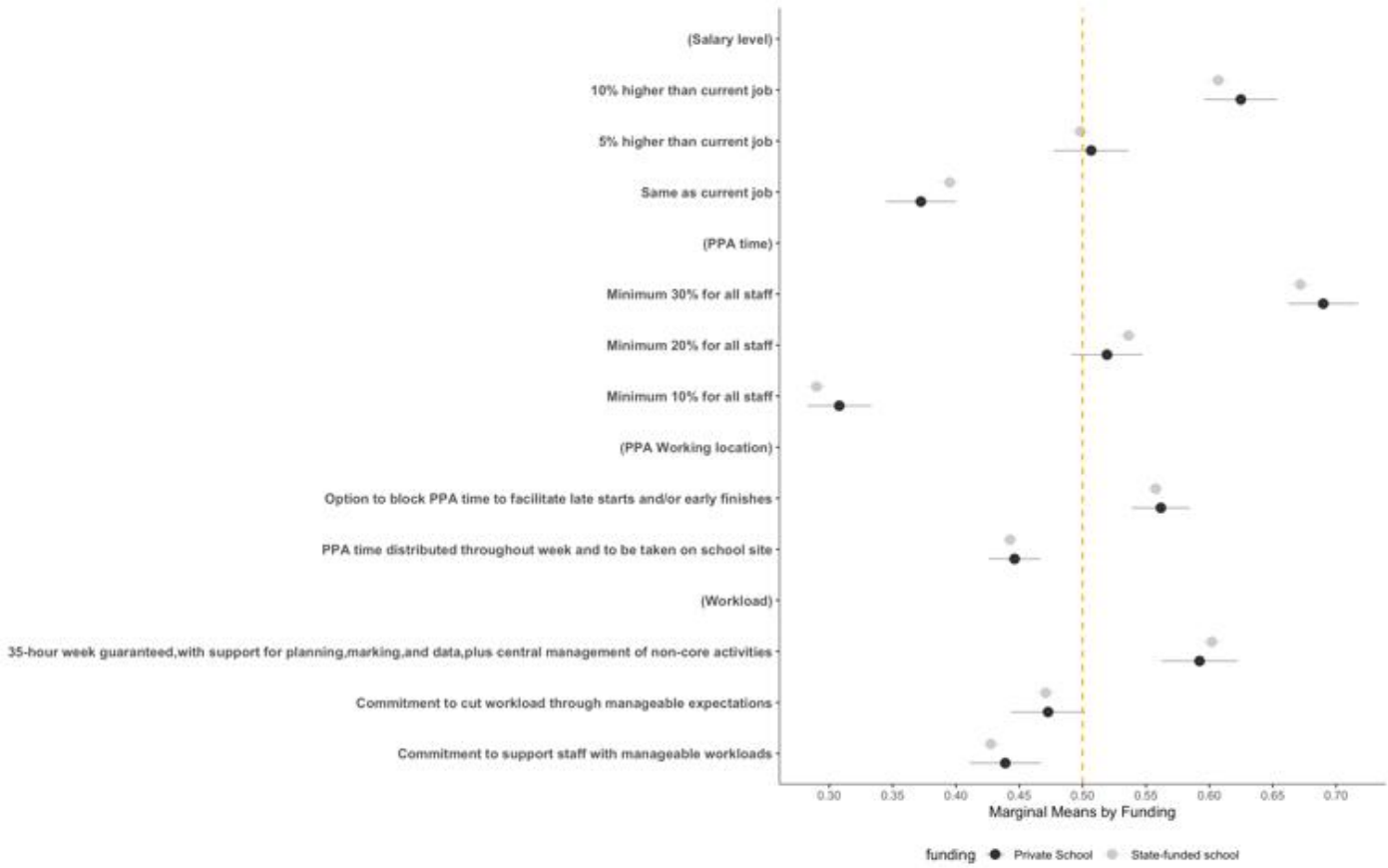

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Phase
phase Primary Secondary

Figure A2: Experiments 1-4, marginal means by school funding






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Funding
funding Private School State-funded school

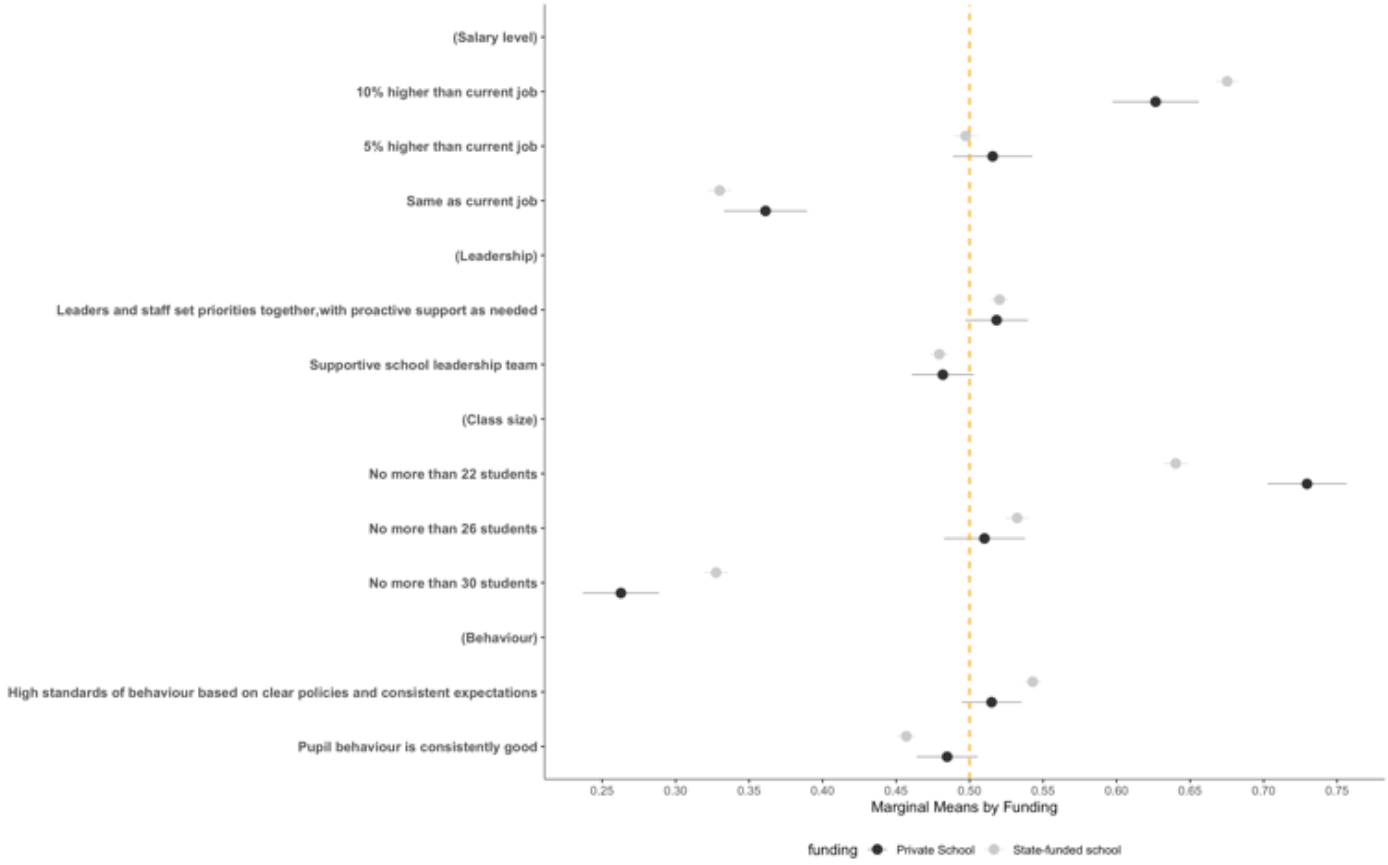
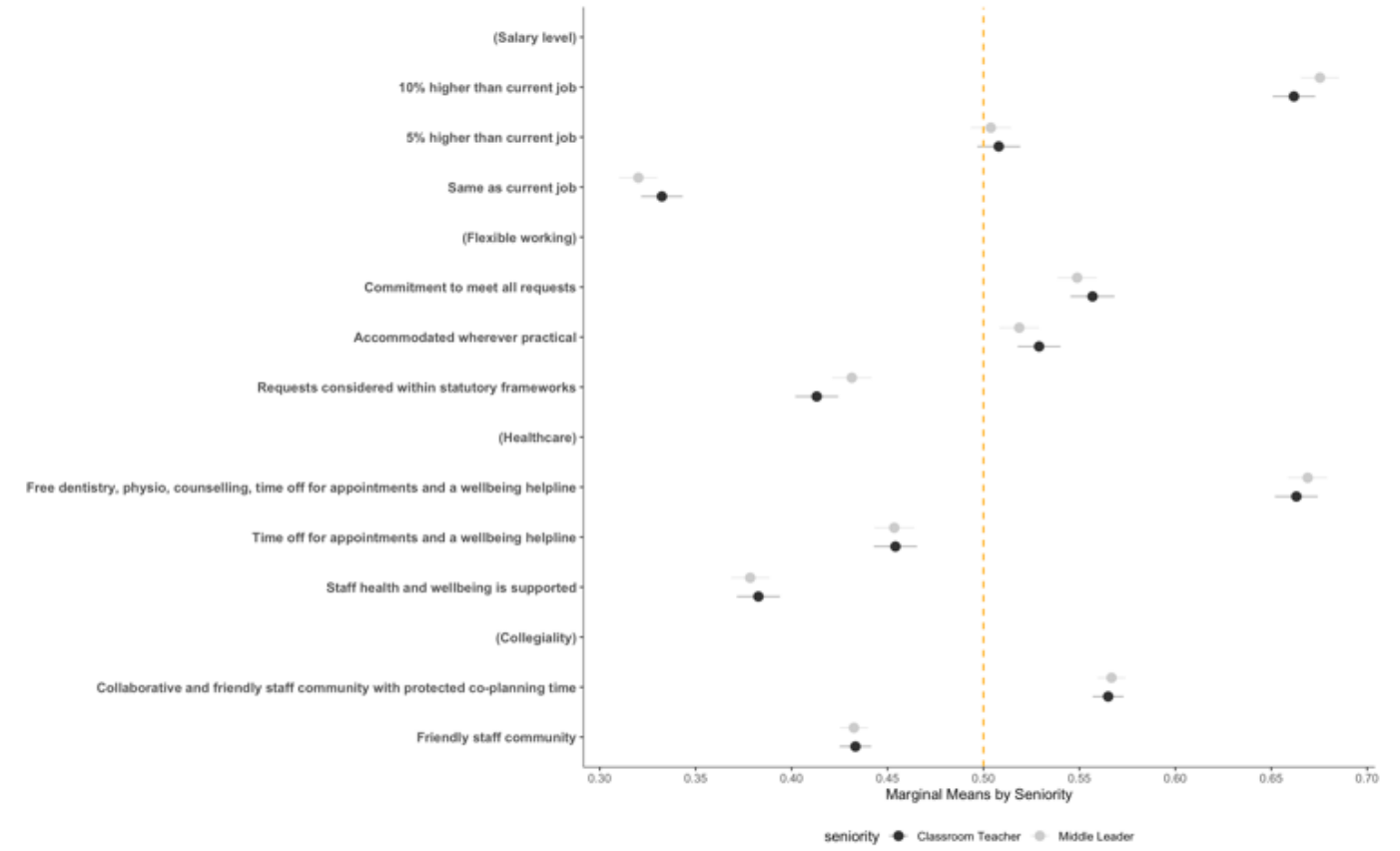
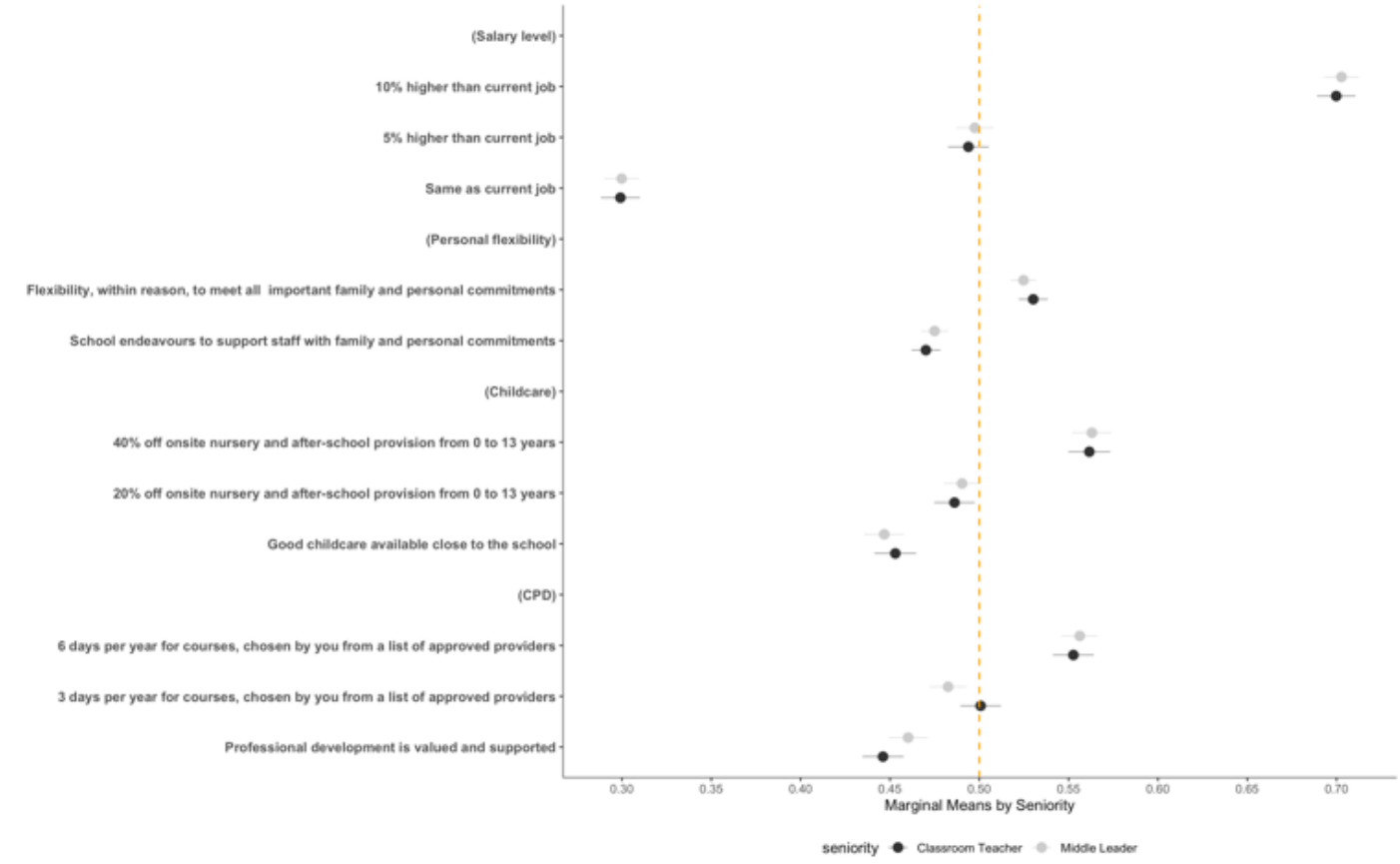
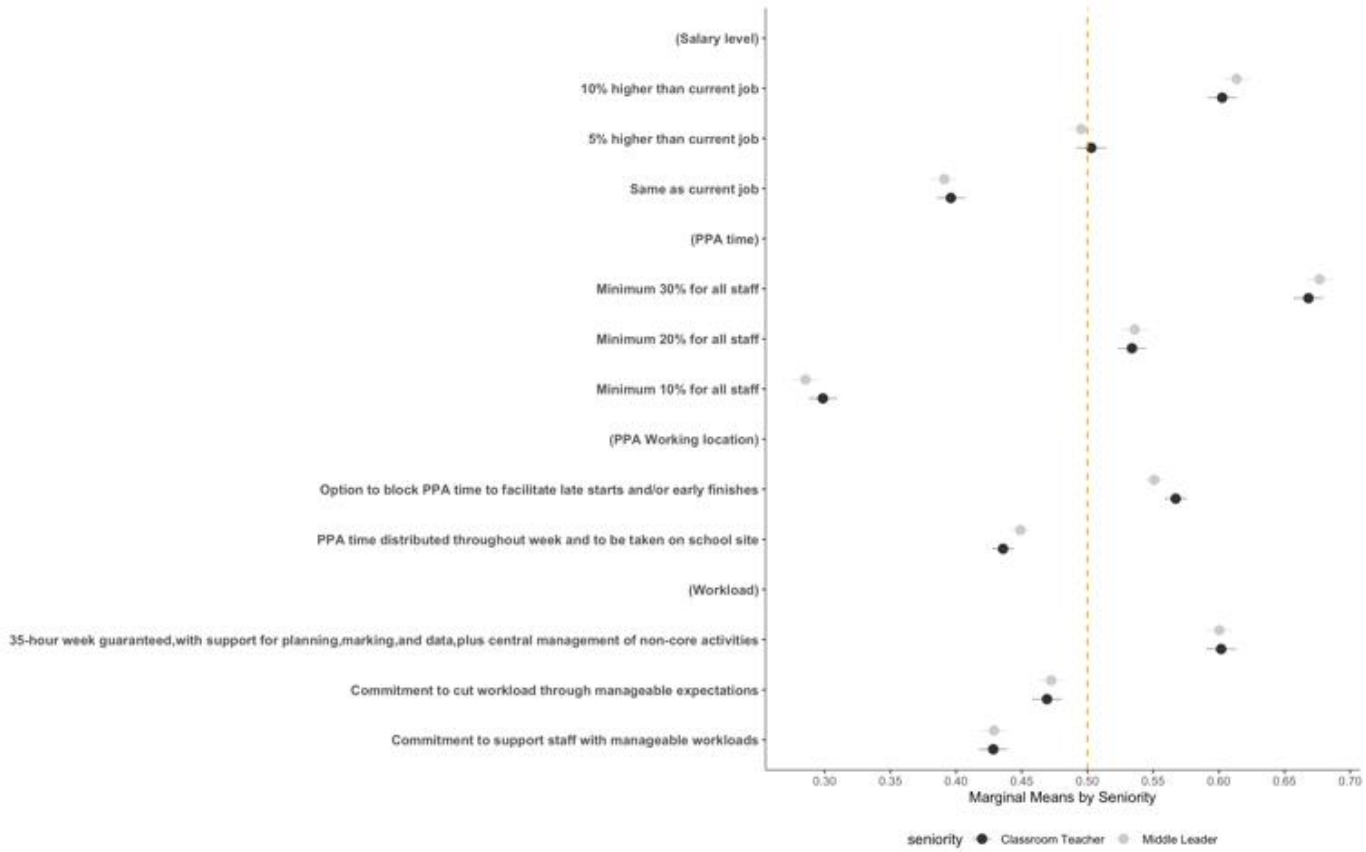

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70 0.75
Marginal Means by Funding
funding Private School State-funded school

Figure A3: Experiments 1-4, marginal means by seniority






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Seniority
seniority
Classroom Teacher
Middle Leader

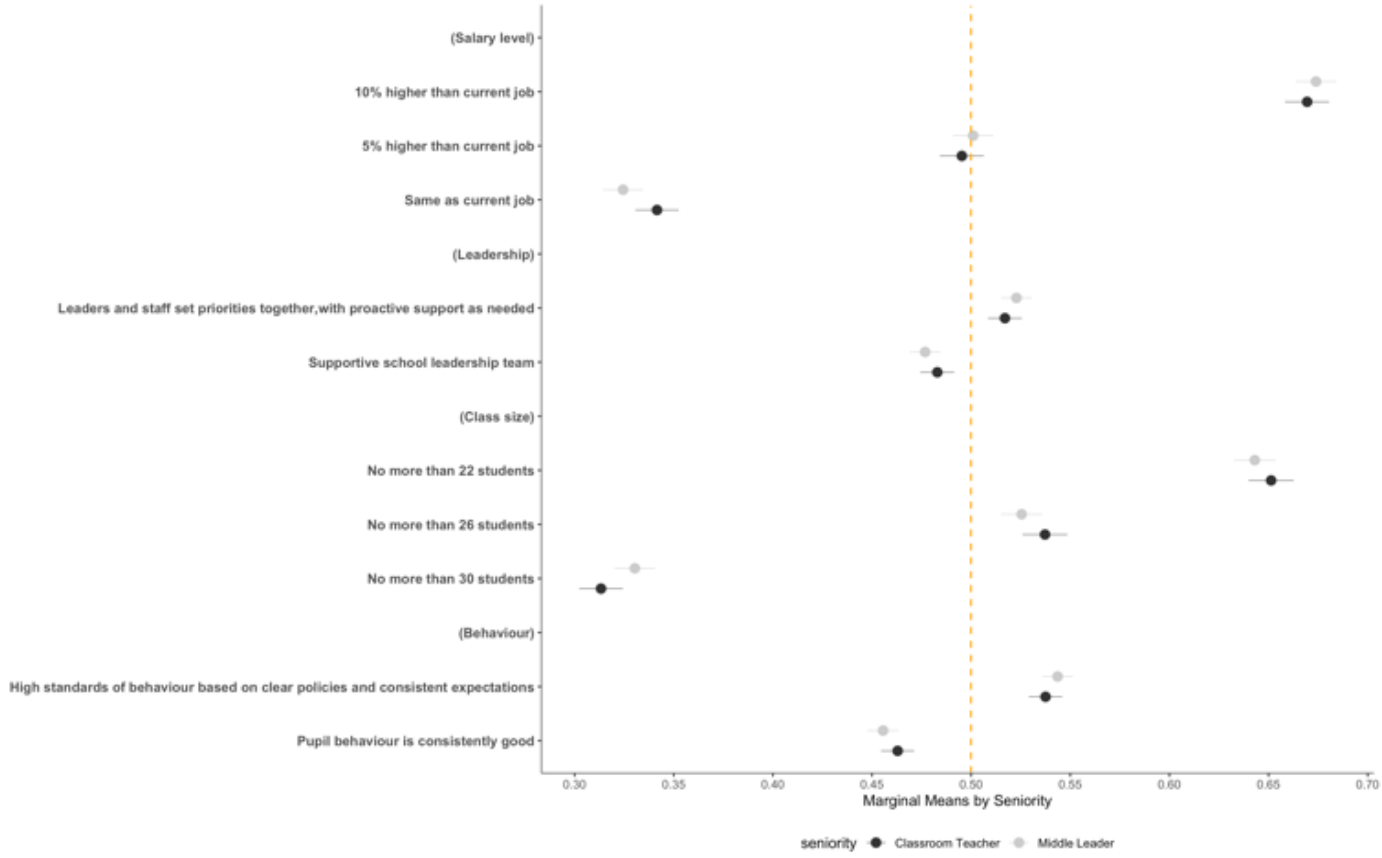
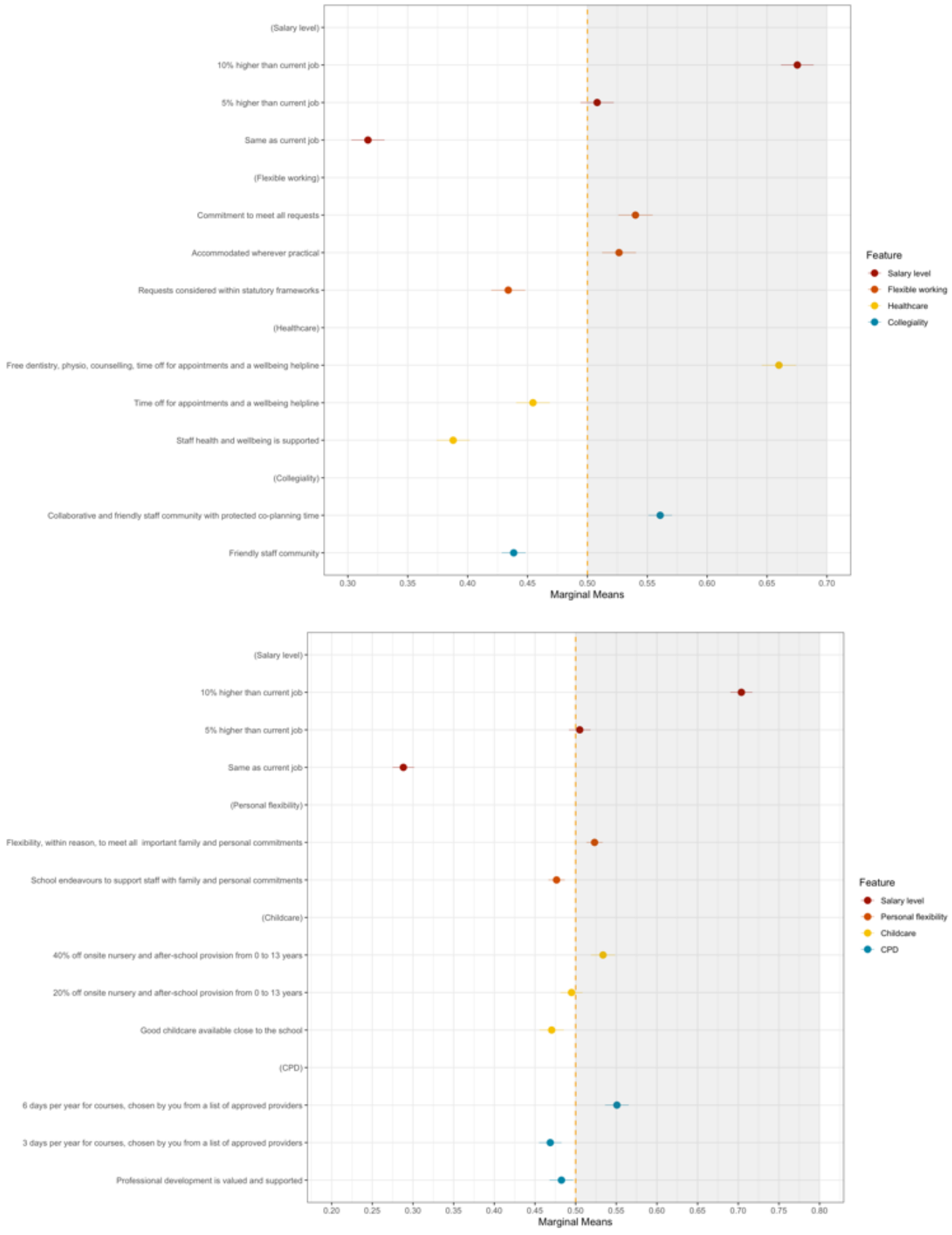
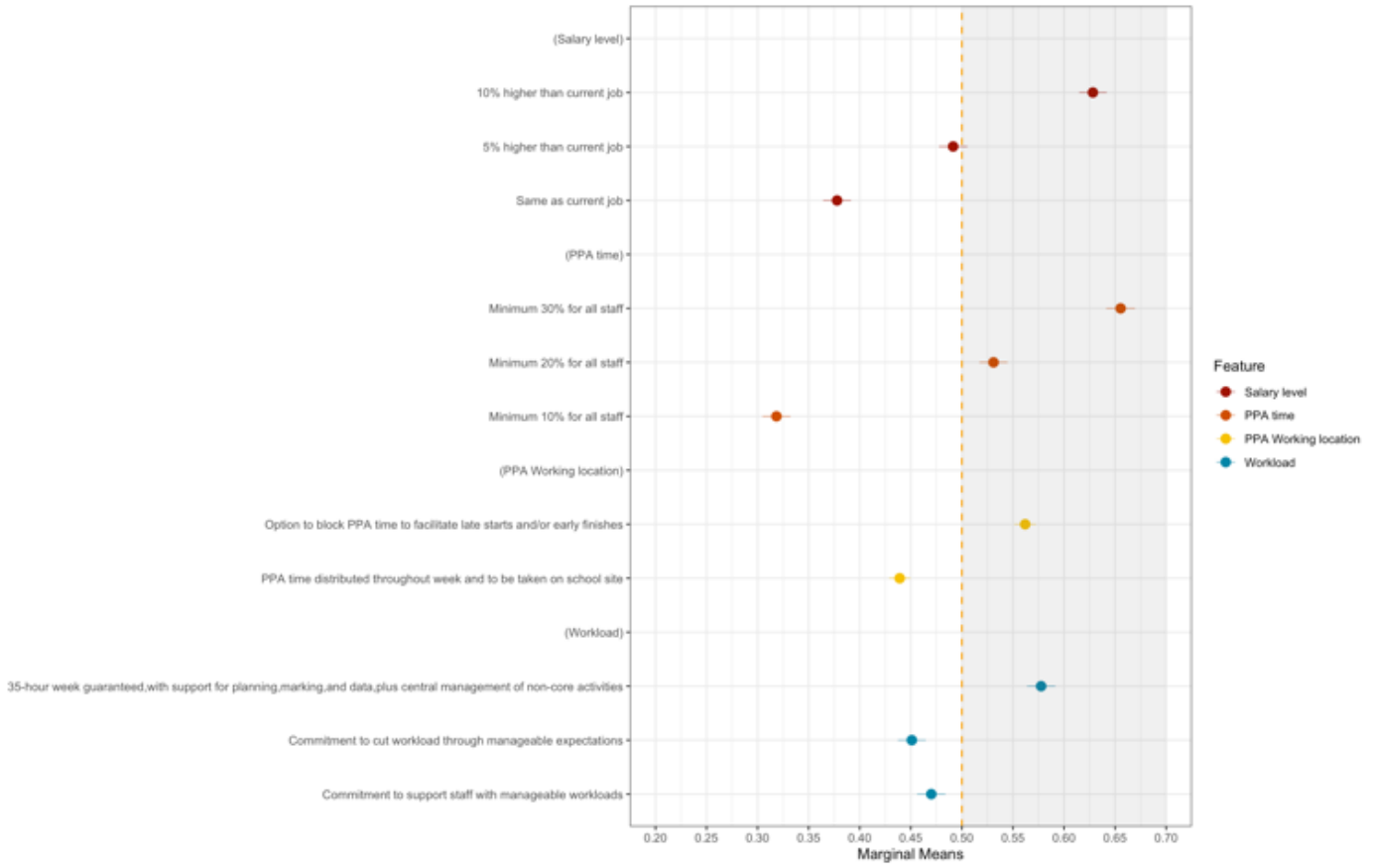

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Seniority
seniority
Classroom Teacher
Middle Leader

Figure A4: Experiments 1-4, marginal means for SLT (excl heads) – not in main sample




(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means
Feature
Salary level
PPA time
PPA Working location
Workload

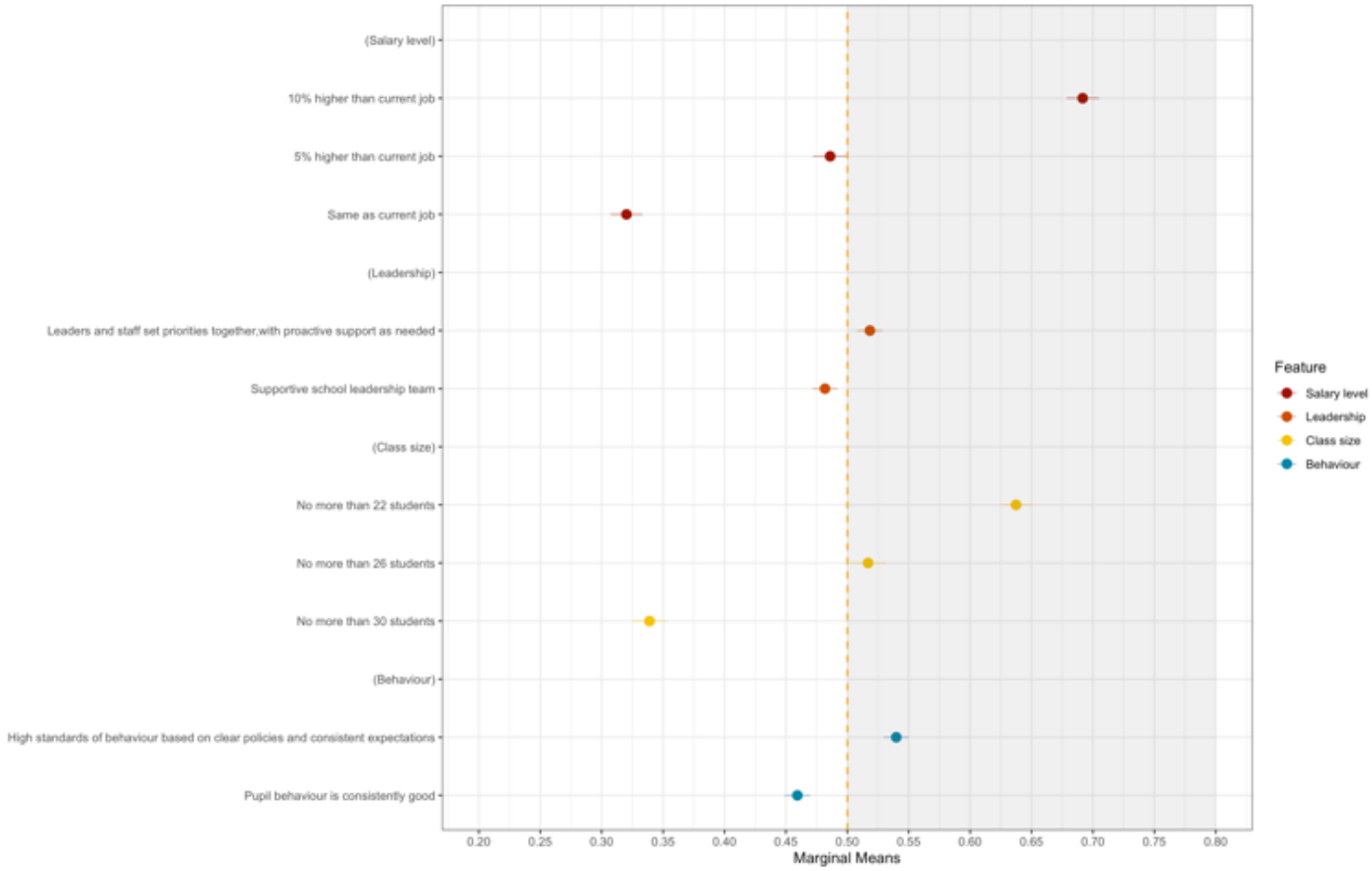
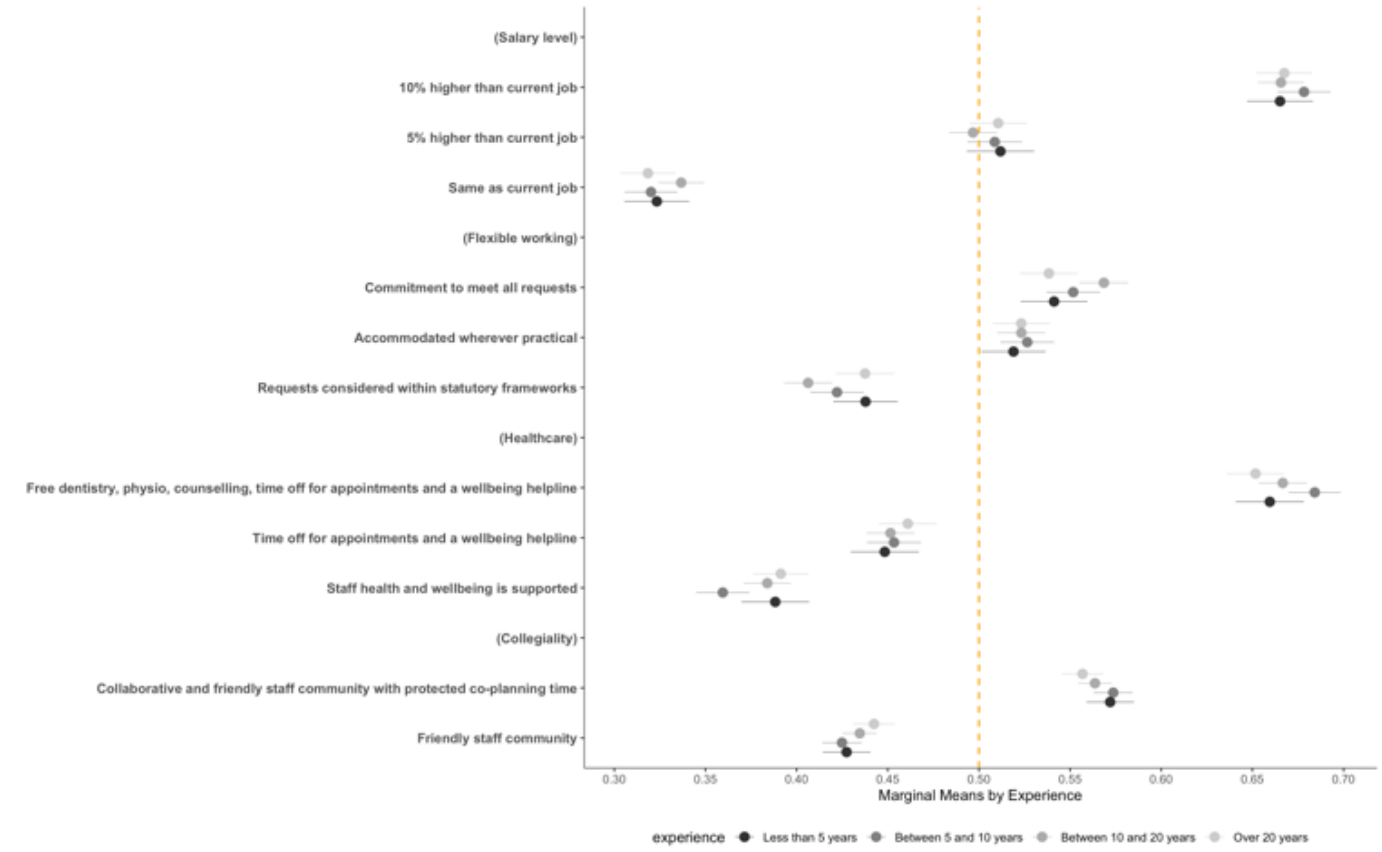
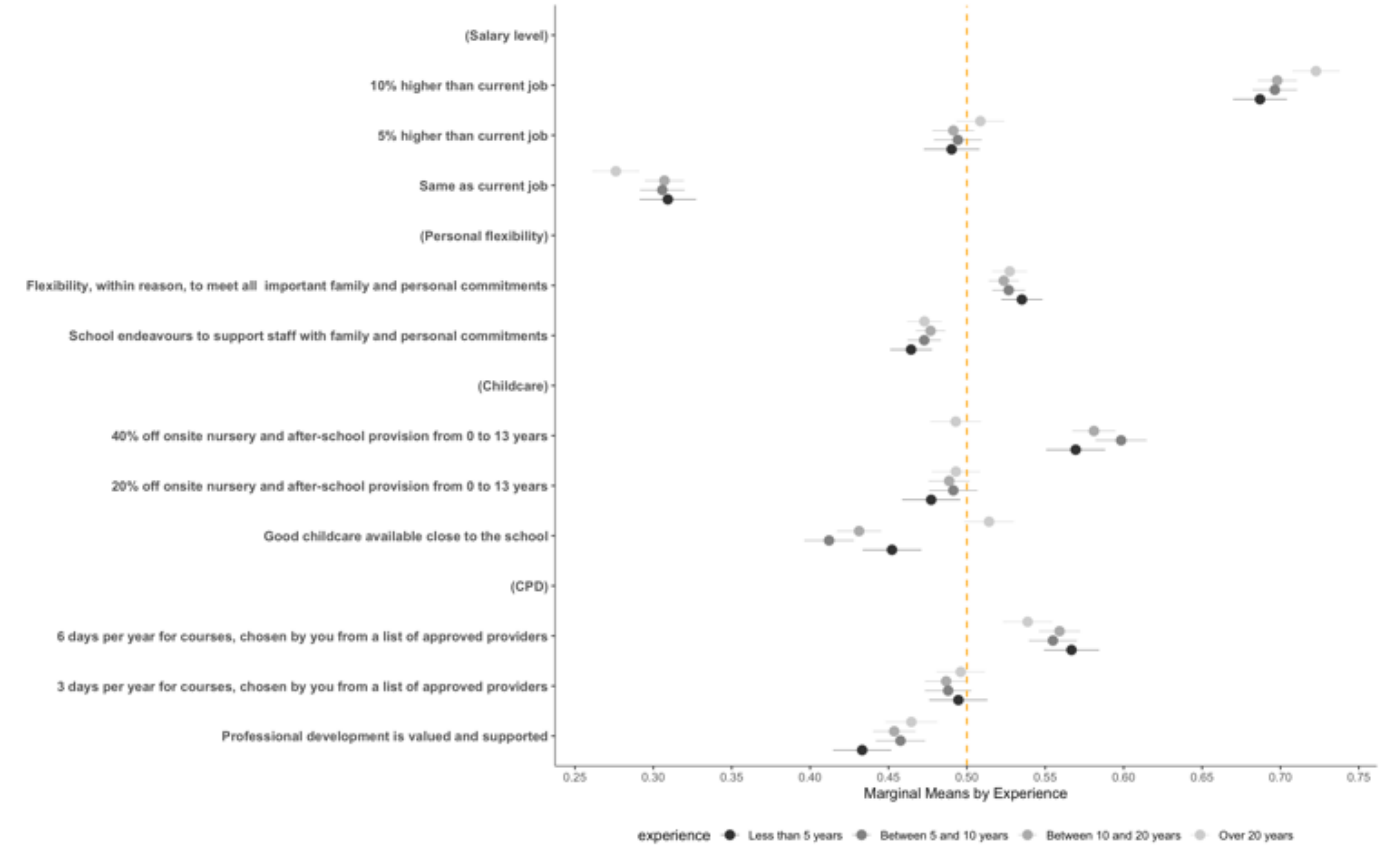
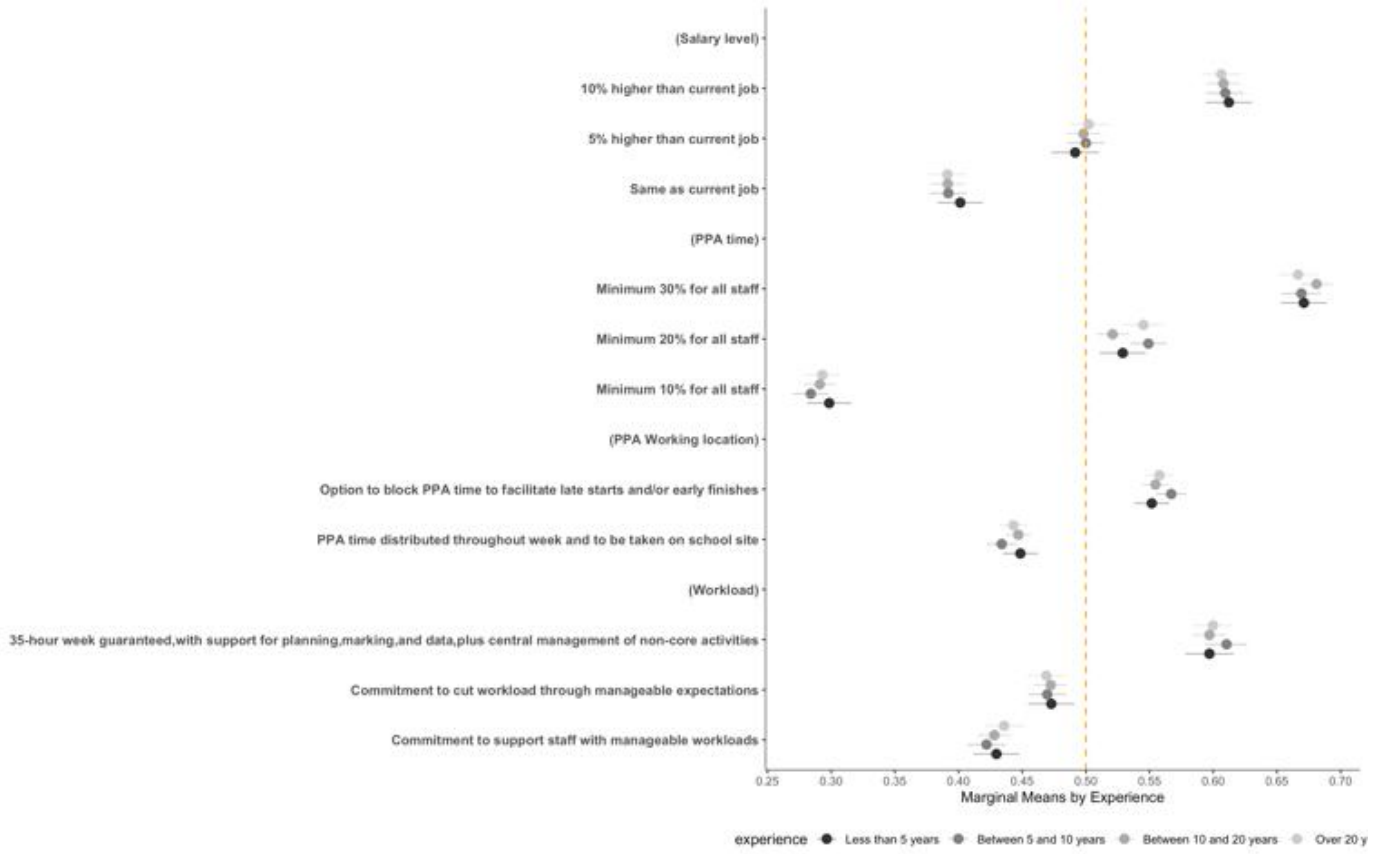

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70 0.75 0.80
Marginal Means
Feature
Salary level
Leadership
Class size
Behaviour

Figure A5: Experiments 1-4, marginal means by years of experience






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Experience
experience
Less than 5 years
Between 5 and 10 years
Between 10 and 20 years
Over 20 y

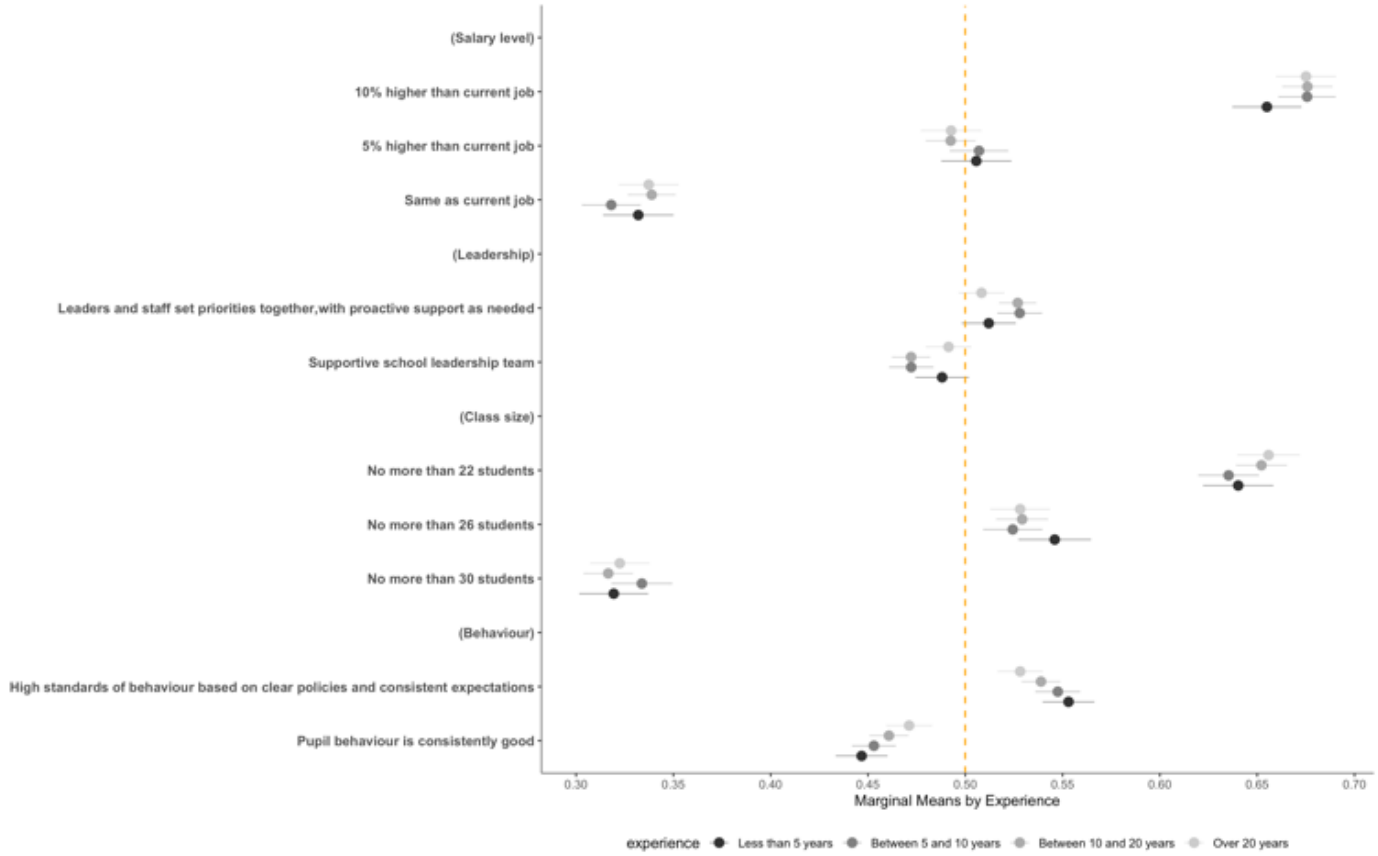
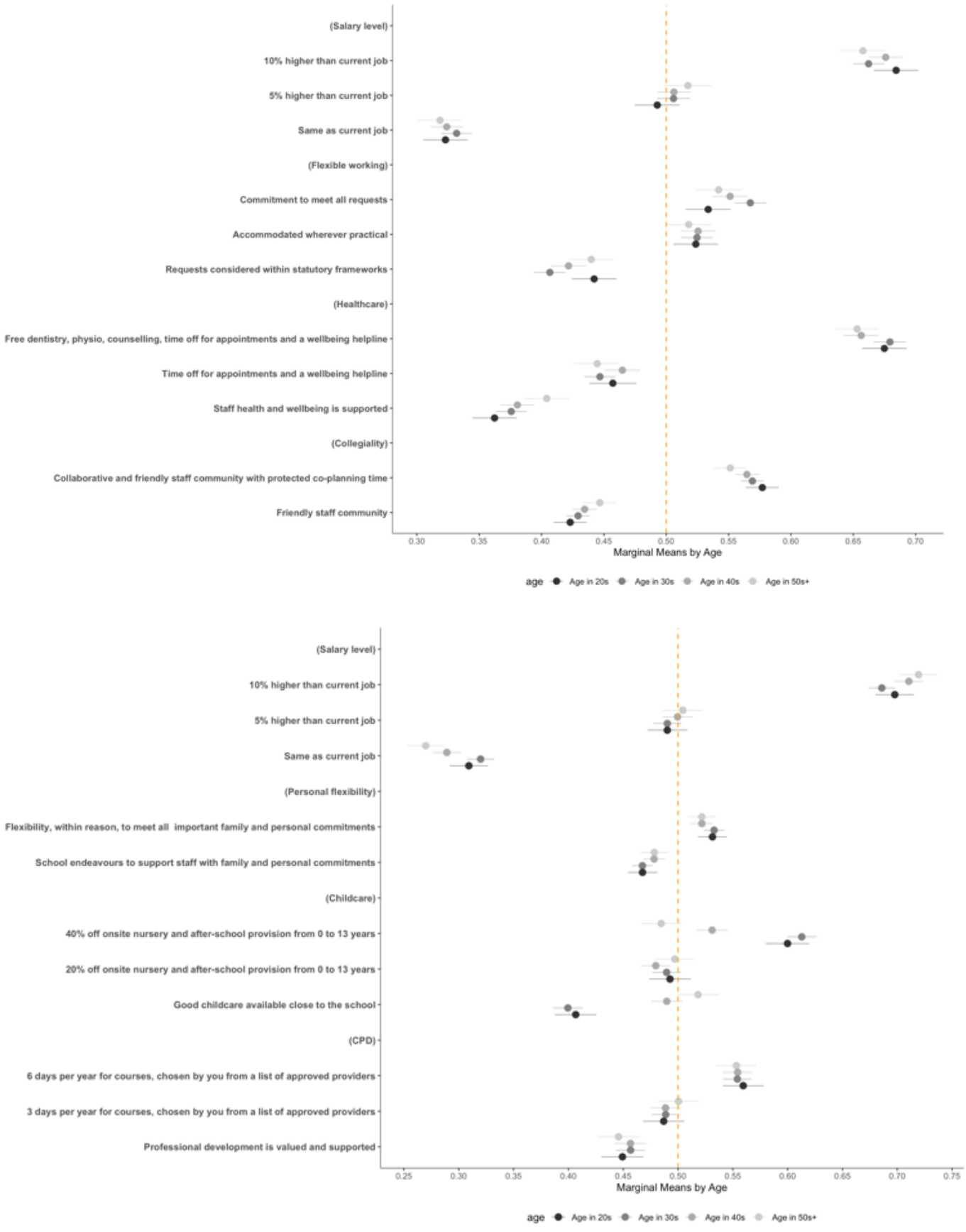
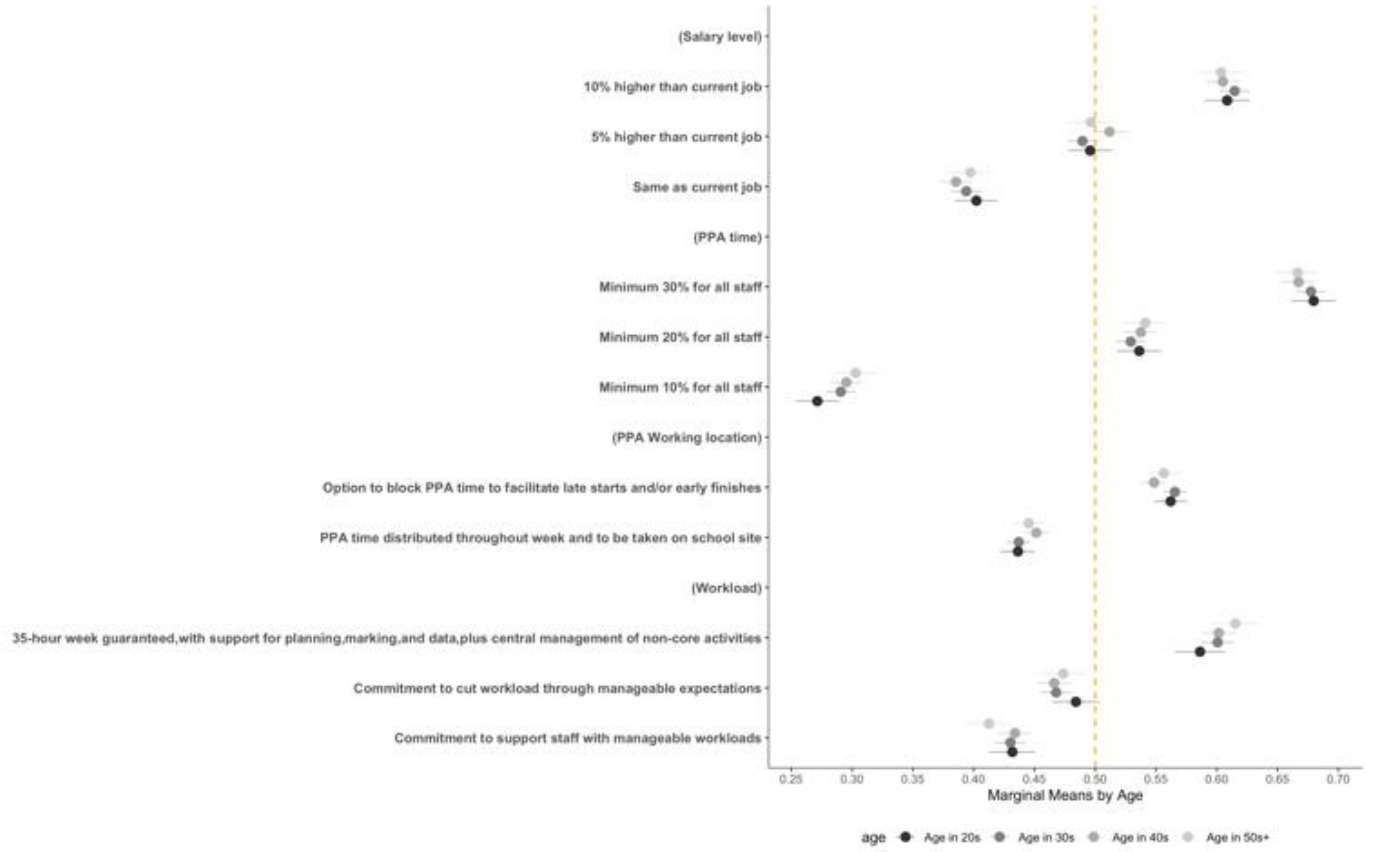

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Experience
experience
Less than 5 years
Between 5 and 10 years
Between 10 and 20 years
Over 20 years

Figure A6: Experiments 1-4, marginal means by age




(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Age
age Age in 20s Age in 30s Age in 40s Age in 50s+

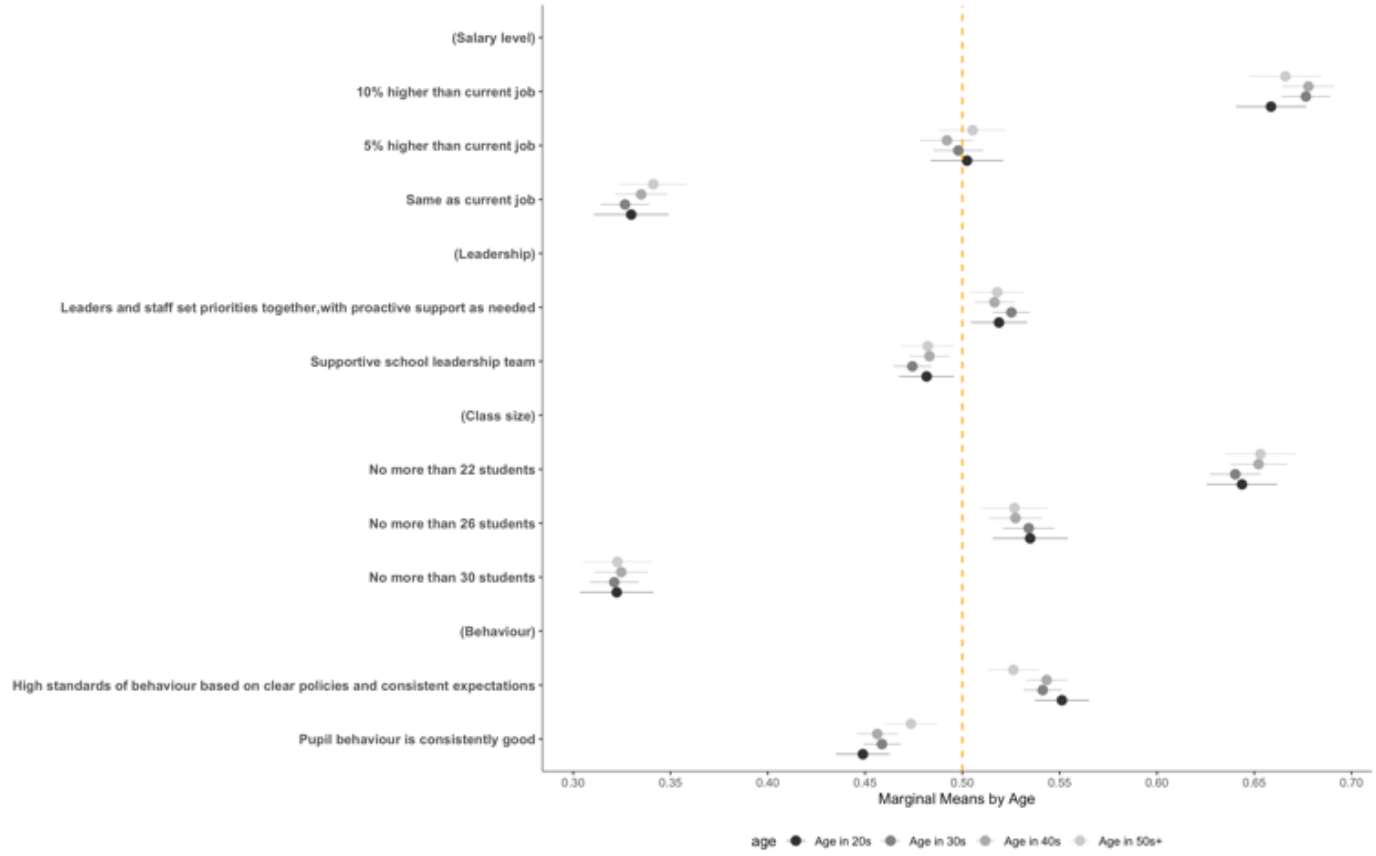
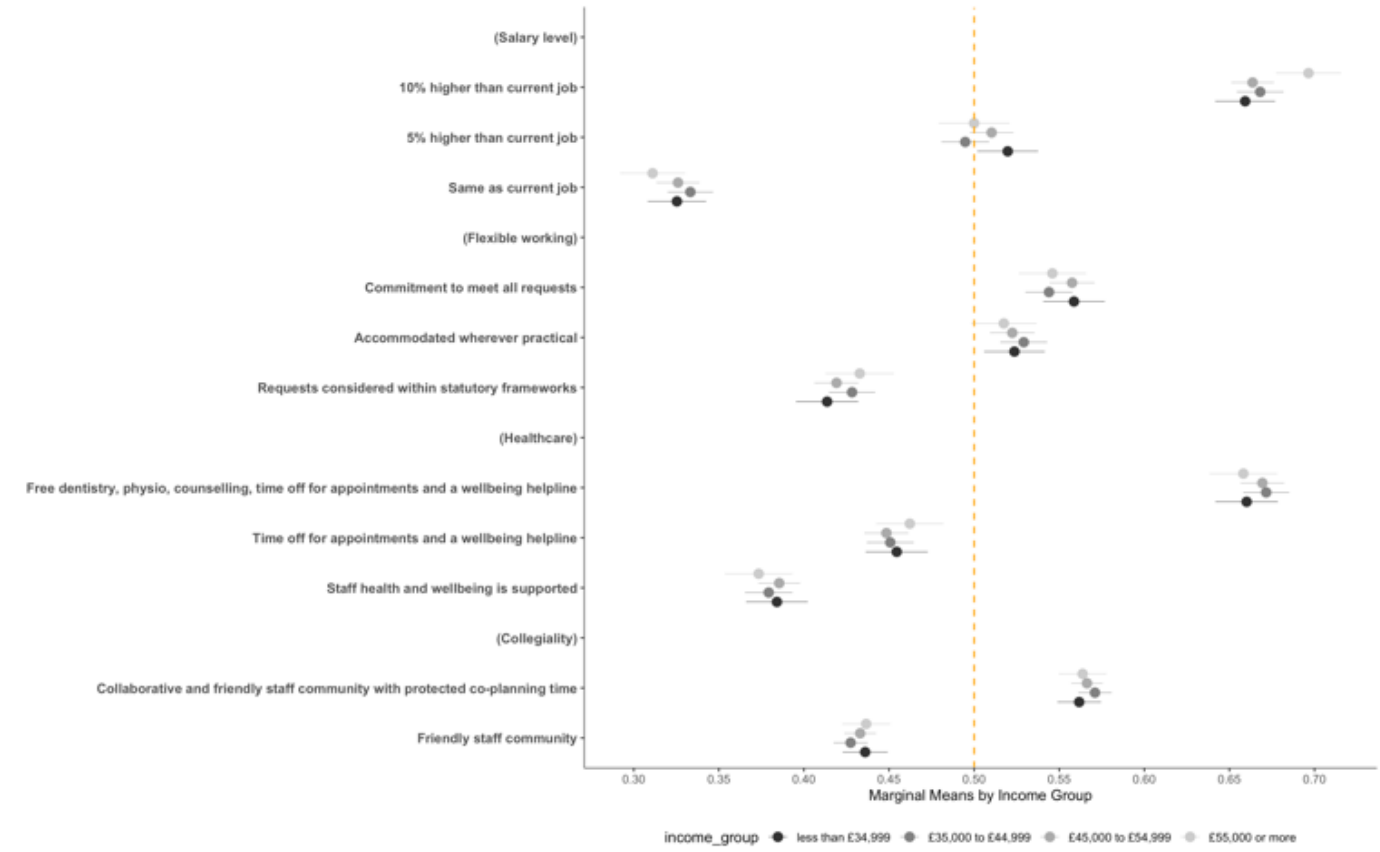
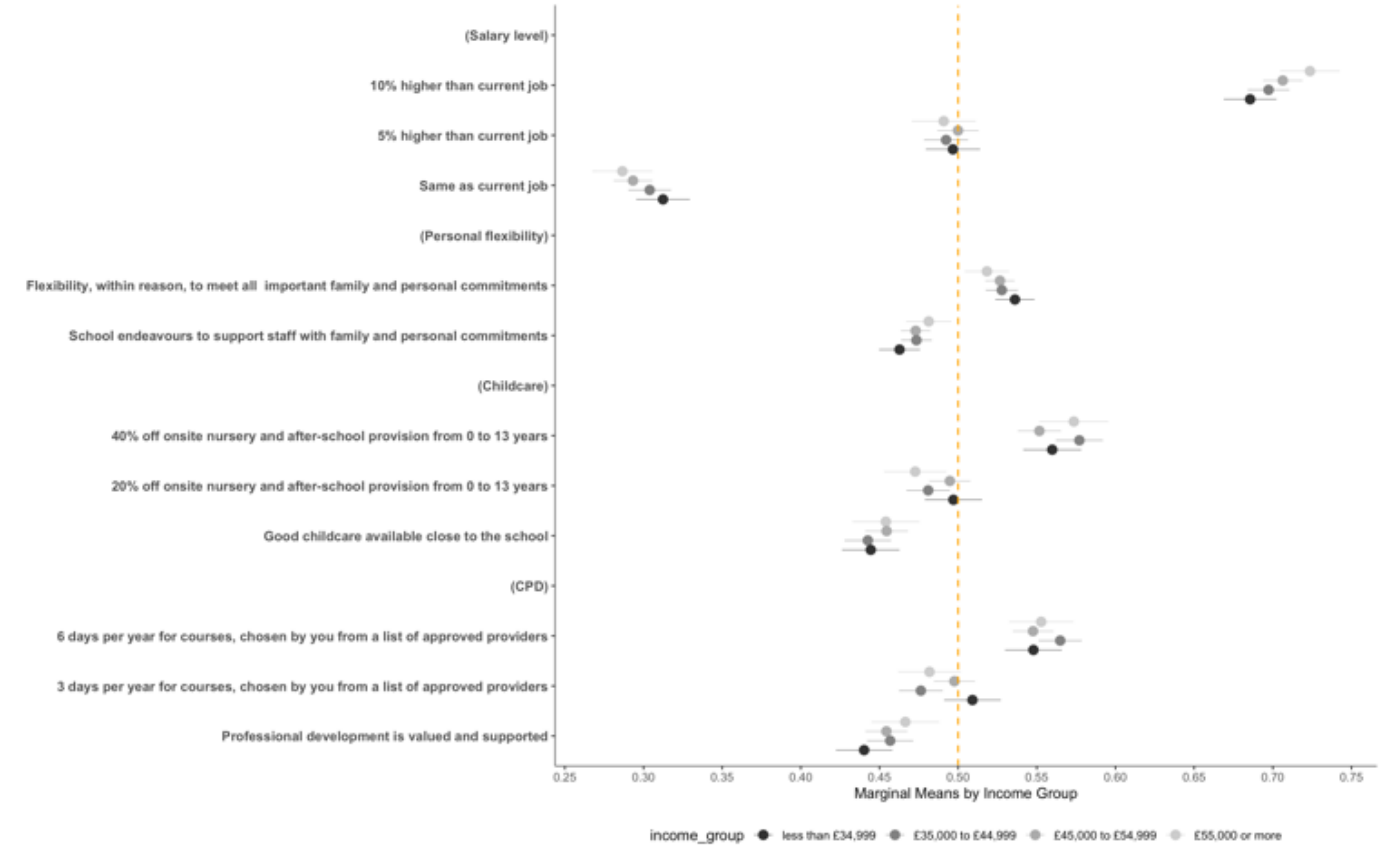
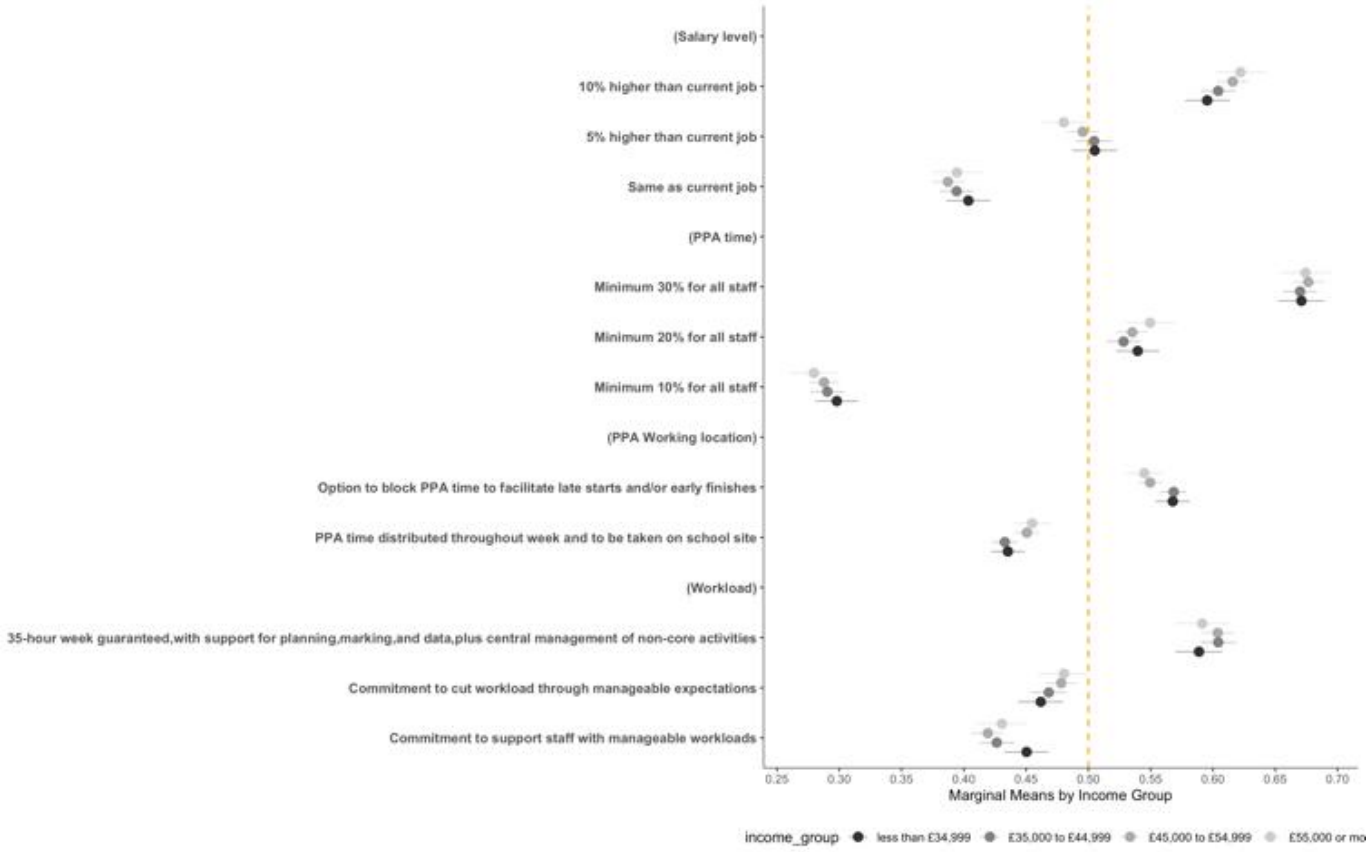

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Age
age Age in 20s Age in 30s Age in 40s Age in 50s+

Figure A7: Experiments 1-4, marginal means by teacher salary






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Income Group
income_group
less than £34,999
£35,000 to £44,999
£45,000 to £54,999
£55,000 or mo

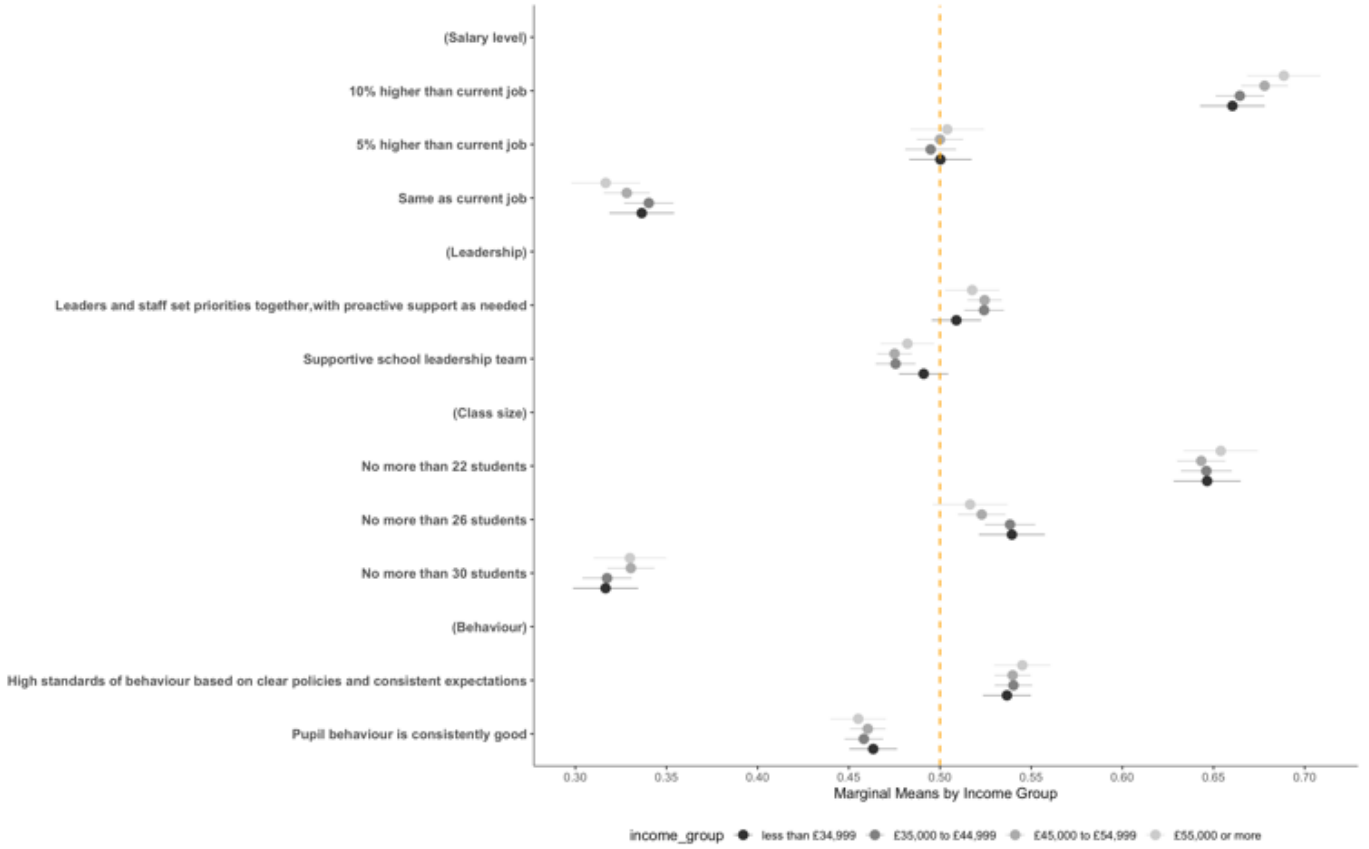
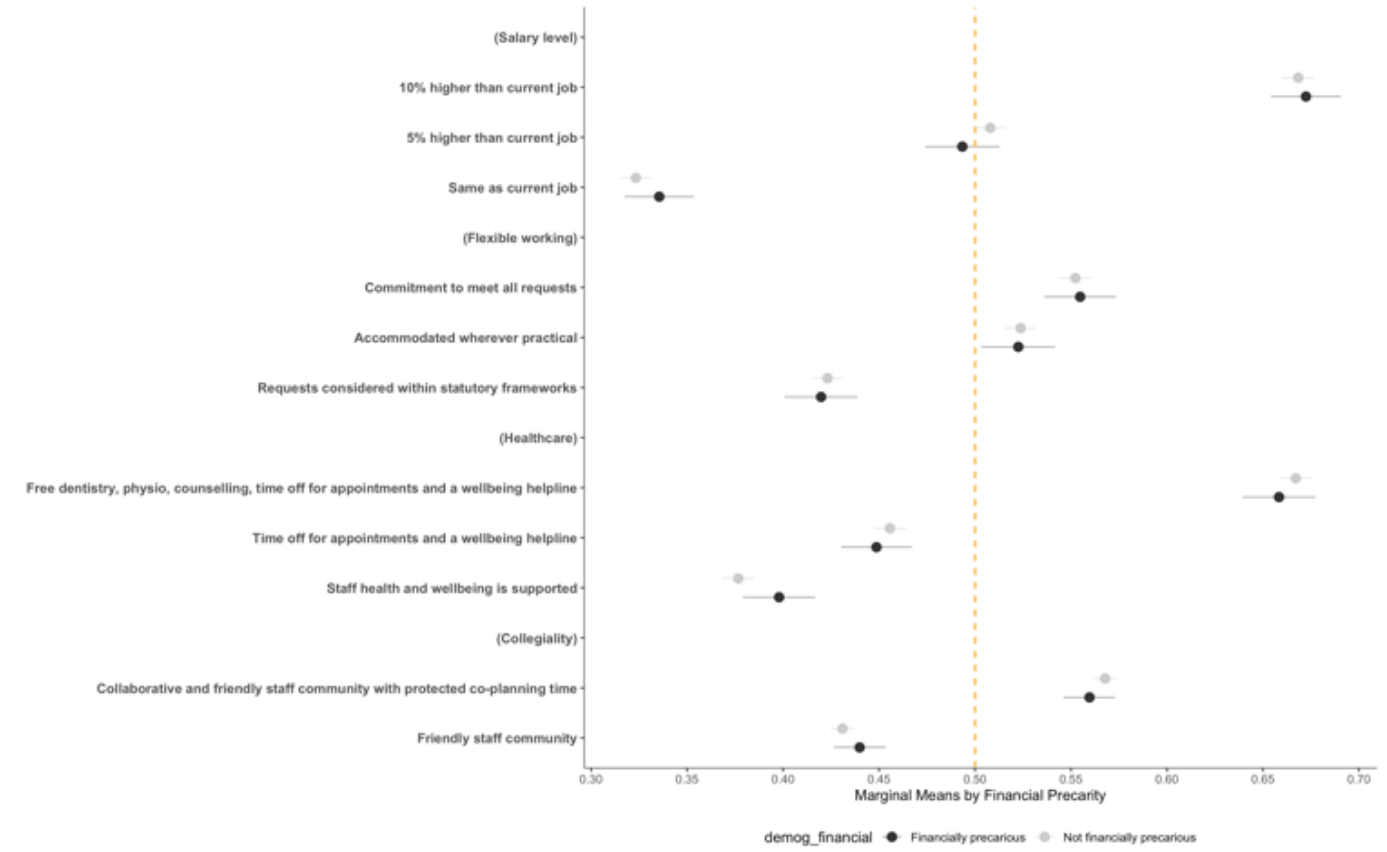
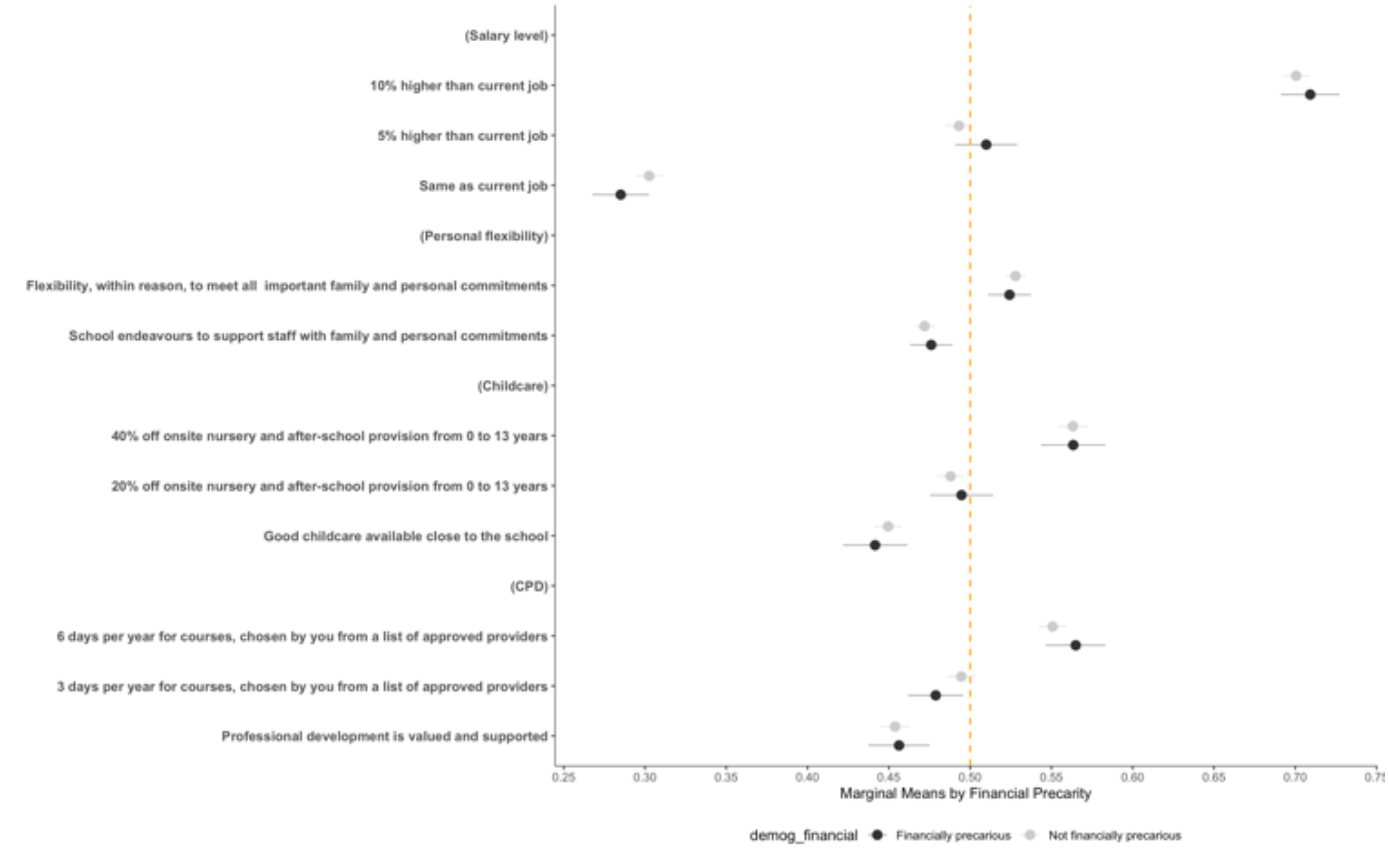
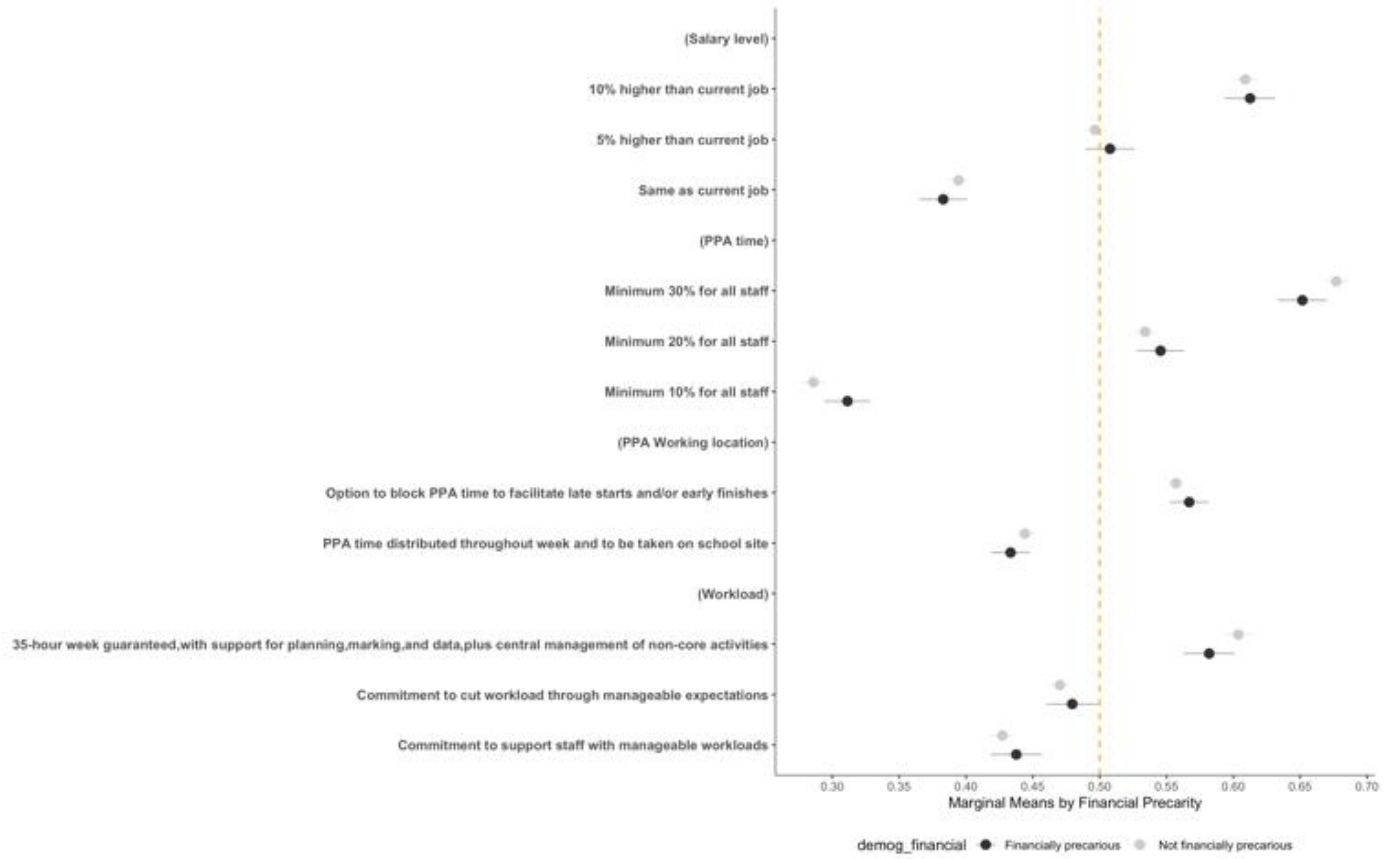

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Income Group
income_group
less than £34,999
£35,000 to £44,999
£45,000 to £54,999
£55,000 or more

Figure A8: Experiments 1-4, marginal means by household financial situation






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Financial Precarity
demog_financial
Financially precarious
Not financially precarious

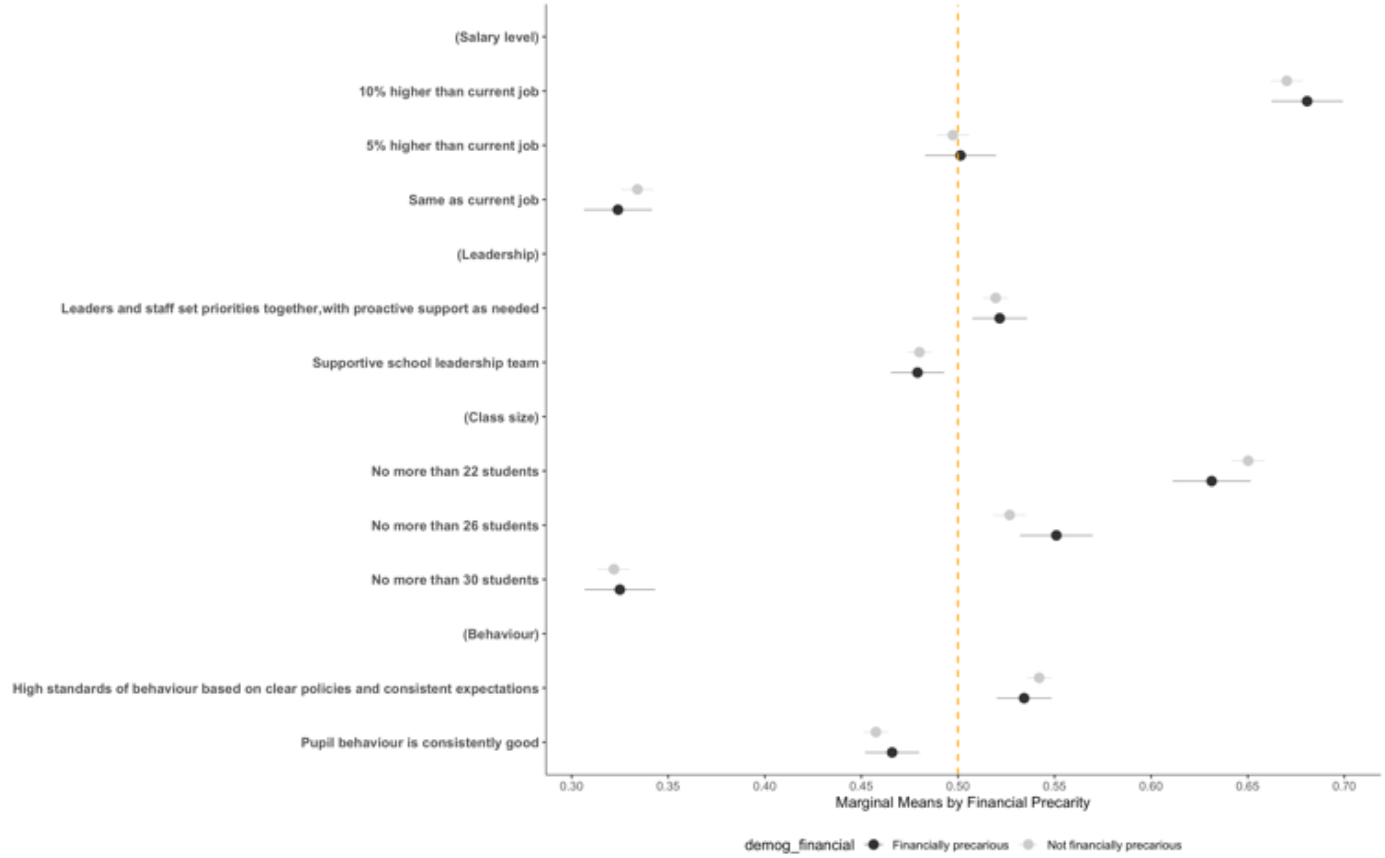
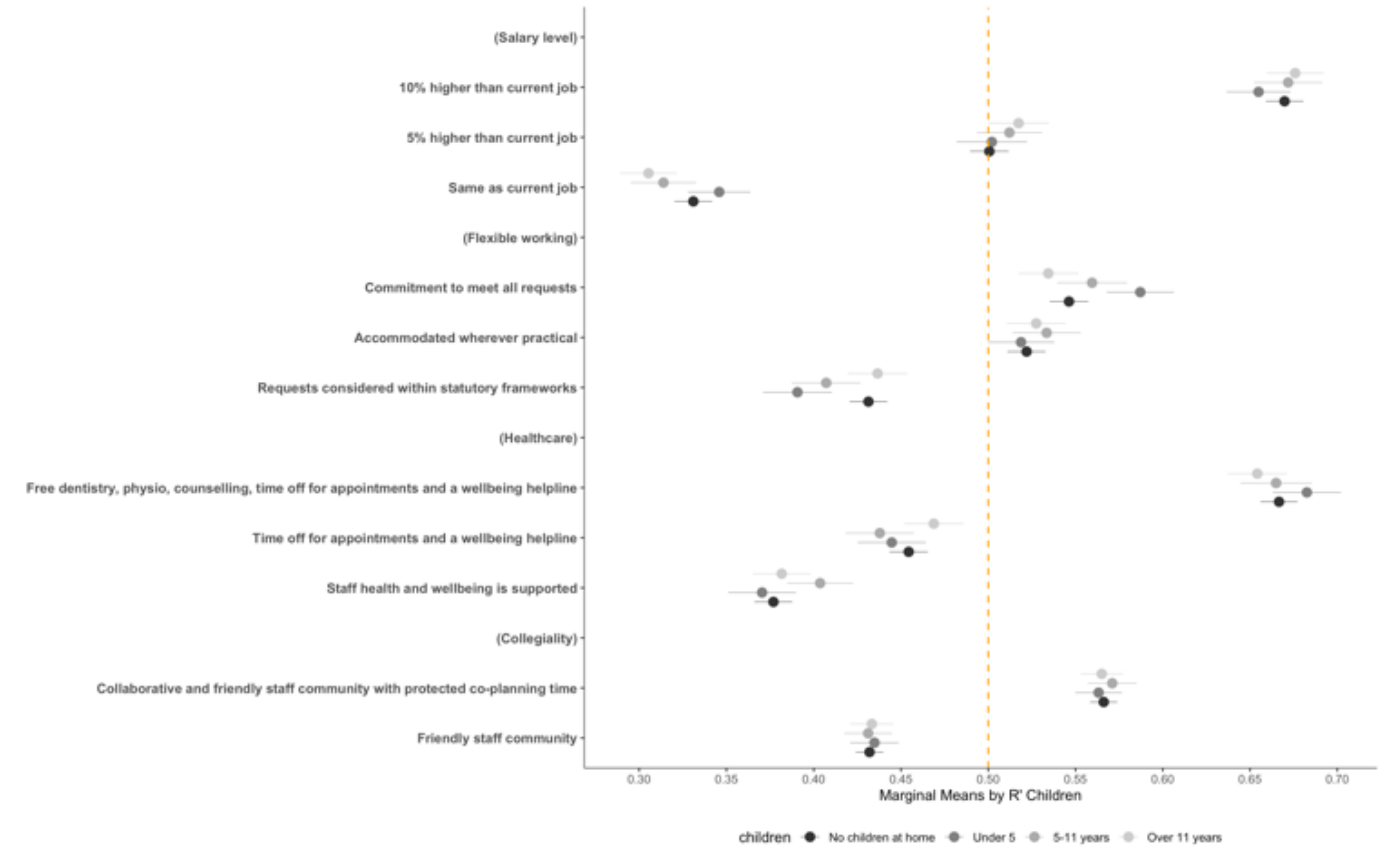
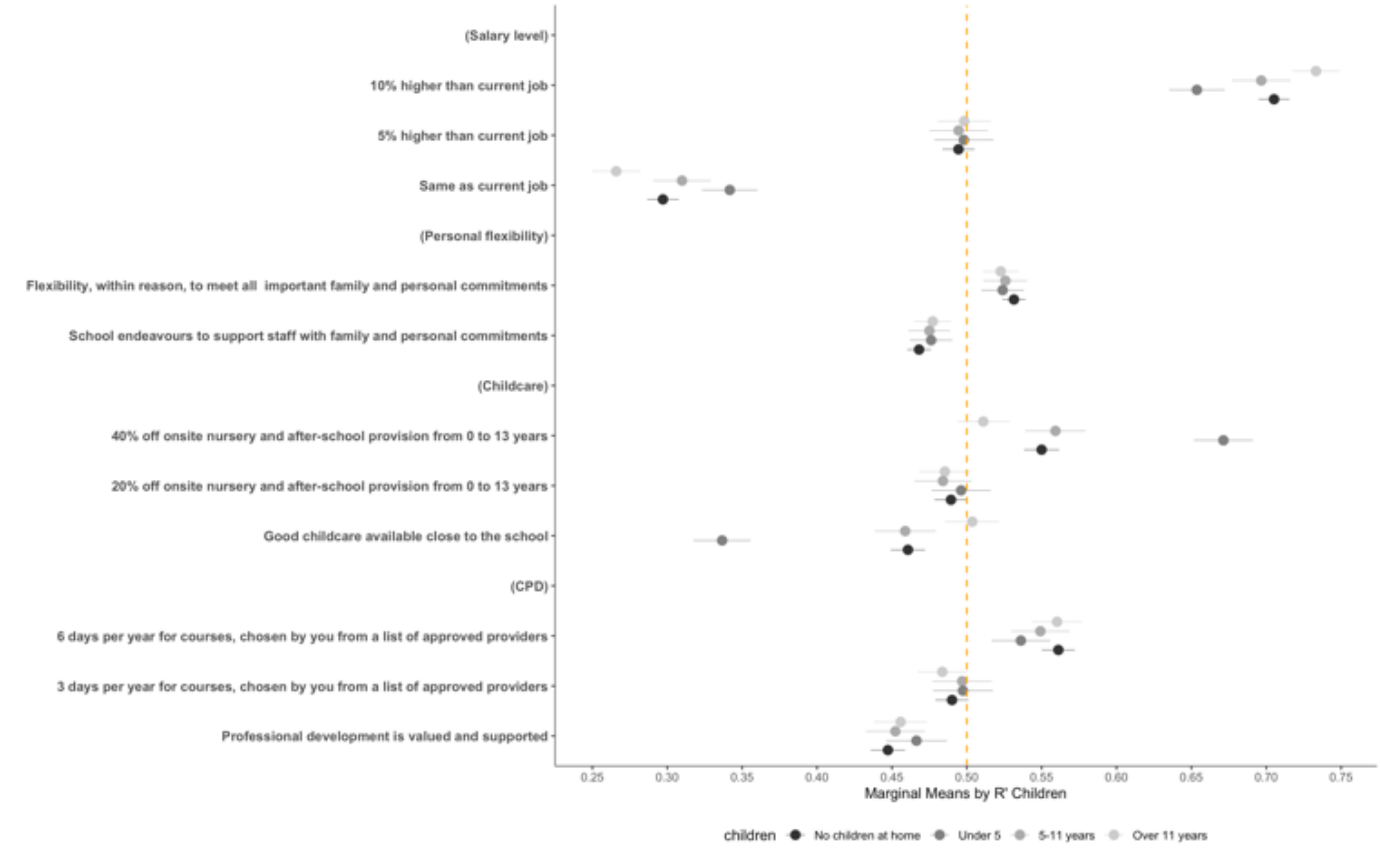
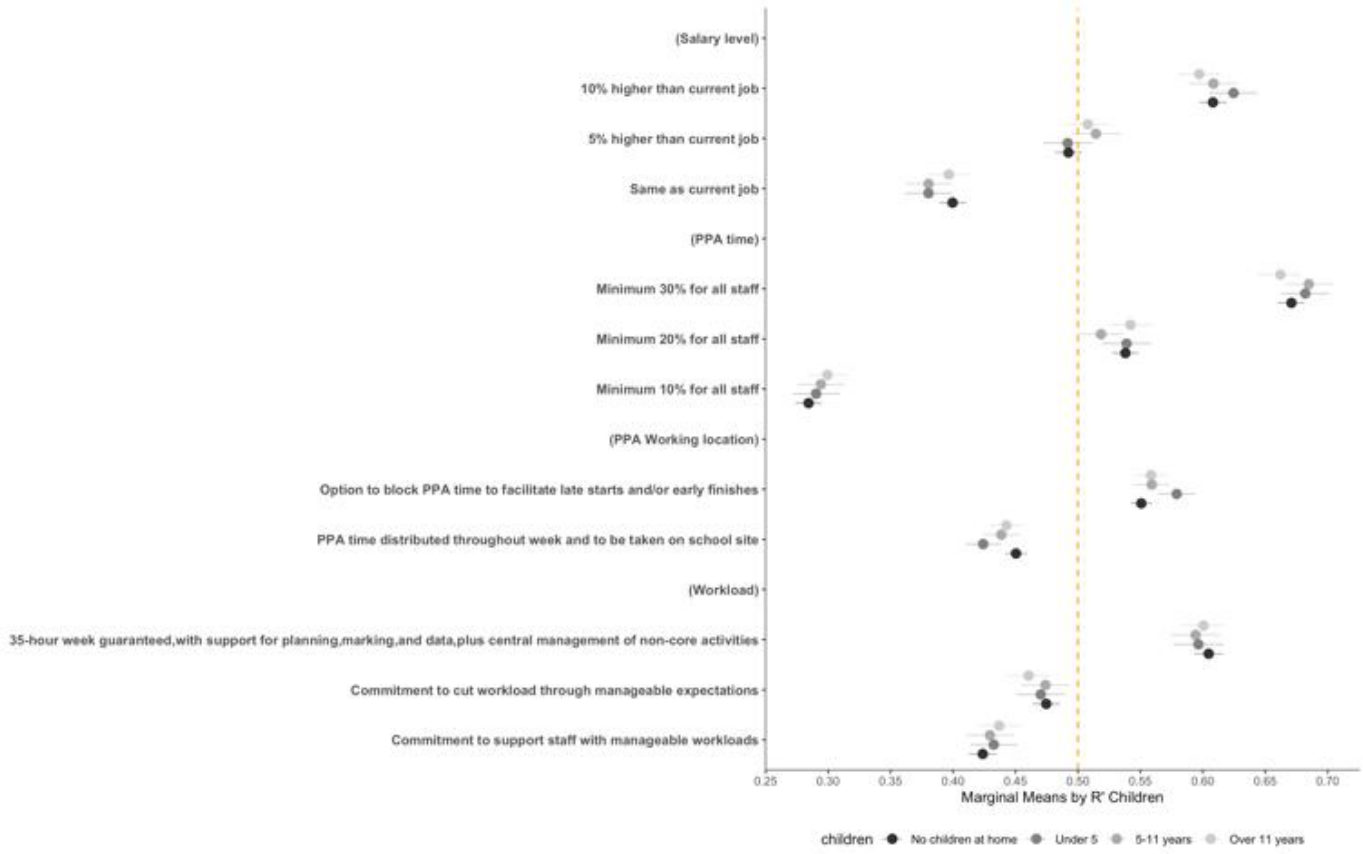

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Financial Precarity
demog_financial
Financially precarious
Not financially precarious

Figure A9: Experiments 1-4, marginal means by children at home






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by R' Children
children No children at home Under 5 5-11 years Over 11 years

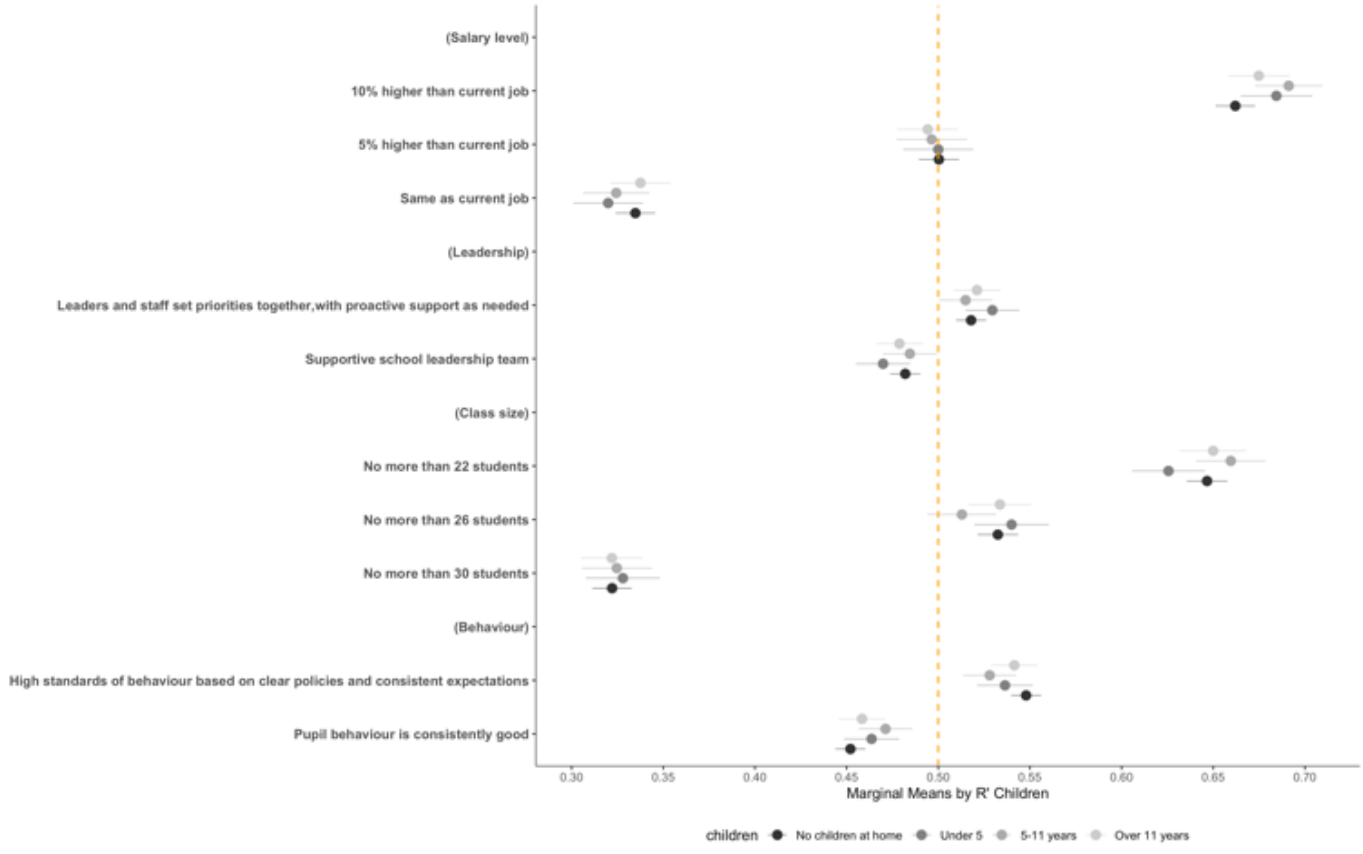
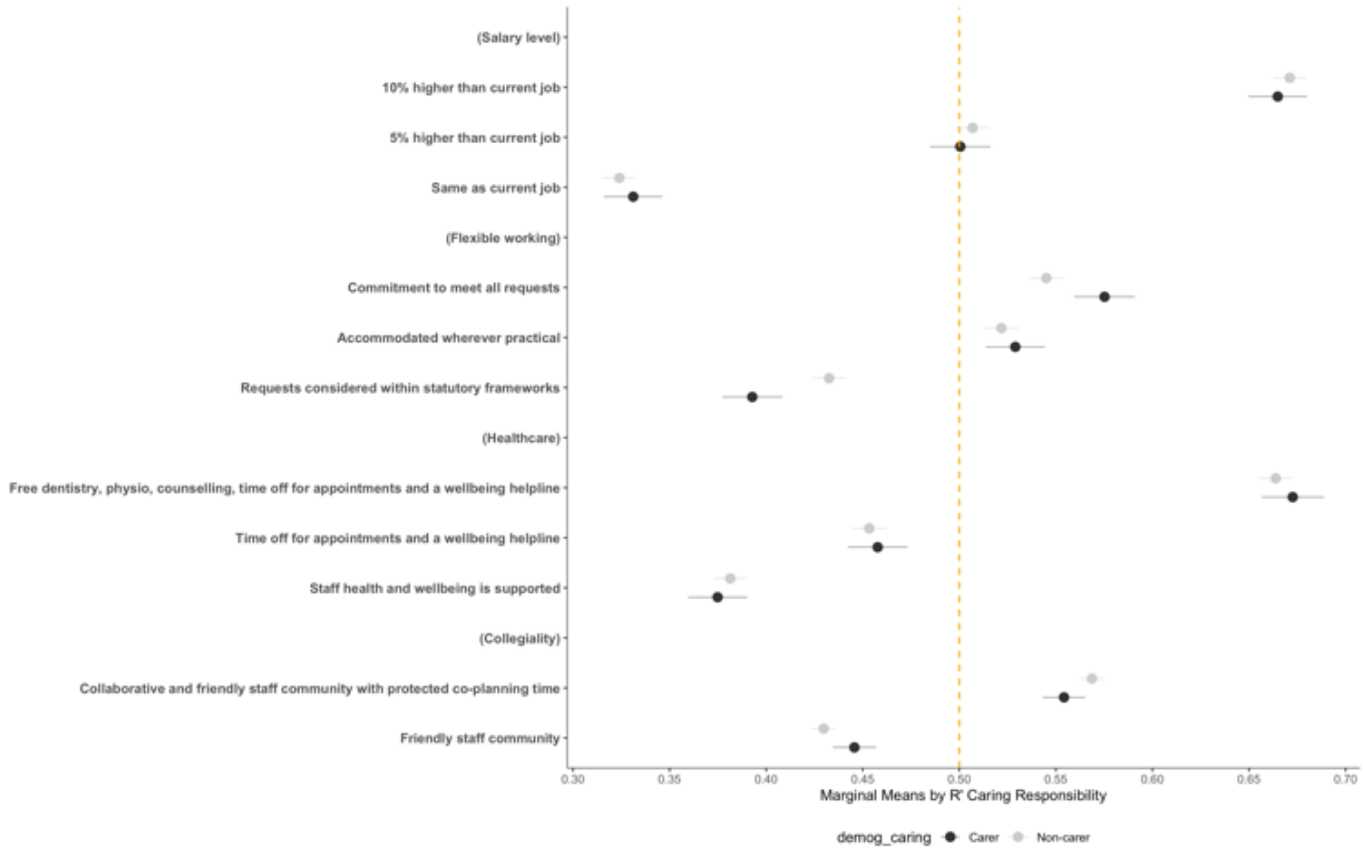
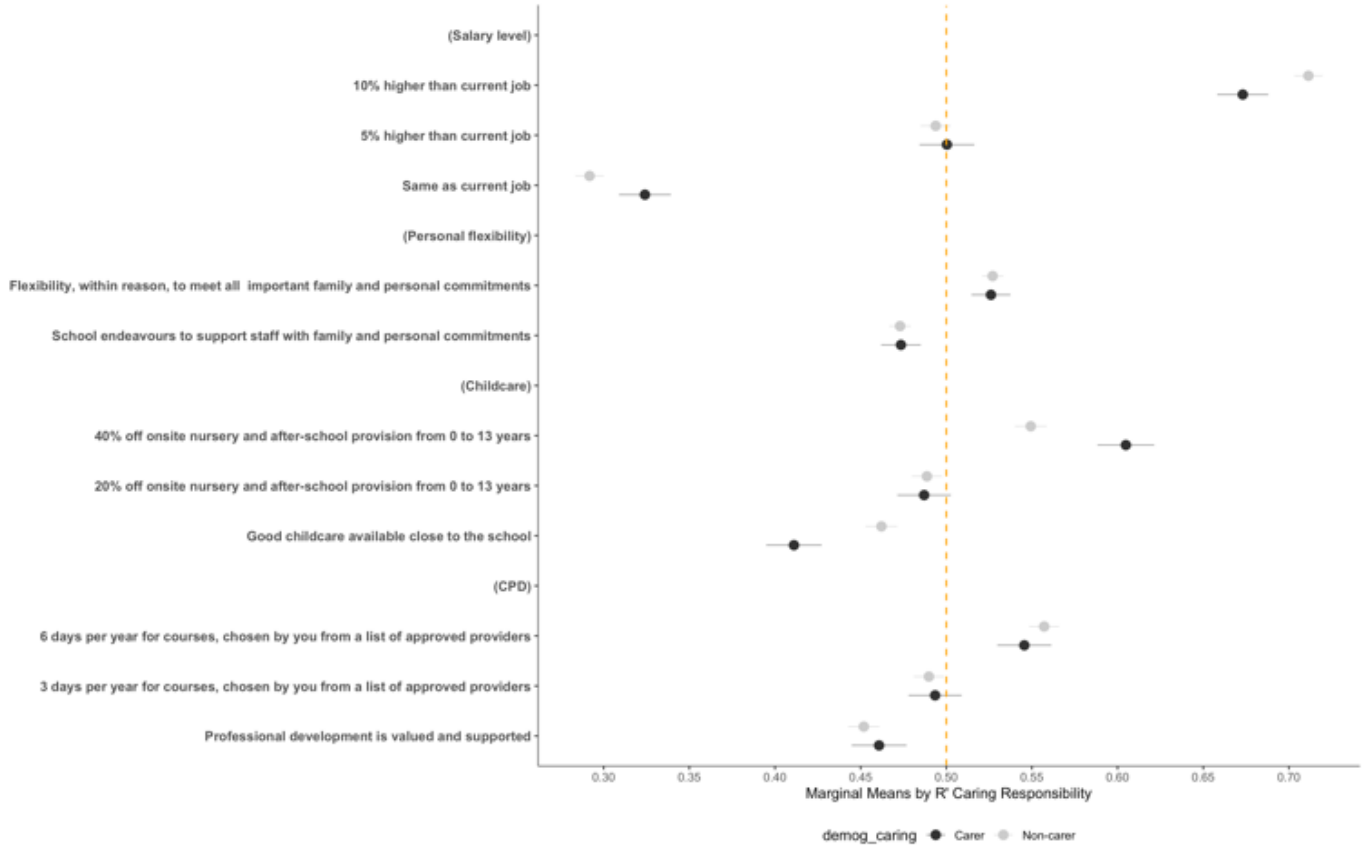
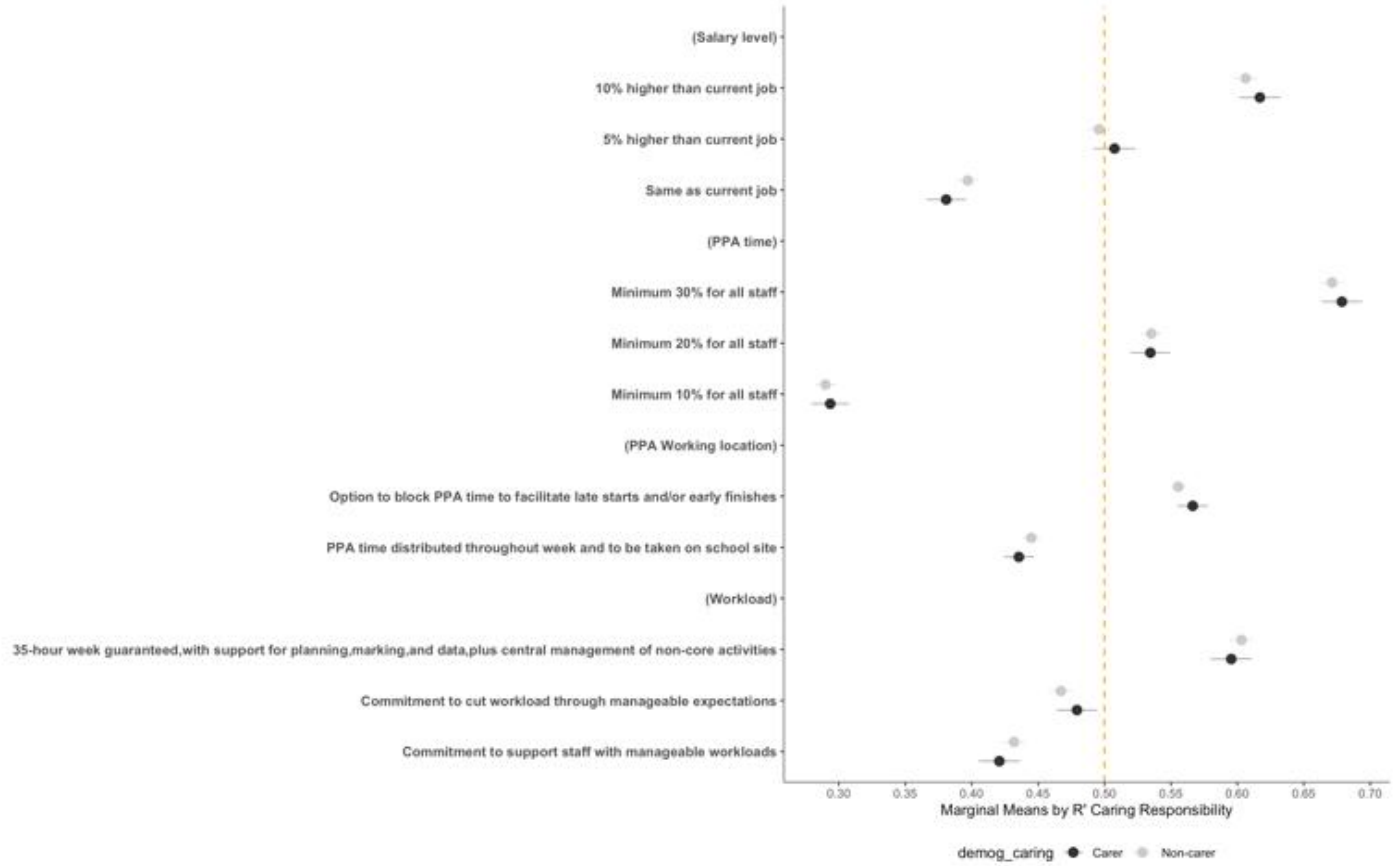

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by R' Children
children No children at home Under 5 5-11 years Over 11 years

Figure A10: Experiments 1-4, marginal means by caring duties






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by R' Caring Responsibility
demog_caring
Carer
Non-carer

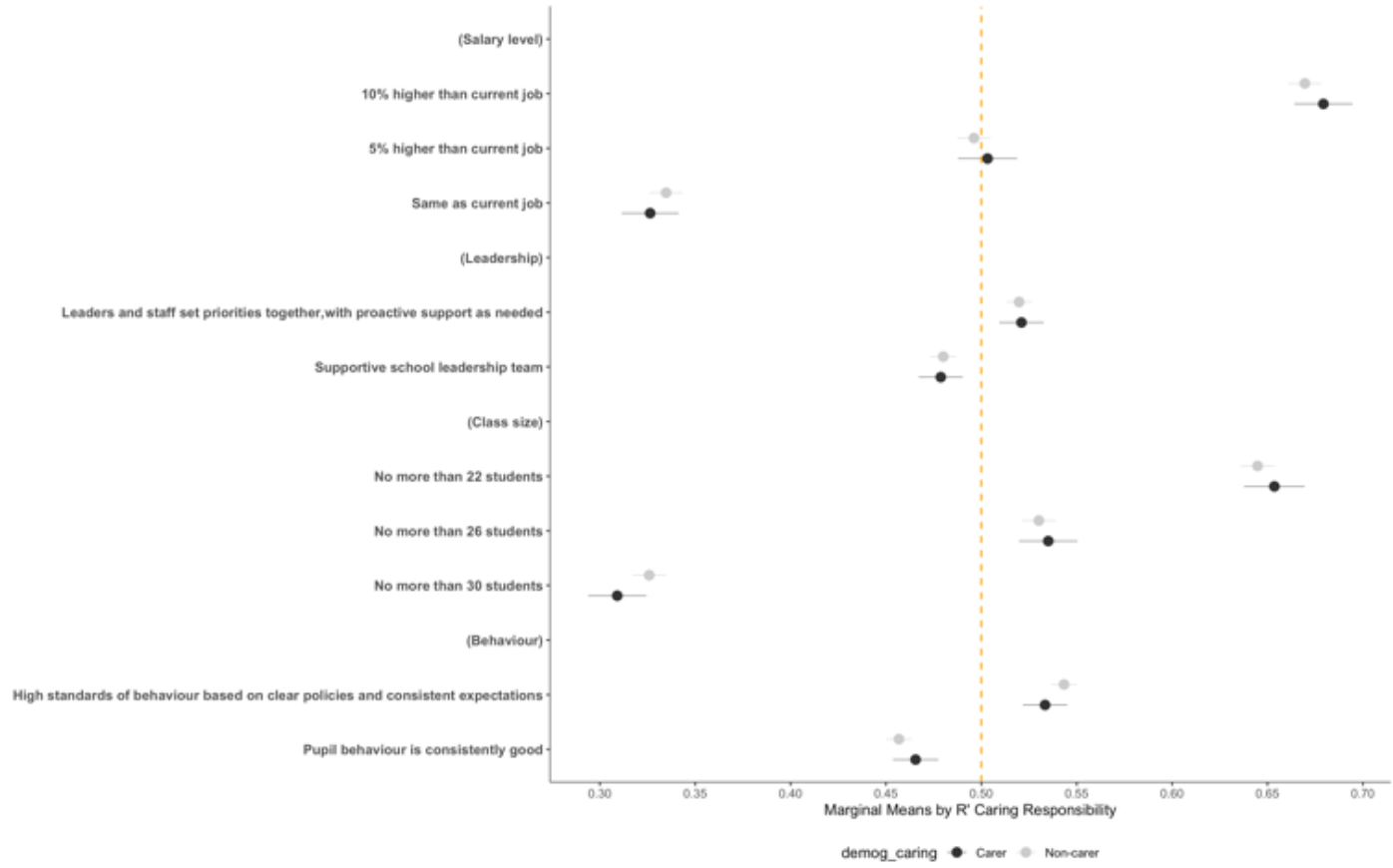
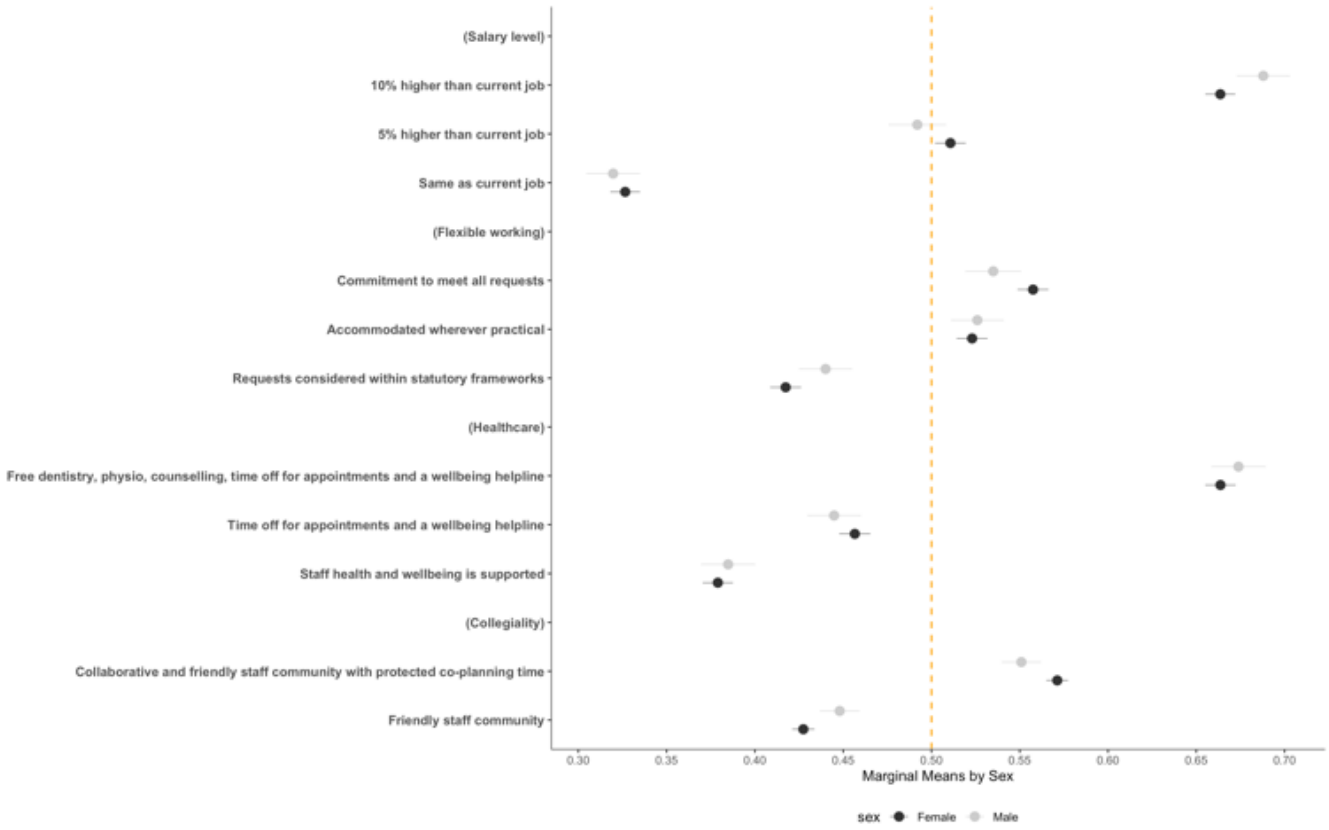
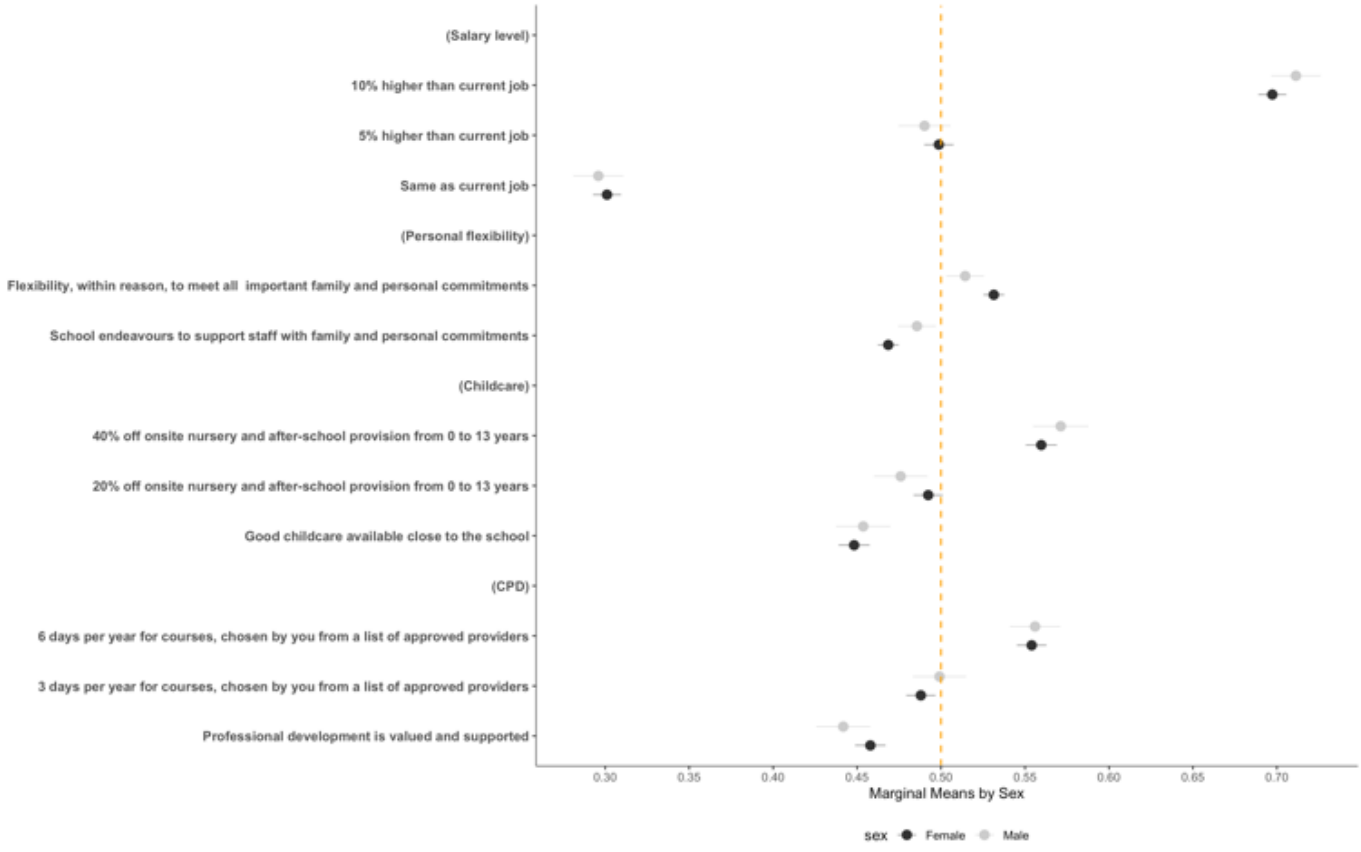
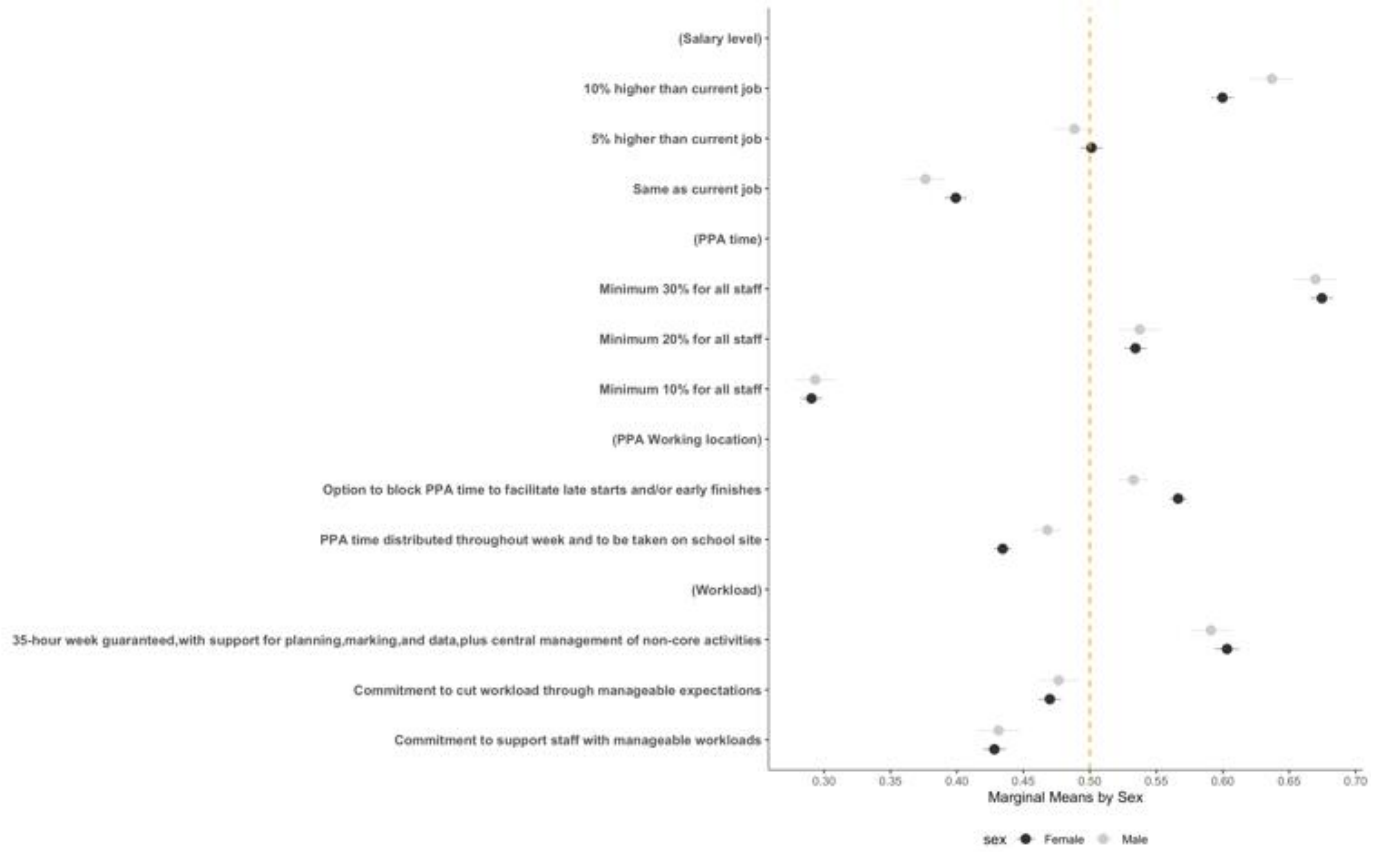

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by R' Caring Responsibility
demog_caring
Carer
Non-carer

Figure A11: Experiments 1-4, marginal means by sex






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Sex
sex
Female
Male

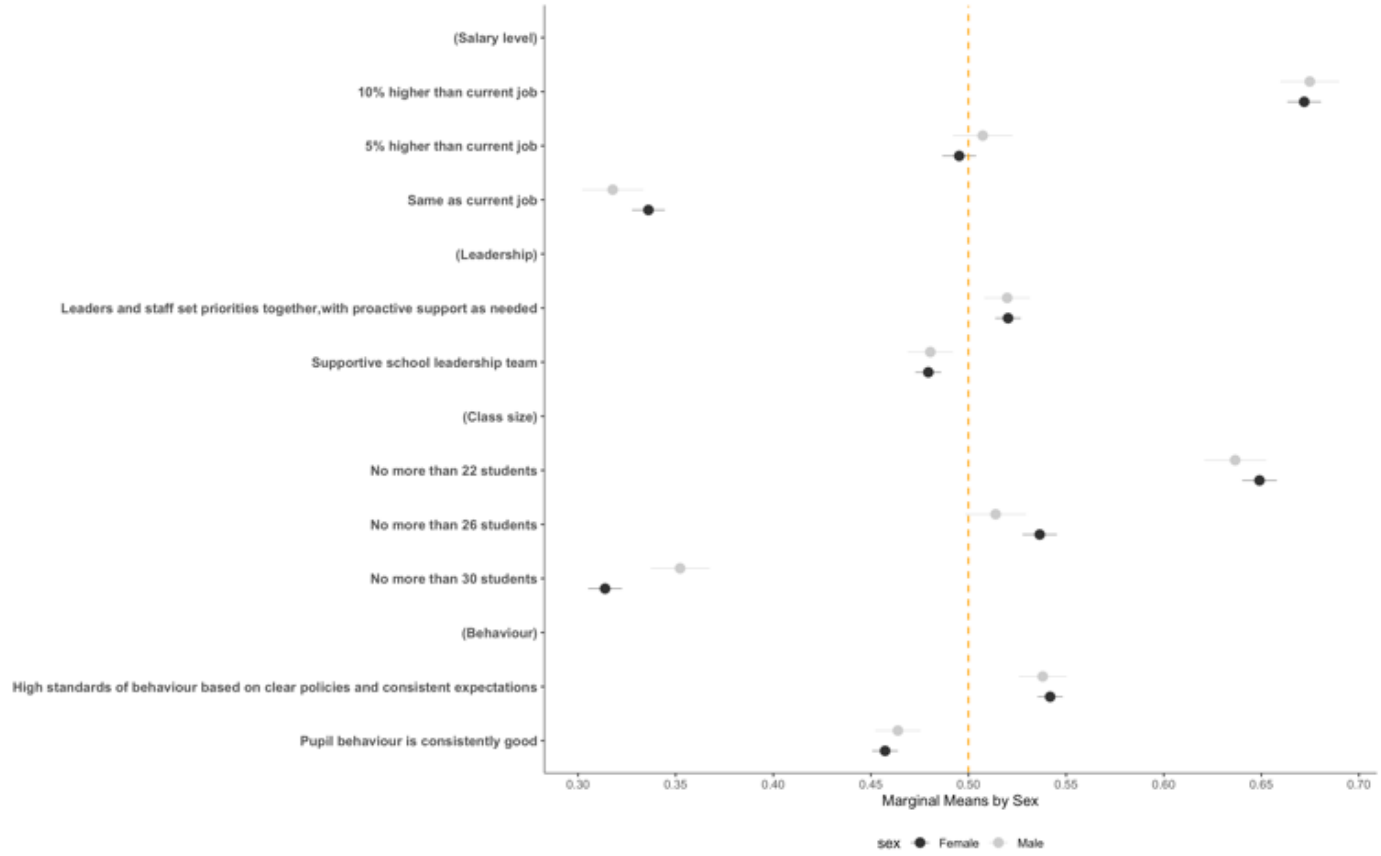
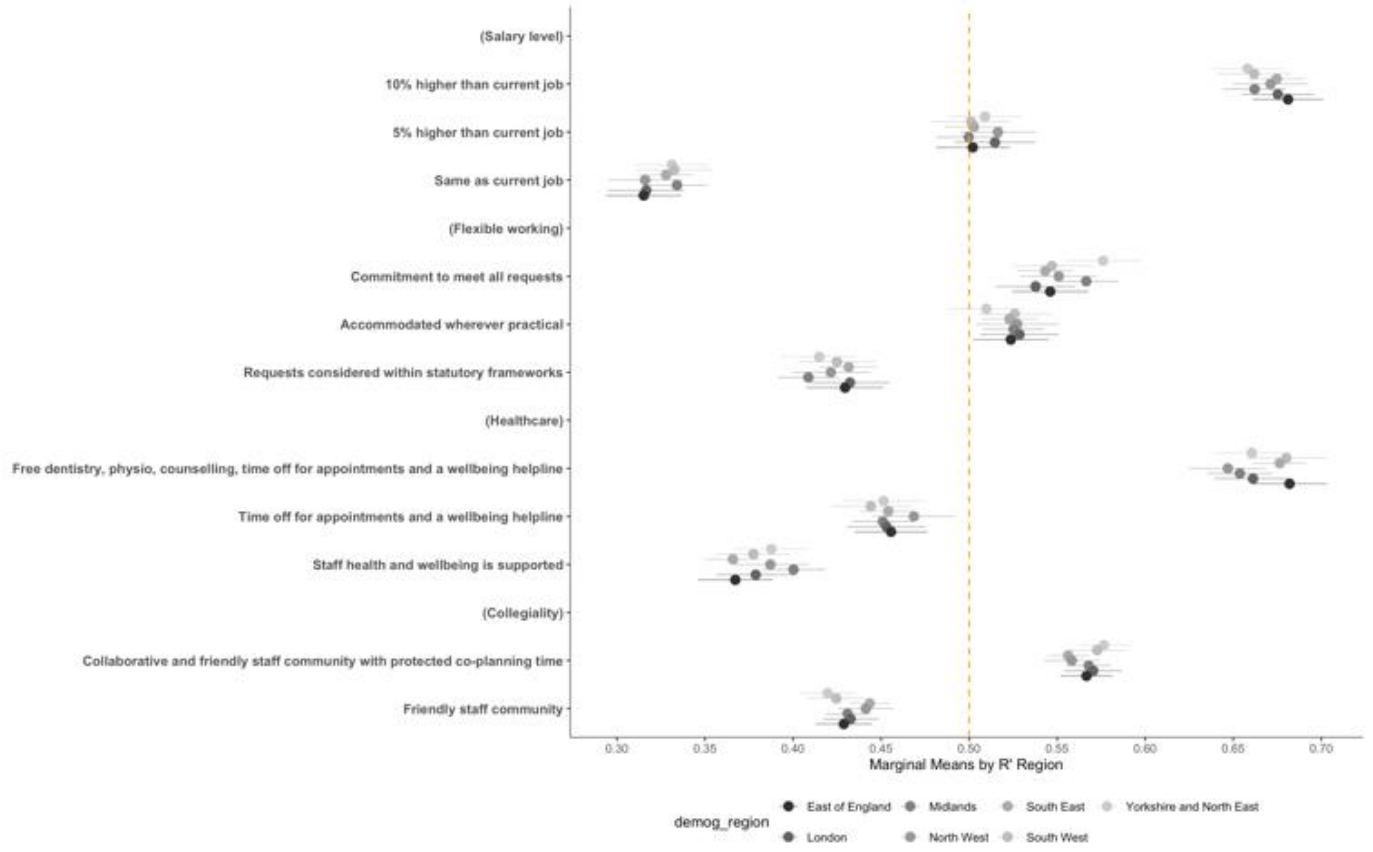
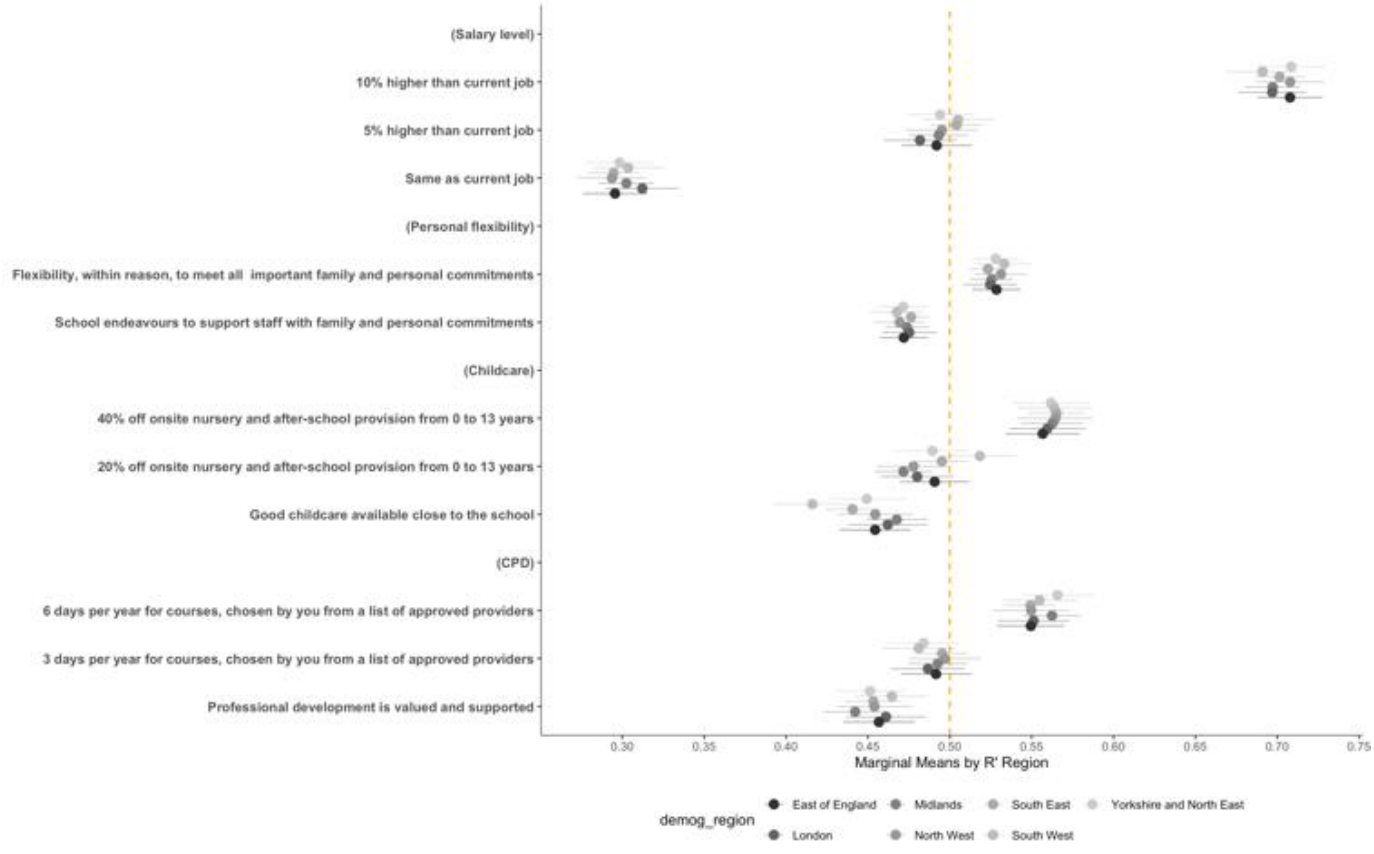
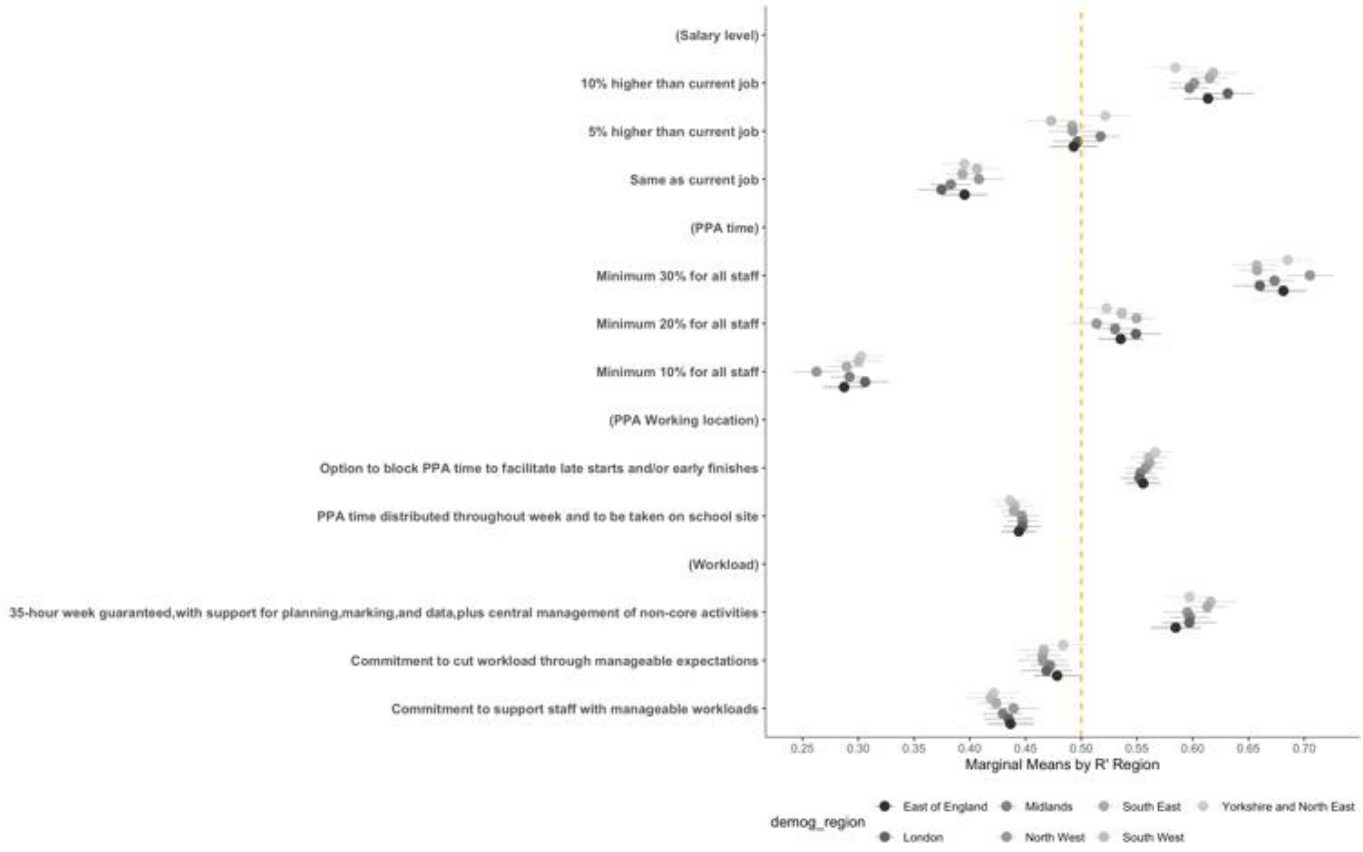
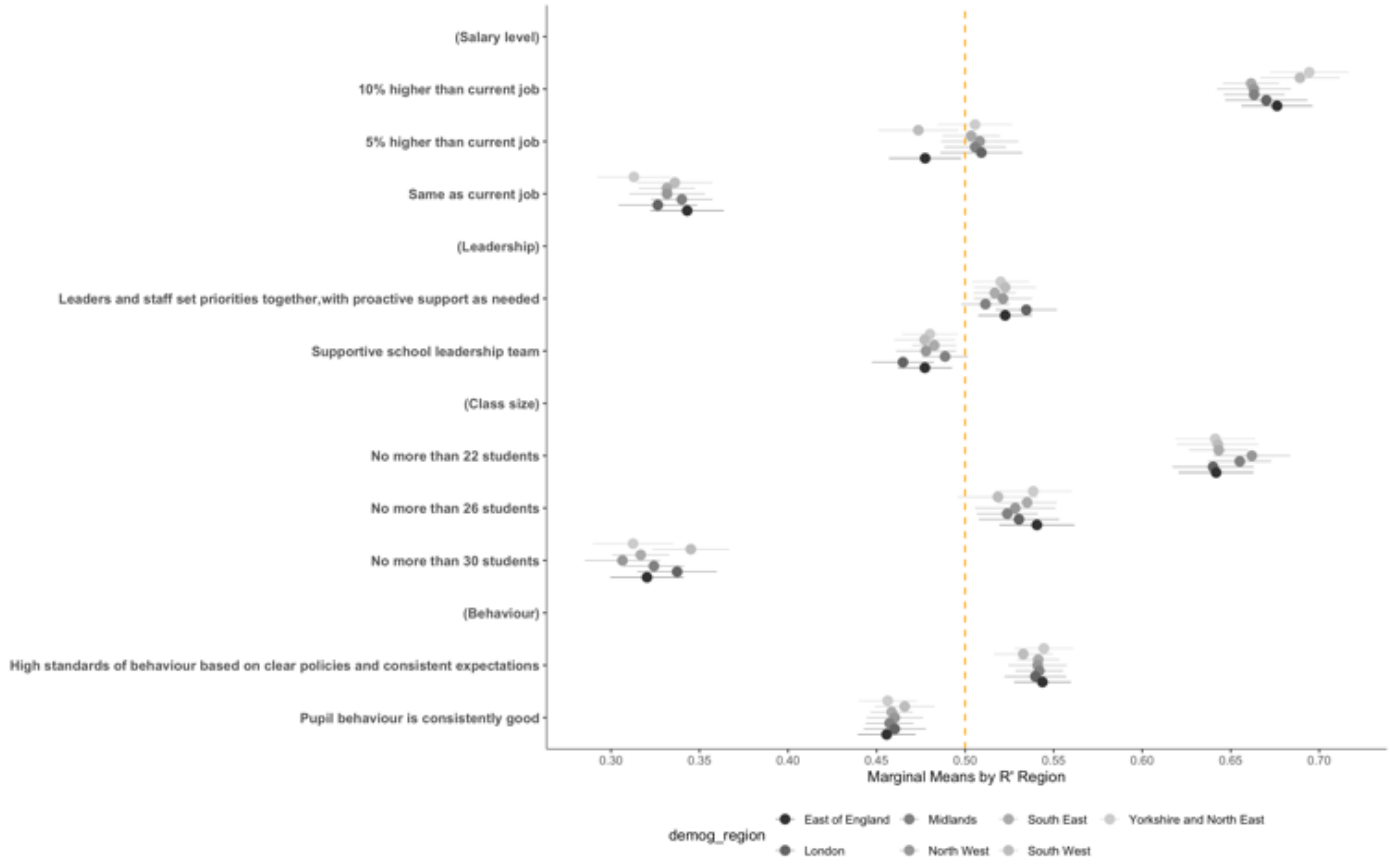

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Sex
sex
Female
Male

Figure A12: Experiments 1-4, marginal means by region



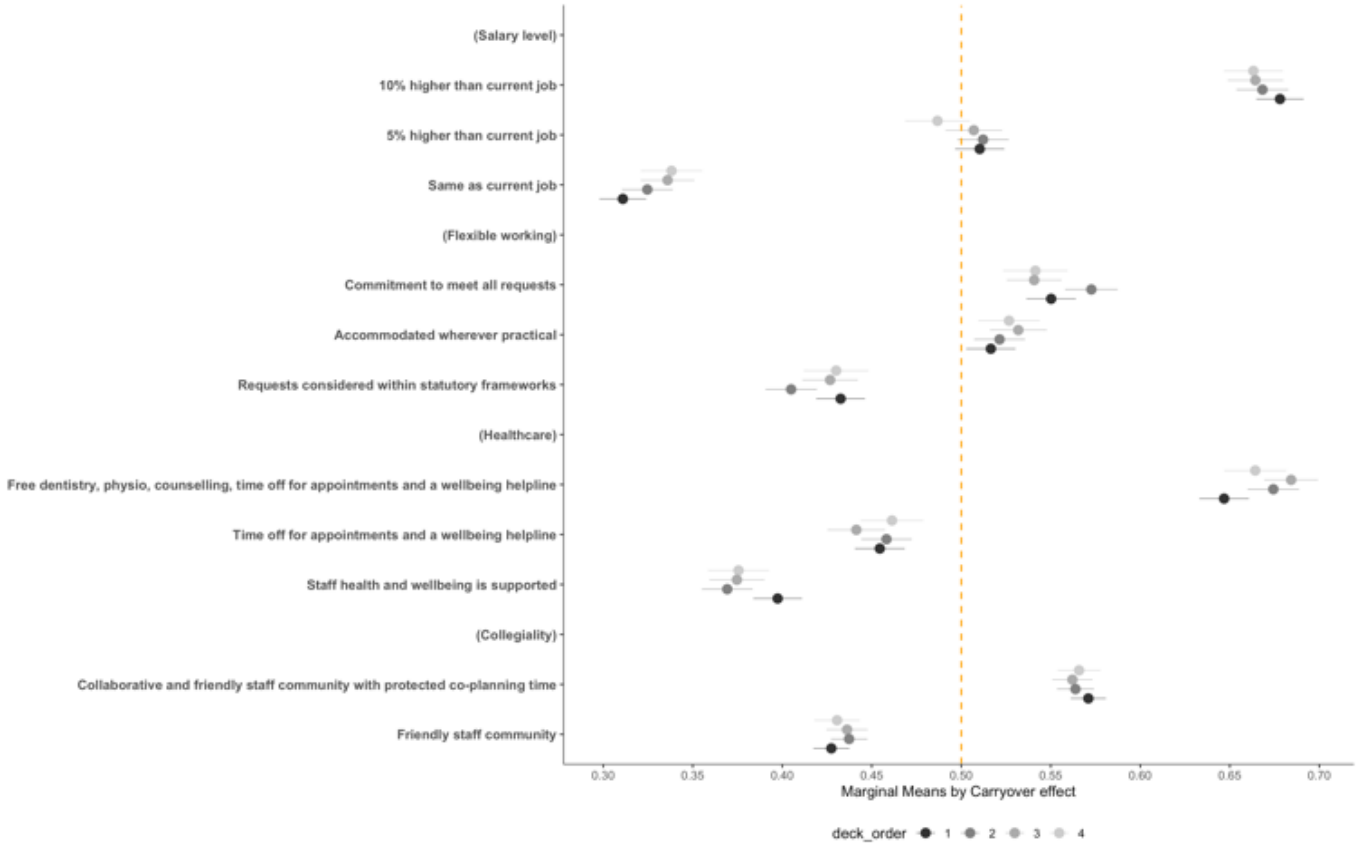


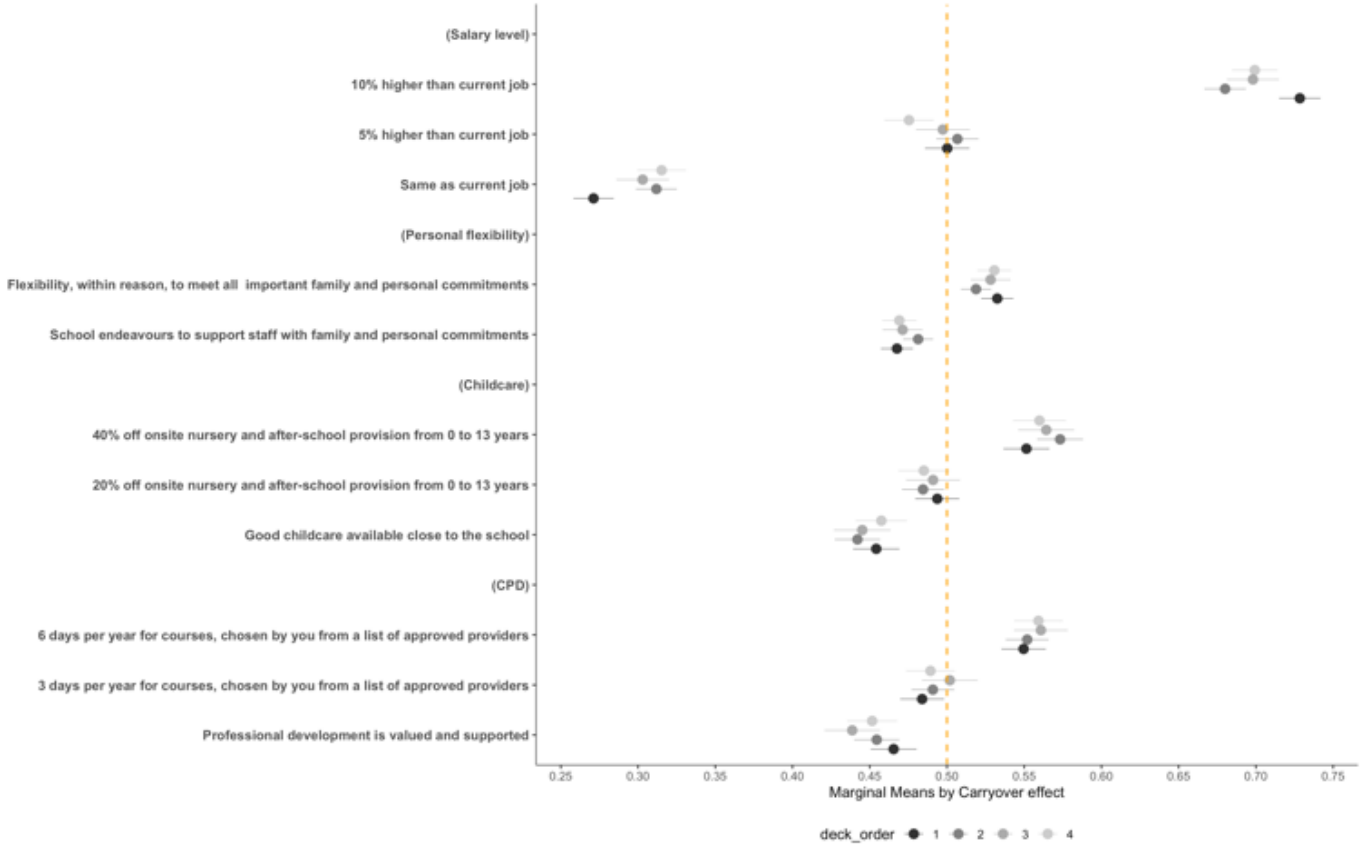

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by R² Region
demog_region
East of England
Midlands
South East
Yorkshire and North East
London
North West
South West

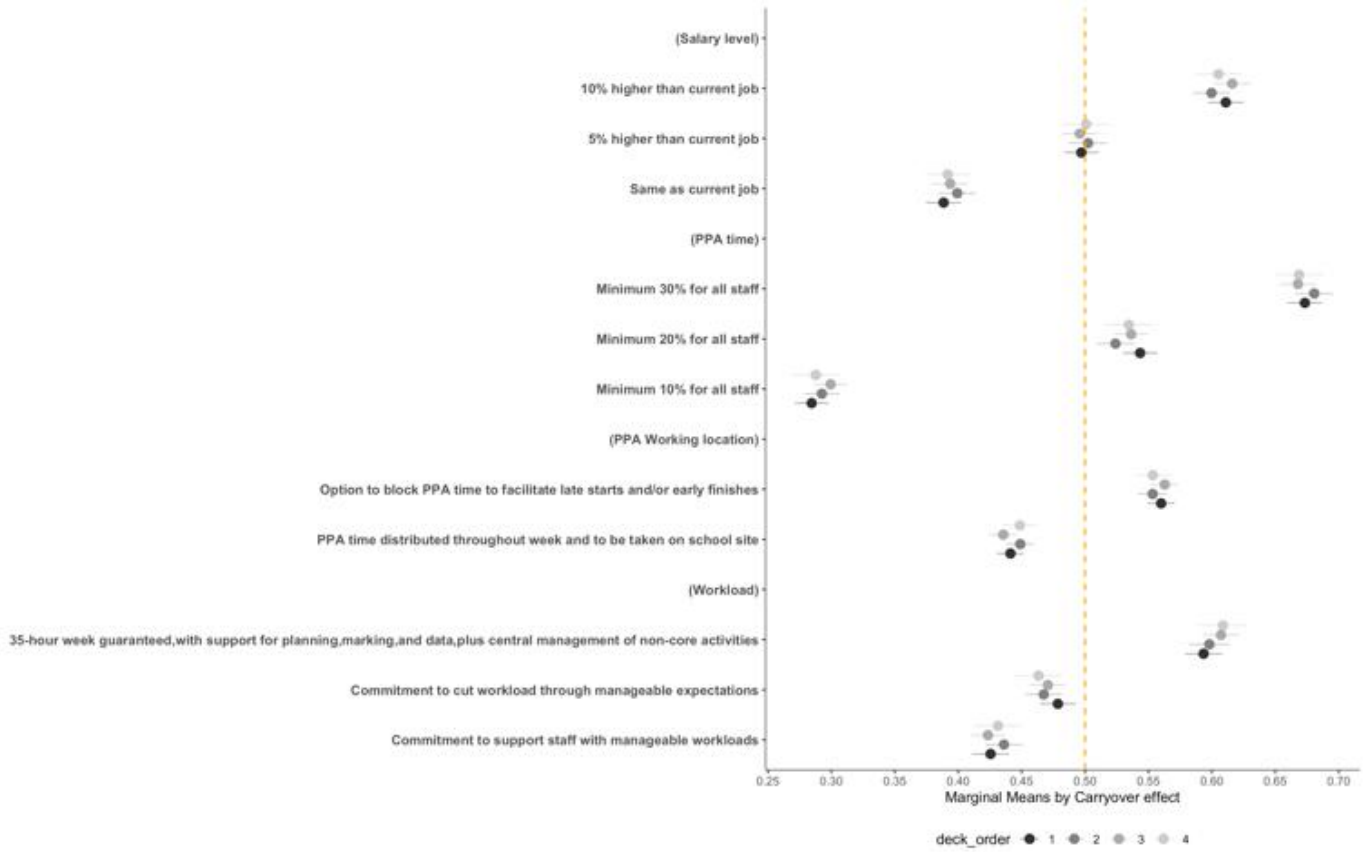

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by R² Region
demog_region
East of England
Midlands
South East
Yorkshire and North East
London
North West
South West

Appendix E - RQ1: Internality validity of experiment

Figure A13: Experiments 1-4, deck order effects






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Carryover effect
deck_order 1 2 3 4

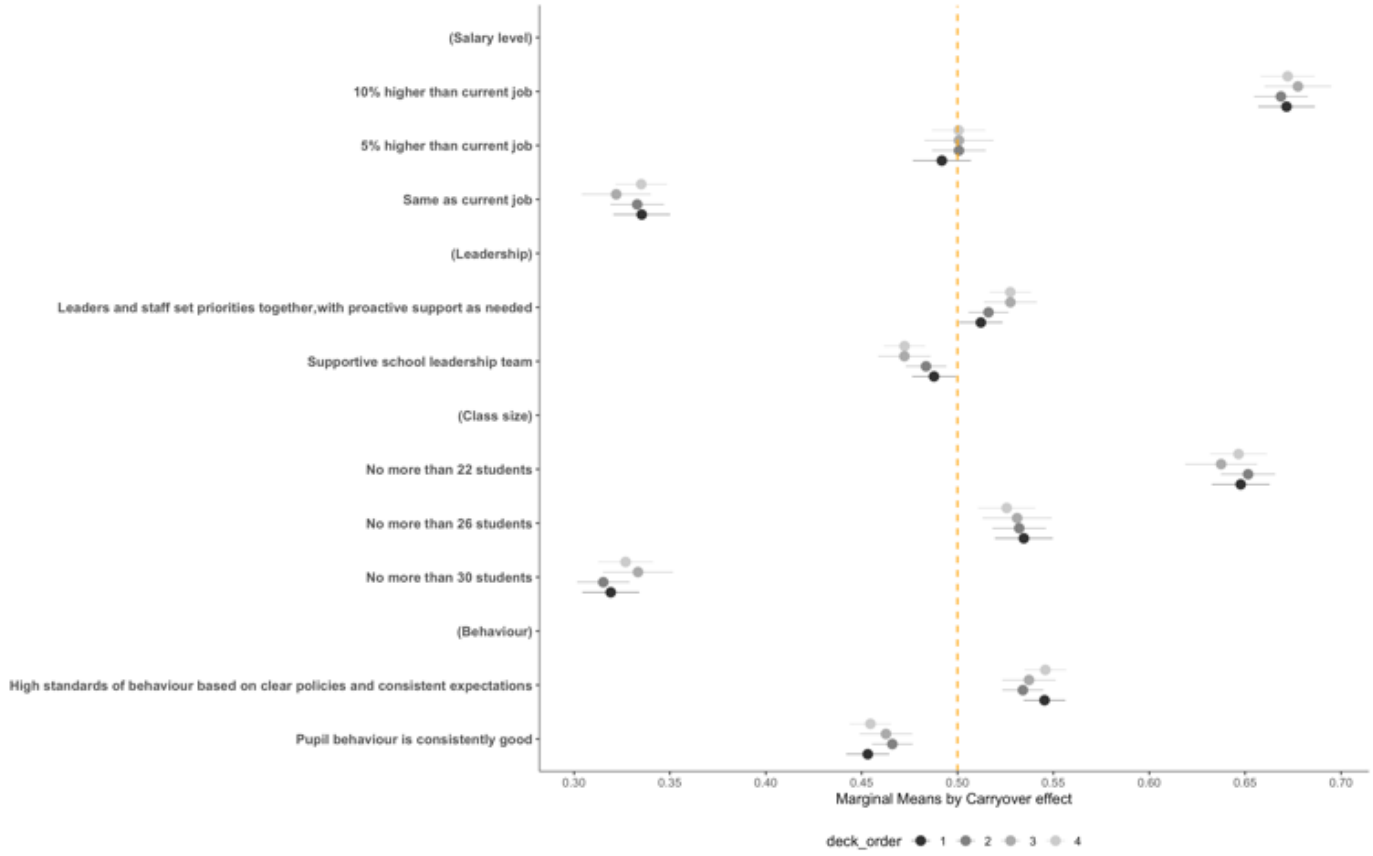
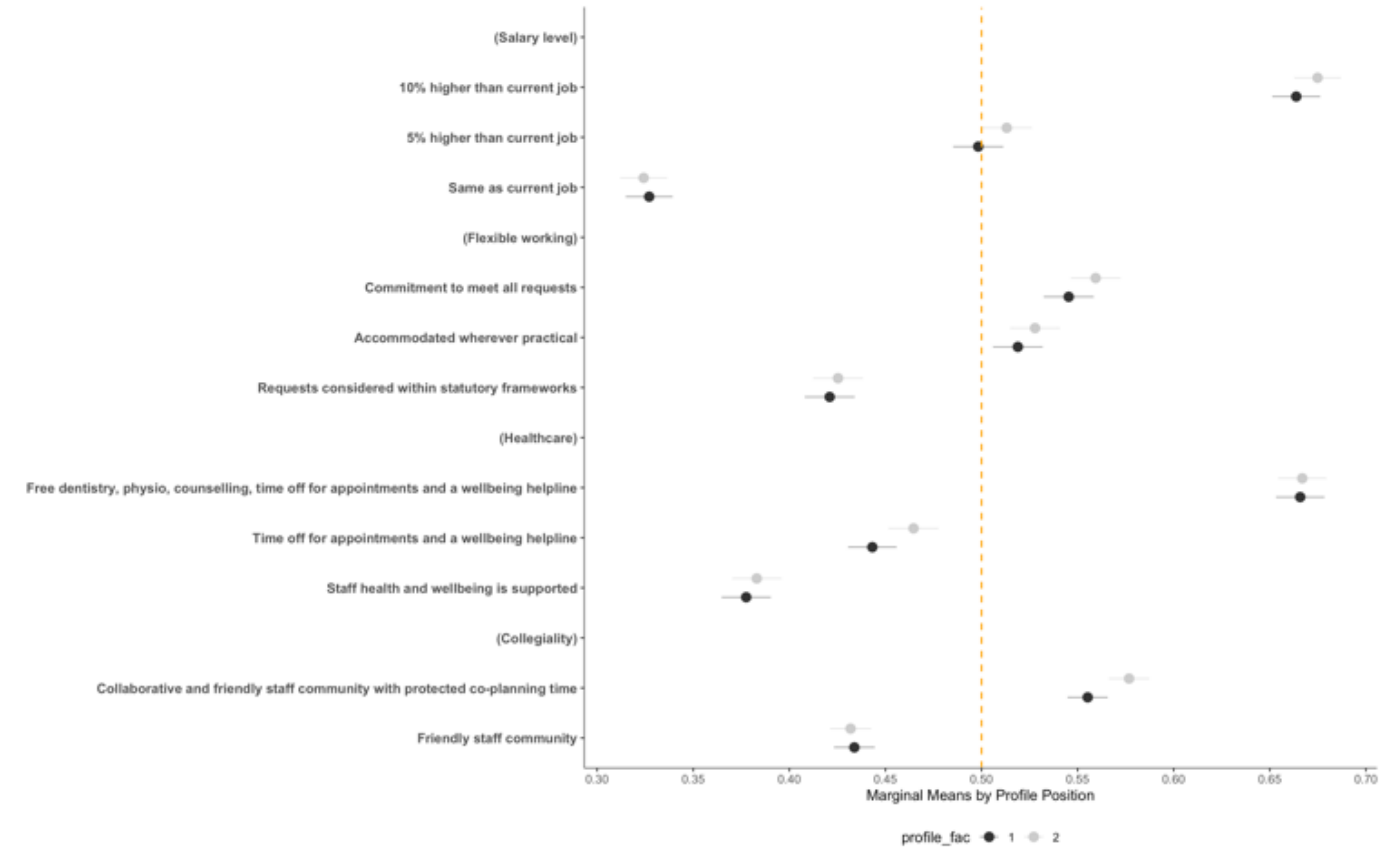
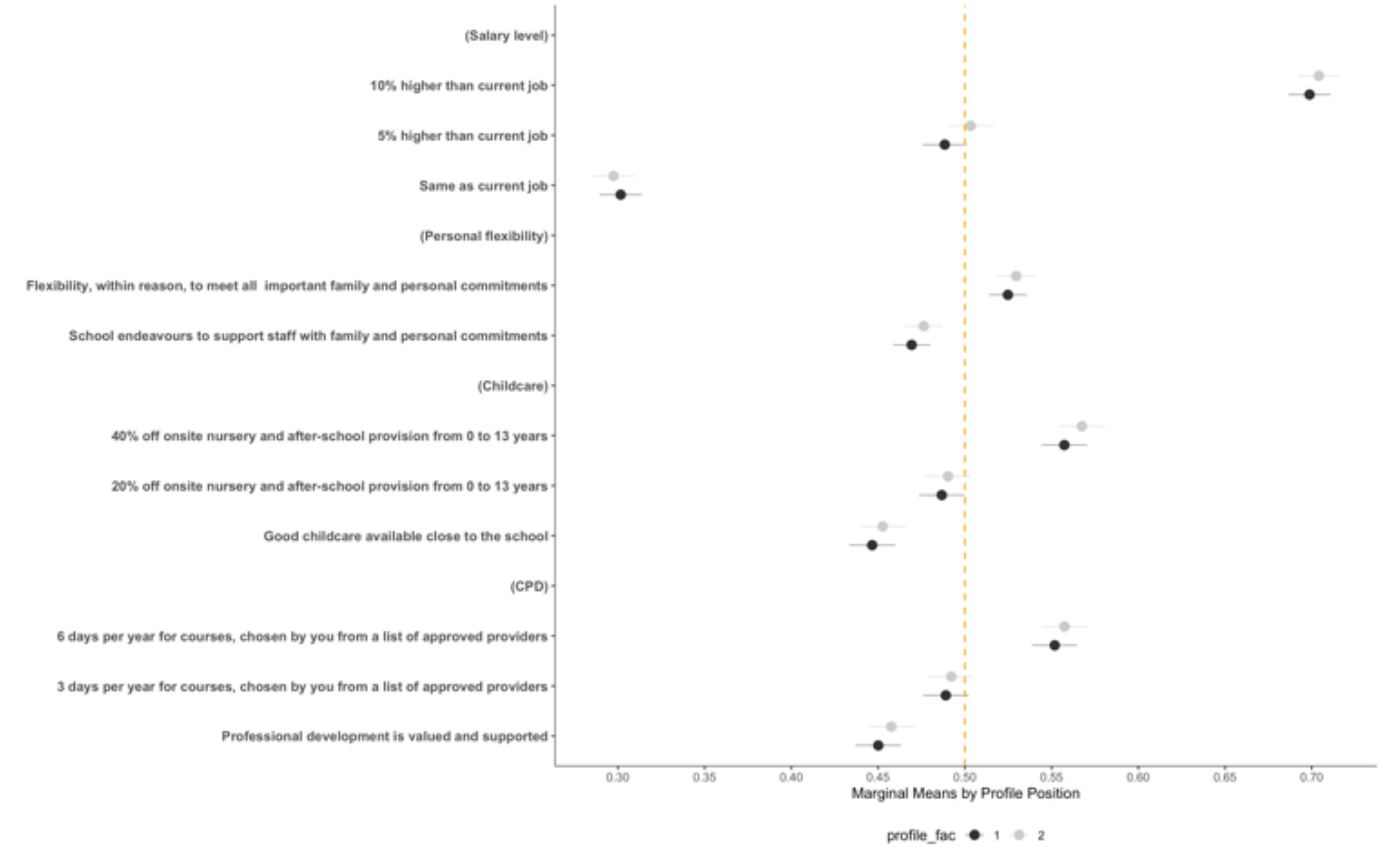
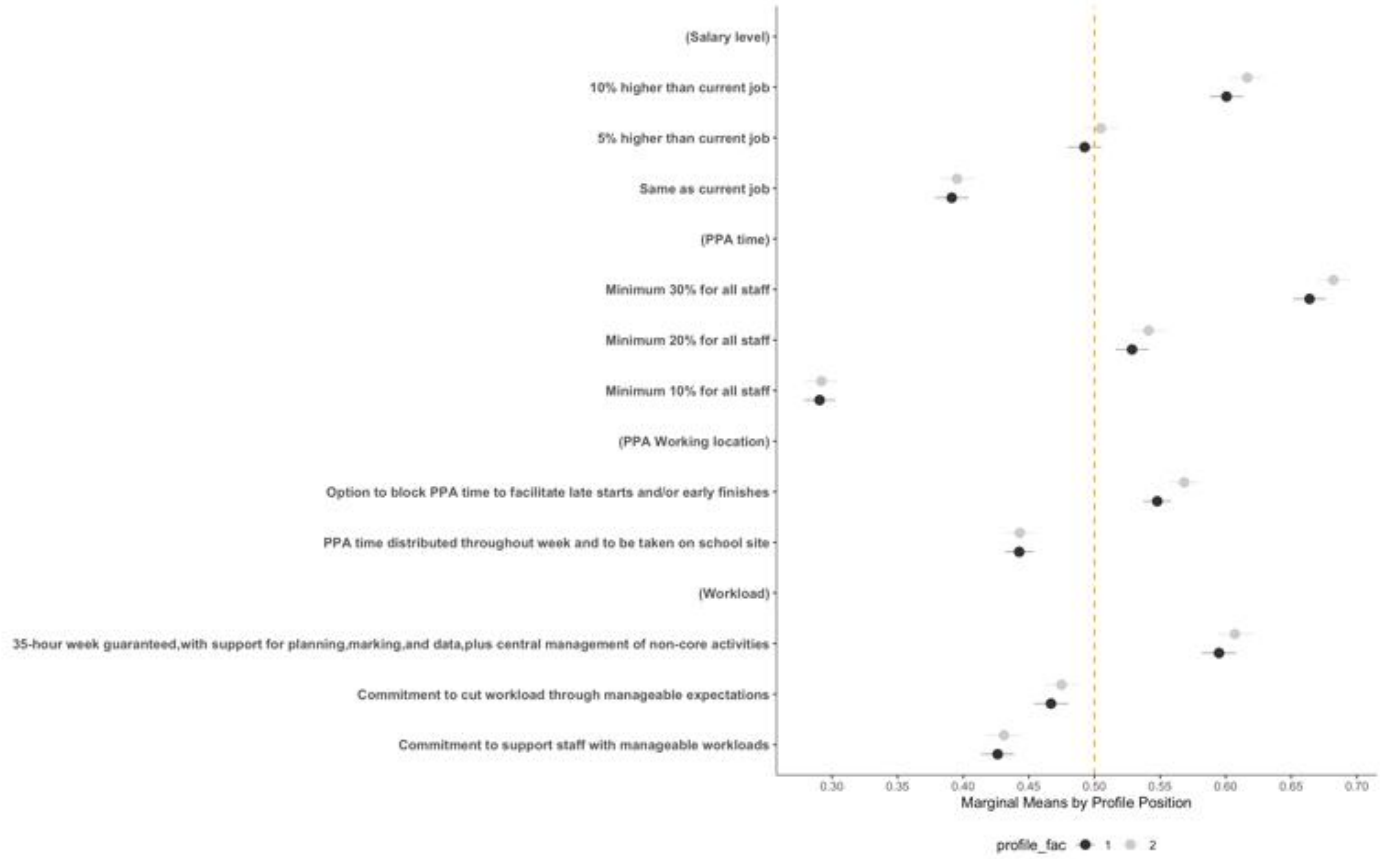

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70
Marginal Means by Carryover effect
deck_order 1 2 3 4

Figure A14: Experiments 1-4, profile order effects






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Profile Position
profile_fac
1
2

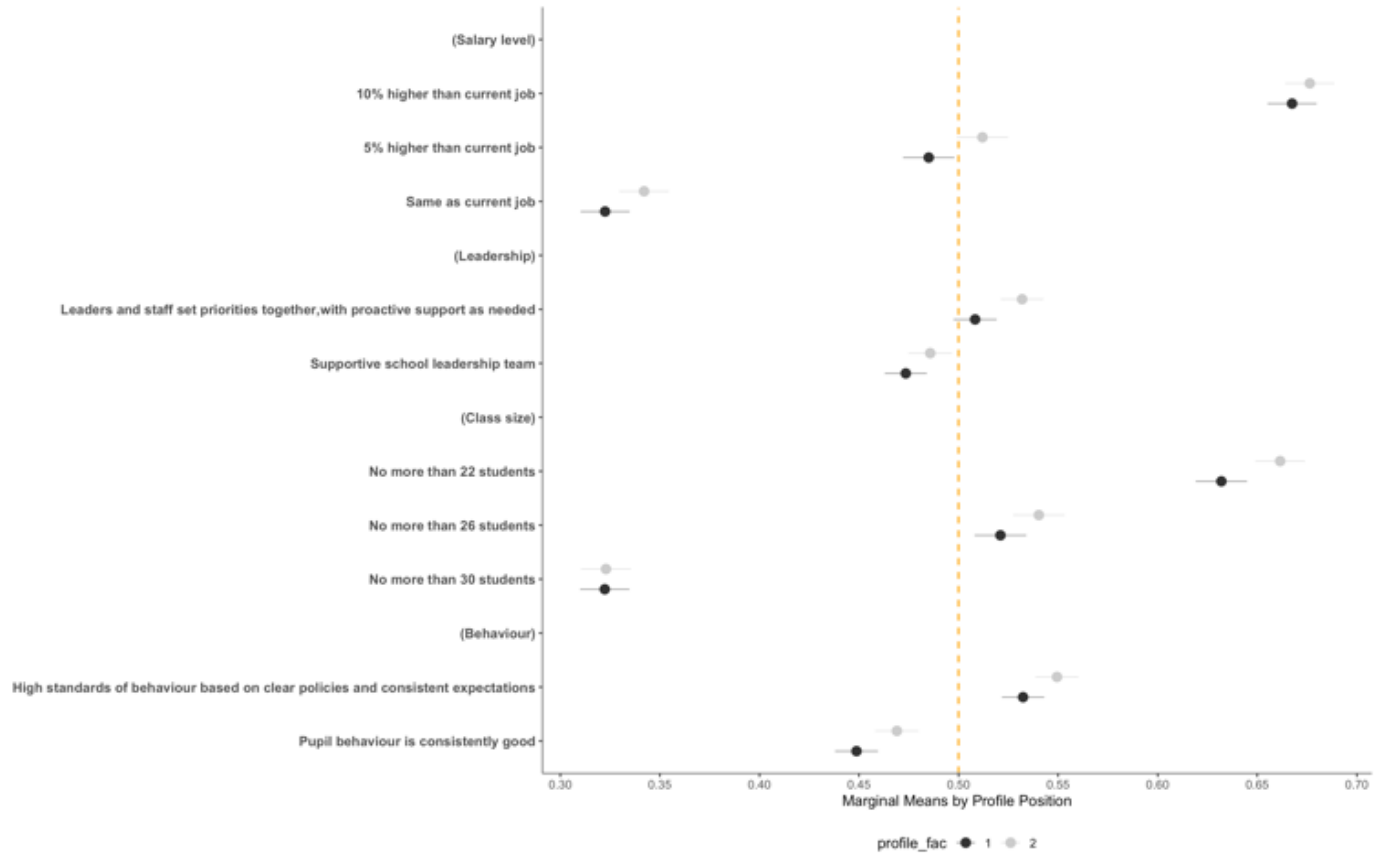
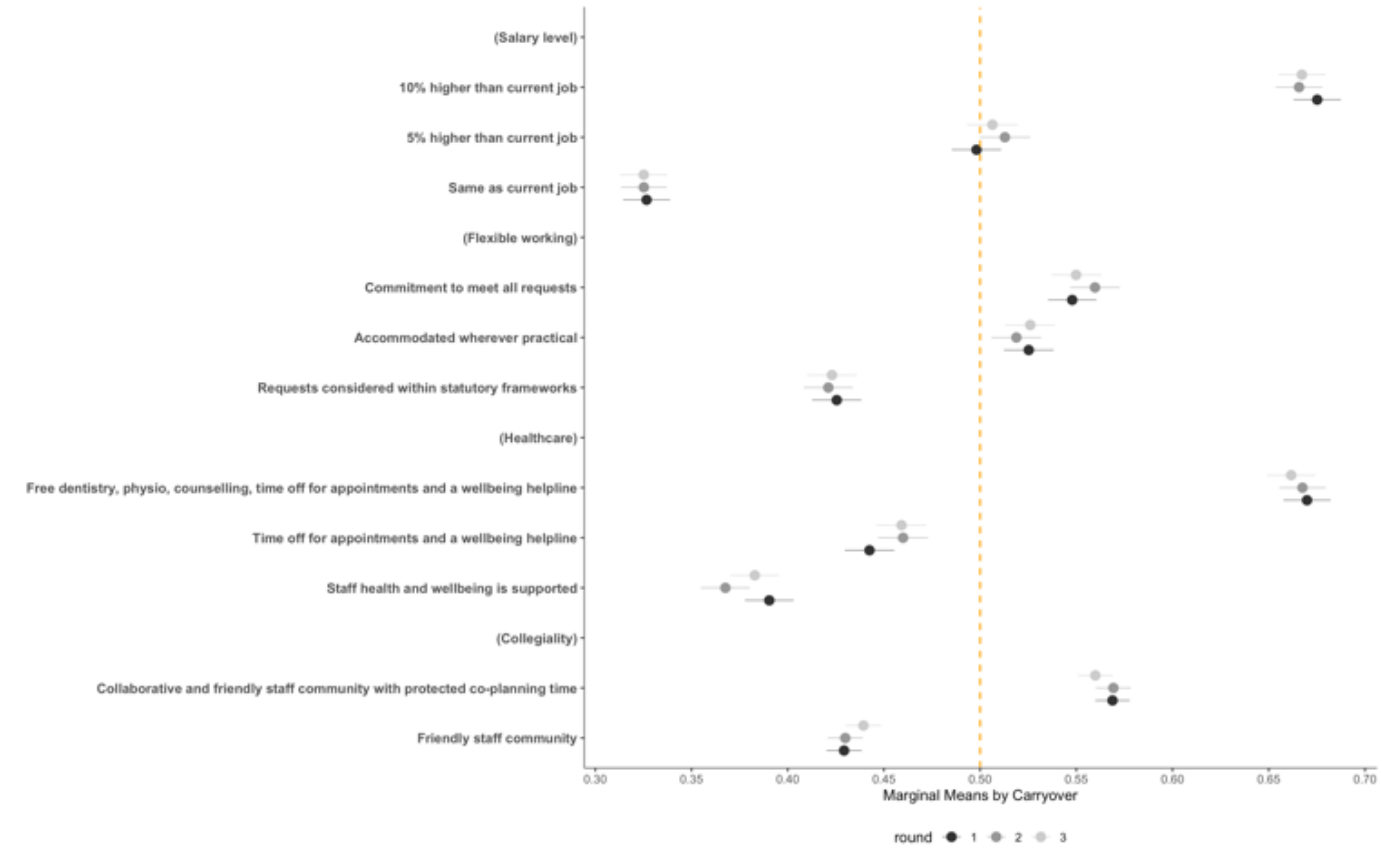
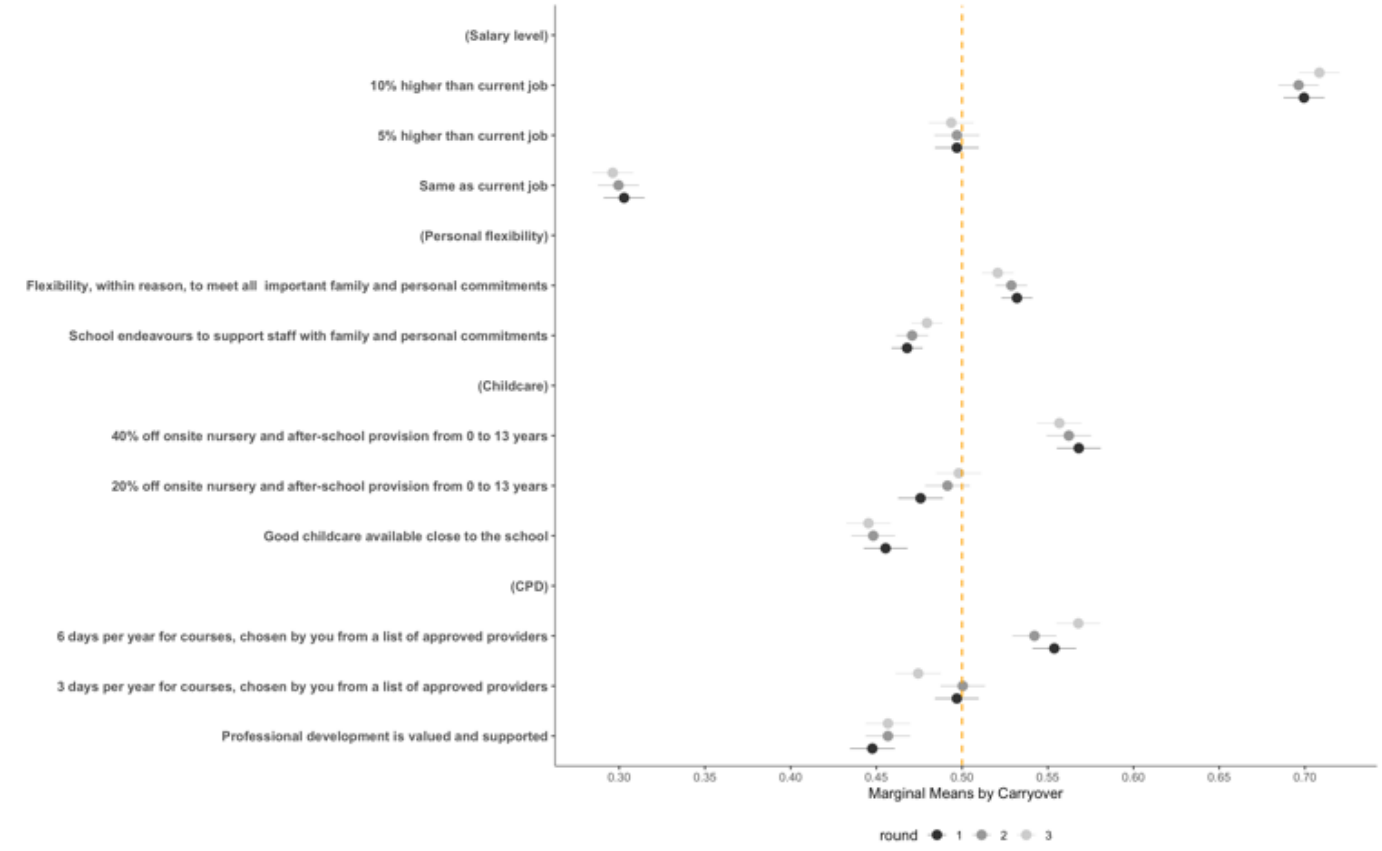
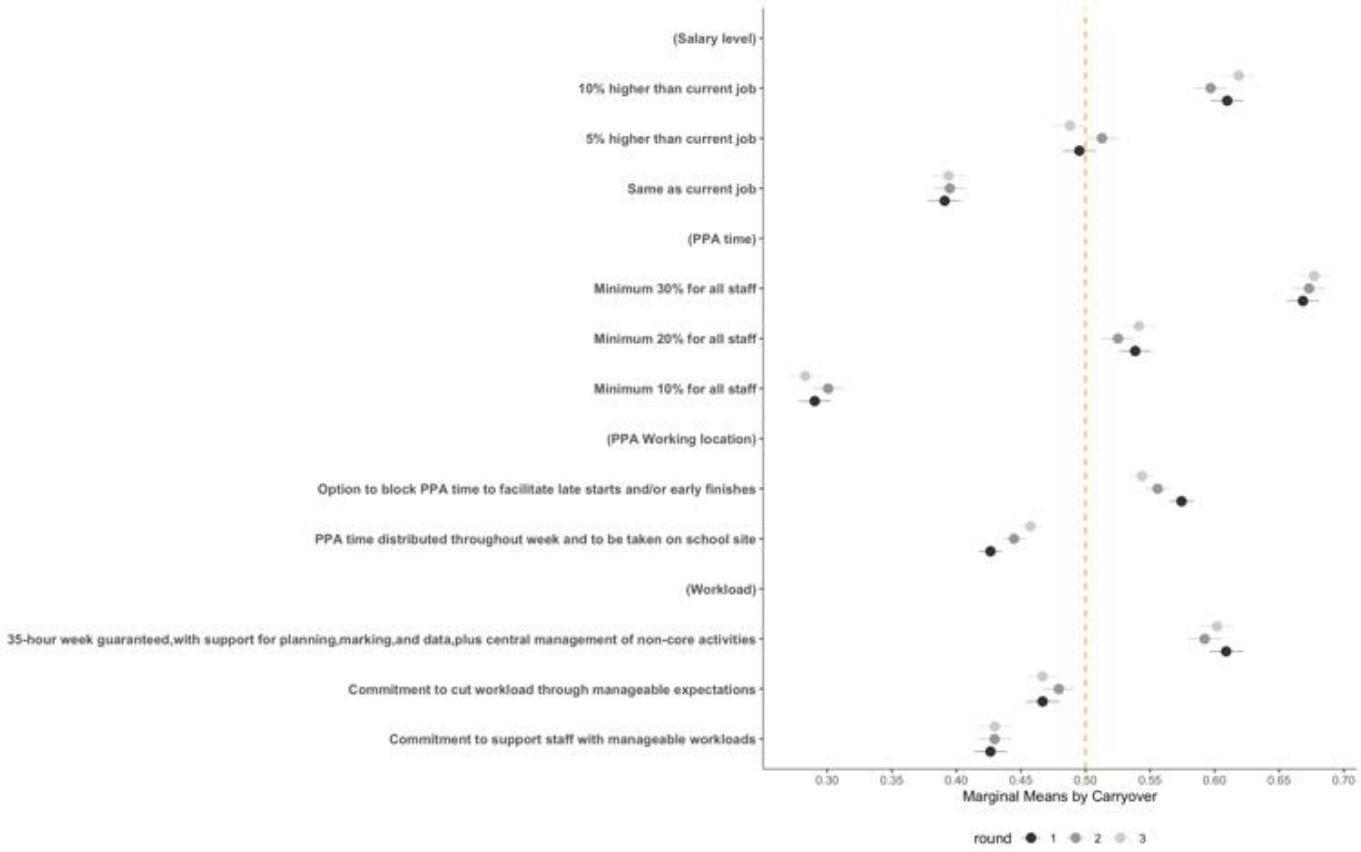
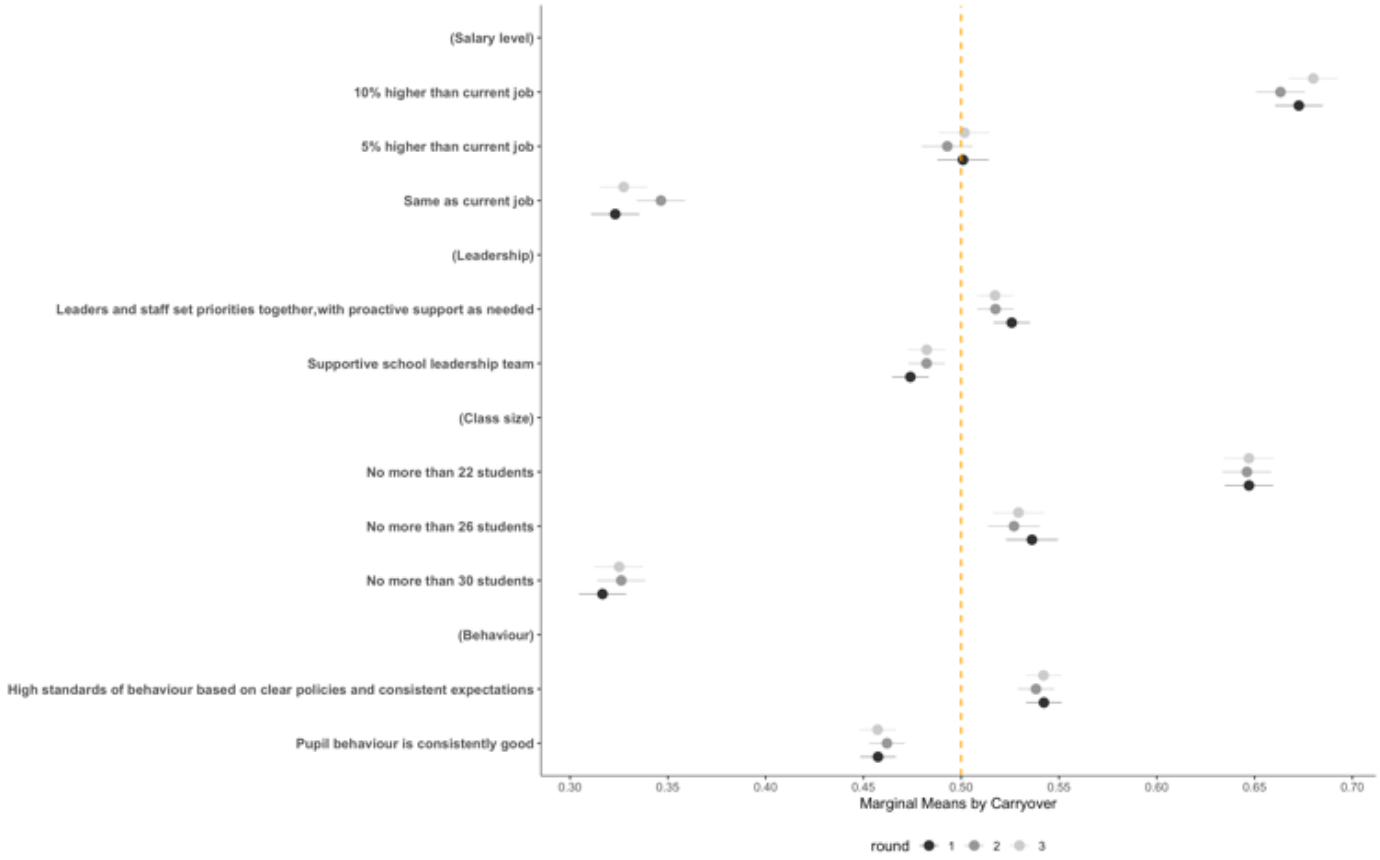

(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Profile Position
profile_fac
1
2

Figure A15: Experiments 1-4, round effects






(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(PPA time)
Minimum 30% for all staff
Minimum 20% for all staff
Minimum 10% for all staff
(PPA Working location)
Option to block PPA time to facilitate late starts and/or early finishes
PPA time distributed throughout week and to be taken on school site
(Workload)
35-hour week guaranteed,with support for planning,marking,and data,plus central management of non-core activities
Commitment to cut workload through manageable expectations
Commitment to support staff with manageable workloads
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Carryover
round
1
2
3


(Salary level)
10% higher than current job
5% higher than current job
Same as current job
(Leadership)
Leaders and staff set priorities together,with proactive support as needed
Supportive school leadership team
(Class size)
No more than 22 students
No more than 26 students
No more than 30 students
(Behaviour)
High standards of behaviour based on clear policies and consistent expectations
Pupil behaviour is consistently good
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
Marginal Means by Carryover
round
1
2
3

Appendix F - RQ1: Open response analysis

There were 2,998 open responses to the question:

> *"Is there anything you'd like to share about the reasons behind your job advert choices? Did you find any part of the decision-making process difficult or unclear?"*

The analysis aimed to explore four key questions:

1. What was the experience of answering the questions like? Did any factors make it easier or more difficult (e.g., ambiguity in wording)?
2. How did participants make their decisions?
3. Did respondents approach their decisions in unexpected ways—particularly in ways that raise questions about the ecological validity of the experiment compared to real-world decision-making?
4. What factors led participants to value a particular attribute or statement?

An AI-assisted coding process was used to categorise responses. Where coding frequency permitted, a sample of 30 responses per category was then manually reviewed to identify themes relevant to the four research questions. This approach was informed by existing literature on human-AI hybrid coding (Perkins & Roe, 2024; Hong et al., 2022; Dai et al., 2023; Yan et al., 2024) and by initial experimentation with the dataset, which helped assess the viability of the proposed methodology.

AI-assisted coding has several well-documented advantages, including its ability to accelerate scientific discovery and enhance the efficiency of research and verification processes (European Commission, 202). Most notably, for a study involving a large volume of responses, this approach allows researchers to focus on interpretation and theory development rather than time-consuming manual coding (Perkins & Roe, 2024).

However, there are important trade-offs. AI-assisted analysis can lack transparency, as identical data and prompts may still produce different results (European Commission, 2024). There is also the risk of replicating biases present in the training data (Perkins & Roe, 2024). Other commonly cited limitations—such as the potential for spurious correlations or an inability to interpret beyond text and images—were less relevant to this study.

The analysis was conducted in three stages, following EU Commission guidance. The specific prompts used at each stage are included in the appendix (European Commission, 2024).

Cleaning

In line with guidance from the existing literature (Bijker et al., 2024; Dai et al., 2023; Holmes & Miao, 2023), the AI was first provided with an initial prompt that included:

- Contextual information about the project
- Simple instructions outlining the task
- Details on the required response format
- Guidance on what to base the response on, including the dataset and the desired coding categories

The full spreadsheet of responses was then uploaded to ChatGPT—henceforth referred to as 'the AI', and it was instructed to clean the dataset.

Previous studies have emphasised the importance of incorporating transparent explanatory mechanisms in AI-assisted research (Yan et al., 2024). To ensure this, the AI was explicitly instructed to report back on its actions throughout the process.

Coding

The primary purpose of coding was text classification, allowing us to identify responses relevant to our research questions for manual review and summarisation. In effect, coding served to stratify a sub-sample for deeper analysis. ChatGPT was considered well suited to this task, as it has been found to be particularly effective in "transforming unstructured datasets into structured formats" (Yan et al., 2024, p.5).

The process combined AI-assisted coding with manual validation, following a 'machine-in-the-loop' rather than 'human-in-the-loop' approach. In other words, the AI supported and amplified the human-driven process, rather than the reverse (Hong et al., 2022). This approach balances "the qualitative richness of human interpretation with the quantitative efficiency and consistency of AI analysis", while also mitigating risks such as AI 'hallucinations' and other limitations of emerging AI technologies (Sabbaghan, 2024, p.132).

AI-assisted analysis was conducted in iterative steps, allowing prompt refinement to increase reliability. This aligns with recommendations that AI-assisted coding should incorporate human-AI feedback loops, iterative functionality, and robust validation mechanisms (Dai et al., 2023; Yan et al., 2024). Given that AI coding quality is highly dependent on the quality of prompts used, these opportunities for refinement were essential (Perkins & Roe, 2024).

The coding process was primarily deductive, using predefined categories based on the attributes within the experiment, namely:

- Salary
- Flexible working
- Healthcare
- Collegiality
- Leadership
- Childcare
- Pupil behaviour
- Professional development (CPD)
- PPA arrangements
- Workload
- Class size

In addition to these categories, an extra coding category was introduced: "comments or criticisms of the questions", to capture feedback related to the first research question. The AI was provided with examples from the dataset for each code to improve accuracy.

While coding was largely deductive, previous research suggests that AI can be effective in developing coding schemes (Bijker et al., 2024) and can identify latent themes, nuances, and subtleties that might otherwise be overlooked (Perkins et al., 2024; Sabbaghan, 2024). To account for this, the AI was also asked to propose additional categories it identified and apply them to the dataset. However, this inductive aspect of the analysis was minimised, as our research questions were primarily structured around a pre-established framework, and most responses were brief and already aligned with the predetermined categories, making extensive interpretation unnecessary.

Following Dai's (2023) four-step process, the inductive coding began by:

1. Providing the AI with an exemplar response.
2. Requesting an initial set of codes.
3. Engaging in a 'code refinement discussion' to clarify and refine categories.
4. Applying and evaluating the final coding scheme (integrated into the review process described below).

Validation and iterative refinement

The initial coding output, covering all three types of codes (attribute-focused, question-critique focused, and additional inductive categories), was then reviewed. To assess reliability, 50 randomly selected responses were manually coded. Reliability was evaluated separately for each coding type and for individual codes, as AI coding reliability can vary between inductive and deductive methods and depend on the complexity of the coding scheme (Bijker et al., 2024).

To measure agreement, a simple percentage agreement metric was used for each of the three coding types. More sophisticated inter-coder reliability measures, such as Cohen's kappa or Krippendorff's alpha, were deemed unnecessary given that the primary aim of coding was categorisation rather than in-depth interpretative analysis. The emphasis was on ensuring rapid and efficient data organisation to facilitate subsequent human-led analysis, rather than prolonging validation.

Based on the first review, additional instructions were given to the AI to improve its interpretation of responses. It is important to note that coding disagreements did not always indicate AI errors. In some cases, after reviewing discrepancies, it was judged that the AI coding better captured the nature of the comment or offered a useful alternative interpretation. This process led to an evolving coding dialogue between human and AI coders.

At each review stage, the AI was asked:

- Whether it understood the feedback.
- Whether any further clarification was required.
- To re-code the dataset incorporating the refinements.

This process was repeated for a second round of coding checks, with further refinements made based on new observations. A third review was conducted, during which errors in two specific codes (which affected a third) were identified. The AI was asked to correct these issues, leading to a third refinement stage (Round 3b).

Final coding review and data familiarisation

Throughout the review process, a total of 200 responses were manually coded (representing 198 unique responses). A key benefit of this iterative review was data familiarisation, which prepared the ground for deeper human interpretation. It also helped identify additional low-frequency codes, ensuring that rare but meaningful themes were included in the final analysis.

Moreover, the manually coded dataset served as a valuable counterpoint to the AI-generated output. While the AI coding was efficient, some inaccuracies persisted, reinforcing the importance of maintaining human oversight.

Table A6 reports percentage agreement by coding type at each stage, while Table A7 provides percentage agreement by individual code. However, it is important to acknowledge that percentage agreement measures have well-documented limitations, particularly for low-frequency codes. Although alternative inter-coder reliability measures (e.g., Geilser & Swarts, 2019) could have mitigated these issues, the priority in this study was rapid categorisation rather than extended validation.

Table A6: Agreement by code-type at each data coding stage

Round 1			Round 2			Round 3			Round 3b		
Attribute categories	Additional categories	Question critique categories	Attribute categories	Additional categories	Question critique categories	Attribute categories	Additional categories	Question critique categories	Attribute categories	Additional categories	Question critique categories
92%	97%	89%	92%	96%	82%	93%	98%	86%	94%	96%	89%

Table A7: Percentage agreement by code (i.e. job attribute or similar)

	Number of comments coded, across all rounds of human coding	N comments coded by AI in the full sample - round 3b	Percentage agreement in coding review - round 3b	Code prevalence in full sample of human coding	Code prevalence in AI coding - full sample round 3c
	(N = 198)	(N = 2,998)	(N = 50)	(N = 198)	(N = 2,998)
Salary	70	836	86%	35%	28%
Health and well-being	18	233	94%	9%	8%
Collegiality	9	94	100%	5%	3%
Job Flexibility	32	175	88%	16%	6%
Childcare	66	873	92%	33%	29%
Professional Development (CPD)	10	192	98%	5%	6%
PPA Arrangements	51	734	100%	26%	24%
Workload	41	195	86%	21%	7%
Leadership	15	122	100%	8%	4%
Pupil Behaviour	12	27	100%	6%	1%
Class Size	33	288	94%	17%	10%
Support for Non-Parental Care Responsibilities	4	79	94%	2%	3%
Commuting and Location Preferences	1	0	100%	1%	0%
Stress Reduction and Work Environment	18	32	90%	9%	1%
Career Progression	2	2	100%	1%	0%
Healthcare Benefits	8	182	96%	4%	6%
Critique of the Question (Overall)	69	212	72%	35%	7%
Negative Perceptions of the Survey Task	12	7	94%	6%	0%
Critique of Specific Descriptors or Phrasings	6	32	98%	3%	1%
Inapplicability to Respondents' Circumstances	43	175	82%	22%	6%
Frustrations About forced Comparisons or similarity	6	0	100%	3%	0%

Thematic analysis

Where there were sufficient responses within a category, 30 responses were randomly selected from the AI-coded dataset, and the main themes were summarised. If a category contained 30 or fewer AI-coded responses, human-coded comments within that category were also reviewed.

Some categories, such as “Commuting and Location Preferences” and “Career Progression”, had largely disappeared from the AI’s coding in the later rounds, with very few instances identified in the human coding. In these cases, responses flagged in earlier versions of coding were reintroduced to ensure they were captured in the final analysis.

Accurately assessing the prevalence and potential impact of respondents’ critiques of the survey questions was particularly important for answering research question (What was the experience of answering the questions like? Did any factors make it easier or more difficult?). To achieve this, the full set of human-coded comments within the “overall critique of question” category (n=69) was manually re-coded, using a combination of inductive and deductive sub-codes.

Finally, the entire human-coded sub-sample (n=198) was reviewed once more. This step was essential to mitigate the risk that the AI had been more effective at identifying some types of responses than others, potentially skewing the findings.

For the most part, the analysis stage did not involve AI, given the importance of human criticality and domain expertise in the meaning-making process (Perkins & Roe, 2024). However, once the human analysis was complete, the AI was asked to generate a narrative summary of each code, illustrated with quotes, to determine whether any additional insights could be identified.

This AI-generated analysis was conducted only after the human analysis to avoid biasing interpretation. Ultimately, the output was disregarded, as it was found to contain numerous AI hallucinations.

Findings

Salary

Salary was the most referenced theme and was mentioned in around a third of comments (35% of the human sub-sample, 28% of the full AI coded dataset). Many comments simply emphasised the importance of salary, by just stating “higher salary” or “money!” while some emphasised the current cost of living or being unable to pay bills:

> *“I’m paid well but with the children of university age and rising costs in every aspect of life I still struggle at the end of every month and this is without having new cars or any foreign holidays and very few meals out. Nonetheless I price flexibility, a positive culture and a genuine desire to accommodate family commitments above salary”*

In other cases, comments on salary focused on explaining that factors (such as lifestyle, work-life balance or working environment and arrangements) were more important than salary.

> *"I think most teachers would most appreciate being given more time and support (e.g. have less pupils to teach, more time for PPA and CPD) rather than more money*
>
> *"A pay rise would be great but other factors are more important. The idea of a class size under 26 is appealing too."*
>
> *A willingness to prioritise other factors over salary was sometimes linked to an individuals’ life circumstances.*
>
> *“School culture is more important to me than a salary hike - I appreciate that this is due to my position/stage of my career."*

A third category of individuals - beyond those who emphasised the importance of salary and those who judged other factors were more important - commented that other factors were tempting but insufficient to outweigh salary. For example, one stated that "*Smaller class is appealing but I manage with 30 already so I'd take the higher salary!*” while

another commented that they "*Liked the flexibility of PPA from home but not when the payoff was 20% less of it*!". This made choosing between adverts difficult for some respondents.

Health, well-being

There was an overlap between the health and well-being category and the workload and stress-reduction/working environment categories. Comments about time off for health appointments also overlapped with job flexibility. Themes in the 'health benefits' category are covered here alongside the health and well-being code.

Many comments emphasised other factors' beneficial impact on well-being or mental health (or lack thereof). For example, many respondents believed that smaller classes would lead to improvements in their 'mental health' or allow them "*to do the job well*" which would in-turn benefit their well-being (we return to this below).

The wording of statements about health and well-being, and whether they involved concrete or intangible commitments appeared to be an important consideration for many respondents. One commenting that:

> *"I seemed to prefer the ones with tangible benefits rather than vague claims that may or may not prove true like 'commitment to health and wellbeing' which many schools will claim similar but what do they actually do?"*

It seems that the novelty of healthcare benefits was a draw for some respondents, although one commented critically, that they had worked for a school offering such benefits and that this had been a way of ticking "*a welfare box*" and acted as an "*excuse*" not to do other things.

Respondents' degree of interest in healthcare benefits sometimes depended on the current state of health, such that some respondents who were in good health felt the offer was unnecessary while others - who had a history of poor health - saw it as particularly relevant:

> *"Don't usually require time off, not particularly worried about health needs so for me the main driver is salary."*

> *"I do have medical conditions that involve dental treatment, physio and appointments that it's hard to juggle with school hours and I currently have to wait longer to avoid time off school. This has affected my selections."*

A few respondents had 'done the maths' and decided that free physiotherapy or dentistry was actually worth more than a proffered pay rise. This may have led to some heterogeneity in responses depending on baseline pay - a possibility which was highlighted during focus groups.

Of the different healthcare benefits offered, dentistry appeared to be the most valued, with respondents emphasising the high cost involved and the difficulties of registering with an NHS dentist and, again, some respondents noting that existing issues made this benefit particularly valuable.

> *"Most CPD courses are rubbish so commitment to CPD doesn't affect me. Not massively interested in flexi working so unless they are offering an extra couple of days off for my own reasons it doesn't change my position. Dentistry is a big offer as it is really expensive."*

> *"Have congenital teeth problems so anything offering free dentistry would be amazing"*

There were also at least two comments expressing ideological opposition to healthcare benefits - which is something that was also flagged strongly in focus groups.

Job flexibility

Many respondents criticised the current status quo and the lack of flexibility available in teaching:

> *"Any one of these would be preferable to what I receive - which is nothing! I would have taken even the most basic improvement to my working conditions! Anything is better than the current situation where I can't have any*

personal time off, any flexible working or any paid health care! Imagine education offering any of those packages! I would bite their hand off. It just shows how poorly we are treated."

Respondents' enthusiasm for job flexibility was often linked to childcare and their personal family situation (for example due to their aging parents - as discussed below). A corollary of this is that many respondents felt that flexibility was not a priority given their circumstances.

Some respondents had negative views on job flexibility and avoided it rather than seeking it out because of its impact on a school and its pupils, or because they felt it was unrealistic:

"I think you should work full time if you're doing a job like this, inconsistency is awful when there are too many pt staff, accountability is impossible. I want pay and time to do an effective job"

"From my experience in a school that offers flexible working, this often comes at the detriment of sustaining an environment for learning."

These negative perspectives were particularly marked for certain school leaders, who felt flexibility for staff would come at a cost to them due to cover implications or difficulties planning staffing.

"My children are grown up so I'm not interested in the childcare. I would be interested to know how flexible working is covered - I would be concerned that this would involve the school providing lots of cover. The free healthcare is very attractive and anything involving a pay increase"

Childcare and other caring responsibilities

Childcare was one of the most referenced categories (33% of human-coded responses, 29% of AI coded responses), but this was partly because so many respondents chose to comment on it not being relevant to their circumstances.

Comments were divided between teachers who had children—who emphasised the value of childcare, and those without—who often said it was unnecessary. Alongside these comments were a small number of respondents for whom childcare was not relevant, but who could see its value to others. Some also anticipated childcare's importance to them in the future once they started a family or recalled that they would have valued it when their children were younger. Some also saw a schools' commitment to childcare as indicative of its wider culture and ethos.

"I don't need the cheaper childcare now but many of my colleagues do and I have seen too many talented female teachers dropping hours and dropping back from leadership due to childcare costs"

"I no longer need childcare. My choices would be different if a young family were a consideration for me."

At least one respondent echoed comments made during focus groups that the disadvantage of an in-school childcare offer was that it might require them to move their children from a current placement they were happy with.

Many teachers without children asked what equivalent benefit would be available to teachers without children, and some expressed resentment—at times vociferously—to the idea that colleagues with children were being given what appeared to be additional perks, with "*no financial consideration for those without*":

"I don't need childcare so the different options don't add value for me. People need to remember not everyone that works in teaching has children. I would need more time off for disabled parents"

"Childcare isn't an issue for me and I actually find it incredibly annoying the amount of flexibility given to parents compared to those who don't have children. It should be allotted amount of time"

Generally, it was respondents with young children of nursery age who saw the greatest value in childcare. Some with older children felt it would no longer be useful, while others believed the benefit would persist:

"20% off childcare would have been huge up to age 13. We still pay really high rates for after school club as our boys' school doesn't run one so it's arranged by a private company"

Parents of older children, including those at university, sometimes referred to the value of being able to attend appointments or important events with their children. References to other, non-parental family responsibilities tended to refer to elderly or aging parents.

Professional development (CPD)

Many respondents were put off by the idea of an approved list of providers, emphasising their desire to choose their own CPD.

"I appreciate CPD but like to choose my own related to my subject as it is useful- the school chosen ones are variable."

Some commented that they were drawn to specifically worded adverts, arguing that comments about CPD being "*valued and supported*" were too "vague"

"Specific information on number of CPD days/courses is better than "we support development"

Respondents often alluded to their dissatisfaction with the current status quo in professional development - either because of its insufficiency, or because 'courses' were not considered to be a valuable approach, or because too much CPD lacked practical focus:

"I don't care about time off for CPD either. Too much educational guff is forced on staff without extra training. Teaching a practical subject, learning new practical skills would be more beneficial. That type of training has not existed since academy trusts took over education. Time off for medical appointments would be useful."

"My CPD is important to me and something I miss since I started this school in May. I am leaving now at the end of Dec. I have been using... my old contacts/ training from the previous LA to keep on top of my CPD (I work in SEND); in spite of the current workload esp. SOW/ Lesson planning time which is drowning out my entire existence. My wellbeing is / has suffered."

A few respondents believed that CPD was less relevant to them because they were at a late stage in their career, or on the cusp of retirement.

Stress reduction and work environment

This category was one of the inductively generated categories created by the AI. Within it, there were occasional comments about the importance of school culture in general, but for the most part, comments in this category related to the potential impact - generally in terms of reduced stress - of some of the measures described in other categories such as workload, pupil behaviour, class size, supportive leadership, or collaborative colleagues. We return to these in turn below.

Numerous comments in this category described current sources of stress. For example, one respondent described argued that stress was a consequence of insufficient specialist support:

"The job wasn't this difficult or stressful 20 years ago, but there was specialist provision available for children with medium/severe behaviour/academic needs. There's no point in me having more money if i'm so stressed out that the first few days/weeks of any holiday is me winding down or being ill. I see staff going off really Ill from stress and pregnant staff really worried about getting hurt, kicked punched."

Another respondent explained that factors like whether a school was an academy; whether it grouped pupils by ability and whether it was single gender or co-educational would be key. There were several other comments that reflected negative perceptions or experiences of academies.

Workload

There was a considerable discrepancy between the proportion of comments coded as 'workload' by the AI and human. Coding checks suggested that there was a lack of consistency of interpretation so human coded excerpts were carefully reviewed alongside the AI sample.

Respondents often used terms like "reasonable" or "manageable" when referring to workload and many made comments about workload outweighing salary in importance.

> *"Whilst an increase in pay would be good this is always weighed against the experience of the work. Workload and wellbeing outweigh increases in pay"*

Many respondents commented that they did not consider a 35-hour week to be credible and one even felt it was undesirable - stating that "I don't think limiting a working week to a set number of hours is reasonable, I would feel more stressed trying to get through stuff in a set time". On the other hand, another described such an offer as "highly attractive" and another noted that:

> *"a guaranteed 35 hour week sounds like fantasy but if it could ever happen I'd love that."*

Many respondents expressed scepticism about intangibly phrased statements on workload. They therefore preferred a "guarantee" of a set number of hours, to a "promise... which we all know would never happen!!"

As we will see, there were also close links between comments on workload, and those referring to PPA arrangements and class sizes, since these were seen as providing tangible steps towards a more manageable workload. Indeed, nearly half (47%) of comments coded 'workload' by the AI were also coded PPA, and a quarter (26%) coded as 'class size' (the same trend held in the human subsample.)

PPA arrangements

After childcare, salary and criticisms of the question/task, PPA arrangements were the most frequently referenced category within respondents' comments (26% of human subsample and 24% of AI coding).

Many responses in this category were remarkably similar, emphasising that factors like PPA (and class sizes - see below), would make a significantly greater difference to teachers' lives, compared to salary.

> *"I would choose smaller class sizes and more PPA over salary increases any day!"*

> *"I'd take Smaller class sizes and more PPA over a pay rise any day"*

> *"PPA time, working hours and class sizes win out over salary!"*

Where respondents explained the value of PPA, they tended to emphasise that - like with reduced class sizes, increased PPA time was a way of reducing the "intense" nature of the job. Indeed, nearly a quarter (22%) of comments that were AI-coded as "PPA," were also coded "class size".

Some respondents expressed preferences regarding the nature of PPA arrangements but there was a lack of consensus regarding preferences. Some teachers said they preferred blocked PPA time, while others preferred it distributed.

> *"I am SLT so PPA in blocks taken off site is not an option for me."*

> *"Blocking the ppa would make work life balance better- it being distributed* through the week doesn't work for me"

In one case, a respondent pointed out that blocked PPA time could indirectly have an impact on salary by allowing them to move into a full time, rather than a part time role:

> "*Blocking PPA was difficult to answer. For me that would mean more pay as I'm currently part time in order to do the school run but if I used my PPA to do it I wouldn't have any time in school for marking and doing things that*

come up during the day. As it is I arrive at school 5 mins before I start teaching and it's incredibly stressful. I can also only stay 30 mins at the end of each day and it's not enough."

However, others felt that blocking PPA meant having several days with no break in teaching and that this would be stressful. They therefore preferred the idea of having it spread out.

"10% PPA and mostly on a Wednesday is killing me this year. I need at least a free per day to feel on top of things and I don't have that on 7 of my 10-day timetable days."

Meanwhile some teachers wanted to be able to work from home during PPA time - seeing this as a route to flexibility, while others liked the idea of getting their work done in school - particularly given the availability of facilities like printers.

"The option to take PPA off site will always be a key factor for me, as is class size."

"I don't mind having my PPA time on site as it helps me keep work at work."

Schools are therefore unlikely to find a PPA arrangement that suits everyone, and one noted that in Scotland they were already free to take PPA time at home, such that this was a "non-question" for them.

Class size

Whereas in other categories the reasons why a job attribute was considered valuable were homogenous, when it came to class sizes, respondents highlighted several different benefits.

As noted above in relation to PPA, many saw class sizes through the lens of workload and job intensity. Indeed, more than half (58%) of comments that were AI-coded as "class size" were also coded "PPA".

"Small class sizes are very appealing and have a real impact on workload"

For other teachers, the benefits of class sizes were linked to efficacy: they felt they could do their job better when class sizes were smaller and that this would be more professionally satisfying:

"Time to do my job well would be lovely for wellbeing, the extra PPA and lower class sizes help with that"

"Smaller class sizes to ensure more effective teaching can take place."

Closely linked, was a sense that smaller class sizes were better for pupils.

"Lower class sizes will always be a draw to the moral teacher in me who wants students to have the best possible education"

While some respondents focused on a particular dimension of class-size's impact, others combined several rationales.

"Class size has a huge knock-on impact to workload I'm primary. On average we would teach 5 lessons a day, that's 150 books to mark for a class of 30 and only 110 for a class of 22. That's more than a whole lesson difference. The job is also more rewarding with a smaller class, as the teacher you aren't spread so thin. You can get to know the children and work with all of them across the day/week. I think it improves pupil results as well."

There were also detractors, with one teacher commenting that they were already teaching larger class sizes and that salary increases could compensate for this:

"I've been teaching classes of 60 this last term, money talks more than class sizes."

Several respondents noted that the importance of class size depended on the nature of the class. For example, one pointed out that teaching a large top set was very different to teaching a large class in which many pupils had complex needs while another said that, as an English teacher, they had a particularly high marking load but that this would be reduced if they had smaller classes.

Pupil behaviour

Pupil behaviour was a low frequency code, so all AI and human coded comments were reviewed together (n=37).

According to one respondent, behaviour creates workload because of the need to phone home and organise detentions. Another explained that with smaller classes it "*would be easier to support students*" but that "*consistently good behaviour is better and more realistic*". Another explained that improved behaviour improved learning and that this supported staff wellbeing. For another teacher, supportive management and PPA would be of primary importance, but behaviour (and smaller classes) would be the "cherry on top."

The importance accorded to behaviour may vary depending on teachers' seniority and experience, with several respondents reporting that behaviour was less of a consideration for them because they were already confident of the behaviour management skills.

Construction of statements on behaviour was a challenge when designing the survey task as it is difficult to communicate a 'soft' attribute like this with credibility. Unsurprisingly there were therefore numerous cynical responses, with teachers questioning the truth of claims made in adverts:

> *"Hi [sic] expectations of pupil behaviour sounds like BS. But who says behaviour is consistently good?"*

> *"pupil behaviour is always good" cannot be accurate and therefore appears ignorant or deliberately misleading. If behaviour is always good there must be a reason behind that."*

Some respondents therefore found statements that gave specific explanations of what steps were in place to manage behaviour more "trustworthy". For example, one noted that it was:

> *"impossible to just trust that pupil behaviour is 'consistently good' - what this means to one person might be different to another, and so I'd prefer supportive leadership and clear and simple behaviour policies."*

There was a link between the credibility of statements on behaviour, and the schools' leadership, as one respondent explained:

> *"Student behaviour and a supportive SLT went hand-in-hand for me. Behaviour is as good as the structures and consistency in place which are set and reinforced by your SLT. ...Smaller class size to support students more effectively is a wish - though a staff with a reduced workload and students with high levels of good behaviour (which the above suggests) would mean we have the capacity to support classes between 22-30. Teachers need more pay but I want the above more."*

Leadership

Comments about leadership primarily emphasised that being "supportive" was important. Unfortunately, comments rarely gave an indication of what supportive leadership would look like in practice. One respondent emphasised the importance of consulting staff and another the leadership being "cohesive". Another drew a distinction between schools in which there are policies, and those in which leaders provide "on-the-ground support" along with flexibility.

Several comments about leadership emphasised the perceived value of flatter, less hierarchical structures, in which decisions are taken together

> *"Salary important but like setting targets with leaders rather than top-down and small class sizes always good"*

On the other hand, at least one respondent felt a flat structure was somewhat idealistic:

> *"Although I like the idea of SLT and staff making priorities together, I think schools need a core leadership team who oversee it all rather than distribute priorities to staff to make. Too many decision makers lead to ineffective policies and change"*

As noted above, there was an overlap between comments on leadership and pupil behaviour, but this is unsurprising given that adverts included statements which explicitly made reference to leadership providing behaviour management support.

Several respondents argued that aspects like leadership (and behaviour) could only truly be judged by visiting a school.

> *"Looking for support from leadership. Smaller classes preferred mainly due to the level of the extra work for the classroom teacher. Makes the expectations in the classroom more achievable. Some of the wording about leadership and behaviour sounds like a line. If these were real jobs, seeing and visiting the school is believing."*

The similarity of statements about leadership that appeared in different adverts gave some respondents difficulty:

> *"I was not sure what the difference would be between a supportive leadership team and one that sets priorities with staff"*

Collegiality

There was some overlap between comments on 'collegiality' and on leadership:

> *"I would prefer to set targets and manage feedback as a staff team with options from SLT rather than targets and feedback being solely given by SLT."*

> *"I found it hard to choose between the leadership options. Supportive leadership vs collaboration with leadership. Seems much the same to me personally. Furthermore, I'm more interested in good behaviour, nice colleagues, enough time to do my work etc."*

Comments on collegiality often focused on co-planning which is–again–unsurprising given the phrasing of the statements. There were also references to the value of a supportive or collaborative working environment.

At least one respondent stated that co-planning time would help to reduce workload because "time with colleagues to plan massively lowers stress and workload". On the other hand, another argued that the relevance of co-planning varied and should only ever be optional. One said they disliked co-planning and would prefer to use their PPA time for their own priorities or to go home.

In one case, the importance of collegiality was used as an argument against off-site PPA:

> *"I'm not sure about the suggestion of PPA being taken off site. I think as part of a team, you have to be on site. In case a behaviour issue arises, there is a fire alarm, or emergency. I just don't think teaching is a job that can be done effectively from home. And any person taking their PPA off site, often means someone on site is having to take up the slack, in terms of pupil/parent queries, or other issues. I believe our perk as teachers is the amount of holidays we get, which is loads. This is our time away from school. But in school hours, we should be on site. We are part of a team/community and should be there."*

Another emphasised that if co-planning was going to happen, it would need to be protected since it would be crowded out otherwise.

As an aside - while it was not our original intention, it was striking that the AI used the code collegiality for altruistic comments in which teachers without children welcomed the prospect of childcare on behalf of their peers, perhaps because of the use of the word 'colleagues', and many of these respondents did indeed seem to be manifesting collegiality through their degree concern for other members of staff:

> *"I don't have children, but when that was a factor in decision-making I chose the most advantageous childcare options, knowing how much this matters to colleagues and the difference it would make."*

Commuting and location preferences

Although this was one of the inductive categories proposed by the AI, very few comments actually fell under it (none were identified in the later rounds of AI coding and only one in the human coding.) The comments that were coded in this category in the first round tended to refer to not being able to afford to live near the school or the impact of commuting on other factors (for example preferring to have PPA time at school due to a long commute or the suitability of childcare arrangements given the school location). One respondent did state that "location and responsibility" were their top priority.

Career progression

There were very few comments in this category. One teacher explained that they had recently moved job for a promotion, another highlighted the importance of CPD for career progression

Comments on the task and prompts

To more accurately assess the prevalence of different concerns about the task or statements, the 69 comments that had been human-coded in this overall category (representing approximately one-third of the sub-sample) were carefully re-coded manually. Table A8 summarises the number of statements assigned to these newly refined categories. The commentary below draws on this sub-sample, as well as AI-coded comments.

Table A8: Manual recode of overall 'critique' comments

Statement irrelevant (childcare)	Statements similar	Statements not credible	Clarity of task	Statements wordy or unclear / repetitive	Technical problem	Advert not how you choose a job
61% (42)	16% (11)	7% (5)	7% (5)	6% (4)	1% (1)	1% (1)

As noted earlier, the most common comments about the task or prompts related to the irrelevance of childcare arrangements, which have already been discussed. This section therefore focuses on other feedback.

Beyond relevance, the most frequent critique was that the statements were too similar to one another. Several respondents noted that there was often only one clear difference between the options, leading them to simply select the obvious distinction rather than engaging in deeper comparison. One respondent specifically stated that they would like to repeat the task with more thought-provoking and contrasting statements.

As noted earlier, a small proportion of comments questioned the realism or credibility of certain statements. These concerns primarily focused on the 35-hour working week pledge, though some respondents also doubted the feasibility of certain PPA or flexible working arrangements. Given that perceived credibility influenced decision-making, this may have slightly affected the quantitative findings, with respondents choosing the advert they most trusted to be realistic, rather than the one they would actually prefer.

> *"Some of the 'promises' were simply unrealistic in the current climate. I tended to go for the option I thought was more credible (or, more correctly, the less incredible)."*

> *"I didn't choose any advert that said guaranteed 35-hour week as I don't believe that is possible."*

Another key theme related to ecological validity. Several respondents noted that real job decisions would be shaped by school visits, interviews, and additional research, rather than just the written job advert. Although only one manually coded comment explicitly mentioned this, it was also highlighted in other responses reviewed during the analysis.

Overall, comments suggested that most respondents engaged thoughtfully with the task, which is unsurprising, as those providing open responses are likely to be more engaged users. However, a small proportion of respondents found the task repetitive or difficult, with one participant eventually selecting at random.

> *"Too many—after the first option, I just randomly chose. Too wordy and too many."*

By contrast, at least one respondent explicitly praised the clarity of the statements:

> *"Nothing unclear. A job offering free dentistry, physio, and counselling would be AMAZING."*

A very small minority also reported technical difficulties, such as job adverts failing to display properly.

Despite some criticisms, several respondents commented positively on the task, particularly noting its engaging and thought-provoking structure.

> *"I liked answering the questions, but I couldn't quite understand what they were sometimes driving at—the trade-offs seemed too easy, at least for me. I can tell you what matters to me: childcare and salary. Flexible working has the potential to trump salary, but the way they were worded in the adverts was too ambiguous for me to trust them as genuine commitments.... An interesting exercise. I would like to do it again with some harder choices."*

Others appreciated the decision-making format and found the task insightful:

> *"I liked the format. Helpful when making decisions."*

One respondent also reflected on how their preferences would have changed earlier in their career and expressed enthusiasm for the study:

> *"My choices would've been very different earlier in my career.... Excellent questions, by the way. I'll be very interested in reading your results this week. Keep up the good work! (Also, if you are recruiting, I love Teacher Tapp and would love to join your team!)."*

Conclusion

Most comments restated preferences rather than providing extensive explanations of respondents' decision-making processes. However, several key insights emerged from the analysis.

Salary often acts as the decisive factor in job selection. For some, it is a highly desirable incentive, while for others, cost of living pressures create an urgent and pressing need for a higher salary. However, this is not universal; for another group of teachers, well-being and lifestyle considerations outweigh salary, suggesting that financial incentives alone may not be the most effective means of attracting all candidates. Some teachers will seek any incremental salary increase, while for others, salary is a hygiene factor—once their financial needs are met, lifestyle considerations become more important.

There is also a clear divide between teachers with children's priorities, and those of teachers without. This raises the possibility that policies favouring parents—such as flexible working arrangements—could alienate some teachers without children if not carefully framed.

Autonomy in professional development is another important factor. Teachers want greater control over their professional growth, rather than rigid, top-down approaches. Similarly, reducing workload and the perceived intensity of the job is a priority, but teachers seek concrete explanations of how this will be achieved—whether through PPA arrangements, class sizes, or other structural changes. However, preferences around PPA arrangements vary significantly, indicating that a one-size-fits-all approach may not be effective.

The appeal of reduced class sizes is driven by multiple factors, including the perceived impact on workload, a sense of enhanced professional efficacy, and a belief that smaller class sizes benefit pupils. Pupil behaviour, workload, and leadership support are also closely interlinked, reinforcing the importance of school culture and leadership style in shaping teachers' experiences.

Overall, respondents were largely positive about the survey task and statements, engaging thoughtfully with the choices presented. However, considerations around ecological validity remain important. Teachers may have based responses on the credibility of statements rather than their ideal preferences, and real-world decisions would ultimately depend on

visiting a school and experiencing its environment firsthand. These factors should be taken into account when interpreting the findings and considering their implications for teacher recruitment and retention strategies.

Limitations, and reflections on the AI aided process

Finally, it is worth reflecting on the AI-assisted coding process. The primary goal of this approach was to save time in response categorisation, allowing greater focus on interpretation and further analysis. However, in practice, the process proved to be highly time-consuming and the human-coded sub-sample was perhaps more useful, given the limitations of the AI coding.

Considerable time was spent attempting to improve coding reliability through multiple rounds of validation, but actual improvements were minimal to non-existent. The process was further complicated by technical issues, including software bugs—on two occasions, the system was unable to generate a downloadable spreadsheet for several hours. Inconsistencies in coding were also problematic, with some categories disappearing entirely in later rounds. The inductive codes generated by the AI provided limited additional insight, while 'hallucinations' in AI-generated outputs required such extensive scrutiny and validation that the expected efficiency gains were further diminished.

For a task like this, it appears that manual coding of a larger sub-sample would have been equally, if not more, effective – and possibly less time-consuming.

While more robust validation methods, such as sophisticated inter-coder reliability measures and additional AI training, might have improved coding quality, they would have required a upfront time investment, further increasing the opportunity cost of using AI-assisted strategies.

Until AI technologies advance further, researchers considering AI-assisted approaches for qualitative research should carefully weigh their relative value against fully human-driven analysis of sub-samples. Key factors to consider include:

- Whether the entire dataset requires coding, or if a representative sub-sample is sufficient.
- The researchers' expertise and familiarity with AI tools, which could streamline the process.
- Whether the task is a one-off analysis or part of a longitudinal study, where coding must be conducted repeatedly—potentially making the upfront investment in AI training more worthwhile.

Appendix G - RQ3: Pilot testing of stepwise questions

During the pilot testing of the survey questions for RQ3, we evaluated two contingent valuation techniques for eliciting teachers' willingness to apply (WTA) for a local job. WTA is measured as a nominal salary increase or decrease from their existing salary and we call this the salary increment. The two survey approaches we evaluated were:

Stepwise (Iterative Bidding) Approach: This method involves presenting teachers with a starting offer (change in salary) and adjusting it iteratively based on their responses until their maximum WTA is identified.

Ladder (Payment Card) Approach: In this technique, teachers select their WTA from a predefined range of salary changes listed on a card.

There are advantages and disadvantages of both approaches. The ladder is simple to implement and quick to answer, has risk of anchoring bias whereby respondents are influenced by the range of values presented. The stepwise approach is more complex to implement and takes longer to answer. However, some studies have shown that participants find it easier to respond to each individual question in turn. So, whilst there already exists a literature on the relative merits of each approach (e.g. Breidert et al. (2006), Schmidt and Bijmolt (2019), Fonta et al. (2010), Uehleke (2016)), we felt our survey question was novel enough to justify a small experiment to compare the merits of the two approaches.

Sample selection and survey questions

We drew a 10% sample of Teacher Tapp panellists, who would participate in the pilot survey questions on 11th December 2024 (and thus be excluded from the main data collection). After answering the questions where they named a local school and were asked how well they knew it, as in the main experiment. The pilot teachers were asked:

> *Suppose the school you named advertised a job with the same role and the **same** salary that you currently receive. Would you consider applying for it?*

After this, the teachers were randomised into two equal-sized groups. The stepwise group received exactly the same stepwise salary questions as set out in the main description of the method for RQ3. The ladder group received one of two questions, depending on their response to the previous question:

> *What is the minimum salary increase that would make you consider applying for the same role you currently hold at the local school you mentioned earlier?* [Response categories from '*£2,500 salary increase over current salary*' up to '*£50,000 or greater salary increase'*, plus '*There is no salary increase at which I would consider applying for a role at that school*' and '*Not relevant / cannot answer*']
>
> *What is the maximum salary decrease that would make you consider applying for the same role you currently hold at the local school you mentioned earlier?* [Response categories from '*£2,500 salary reduction over current salary*' up to '*£50,000 or greater salary reduction'*, plus '*I would be prepared to apply to work at the school for no salary*' and '*Not relevant / cannot answer*']

The salary ladder values were designed to match those in the stepwise questions. We considered the ladder to be sufficient to cover most plausible WTA values. For the minority of teachers outside these values (i.e. responding £50,000 or greater salary increase/reduction), we gave them an additional salary ladder with values from £50,000 up to £100,000 or above in £10,000 increments.

Differences in completion rates and response times

We judge both completion rates and response times by observing each teacher's response to the first salary question and to the question about their knowledge of the characteristics of the named school, which followed the salary increment questions in both pathways. Table A9 below shows that, not surprisingly given the greater number of questions, more teachers dropped out of the stepwise pathway. However, with dropout rates very low at below 2% we do not consider this

to be a substantive issue. The stepwise approach also led to a slightly longer response time, but only by around 6 seconds (24% increase).

Table A9: Differences in completion rates and response times for RQ3 pilot

	Number starting questions	Number completing questions	Dropout rate	Mean response time (secs)	Min response time (secs)	Max response time (secs)
Ladder	252	255	1.18%	26.8	6	115
Stepwise	248	253	1.98%	33.2	11	119

NB. Average and max response times exclude 7 respondents with excessively long times (suggesting they did something else and returned to the question).

Pattern of survey responses

In both question presentations, we establish a salary range at which they would be willing to apply to the named local school. We present these ranges alongside the number of teachers selecting that option for the ladder versus Stepwise approach in Table A10. It shows a few key disadvantages of the ladder approach to responding the questions:

1. 20 teachers (10% of sample) selected "Cannot answer". Given that these teachers had both named a school and said they knew something about it, it seems plausible that they would in theory have a valid response to this question. It is likely that the complexity of the ladder question overwhelmed them, producing this response.
2. 50 teachers (24% of sample) said there was no salary at which they would consider applying for a job at the local school they had named. This is a natural instinctive response, but it doesn't seem that plausible in practice, relatively to simply demanding an extremely high salary increase to compensate for their perceived downsides of the school.
3. 4 teachers (2%) said "I would be prepared to apply to work at the school for no salary". Again, this does seem surprising because it suggests they value working at the school equally to having no job at all!
4. Among the more plausible responses, there is evidence of round number bias, with a greater concentration of responses at values like £0, ±£5,000, and ±£10,000, compared to intermediate values such as £2,500 or £7,500.

We also reviewed the open-text responses by hand to see whether any teachers mentioned difficulties in responding to the questions. However, nearly all the responses were explaining why they made the choice they did. Just one teacher on the ladder pathway said she was confused by the question.

Despite the marginal lower completion rate and slightly longer completion time, it was clear to us that the response profile of those answering the Stepwise questions was far more likely to reflect their true and considered responses to what salary increment it would take to move schools.

Table A10: Response to salary increment questions, Ladder versus Stepwise questions

Salary increment over existing salary	Ladder	Stepwise
Would work for free	4	0
£10,000 - £12,499 decrease	0	1
£5,000 - £7,499 decrease	0	4
£2,500 - £4,999 decrease	3	2
£0 - £2,499 decrease	22	27
£1 - £2,500 increase	6	22
£2,501 - £5,000 increase	31	28
£5,001 - £7,500 increase	9	33
£7,501 - £10,000 increase	28	12
£10,001 - £12,500 increase	6	9
£12,501 - £15,000 increase	11	14
£15,001 - £17,500 increase	1	6
£17,501 - £20,000 increase	13	7
£20,001 - £25,000 increase	1	3
£25,001 - £30,000 increase	1	3
£35,001 - £40,000 increase	0	4
£40,001 - £45,000 increase	0	1
£45,001 - £50,000 increase	0	4
£50,000 increase	2	0
No salary high enough	50	18
Cannot answer	20	3

Appendix H - RQ3: Supplementary tables

Table A11: Rank of school named by teacher across regions in England

		East of England	London	Midlands	North West	South East	South West	Yorkshire and North East
		N=396	N=347	N=562	N=319	N=6658	N=387	N=371
Ranked distance of local school:								
1	N=1183	39%	35%	37%	39%	43%	37%	42%
2	N=555	20%	17%	19%	19%	19%	18%	15%
3	N=361	13%	13%	13%	9%	10%	13%	12%
4	N=219	7%	6%	6%	11%	7%	7%	7%
5	N=178	6%	7%	6%	5%	6%	7%	5%
6	N=114	3%	3%	4%	3%	4%	4%	4%
7	N=85	2%	3%	3%	3%	3%	3%	3%
8	N=57	1%	3%	3%	2%	1%	2%	2%
9	N=45	1%	2%	2%	1%	2%	1%	2%
10	N=44	1%	2%	1%	2%	1%	2%	2%
11-20	N=165	6%	6%	6%	5%	4%	5%	7%
21-30	N=25	1%	2%	0%	0%	1%	1%	1%

Note: Teacher Tapp groups some small Government Official Regions to maintain adequate sample size within groups

Table A12: Salary increment, by teacher demographic characteristics

Demographic	Group	N	Salary increment (mid £)	Salary increment (mid % of salary)
Age	Age in 20s	274	£7,847	20%
	Age in 30s	997	£9,025	18%
	Age in 40s	1115	£9,346	17%
	Age in 50s+	647	£10,392	19%
Gender	Female	2232	£9,671	19%
	Male	789	£8,390	16%
Experience	Less than 5 years	329	£7,245	20%
	Between 5 and 10 years	623	£8,740	18%
	Between 10 and 20 years	1184	£9,400	17%
	Over 20 years	861	£10,436	18%
Seniority	Classroom Teacher	898	£8,487	20%
	Middle Leader	1265	£9,214	18%
	SLT (excl head)	628	£9,463	15%
	Headteacher	165	£14,742	18%
Youngest child at home	Under 5	435	£7,894	15%
	Aged 5-11	585	£9,611	19%
	Aged 11-16	460	£8,633	16%
	Aged 16+	280	£10,888	20%
	No children at home	1259	£9,562	19%
Tenure at current school	I started this academic year	226	£8,296	16%
	I started during last academic year (i.e. less than two years)	353	£7,886	17%
	Over two and less than three years	274	£8,905	19%
	Over three and less than four years	190	£8,553	17%
	Over four and less than five years	211	£7,666	17%
	Over 5 and less than 10 years	829	£9,406	18%
	Over 10 and less than 15 years	442	£10,271	19%
	15 years or more	491	£11,092	20%

Table A13: Regression of salary increment on teacher demographic characteristics

Variable	Coefficient	Std. Error	t-stat	p-value
Constant	4388	2060	2.13	0.03
Salary mid value	0	0	1.58	0.11
Salary mid value ^2	0	0	-0.93	0.35
Age (reference = Age in 20s):				
Age in 30s	394	827	0.48	0.63
Age in 40s	-413	911	-0.45	0.65
Age in 50s+	-240	980	-0.24	0.81
Gender (reference = Female):				
Male	**-1252**	449	**-2.79**	0.01
Experience (reference = Less than 5 years):				
Between 5 and 10 years	232	809	0.29	0.77
Between 10 and 20 years	393	853	0.46	0.64
Over 20 years	792	933	0.85	0.40
Seniority (reference = Class teacher):				
Middle Leader	-181	527	-0.34	0.73
SLT (excl. head)	-578	714	-0.81	0.42
Headteacher	**4228**	1222	**3.46**	0.00
Age of youngest child at home (reference = No children):				
Under 5	**-1585**	624	**-2.54**	0.01
Aged 5-11	-167	576	-0.29	0.77
Aged 11-16	**-1603**	637	**-2.52**	0.01
Aged 16+	710	761	0.93	0.35
Tenure (reference = first year at school):				
Over one and less than two years	-122	886	-0.14	0.89
Over two and less than three years	677	937	0.72	0.47
Over three and less than four years	-34	1027	-0.03	0.97
Over four and less than five years	-443	1001	-0.44	0.66
Over 5 and less than 10 years	869	795	1.09	0.27
Over 10 and less than 15 years	**1797**	872	**2.06**	0.04
15 years or more	**2369**	887	**2.67**	0.01

N=2918; R-squared=3.8%

Table A14: Demographic sorting of teachers across schools by FSM quintiles

	FSM quintile				
	1 (= low FSM)	2	3	4	5 (=high FSM)
	0%-10.4% FSM	10.5%-16.8% FSM	16.9%-23.6% FSM	23.7%-33.7% FSM	33.8%-71.0% FSM
% Age in 20s	7.2	8.2	8.7	10.7	10.2
% Age in 50s	22.7	20.4	22.1	22.2	19.1
% Female	71.5	75.5	75.3	71.2	74.0
% Less than 5 years' experience	9.5	10.2	9.9	12.0	12.5
% Started at school 0-2 years ago	17.6	17.3	18.9	21.3	20.3
Mean salary	£52,202	£53,149	£53,076	£53,275	£54,708

Table A15: Salary increment by type of current school

Demographic	Category	N	Mean salary increment	Mean salary increment as % of salary
Current school Ofsted	Outstanding	391	£10,818	20%
	Good	2049	£9,295	18%
	Requires improvement	306	£7,226	14%
Current school governance	Independent	177	£10,028	19%
	LA community	476	£8,797	17%
	LA non-community	404	£9,688	19%
	Large MAT (N>3)	1422	£9,084	18%
	Small MAT (N<=3)	142	£9,252	18%
	Stand-alone Academy	415	£10,142	20%
Current school size quartile	Q1 Small	807	£8,779	18%
	Q2	787	£9,322	18%
	Q3	765	£9,482	19%
	Q4 Large	668	£9,704	18%
Current school religious denomination	None	2384	£9,188	18%
	Religious RC	206	£11,899	22%
	Religious non-RC	447	£8,887	18%

Appendix I - Cost calculations for attributes

This appendix sets out the assumptions for calculating the costs of providing each job attribute. In each case, we express the cost in per teacher nominal terms and as a % of average salary.

Current average salary

This outlines the calculations for the average salary of all teachers in England, including employer costs such as National Insurance and pension contributions. (Note: This group includes teachers with reduced teaching loads, such as headteachers. While this affects later calculations, we will set this aside for now.) Median pay is slightly higher in secondary schools due to more middle leadership roles and is also higher in Greater London.

We use the median teacher pay estimate of £46,525 from the Teacher Pension Scheme records for the academic year 2023/24 (Department for Education, 2024c). (The School Workforce Census for the same academic year reports a lower figure because it was unlikely to fully reflect the recent pay rises (Department for Education, 2024a)).

Employer National Insurance contributions will rise from 13.8% to 15% in April 2025 (House of Commons Library, 2024). We assume that all incremental salary associated with these policy options will attract full NI.

Teachers' Pension Scheme employer contributions are now 28.68% and we assume that all teachers participate in this or a scheme of equivalent cost (Teachers' Pensions, 2024).

Therefore, total employer costs of a median salary are:

£46,525 (median salary) + £6,979 (employer NI) + £13,343 (employer pension) = **£66,847**

Cost of raising salaries by 5 or 10%

The all-in costs of raising salaries are:

- 5% = £3,342
- 10% = £6,685

Flexible working

The job attribute "*Commitment to meet all requests for flexible working*" is difficult to cost, as it requires assumptions about how many teachers would request reduced working hours or late starts/early finishes. In January 2025, Teacher Tapp asked teachers what working pattern they would choose—taking into account any associated pay cuts—if their employer could not refuse the request.

Among full-time teachers only, just 50% said they would choose to remain full-time if given the option. A further 9% said they would prefer to work 4.5 days, 32% would opt for 4 days, and 6% would choose to work 3 or 3.5 days.

In theory, this shift may carry no direct cost, provided schools could hire counterpart teachers to make up the lost hours. In practice, however, this is difficult to achieve. Part-time teachers often require their non-working time to be grouped into a single block, making timetabling more complex. This is likely to be a greater issue in secondary schools, where timetable inefficiencies (i.e. periods when a teacher is available but unscheduled to teach) are already more common.

To assess the likely cost to schools, we posed the following question to headteachers:

> *If half of your full-time teachers requested to work four days a week (with flexibility on which day off), and you recruited additional teachers to cover their lost hours, what would be the likely financial impact on your school's staffing costs?*

Responses are shown in Table A16 below. Around one-third of headteachers felt unable to answer the question immediately. Among those who did respond, the typical view—across both primary and secondary phases—was that a 5–10% increase in staffing costs would be required. We therefore estimate an increased cost equivalent to 7.5% of the staffing budget.

In open-ended responses, several headteachers noted additional costs associated with part-time staff, such as higher training costs and increased complexity in scheduling INSET and meetings. Secondary headteachers also highlighted that more complex timetables and an increase in split classes would likely reduce the quality of the student experience.

Table A16: Headteacher responses to survey question about accomodating part-time teachers

	Primary		Secondary	
	Including 'Not sure'	Excluding 'Not sure'	Including 'Not sure'	Excluding 'Not sure'
	N = 329	N = 223	N = 80	N = 52
No significant additional cost (e.g. flexible staffing could absorb this without major expense)	6%	9%	5%	8%
Minimal increase (less than 2% of staffing budget)	5%	8%	1%	2%
Moderate increase (2–5% of staffing budget)	18%	26%	16%	25%
Significant increase (5–10% of staffing budget)	22%	32%	21%	33%
Major increase (10-20% of staffing budget)	10%	14%	14%	21%
Very substantial increase (More than 20% of staffing budget)	7%	10%	8%	12%
Not sure – too many unknown factors	32%		35%	

Protected PPA time

This job attribute makes no assumptions about how these additional PPA hours are distributed, so we assume the costs are as follows:

- Reduction in teaching load from 90% to 80% = increase in salary costs of 12.5% of current salary = £8,356
- Reduction in teaching load from 90% to 70% = increase in salary costs of £16,712

Of course, there are some ambiguities here, particularly around how this would affect the teaching load of Early Career Teachers and Senior Leaders. So, this cover estimates may be an overestimate.

PPA working location

To assess the potential costs of placing restrictions on where and when PPA time can be taken, we asked headteachers the following question:

> *If the government mandated that all PPA time must be scheduled as a single block and include either morning registration or end-of-day dismissal (i.e. allowing for late starts or early finishes), could your school accommodate this within your existing budget?*

Responses are shown in Table A17 below. In the primary sector, this is already standard practice in most schools, so a zero-cost assumption is appropriate. In the secondary phase, however, where the impact on timetabling is more substantial, cost estimates varied considerably.

If we assume a 15% increase in staffing costs in the secondary phase and 0% in the primary phase, this results in an estimated average increase of 7.5% across both phases.

In open-ended responses, many secondary headteachers noted that such a policy could lead to a loss of stable form tutors, an increase in split classes, and difficulties in finding specialist teachers to fill resulting gaps in the timetable. While these concerns are not direct financial costs, they may nonetheless have a negative impact on teaching and learning.

Table A17: Headteacher responses to survey question about PPA blocking

	Primary	Secondary
	(N = 322)	(N = 78)
Yes – and we already do this	63%	1%
Yes – we could implement this within our existing budget	9%	6%
No - staff costs would rise a small amount (e.g. less than 5%)	4%	6%
No - staff costs would rise a moderate amount (e.g. between 5 and 15%)	11%	31%
No - staff costs would rise a significant amount (e.g. more than 15%)	7%	33%
No - even with additional budget this would be too disruptive	6%	22%

Class sizes

Class sizes are not currently fixed at 30, but this is generally the maximum observed in schools. If the government were to mandate a maximum class size of 26, we assume that the distribution of class sizes would shift downwards by four pupils. Given the current decline in the student population, we do not factor in constraints related to staffing availability or building capacity. We also assume no increase in the need for teaching assistants, although this assumption is more questionable for EYFS and Key Stage 1.

Under these assumptions:

- For a maximum class size of 26, the percentage increase in the number of classes is calculated as:
  (1/26 − 1/30) / (1/30) = 15.4%
- For a maximum class size of 22, the percentage increase in the number of classes is:
  (1/22 − 1/30) / (1/30) = 36.4%

Personal flexibility

The job attribute in question states: *"Flexibility, within reason, to meet all important family and personal commitments."* For the purposes of costing, we assume this equates to an average of three days' leave per year, covered either by supply agency teachers or internal cover. Current Teacher Tapp survey data suggests that internal cover is far more commonly used than agency staff.

The cost of using internal cover is estimated as 1/195 of the full employment cost per teacher, which equates to £343 per day, based on earlier calculations.

It is reasonable to assume that the cost of using supply agency teachers is broadly similar. On the one hand, supply teachers are often paid less than directly employed teachers, particularly as they are not eligible for responsibility points and may struggle to reach the higher ends of the pay scale. Pension contributions may also be lower (Impact Teachers, 2024). On the other hand, schools must pay an agency mark-up. While this figure was historically as high as 38%, it has fallen in recent years (Crown Commercial Service, 2019).

Healthcare

If staff are given time off for appointments and a wellbeing helpline, these two parts can be costed separately.

First, let's assume that time off for appointments amounts to one full day, on average per year = £343 per day based on calculations above.

Many organisations provide wellbeing helpline services, typically called Employee Assistance Programmes (EAP). None of these services are expensive on a per employee basis, relative to time off work, and a figure of £15 per employee per year is likely to cover smaller organisations such as schools (Engage Health Group, 2025).

The provision of "free" dentistry, physiotherapy and counselling services falls under insurance products specifically designed to cover ancillary healthcare services such as dentistry, physiotherapy, and counselling, excluding standard NHS-provided care. These are commonly known as health cash plans. Health cash plans allow individuals to claim back a portion of the costs incurred for routine healthcare expenses. For a regular monthly premium, these plans reimburse policyholders for services including dental treatments, optical care, physiotherapy, and counselling. The reimbursement amounts and covered services vary depending on the provider and the chosen plan level.

The cost of health cash plans varies based on the level of coverage and the provider. For example, some plans start at around £7 per month, offering basic coverage, while more comprehensive plans can cost up to £28.50 per month. These premiums correspond to annual benefits with specified limits for each type of service. For example, for £15 a month the Unite Health Cash plan covers up to £150 of dental costs, £350 of physio and £150 of optical (Union Insurance, 2025). Given the proposal was for "free" ancillary health services, it is prudent to assume a cost of £30 per teacher per month, or £360 per teacher per year.

Childcare

To estimate the cost of offering childcare subsidies, we first need to estimate how many teachers are likely to use them. Recent Teacher Tapp survey data indicate the following groups are most likely to require childcare:

- Teachers with their youngest child under 5: 14% of teachers, of whom 84% report spending money on childcare.
- Teachers with their youngest child aged 5–11: 15% of teachers, of whom 68% spend money on childcare.
- Teachers with their youngest child aged 11–16: 13% of teachers, but only 9% report childcare spending. As most of this falls outside the scope of the proposed job attribute, we assume only 4% of teachers in this group are relevant.

Teacher Tapp does not collect data on family size. Drawing on government statistics—and assuming that those with younger children have not yet completed their families—we estimate the following average number of children per teacher for cost purposes (Office for National Statistics, 2023):

- Those with a youngest child over 5: 1.7 children
- Those with a youngest child under 5: 1.4 children

For simplicity, we assume all children fall within the same age band when calculating costs. We estimate two separate costs: one for pre-school childcare and one for school-aged children, both based on term-time only.

- Pre-school childcare: £305.11 per week for children under 2 and £290.77 for children over 2 (Albert, 2025). Taking an average and allowing for 40 weeks of care results in an annual cost of £11,918. Therefore, a 20% or 40% school discount would cost £2,484 and £4,767, respectively.
- After-school childcare: £69 per week on average, amounting to £2,760 per year (Handley, 2024). A 20% or 40% discount would cost £552 and £1,104, respectively.

Bringing this together, the expected average per-teacher cost of offering a 20% childcare discount is calculated as follows:
(0.14 * 0.84 * 1.4 * 2484) + (0.15 * 0.68 * 1.7 * 552) + (0.04 * 0.09 * 1.7 * 552) = £508

CPD

There are three components to the per-day cost of CPD:

- Cover teacher costs – As noted above, this is estimated at 1/195 of a full salary, which equates to £343 per day.
- Course costs – Teacher Tapp maintains an events database via its app, showing that course fees vary widely, from free to over £350 per day. The mean cost for paid one-day courses is £228, while the average including free courses is £156. We assume an average course cost of £200 per day.
- Travel costs – An allowance of £30 per day covers typical travel expenses, including short train journeys or longer car journeys (over 50 miles) at standard HMRC mileage rates.

Taken together, the estimated total per-day cost of CPD is: £343 (cover) + £200 (course) + £30 (travel) = £573

Workload

We assume that the job attribute *"35-hour week guaranteed, with support for planning, marking, data, and non-core activities"* would require additional teaching assistant (TA) time—but the key question is: how much?

To explore this, we asked two questions to (non-head) teachers:

1. *If a teaching assistant were available to support you with planning, preparation, marking, and administration from 3:30pm onwards, how many hours of your work would you realistically be able to delegate to them each day?*

2. *After delegating these tasks to the teaching assistant, how many hours would you still need to work each evening to prepare for the next day's teaching?*

While primary teachers were generally more optimistic about the ability to delegate tasks to a TA, the modal response in both phases was an hour or less of work per day (see Table A18). If such delegation occurred, most teachers reported they would still have 1–2 hours of work remaining in the evening. One additional hour could feasibly fit within a 35-hour working week, but two hours would not—especially once ancillary duties such as break duties, registration, and meetings are taken into account.

In open-ended responses, many teachers expressed concern that TAs lack the pedagogical and subject-specific expertise required for effective lesson planning or meaningful marking. Secondary teachers, in particular, highlighted that subjects such as maths, science, languages, and English require specialist knowledge for accurate planning and assessment—something most TAs are not trained to provide. Teachers also emphasised the importance of retaining control over lesson planning and marking, citing personal preferences and the central role of marking in assessment for learning. Many felt that outsourcing this work would reduce the insights they gain into pupil understanding.

In reality, implementing a 35-hour work week would likely require broader structural changes beyond simply allocating TA time. However, for the purposes of this costing exercise, we assume that after-school preparation work equates to one-third of a TA role per teacher.

Table A18: Teacher responses to survey question about delegating work to a teaching assistant

		Primary	Secondary
		N = 2991	N = 2903
Offer for a TA to delegate after-school prep to	I would not be able to delegate any of my work	8%	28%
	I could delegate up to 1 hour of work each day	64%	48%
	I could delegate 1–2 hours of work each day	24%	19%
	I could delegate 2–3 hours of work each day	4%	3%
	I could delegate 3 or more hours of work each day	1%	2%
Workload it is impossible to delegate	I would be able to delegate everything and would not need to do additional preparation	6%	8%
	I would still need to work up to 1 hour each day	38%	33%
	I would still need to work 1-2 hours each day	40%	40%
	I would still need to work 2-3 hours each day	13%	14%
	I would still need to work 3+ hours each day	3%	5%

Teaching assistant costs are difficult to estimate precisely, as employment terms vary considerably. Many TAs are employed on term-time-only contracts, meaning they are paid only for the weeks worked during the school year. As a result, actual take-home pay may be lower than annualised figures suggest.

According to Unison research, the average annual salary for a teaching assistant is approximately £12,081, though this likely includes part-time roles. A reasonable estimate for a full-time teaching assistant is around £18,000, plus employer National Insurance and pension contributions (Think Teaching, 2025). Employer pension contributions typically range from 14% to 18% (Department for Education, 2024d). Taken together, this results in an estimated all-in average cost of approximately £23,400 per full-time teaching assistant.

Collegiality

In focus groups, headteachers agreed that the attribute *"Collaborative and friendly staff community with protected co-planning time"* would be difficult to deliver without additional PPA time specifically arranged to coincide with colleagues' schedules. If we assume this requires approximately 1/20th of a teacher's weekly timetable, the estimated cost of this attribute is £3,342 per teacher.

Behaviour

Our headteacher focus group discussions highlighted that the costs of delivering the attribute "*High standards of behaviour based on clear policies and consistent expectations*" vary significantly depending on the type of school. However, leaders of inclusive and comprehensive secondary schools agreed that it is not possible to maintain these standards without specific infrastructure—such as on-call facilities, a 'cool down' or isolation room, and centralised detentions.

Assuming these systems are staffed by two teachers and one teaching assistant, and the costs are shared across 75 teachers within a typical secondary school, the per-teacher cost is estimated at £2,095. This figure is likely to be an underestimate for secondary schools with more challenging intakes, and possibly an overestimate for primary schools or more socially selective secondary schools.

Appendix J - Implementation focus groups

Four online focus groups with head teachers were held after completing all stages of survey data collection (two with primary school Heads and two with secondary school Heads). The purpose of these discussions was two-fold: firstly, to contribute to dissemination of emerging findings and to support the project's impact by facilitating an action-focused peer-to-peer discussion; and secondly, to contribute to the final stages of the study itself.

Method

A total of 717 state-school headteachers were invited to participate in a 45-minute roundtable-style focus group, on one of two dates. In order to receive the invite, the headteacher either had to be in a school with a disadvantaged intake (primary 19.2%+ FSM, secondary 21.2%+ FSM), or be deemed 'innovative', on the basis that they had responded that they were doing one of the following things to improve teachers' working conditions: adjusting PPA time to allow full days from home, late starts or early finishes, using budget to increase PPA time or implementing significant workload reduction policies.

On signing up, potential participants were informed of the purpose of the discussion, that it would be recorded, and how their data would be used. 75 signed up, (48 primary, 27 secondary), and 38 eventually participated (22 primary; 16 secondary). Some participants were leaders of multiple schools as part of a trust, though we refer to all participants as Heads for the purposes of this appendix.

The discussion was intended to contribute to the research in two main ways. Firstly, to explore the practical realities of implementing the strategies explored in the survey, including by casting further light on some of the non-financial costs involved (also explored in the RQ2 open-response analysis). Secondly, to support the cost estimates used in the value for money calculations, particularly in relation to 'soft' attributes such as collegiality and behaviour.

On joining the meeting, attendees were reminded of the information about their participation that had been provided to them. They were also encouraged to change the name in the video-conferencing software if they preferred to remain fully anonymous during the discussion. They were also told they could leave at any point without giving a reason. The discussion then proceeded in three parts, beginning with a short (5-10 minute) summary of the survey findings. This included a PowerPoint with summary graphs, focusing on teachers' preferences for different job attributes; trade-offs between salary and other attributes; initial cost estimates; implications for cost-benefit; as well as findings from the Head teacher survey and a selection of open-responses. These responses were used to segue into the second part of the discussion, in which participants were asked to share any ways they had 'innovated', by putting into practice the strategies described in the analysis. Most participants described approaches to blocked PPA time or reduced timetables. This part of the discussion served as an opportunity for Heads to learn from each other, as well as to highlight the practical steps involved in implementing the strategies and the challenges they had faced. The discussion was largely unstructured, though the following prompts were used as necessary:

- What did you do?
- Why did you decide to do it?
- How did you go about it?
- What impact do you think it had?
- What challenges did you face?
- What advice would you give to someone else considering it?

The third part of the discussion tended to follow naturally and overlap with this discussion, but involved more focused discussion about the financial costs involved in two of the 'softer' attributes (collegiality and behaviour). The first pair of discussions focused more on the former, and the second pair on the latter.

Notes were taken throughout the discussion and recordings were also auto-transcribed. Following the focus groups, transcripts and field-notes were re-read and further notes taken under thematic headings, initially relating to the different attributes, with additional headings added inductively.

Findings

Most participants had tried at least one approach linked to the strategies explored in the main study, and they were strikingly enthusiastic about the impact of these changes. Several headteachers cited high numbers of quality applicants for recent vacancies, often contrasting this with previous recruitment experiences.

> *"I've been here four and a half years and this was the first year I've had to recruit a teacher and I had 200 applicants for my post. So yeah, it [providing a morning of PPA time at home each week, as well as subject leader time] does make a big difference."*
>
> Primary Head

Recruitment and retention were closely connected in the experience of headteachers, with most 'innovators' reporting reduced staff turnover alongside improved application rates.

> *"It [20% PPA time] was hugely popular and it's obviously very expensive. But we had no staff leave. We put out an ad for a teaching post and we got 18 applications for a one class teacher post. All our TAs as well or quite a few of them are going forward and training for unqualified teacher status and want to go into teaching as well...so it's had a knock-on effect across the school for growing future teachers...If there was anything I could do, I'd run it again continuously because it's the biggest impact on staff morale"*

Heads also reported positive spillover effects that should be considered when weighing the costs and benefits of different approaches. One noted that introducing a nine-day fortnight had improved staff attendance. Another observed that, when teachers were given blocked PPA time to use flexibly off-site, many used it to schedule medical appointments or other personal commitments, thereby reducing the need to request additional leave. One headteacher summed it up by saying that the goodwill generated by accommodating personal needs "was incredible".

Collegiality

The teacher survey focused on co-planning time as a route to collegiality, and there was strong consensus among headteachers that creating opportunities for teachers to plan together is indeed a key way to foster a collegial working environment. Some heads also highlighted additional factors that contributed to collegiality, such as the stability of the staff team, the way shared challenges in difficult schools can bring staff together, and the informal benefits of teachers spending time with one another during playground duty. However, the majority of discussion around collegiality centred on the importance of providing additional or explicitly designated non-teaching time for collaborative work. The theme of collegiality is therefore best understood in relation to broader considerations about timetabling and the allocation of PPA time.

Timetabling and PPA time

The discussion regarding timetabling and PPA time fell into two interlinked categories: volume and allocation.

Increasing the quantity of non-teaching time

Many participants had either increased PPA time or expressed a desire to do so. One primary school had temporarily raised PPA to 20% in the lead-up to an Ofsted inspection, due to the scale of improvements required. However, most headteachers viewed a 20% allocation as desirable but unaffordable. Some schools had previously moved to a 15% allocation but were now reducing this in response to budget pressures. One school that had been unable to offer any enhancement noted that this created difficulties, as neighbouring schools with stronger financial positions could do so, creating competitive pressure in the local recruitment market.

Schools described a range of strategies for creating additional non-teaching time, each with different cost implications. In secondary schools, heads reported deliberately overstaffing larger departments to build in flexibility. In primary schools, a common approach was to use specialist teachers to cover subjects such as art, languages, and PE—either brought in from external organisations, in one case using a carousel model. Others made use of Forest School, delivered by specially

trained Higher Level Teaching Assistants (HLTAs). One head suggested that, with additional funding, more HLTAs could be trained to deliver specialist subjects such as art and PE in order to free up teachers' time.

> *"We do forest schools on a Friday afternoon. We find it works best because a lot of our parents are around to support us on a Friday afternoon and so each class gets a slot of that and I've encouraged them really basically to go home or to take the PPA and when they've got a weekend away or whatever work it out so they can actually go on holiday or away for the weekend early."*
>
> Primary Head

Some schools were turning to HLTAs out of financial necessity. One head described moving away from external specialist teachers due to budget constraints, despite believing they were preferable.

> *"The way previously the school had been enabling all three teachers in their year group to have their PPA together and therefore cover all three classes has been to bring in outside providers to deliver parts of the curriculum [like] music. It's great because teachers get their PPA at the same time ... However, I've also inherited a great big deficit budget. And that provision, the music and the sport all in is costing about £68,000 a year. So, that's got to go, I'm afraid. And it's purely a financial decision."*
>
> Primary Head

HLTAs were also being used in more creative ways. One primary head had freed up an additional hour of PPA per week for each teacher by paying teaching assistants to cover pre-lesson time, including welcoming pupils and taking the register.

In one multi-phase trust, secondary teachers were deployed to provide specialist cover in the primary phase—particularly for Key Stage 2. This was combined with staggered late starts for Years 10 and 11. However, the school acknowledged that such strategies could undermine collaboration and collegiality by reducing shared time.

It was clear from participants' comments that the calibre of the staff providing cover was critical to maintaining the quality of learning, and that this had implications for staff, parent, and teacher confidence. This in turn has cost implications, as staff delivering cover need to be at least HLTA level and may require additional training. However, where schools brought in more skilled staff, they often gained wider benefits. For instance, one school employed a specialist Spanish teacher two days a week, who also supported pupils in the school's dedicated SEND space.

Several heads saw enhanced subject leadership time (in primary schools) or department time (in secondary schools) as a more affordable alternative to increased PPA. In some cases, this was planned to support collaboration. For example, one secondary school timetabled a weekly department meeting slot on top of the standard 10% PPA time. This was seen as beneficial for collegiality and, when used for activities like moderation, freed up time elsewhere for more personalised professional development. The head noted challenges in larger departments such as Maths, where teaching loads were high, but explained that this was resolved by scheduling department time during the delivery of the personal development curriculum, which did not involve Maths staff.

Another head had blocked subject leadership time to give each subject-lead a full week to focus on subject knowledge development. In another example, a school had moved away from traditional subject leadership structures and instead formed a cross-subject learning leadership team. This team could 'bid' for time to work on specific priorities, such as curriculum redesign. These examples suggest that schools may be able to create collaboration time through targeted leadership allocations rather than global increases in PPA—though the cost and logistical implications are likely to be similar.

Not all schools were moving in the same direction, however. One was considering abolishing subject leadership time and planning days altogether, instead increasing PPA to 20%. They believed this would offer greater predictability and stability for staff, as well as reducing disruption.

Allocation of non-teaching time

Headteachers reported balancing several considerations when allocating PPA time—most notably, teachers' desire for the flexibility to work from home during a blocked period and the benefits of collaborative planning. This tension was highlighted in open responses to both the teacher and headteacher surveys, as well as in initial focus groups.

Many participants were actively managing this trade-off by explicitly planning for both types of time, often in alternating weeks. In some schools, teachers had a flexible slot they could use at home one week, followed by a week with a shared session for collaborative activities. These arrangements were often facilitated by two-week timetables.

> *"Teachers actually told us how much they valued the hybrid- that they did get some time away from school at home but they also really valued the collegiate working with their team the other PPA…they always tell me they feel they have the best of both worlds."*
>
> Primary head

At secondary level, some heads noted that teachers preferred their PPA time to be spread across the week to provide breathing space, while primary teachers tended to favour blocked time. This is reflected in the phase-by-phase subgroup analysis of the survey results. One secondary head offered teachers a choice, with around half opting for the blocked approach.

Another secondary head allowed staff to choose whether they wanted off-site time in the morning or afternoon. Those with children tended to choose the afternoon to allow for school pick-up, while younger staff without caring responsibilities preferred mornings—for example, to go to the gym. The school used assembly time to facilitate flexibility in the mornings, and another head noted that their four-period day structure made it easier to block time for at-home PPA.

> *"We're down on staff for form time. However, we've overcome that by manipulating when assemblies are so that the staff that have got it off in the morning, it happens to be their year groups assembly day."*
>
> Secondary head

Several heads emphasised the need to ensure allocations were perceived as fair—avoiding, for example, the same individuals being given blocked PPA time on Mondays or Fridays each year.

Other strategies for promoting collaboration and collegiality included starting the school day earlier and finishing early on one day each week to allow for joint planning. Two schools operating a nine-day fortnight followed a similar model. Other schools provided less frequent but regular and structured co-planning or peer learning time. One school, for instance, held Teacher Learning Community sessions every four Mondays, while a large primary school allocated one or two days each year for year teams to work together. These were held off-site in a hotel, with lunch provided—an approach the head felt contributed positively to staff morale, though they noted that the number of such days was being reduced due to budget pressures.

The two most commonly cited barriers to scheduling collaborative time and implementing innovative timetabling were the technical complexity of writing the timetable, and the challenge of aligning schedules for part-time and job-share staff.

> *"I made the offer before I asked my timetabler. So, whether we do make it work or not, I don't know. But I figured if I asked the timetabler first, it would be dead before it started".*
>
> Secondary Head

One head, responsible for two schools across separate sites, noted that this constraint made it more difficult—and more expensive—to provide additional teaching capacity.

Some schools found that AI timetabling software was helping to address the first of these challenges. One reported having the first draft of the September timetable ready by March, allowing time for feedback and iteration. However, this may be difficult in schools with high staff turnover or significant staffing uncertainty.

The most common solution to managing collaboration for part-time staff was to designate a day each week when all staff were present, and to schedule departmental, subject leadership or CPD time on that day. Achieving this level of overlap sometimes required modest overstaffing, which some heads viewed as a necessary cost. One head reported difficulties due to not having any overlap. Another found that shared time among job-share teachers created useful capacity to cover subject leadership responsibilities.

> *"Nobody can request flexible working on a Wednesday, so everybody is in school on a Wednesday... [we've worked hard at] really putting the mechanisms in so the meeting schedules support all staff to share information - and being really conscious with the design of that. I think we've gone through four iterations... and we've learned a lot along the way"*
>
> Secondary Head

At primary level, some heads of larger schools said their scale made it easier to offer flexibility and blocked PPA time. However, another head of a smaller, single-form entry school felt it was easier to provide off-site PPA in their context, as there was no need to coordinate shared planning across year groups.

In some schools serving large Muslim communities, Friday arrangements were already adapted in ways that supported innovation. One school had adopted a core timetable on Friday afternoons similar to its remote learning model during lockdowns.

> *"We close at one PM on a Friday, and we do it for a number of reasons. We're 60% Muslim and it goes down well with Friday prayers. But also, it allows all staff to meet at the same time"*
>
> Primary Head

Another head acknowledged that reduced contact time could meet with resistance from parents, but reported success in gaining support by emphasising the potential to improve pupil outcomes—framing the change not as a wellbeing initiative for teachers, but as a means of securing and retaining high-quality teaching.

Reducing workload

Workload reduction strategies often overlap with increased PPA time, as the latter frequently contributes to the former. However, when discussing workload reduction more broadly, headteachers typically focused on either the deployment of additional staff or the adaptation of policies and processes.

Additional staffing to reduce workload

Most additional staffing came from teaching assistants, ideally those at HLTA level. In addition to responsibilities noted above—such as greeting pupils and taking the register—these staff members were often given responsibility for preparing classroom resources (e.g. laminating, labelling, photocopying) or supporting the SENDCO with administrative tasks. One head noted that this had the added benefit of freeing the SENDCO to spend more time directly with pupils. Most, if not all, references to these types of approaches came from primary schools.

Once again, the level of specialist expertise was seen as critical in ensuring that quality did not suffer and that teachers—who, according to one head, can be "control freaks"—were willing to delegate tasks. One school with 265 pupils had even employed a fully qualified teacher, who was happy to work on an HLTA salary in exchange for reduced responsibilities, to assist with marking and tasks such as question-level analysis.

> *"It means that a lot of those administration tasks that you couldn't necessarily or... you might not necessarily hand over to teaching assistants are performed by her. That's been really helpful."*
>
> Primary Head

Several heads referred to a teaching assistant-to-teacher talent pipeline, noting that having support staff who were training—or planning to train—as teachers made it easier to give them more responsibility. These arrangements also had the added benefit of nurturing 'home-grown' talent to meet future recruitment needs.

> *"We were also in the very lucky position that we had - costing-wise - We had four of our teaching assistants...who were training to be teachers and so wanted that experience and qualified teachers experience. they were at university one day two days a week and then they needed the additional experience. So, they were the ones that were predominantly our cover teachers"*
>
> Primary Head

At secondary level, non-teaching staff were sometimes used to reduce teacher workload related to managing pupil behaviour and liaising with parents—an area we return to below.

Adapted policies and processes to reduce workload

Several schools were using computer- or AI-assisted marking in a bid to reduce workload. Headteachers highlighted a range of potential benefits and risks, both in terms of time savings and the wider, non-financial implications.

One head noted that AI marking was useful for basic assessments where only a mark was required, but less appropriate when teachers needed to understand pupil misconceptions or inform future teaching. Others disagreed, suggesting that AI tools could support moderation, improve consistency, and even offer helpful insight into misconceptions.

> *"We're trialling AI marking and we've been doing that for English, and it's been remarkable. It's been really good as a moderation tool, but it's also been really good to have the language papers all marked by AI, and a certain percentage marked by our teachers. We then compare to see whether we're on the same track as AI to make sure we're quality assuring it"*
>
> Secondary Head

Another head, working in a trust using AI-assisted marking across multiple schools, found the approach useful for identifying patterns in pupil misconceptions across the trust. However, they noted that *"the jury's still out"* on whether teachers engage as deeply with AI-generated insights as they do with their own marking. They emphasised that their trust was limiting AI to simple tests and low-stakes assessments. They also pointed out logistical challenges, including the need to use computer rooms for such assessments, which introduced resourcing pressures and implied potential capital costs if such systems were scaled up.

One school had opted to purchase curriculum packages—including lesson plans, resource slides, and an assessment system—from an external provider. While some heads saw value in this, others expressed discomfort with reducing teachers' role in planning, which they felt was a central part of the profession. Opportunities to delegate other tasks, such as letter writing or data analysis, were also raised.

At the primary level, there were various examples of staff redeployment to reduce workload. In one three-form entry school, the same teacher taught a specific content area across multiple classes, which the head believed enhanced consistency, ensured pupils were taught by the most expert teacher in that area, and reduced planning time. Teachers reportedly enjoyed this arrangement, particularly when it aligned with their subject interests. A similar approach was reported in another school, where subject leaders were responsible for planning across year groups and this approach may be suited to smaller schools.

> *"They're finding that they're enjoying that because it just so happens for that team that they're teaching subjects they're really passionate about, which is great for the kids. and it's reduced their workload because they're only planning and delivering for them at the minute just over a third of the curriculum."*
>
> Primary Head

Finally, one school—unable to offer the two-week October half-term provided by other schools in their area—chose instead to give teachers an additional day off following an open evening. Although the statutory requirement for a fixed number of school days posed a challenge, the head felt this smaller gesture had a significant positive impact.

> *"The teachers have been absolutely over the moon about it. and so we've managed to do that. It's a small thing, but it's been a big win in our community"*
>
> Secondary Head

Behaviour

The costs associated with offering high standards of behaviour (or *"disruption-free learning"*, as one head described it) varied significantly between primary and secondary schools. We therefore consider the two phases separately.

Secondary school behaviour management

As noted in the initial teacher roundtables, headteachers—particularly those in secondary schools—highlighted that statements about behaviour in job adverts could be interpreted as statements about pupil intake.

> *"When I see those adverts that say, 'Come and work in our lovely school with our lovely children,' I don't think they're describing my school, even though my school is lovely and my children are lovely."*
>
> Secondary Head

There was also a wider perceived tension between behaviour and inclusivity. While high standards of behaviour can be achieved at relatively low cost by skewing a school's intake or through exclusion, achieving such standards within an inclusive model was seen as significantly more expensive.

> *"You'll see job adverts with 'high standards of behaviour'. What does that mean? Everybody has high standards of behaviour.... it's whether the support is there to actually achieve that... and if you're an inclusive school - and we're all supposed to be inclusive, then you're going to have challenges within a classroom and within your school environment, there's always tension there between the two things"*
>
> Secondary Head

Heads described two main approaches to achieving high behavioural standards with more challenging intakes: enhanced staffing for centralised behaviour management, and greater investment in supportive or alternative provision. Both approaches carry distinct cost implications, and most heads reported using a combination of the two. One head noted that different strategies—with different cost profiles—were needed to reach a "basic" standard of behaviour compared to achieving an "exemplary" standard. The former, they argued, could be attained largely through clear, consistent systems.

Whole-school detention systems and the deployment of additional pastoral or leadership staff were common strategies for centralising behaviour management. For example, one school employed non-teaching pastoral staff to manage detentions and deal with parental queries. The head described this as a low-cost model: detentions were run in the hall, avoiding the need for multiple teachers supervising small groups in different rooms. Initial training for pastoral staff was required to ensure they had the confidence and authority to fulfil their roles. The head also reported additional benefits, noting that time freed from detention duties allowed teachers to focus on planning, marking, and informal collaboration.

> *"We have centralised detentions. We have a huge non-teaching pastoral team, and no teacher sits in detentions after school. We do that centrally with pastoral support workers and I think that creates the time for people to have those informal conversations after school and not be impacted by having to deal with that... It doesn't cost a huge amount to be honest we run all the detentions in the hall so they're all in there together... we have four or five support staff staying until 4:15 after school"*
>
> Secondary Head

In another school, a member of staff was employed for two hours each evening to oversee the detention system—for example, rolling over missed detentions to lunchtime the next day. Another school had three pastoral assistants running the school's internal exclusion space.

Several heads reported deploying middle or senior leaders for 'call outs'. One head explained that additional leadership capacity could also support more proactive behaviours, such as walkabouts and maintaining an open-door culture, which contributed to collegiality and inclusive leadership.

> *"So, we've got praise going on for staff and students during that time which is really valued by the staff. and it all leads to the belonging of staff. So, our staff say that they feel that they belong and that I have this open door*

> *policy…They can come and tell me whatever they want to tell me and they feel that they can do that and I think all of that contributes to our retention. We have very low turnover."*
>
> Secondary Head

Another noted that middle leaders were essential to sustaining the system: without them, the senior leadership team would be entirely absorbed in responding to call-outs.

Secondary heads emphasised that achieving high behavioural standards for pupils with complex needs required significant additional support beyond punitive approaches. One school reported employing three additional pastoral support assistants to support behaviour and attendance post-pandemic. Another listed a range of specialist staff in place:

> *"Two attendance people… two reading people… one attachment and LAC (looked after children) mentor, one SEMH (Social, Emotional and Mental Health Difficulty) key worker, two SEMH instructors, two SEMH TAs … vin my unit, one pastoral support worker that does all the data and management and infilling of things and more latterly an assistant SENCO because my SENCO was sinking, as well as a private part-time EP (educational psychologist and a private part-time EWO (education welfare officer)."*
>
> Secondary Head

A head who had invested in centralised detentions, overnight support, and an internal exclusion team estimated that this—alongside a support unit for SEMH pupils and an on-call leadership rota—cost approximately half a million pounds. They framed this as the cost of pursuing high standards *with* inclusion, noting that the cheaper alternative would be to exclude pupils needing intensive support.

Not all heads felt that an inclusive model was viable within their budgets. Some were relying instead on external alternative provision (AP), despite rising costs.

> *"I hear people talking about setting up inclusion units and things like that and that's what we're striving and aiming to do but at the moment we haven't got the funding to even begin to start to do that. So, we are spending money on AP to provide the high standards of behaviour."*
>
> Secondary head

Primary school behaviour management

One headteacher who led schools across different phases remarked that behaviour issues—and the strategies required to manage them—were markedly different in secondary compared to primary and early years settings, noting, *"I don't have conversations about behaviour with my primary leads, but I do in my secondary."* This distinction was evident when comparing discussions with primary and secondary heads, and it underscores the need for subgroup-level analysis by phase.

At primary level, heads tended to emphasise the importance of additional leadership capacity to support high standards of behaviour—for example, being present on the playground to address issues early and build relationships. One head echoed this emphasis, noting that playground presence *"cuts down a lot of the nonsense that happens, which is wearing for teachers."* Their strategy included deploying additional staff on the playground, setting up buddying schemes with older pupils, and employing a sports coach from a local football team. The coach was paid for an extra hour each day, although costs were kept low due to a subsidy from the football association. Another school had implemented the Outdoor Play and Learning (OPAL) programme, which the head believed had significantly reduced the number of issues arising from lunchtime, at a cost of approximately £6,000.

Two headteachers mentioned external programmes or training as a means of establishing consistent behaviour expectations. One had participated in a local initiative based on Paul Dix's work, while another adopted an approach informed by *Teach Like a Champion*. The former did not report any direct costs—possibly because these were covered by the local authority—aside from some of their own time spent meeting pupils one-to-one to build relationships over hot chocolate. The latter had purchased thirty-five copies of *Teach Like a Champion* (£25 each) and invested in instructional coaching via StepLab, at a cost of approximately £2,000 per year, in addition to some INSET time.

Finally, some heads referred to behaviour strategies tailored to support pupils with SEND, including "*sensory circuits, movement breaks, and nurture*." One head believed that having a dedicated SEND class had improved behaviour overall, as it reduced the number of children placed in environments that didn't meet their needs.

Class sizes

Only two headteachers reported operating with reduced class sizes. In one case, this was an alternative provision (AP) school where class sizes had been capped at 10 until recently—though financial pressures had led to an increase to 12. The head believed that smaller class sizes were indeed a draw for recruitment and noted the additional benefit of reduced marking workload.

The second head, from a mainstream secondary school, felt that smaller classes were essential for making mixed-ability teaching effective. Their school was relatively new, and they were finding the approach increasingly difficult to sustain as the school expanded—particularly due to space constraints. This highlights the potential capital cost implications of reducing class sizes, which may be less of an issue in areas with falling pupil numbers, but more problematic in regions where school populations are growing.

> *"we've got class sizes of 25. They just gone up to 26 for year seven and eight. The challenge is rooming. Now we're ... up to our second year of year 13 we sometimes have lessons in the canteen... so that's the challenge."*
>
> Secondary Head

Other attributes

We did not focus specifically on CPD in the focus groups, but several headteachers expressed surprise that it was such an unpopular attribute in the survey. Where CPD was mentioned, heads tended to describe more self-directed approaches—contrasting with the phrasing used in our survey ("6 days per year for courses, chosen by you from a list of approved providers"). For example, one head highlighted the value of staff being able to "go off on tangents about something that personally interests them professionally."

Several heads also reported deliberately appointing more expensive senior staff, which they believed had a significant impact on reducing turnover. One head felt that opportunities for progression were an important job attribute that had not been fully captured in the survey, while another believed they had benefited from being part of a well-known and respected school trust that teachers were actively seeking to join.

References for Appendix

Albert, Angeline (2025, February 18). *Childcare costs: How much do you pay in England 2025?* Retrieved March 3, 2025, from https://www.daynurseries.co.uk/advice/how-much-does-childcare-cost-in-england-2024.

Allen, R., Ford, I., & Hannay, T. (2023). *Teacher Recruitment and Retention in 2023: Teacher views on coping with shortages, job attachment and flexible working*, Teacher Tapp Research Report, https://teachertapp.co.uk/app/uploads/2023/06/Teacher-Recruitment-and-Retention-in-2023-TT_SchoolDash-Final.pdf.

Allen, R., Ford, I., Hallahan, G. & Hannay, T. (2024). *Teacher Recruitment and Retention in 2024: An exploration of recruitment challenges in disadvantaged schools*, Teacher Tapp Research Report, https://teachertapp.co.uk/publications/recruitment-and-retention-in-2024/.

Allen, R. & McInerney, L. (2019). *The Recruitment Gap: Attracting teachers to schools serving disadvantaged communities*. Sutton Trust Research Report, https://www.suttontrust.com/wp-content/uploads/2019/07/The-Recruitment-Gap.pdf

Allen, R. & Sims, S. (2018). Do pupils from low-income families get low-quality teachers? Indirect evidence from English schools. *Oxford Review of Education*, 44 (4), 441-458.

Adams, L., Coburn, S., Sanders-Earley, A., Harris, H., Taylor, J. & Taylor, B. (2023). *Working lives of teachers and leaders – wave 1*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1148571/W o rking_lives_of_teachers_and_leaders_-_wave_1_-_core_report.pdf (Accessed: 17 August 2023).

Benhenda, A., & Sims, S. (2022). *The effect of financial incentives on the retention of shortage-subject teachers: Evidence from England*, (CEPEO Working Paper 22-04). UCL Centre for Education Policy and Equalising Opportunities (CEPEO).

Bijker, R., Merkouris, S. S., Dowling, N. A., & Rodda, S. N. (2024). ChatGPT for Automated qualitative research: content analysis. *Journal of medical Internet research*, *26*, e59050.

Breidert, C., Hahsler, M., & Reutterer, T. (2006). A review of methods for measuring willingness-to-pay. *Innovative marketing*, *2*(4).

Burge, P., Lu, H. & Phillips, W. D. (2021). *Understanding Teaching Retention: Using a Discrete Choice Experiment to Measure Teacher Retention in England*, RAND Research Report, https://www.rand.org/pubs/research_reports/RRA181-1.html.

Campbell, S. (2009). *Cognitive Testing in Survey Questionnaire Design*. Scottish Government. https://www.gov.scot/binaries/content/documents/govscot/publications/advice-and-guidance/2009/12/social-research-methods-guides/documents/cognitive-testing-in-survey-questionnaire-design/cognitive-testing-in-survey-questionnaire-design/govscot%3Adocument/.

Camelo, R., & Ponczek, V. (2021). Teacher Turnover and Financial Incentives in Underprivileged Schools: Evidence from a Compensation Policy in a Developing Country, *Economics of Education Review*, Volume 80, 102067.

Campoamor, N. B., Guerrini, C. J., Brooks, W. B., Bridges, J. F., & Crossnohere, N. L. (2024). Pretesting discrete-choice experiments: a guide for researchers. *The Patient-Patient-Centered Outcomes Research*, 17(2), 109-120.

Campoli, A. K., & Conrad-Popova, D. (2017). Invisible threads: Working conditions, interpersonal relationships, and turnover among Black female teachers. In *Black female teachers* (Advances in Race and Ethnicity in Education, Vol. 6, pp. 117-134). Emerald Publishing Limited. https://doi.org/10.1108/S2051-231720170000006007

Chong, S.W., Oxley, E., Negrea, V., Bond, M., Liu Q., and Kong, M.S. (2024). *Teacher recruitment and retention in schools in socio-economically disadvantaged areas in England - Review of Practice*, Education Endowment Foundation Research Report.

Clotfelter, C., Glennie, E., Ladd, H. and Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina, *Journal of Public Economics*, Volume 92, Issues 5–6.

CooperGibson Research. (2018). Factors affecting teacher retention: qualitative investigation. Department for Education Research Report. https://assets.publishing.service.gov.uk/media/5aa15d24e5274a53c0b29341/Factors_affecting_teacher_retention_-_qualitative_investigation.pdf

Crown Commercial Service. (2019, December 3). *Supply teachers: Agency mark-up and the impact on worker pay*. Crown Commercial Service. https://www.crowncommercial.gov.uk/news/supply-teachers-agency-mark-up-crown-commercial-service

Dai, S. C., Xiong, A., & Ku, L. W. (2023). LLM-in-the-loop: Leveraging large language model for thematic analysis. *arXiv preprint arXiv:2310.15100*.

Davis, B. K., Ingle, W. K., & Choi, N. (2021). Preferred teacher applicant characteristics among school leaders in twelve Kentucky rural and suburban school districts. *Leadership and Policy in Schools*, *20*(3), 368-383.

Department for Education. (2024a). School workforce in England: Reporting year 2023. Explore Education Statistics. Retrieved 28/02/2025, from https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england

Department for Education. (2024b). *Working lives of teachers and leaders: wave 2 summary report*. https://www.gov.uk/government/publications/working-lives-of-teachers-and-leaders-wave-2/working-lives-of-teachers-and-leaders-wave-2-summary-report#:~:text=The%20working%20lives%20of%20teachers,wellbeing%20and%20flexible%20working%20arrangements

Department for Education. (2024c). Median teacher pay using Teacher Pension Scheme data: 2023-24. Explore Education Statistics. https://explore-education-statistics.service.gov.uk/find-statistics/median-teacher-pay-using-teacher-pension-scheme-data/2023-24

Department for Education. (2024d, December 16). Local Government Pension Scheme (LGPS). GOV.UK. https://www.gov.uk/government/publications/academy-trust-financial-management-good-practice-guides/local-government-pension-scheme-lgps

Engage Health Group. (2025). *Employee assistance programmes.* Engage Health Group. Retrieved March 3, 2025, from https://www.engagehealthgroup.co.uk/employee-assistance-programmes/

Engel, M. (2013). Problematic preferences? A mixed method examination of principals' preferences for teacher characteristics in Chicago. *Educational Administration Quarterly*, *49*(1), 52-91.

Engel, M., & Curran, F. C. (2015). *Toward Understanding Principals' Hiring Preferences*. Working paper.

European Commission, Directorate-General for Research and Innovation. (2024). *Living guidelines on the responsible use of generative AI in research* (First edition). https://op.europa.eu/en/publication-detail/-/publication/5cf2cb59-e631-11ee-8a44-01aa75ed71a1

Fonta, W. M., Ichoku, H. E., & Ogujiuba, K. K. (2010). Estimating willingness to pay with the stochastic payment card design: further evidence from rural Cameroon. *Environment, Development and Sustainability*, *12*, 179-193.

Feng, L., & Sass, T. (2018). The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management*, 37(1), 112-135.

Fuchsman, D., McGee, J.B., & Zamarro, G. (2020). *Teachers' Willingness to Pay for Retirement Benefits: A National Stated Preferences Experiment*. Brown University Working Paper 20-313.

Giersch, J., & Dong, C. (2018). Principals' preferences when hiring teachers: A conjoint experiment. *Journal of Educational Administration*, 56(4), 429-444.

Gundlach, H.A.D, Slemp, G.R. & Hattie, J. (2024). A meta-analysis of the antecedents of teacher turnover and retention. *Educational Research Review,* 44. https://doi.org/10.1016/j.edurev.2024.100606

Handley, S. (2024, October 4). *What are the average childcare costs in the UK in 2025?* GoodTo. https://www.goodto.com/money-news/average-childcare-costs

Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

Hong, M. H., Marsh, L. A., Feuston, J. L., Ruppert, J., Brubaker, J. R., & Szafir, D. A. (2022). Scholastic: Graphical human-AI collaboration for inductive and interpretive text analysis. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology,* (October, pp. 1-12).

Horng, E. L. (2009). Teacher Trade-offs: Disentangling teachers' preferences for working conditions and student demographics. *American Educational Research Journal*, 46(3), 690-717. https://doi.org/10.3102/0002831209348979

House of Commons Library. (2024*). Teacher recruitment and retention in England* (Briefing Paper No. CBP-10189). UK Parliament. https://commonslibrary.parliament.uk/research-briefings/cbp-10189/

Impact Teachers. (2024, September 24). *Understanding supply teachers pay rates 2024/2025: What you need to know*. Impact Teachers. https://impactteachers.com/blog/understanding-supply-teachers-pay-rates/

Ingle, K., Rutledge, S., & Bishop, J. (2011). Context matters: principals' sensemaking of teacher hiring and on-the-job performance. *Journal of Educational Administration*, *49*(5), 579-610.

Jabbar, H. (2018). Recruiting "talent": School choice and teacher hiring in New Orleans. *Educational Administration Quarterly*, *54*(1), 115-151.

Jabbar, H., & Holme, J. J. (2025). Teacher Turnover, Social Capital, and Improvement: How Instability Disrupts Schools. *Educational Evaluation and Policy Analysis*, 0(0). https://doi.org/10.3102/01623737241311304

Jerrim, J. (2024). The link between teacher buy-in and intentions to continue working in their current school. *British Educational Research Journal*, 50, 1084–1105. https://doi.org/10.1002/berj.3961

Jerrim, J. & Sims, S. (2019). *The Teaching and Learning International Survey (TALIS) 2018*. Department for Education.

Jerrim, J. & Sims, S. (2024). Interpersonal relationships and organisational commitment. Longitudinal evidence from school staff in England. *Working paper*.

Johnson, A., Lee, M. H., & Cheng, A. (2024). Which characteristics do religious school administrators value in teachers? Experimental evidence from the global Christian school sector. *Journal of Religious Education*, *72*(1), 91-108.

Johnston, A.C. (2021). *Preferences, Selection and the Structure of Teacher Pay*, IZA-Institute of Labor Economics.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant* 47, 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Kukla-Acevedo, S. (2009). Leavers, Movers, and Stayers: The Role of Workplace Conditions in Teacher Mobility Decisions. *The Journal of Educational Research*, *102*(6), 443–452. https://doi.org/10.3200/JOER.102.6.443-452

Kurtz, M., & St. Maurice, H. (2018). Teacher retention: Why they stay. *Advance*. https://doi.org/10.31124/advance.7403942.v1

Lentini, V., Gimenez, G., & Valbuena, J. (2024). Teachers' preferences for incentives to work in disadvantaged districts: A discrete choice experiment in Costa Rica. *Economic Analysis and Policy*, 82, 831-845.

Lovison, V. S., & Hyunjung Mo, C. (2024). Investing in the Teacher Workforce: Experimental Evidence on Teachers' Preferences. *American Educational Research Journal*, 61(1), 108-144. https://doi.org/10.3102/00028312231208956

Lynch, D., Peddell, L., McGuigan, H., Willis, R., Yeigh, T. & Marcoionni, T. (2024). School leader perception of control and professional learning decision-making influences, *Professional Development in Education*, 50(1), 129-143.

McLean, D., Tang, S., & Worth, J. (2023). *The impact of training bursaries on teacher recruitment and retention: An evaluation of impact and value for money*. National Foundation for Educational Research.

McLean, D., Worth, J. and Smith, A. (2024). *Teacher Labour Market in England Annual Report 2024*, https://www.nfer.ac.uk/publications/teacher-labour-market-in-england-annual-report-2024/

Menzies, L., Parameshwaran, M., Trethewey, A., Shaw, B., Baars, S., & Chiong, C. (2015). *Why Teach?* http://dx.doi.org/10.13140/RG.2.2.12227.86567

Nguyen, D., See, B.H., Brown, C. & Kokotsaki, D. (2023). *Reviewing the evidence base on school leadership, culture, climate and structure for teacher retention: Rapid evidence assessment*. EEF Research Report.

Office for National Statistics. (2023, May 18). *Families and households in the UK: 2022*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2022

Patel-Syed, Z., Becker, S., Olson, M., Rinella, H., & Scott, K. (2024). What do you think it means? Using cognitive interviewing to improve measurement in implementation science: description and case example. *Implementation Science Communications*, *5*(1), 14. https://doi.org/10.1186/s43058-024-00549-0

Perkins, M., & Roe, J. (2024). Generative AI tools in academic research: Applications and implications for qualitative and quantitative research methodologies. *arXiv preprint arXiv:2408.06872*.

Perrone, F., & Meyers, C. V. (2023). A Review of the Research on Teacher Hiring in the United States (2001–2021). *Journal of Education Human Resources*, (aop), e20220041.

Perryman, J. & Calvert, G. (2020). What Motivates People to Teach, and Why Do They Leave? Accountability, Performativity and Teacher Retention, *British Journal of Educational Studies*, 68 (1), pp. 3–23. https://doi.org/10.1080/00071005.2019.1589417

Prost, C. (2013). Teacher Mobility: Can Financial Incentives Help Disadvantaged Schools to Retain Their Teachers? *Annals of Economics and Statistics*, 111/112, 171–191. https://doi.org/10.2307/23646330

Robinson, V. & Gray, E. (2019). What difference does school leadership make to student outcomes?, *Journal of the Royal Society of New Zealand*, 49:2, 171-187, DOI: 10.1080/03036758.2019.1582075

Royal Society of Chemistry. (2023). *The Science Teaching Survey 2023: A summary of the findings*. Royal Society of Chemistry. Retrieved from https://www.rsc.org/policy-evidence-campaigns/chemistry-education/education-reports-surveys-campaigns/the-science-teaching-survey/2023/

Sabbaghan, S. (2024). Exploring the synergy of human and AI-driven approaches in thematic analysis for qualitative educational research. *Journal of Applied Learning and Teaching*, *7*(2).

Schmidt, J., & Bijmolt, T. H. (2020). Accurately measuring willingness to pay for consumer goods: a meta-analysis of the hypothetical bias. *Journal of the Academy of Marketing Science*, *48*, 499-518.

See, B. H., Morris, R., Gorard, S., Kokotsaki, D. & Abdi, S. (2020a) Teacher Recruitment and Retention: A Critical Review of International Evidence of Most Promising Interventions, *Education Sciences*, 10 (10), p. 262. https://doi.org/10.3390/educsci10100262

See, B., Morris, R., Gorard, S., & Soufi, N. (2020b). What works in attracting and retaining teachers in challenging schools and areas?. *Oxford Review of Education*, 46, 678 - 697. https://doi.org/10.1080/03054985.2020.1775566.

Sharp, C., Smith, R., Worth, J. & Van den Brande, J. (2019). *Part-Time Teaching and Flexible Working in Secondary Schools*. National Foundation for Educational Research: https://www.nfer.ac.uk/publications/part-time-teaching-and-flexible-working-in-secondary-schools/

Sibieta, L. (2020). *Teacher shortages in England: analysis and pay options*. https://epi.org.uk/wp-content/uploads/2020/02/Teacher-shortages-and-pay_2020_EPI.pdf

Sims, S. & Jerrim, J. (2020). *TALIS 2018: teacher working conditions, turnover and attrition*. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/873922/Teaching_and_Learning_International_Survey_2018_March_2020.pdf

Statham, R., Freedman, S., & Parkes, H. (2022). *Delivering a Childcare Guarantee*. https://www.ippr.org/files/2022-12/delivering-a-childcare-guarantee-dec-22.pdf

Teacher Tapp (2023a). Pay vs PPA, Enjoying *Work, & Pupil Premium Changes*. https://teachertapp.co.uk/articles/ppa-pay-enjoying-work-contentedness-pupil-premium/

Teacher Tapp (2023b). *Measuring underlying traits about teachers*, Blog post 21st November 2023. https://teachertapp.co.uk/articles/measuring-underlying-traits-about-teachers/

Teacher Tapp (2024a, November 26). *What kind of school would you really like to work in?* Teacher Tapp. https://teachertapp.com/uk/articles/what-kind-of-school-would-you-really-like-to-work-in/

Teachers' Pensions. (2024, February). *Updates to contribution rates*. Retrieved 28/02/2025, from https://www.teacherspensions.co.uk/news/employers/2024/02/updates-to-contribution-rates.aspx

Think Teaching. (2025). *Teaching assistant salary*. Think Teaching. Retrieved March 3, 2025, from https://www.thinkteaching.co.uk/careers-advice/teaching-assistant/salary/

Uehleke, R. (2016). The role of question format for the support for national climate change mitigation policies in Germany and the determinants of WTP. *Energy Economics*, 55, 148-156. https://doi.org/10.1016/J.ENECO.2015.12.028.

Union Insurance. (2025). *Health cash plan*. Union Insurance. Retrieved March 3, 2025, from https://www.unioninsurance.co.uk/health-cash-plan/

Viano, S., Pham, L. D., Henry, G. T., Kho, A., & Zimmer, R. (2021). What teachers want: School factors predicting teachers' decisions to work in low-performing schools. *American Educational Research Journal*, 58(1), 201-233.

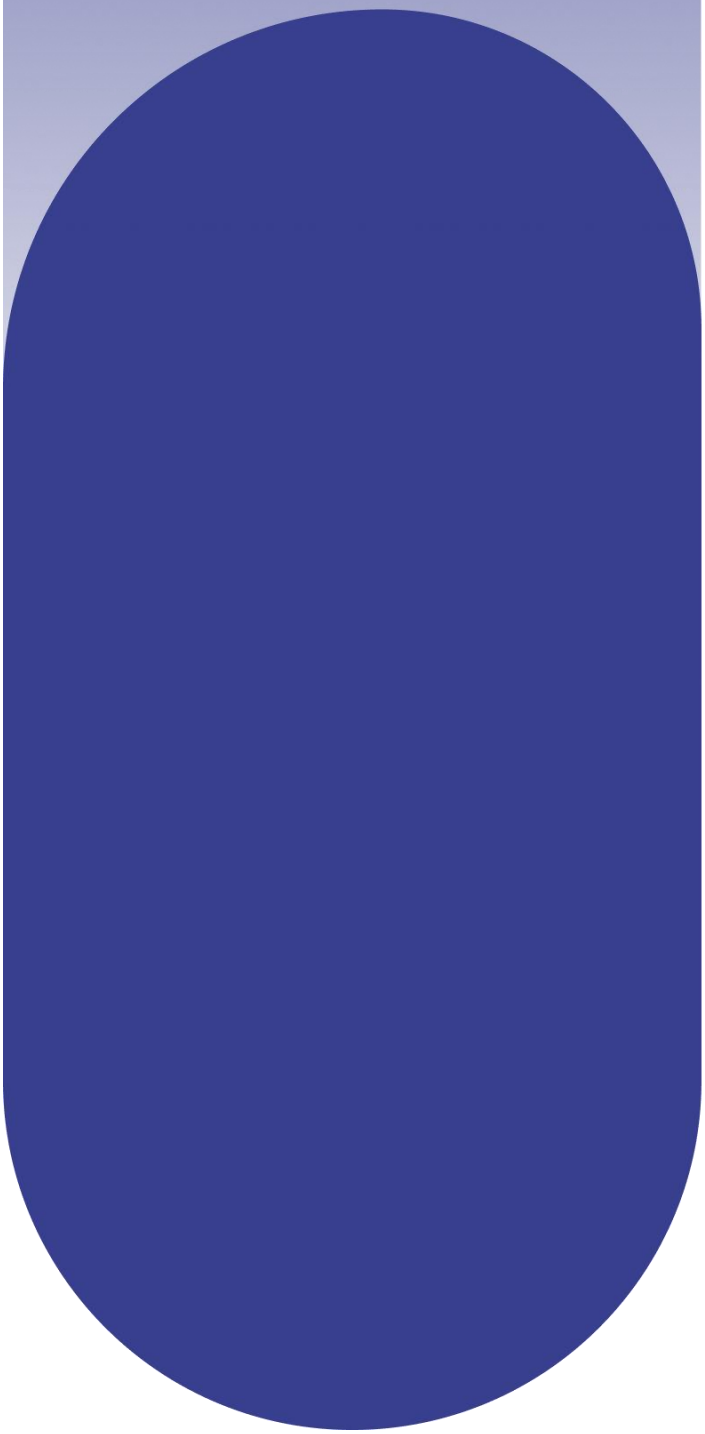
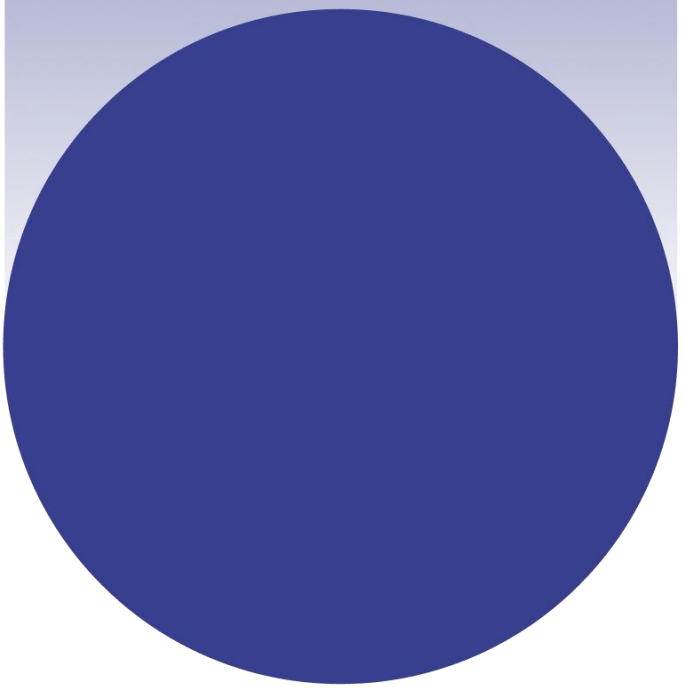
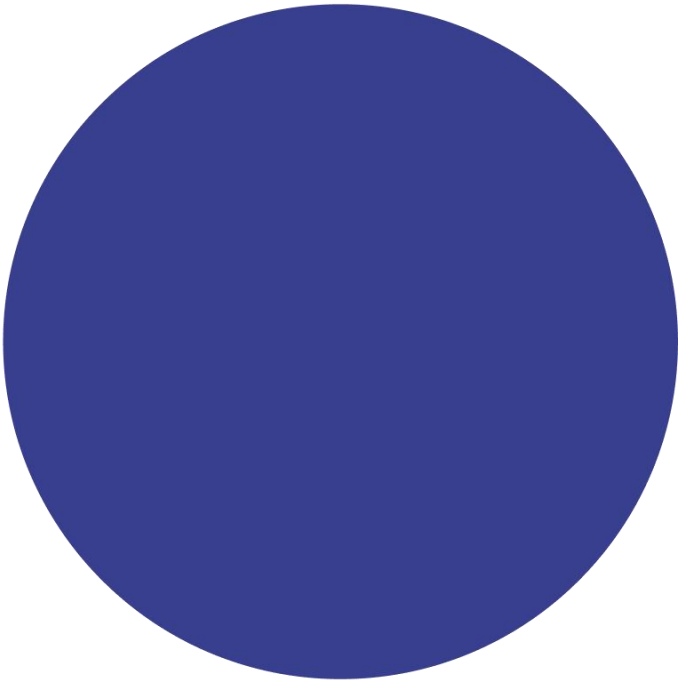
Winter, P.A., & Melloy, S.H. (2005). Teacher Recruitment in a School Reform State: Factors That Influence Applicant Attraction to Teaching Vacancies. *Educational Administration Quarterly*, 41, 349 - 372.

Worth, J., Lynch, S., Hillary, J., Rennie, C. & Andrade, J. (2018). *Teacher Workforce Dynamics in England*, National Foundation for Educational Research, https://www.nfer.ac.uk/publications/teacher-workforce-dynamics-in-england/.

Worth J. & Faulkner-Ellis, H. (2022). *Teacher supply and shortages: implications of teacher supply challenges for schools and pupils*. Slough: NFER.

Yan, L., Echeverria, V., Fernandez Nieto, G., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2024). Human-AI collaboration in thematic analysis using ChatGPT: A user study and design recommendations. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA. ACM. https://doi.org/10.1145/3613905.3650732

This document is available for download at https://educationendowmentfoundation.org.uk

Education Endowment Foundation

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk
@EducEndowFoundn
Facebook.com/EducEndowFoundn