RAISE YOUR WORDS PROGRAMME:

2024/2025 IMPACT EVALUATION REPORT

Attainment Raising Programmes

Network for East Anglian Collaborative Outreach (neaco)

Project description

Raise Your Words is a 9-week programme designed for Year 9 students. Participants were selected by their schools based on their receipt of Free School Meals (FSM) and/or being from underrepresented groups (UGR). Moreover, since the evaluation of this programme included a control group (see more in the Evaluation approach section), two more selection criteria were added: (i) same gender distribution in intervention and control groups, and (ii) same prior attainment across both groups. Regarding the latter, the target students for this programme were students who were at the risk of getting grades 3/4s. To ensure that the control and intervention groups were evenly matched academically, Year 8 English and Maths tests were used.[1] The programme was delivered across 9 schools in East Anglia, with groups of 13 to 15 students in each session. A total of 203 students participated – 102 in the intervention group and 101 in the control group –, with one school unable to attend the intervention final event at Cambridge University. We partnered with Speakers Trust to deliver oracy workshops at both the local and Cambridge campus events, and to adapt their resources for use in the in-school sessions. Speakers Trust are a non-for-profit organisation with expertise in working with young people to improve their speaking and communication skills and to give them the skills to become more confident communicators. The programme was delivered in partnership with our Higher Education Champions (HECs) based in schools in the East of England, and our Speakers Trust colleagues.

The main aim of the programme is to increase students' attainment levels by developing their effective spoken language and listening, and non-verbal communication skills, as well as enhancing their self-confidence, critical thinking, research, and teamwork abilities. On average, oracy skills development has a high impact on pupil outcomes, equivalent to 6 months' additional progress (EEF Toolkit). The oracy skills fostered through the programme are expected to benefit all subjects across the curriculum, which presents an opportunity for broader student development and potential improvements in GCSE results. Moreover, the programme helps students build a set of transferable skills, such as reasoning, collaborative problem-solving, comprehension and evidence-building, that support both improved attainment and future employability, including in interviews, meetings or work life more generally.

The programme consists of 10 sessions, with their respective delivery type, content and desired outcomes summarised in Table 1 below:

[1] Please note that the learners did not necessarily be graded 3/4 borderline for these subjects specifically.

Session	Delivery type	Focus/Content	Outcomes
0	In-school, HEC delivery	To introduce the students to the programme	• Completing of pre-programme survey and assessment • Setting down ground rules about the programme
1	In-school, HEC delivery	To explore the importance of oracy and public speaking skills by discussing what makes a good speech and its components	• Learning the fundamentals of making effective arguments • Learning the components of a (good) speech • Developing their ability to think critically and creatively
2	In-school, HEC delivery	To brainstorm and identify the topic and key message for the students' oral presentations	• Applying their ability to think critically and creatively
3	In-school, HEC delivery	To craft engaging content	• Learning about persuasive language and content • Developing an understanding of facts vs opinions • Learning about how to research information
4	In-school, HEC delivery	To explore the structure and organisation of speeches	• Building confidence in speaking

			• Creating clearly structured and supported arguments • Practising teamwork
5	Local university campus, Speakers Trust run	To receive a Public Speaking workshop in the morning and to apply and practice their developed oracy skills through activities in the afternoon	• Enhancing and applying the oracy skills and speech-delivery knowledge gained • IAG exposure (from campus setting)
6	In school, HEC delivery	To explore ways of expressing oneself while delivering a speech	• Understanding and finding their own 'style' • Building confidence in speaking • Practising listening and responding skills
7	In school, HEC delivery	To apply and practice what they learnt in lessons 1-6 in the classroom setting	• Applying the skills and knowledge gained • Demonstrating progress and improvement in confidence • Preparing for the final event • Giving and receiving feedback from peers • Practising listening and responding skills
8	Cambridge University campus,	To present their speeches at the final event and to receive feedback from Speakers Trust	• Applying and practise the oracy skills

	Speakers Trust run		acquired throughout the programme •Showcasing improvement in vocabulary, speaking confidence and self-efficacy • Practising listening and responding skills • IAG exposure (from campus setting and student ambassador activities)
9	In-school, HEC delivery	To conclude the programme	• Completing of post-programme survey and assessment

Table 1: Session outline of the Raise Your Words programme.

This session outline was completed by the intervention group. The control group only participated in Sessions 0 and 9, which involved completing the surveys and assessments, and separately from the intervention group. Once these sessions were completed, the control group was invited to a campus event at Cambridge University to take part in oracy workshops facilitated by Speakers Trust. This approach allowed the control group to benefit from the study by providing a smaller-scale oracy learning experience, ensuring their participation was still meaningful within the constraints of this year's study design (see more in the Recommendations section).

Evaluation approach

The programme was underpinned by a Theory of Change. All activity was logged on the Higher Education Access Tracker (HEAT) and made use of the HEAT Attainment Raising (AR) Typology to code activity. The evaluation focused on a controlled pre-and-post design, combined with an oracy assessment.

The pre-and-post surveys looked at student oracy skills (and how these affected the learners' confidence), cognitive strategies, academic self-efficacy and sense of belonging. Survey questions used were based on TASO's Access and Success Questionnaire (ASQ). Additionally, some open-ended qualitative questions were included to capture the learners' main takeaways from the project, allowing them to

reflect on their experiences more freely. The oracy assessment evaluated the students' physical, linguistic, cognitive, social and emotional oracy skills. This assessment was developed by Oracy Cambridge, based on the Oracy Skills Framework, and adapted by the Faculty of Education of the University of Cambridge. The pre-and-post surveys as well as the oracy assessment were completed by both the intervention and the control group, which allows for comparison between the two groups with regards to the effect of the intervention. This evaluation tracked changes in these specific skills and outcomes before and after the intervention, collected information on the learners' perceived impact of the project, and included an objective measure of students' oracy skills at both time points.

Pre- and post-programme surveys and assessments were conducted on 203 Year 9 students in 9 schools of East Anglia before and after the Raise Your Words programme took place. Each school contributed 13-15 students if it included only the intervention or the control group, or 26-30 students if it included both groups. Surveys were available in either electronic or paper format, with paper format being the preference – this helped mitigate issues around access to technology in the classroom and support a higher return rate. Assessments were carried out in groups of three students by the designated HEC at each school.

This amounts to an OfS Standards of Evidence Type 2 approach that generates empirical evidence but cannot provide an insight into the specific causal impact of the project. The inclusion of a control group strengthens the Type 2 evidence by allowing comparisons between groups. However, allocation was conducted by teachers, meaning that randomisation and control for confounding factors did not occur. Moreover, no calculation was made to ensure a sufficient sample size or statistical power to confidently detect meaningful differences between groups and reduce the chance of Type I/II errors. As a result, while the current design can show differences in outcomes, it cannot confidently attribute these differences to the intervention itself, as required for Type 3 causal evidence.

To analyse the impact of the intervention, a mixed-effects ordinal logistic regression (CLMM) was conducted on the pre- and post- survey responses. Fixed effects included Time (pre vs. post), Group (intervention vs. control), their interaction, and Gender (male vs. female). Random intercepts were included for School and Student to account for the nested structure of the data (i.e., students nested within schools and repeated measures over time), allowing for an assessment of both the overall intervention effect and variation across schools and students. In addition, pre- and post- assessment scores were analysed using a Linear Mixed-Effects Model (LMM) with the same fixed and random effects structure. Due to the moderate sample size, the conclusions drawn from these analyses should be interpreted with caution.

Results

Participants

A total of 203 students participated in this programme, split into 102 in the intervention group and 101 in the control group. Overall, 188 participants completed the pre-programme survey (92.6% response rate) and 152 completed the post-programme survey (75% response rate). For the oracy assessment, 188 students completed it at the pre-programme stage (92.6% completion rate) and 143 at the post-programme stage (70.4% completion rate). In total, 134 students completed both pre-and-post assessments as well as the pre-and-post surveys, resulting in an overall matched response rate of 66%.

One school was excluded from the analysis, as students from both the control and intervention groups were unable to attend the Cambridge visit event on Week 8 of the programme. As a result, 121 students (55 from the intervention group and 66 from the control group) were included in the final sample, resulting in a 59.6% final sample rate.

A breakdown of these figures for the intervention and control groups is provided below:

Intervention group (n = 102)

- 93 completed the pre-programme survey (91.2% response rate)
- 68 completed the post-programme survey (66.7% response rate)
- 96 completed the pre-programme oracy assessment (94.1% completion rate)
- 63 completed the post-programme oracy assessment (61.8% completion rate)
- 59 students completed all pre- and post-surveys and assessments (57.9% overall matched response rate for this group)

Control group (n = 101)

- 95 completed the pre-programme survey (94% response rate)
- 84 completed the post-programme survey (83.2% response rate)
- 92 completed the pre-programme oracy assessment (91% completion rate)
- 80 completed the post-programme oracy assessment (79.2% completion rate)
- 75 students completed all pre- and post-surveys and assessments (74.2% overall matched response rate for this group)

Findings and discussion

Overall impact

The figures below are constructed from the 121 matched pre-and-post assessments and surveys. Starting with data from the oracy assessment, the results of both intervention and control groups lead to one of the main key findings of the programme:

KEY FINDING 1: Intervention group learners' oracy skills were significantly improved after taking part in the Raise Your Words programme, compared to the control group.

At the start of the programme, assessment scores were similar between both groups. However, while the control group did not show an improvement in scores over time, the intervention group demonstrated a positive progress from pre- to post- assessment. Statistical modelling (LMM) confirmed that this improvement was significant compared to the control group ($p^2 < 0.001$), indicating that the programme had a positive impact on learners' oracy skills. This positive effect is illustrated in Figure 1.

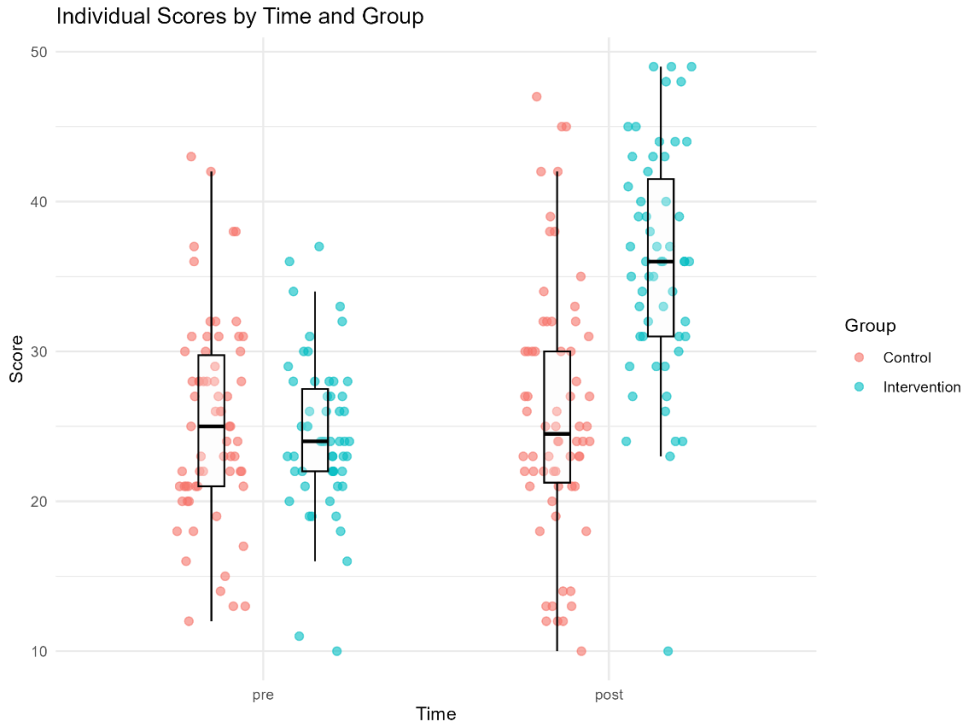

Figure 1: Individual assessment scores by time point and group.

[2] Statistical testing generates a p-value that tells us the probability of getting these results if there was no actual difference between the groups. The threshold for statistical significance is typically $p < 0.05$, which has also been adopted in this report. This means that when a result is called significant, there is a less than 5% probability that it could have happened by chance.

In addition to the oracy assessment to test the progress in the participants' oracy skills, the pre-and-post surveys that learners completed included questions on the different aims and intended outcomes of the programme. The analysis of the responses of both groups led to a second key finding:

KEY FINDING 2: Students in the intervention group showed significantly more positive changes in their self-reported perspectives on cognitive and oracy skills, and sense of belonging after taking part in the Raise Your Words programme, compared to the control group.

When analysing all the blocks of questions, participants in the intervention group reported more positive responses across all outcome categories compared to the control group, as indicated by odds ratios (estimates) greater than 1 – see Table 2. However, only the differences in cognitive and oracy skills, and sense of belonging were statistically significant (at $p < 0.05$), suggesting reliable improvements in these specific outcomes.

Outcome category	Estimate (OR)	p-value	Interpretation
Cognitive skills	1.09	0.003	Significantly more positive in intervention
Listening skills	0.45	0.25	Not statistically significant
Speaking skills	2.21	< 0.001	Significantly more positive in intervention
Presentation skills	2.45	< 0.001	Significantly more positive in intervention
SE (post-16)	0.68	0.201	Not statistically significant
SE (HE)	0.50	0.342	Not statistically significant
HE expectation	1.33	0.102	Not statistically significant
HE knowledge	0.80	0.06	Not statistically significant
SoB	1.47	0.009	Significantly more positive in intervention

Table 2: Between-group comparisons of survey outcome categories.

This aligns with existing literature suggesting that programmes are most effective when their learning objectives are made explicit (Aubin, 2023). In the case of Raise Your Words,

cognitive, speaking and presentation skills are overtly embedded in the programme design. These elements are clearly signposted, which likely made it more evident to students that they were expected to develop in these areas, potentially explaining the confidently perceived improvements. Similarly, the campus visits and the involvement of student ambassadors were salient elements of the programme and may have contributed to the positive impact observed in students' sense of belonging.

In contrast, listening skills, while integral to the programme (e.g., when students attended their peers' speeches and provided feedback (see Table 1)) were not explicitly taught with a standalone focus. This lack of direct instruction may account for the comparatively smaller reported gains in this area. Likewise, although students visited university campuses on two occasions, they were not directly provided with structured information, advice, and guidance (IAG), which may explain the more modest changes in Higher Education (HE) expectation and knowledge. Finally, changes in self-efficacy (e.g., grade improvements) are a longer-term outcome than changes in skills or knowledge. As such, it is not surprising that students self-reported fewer perceived improvements in this area (see also Figure 10 in Differences in impact by programme group section).

Figures 2-10 below show the distribution of learners' responses by programme group across outcome categories, comparing intervention and control groups. These visualisations reflect the statistical findings and discussion presented above.

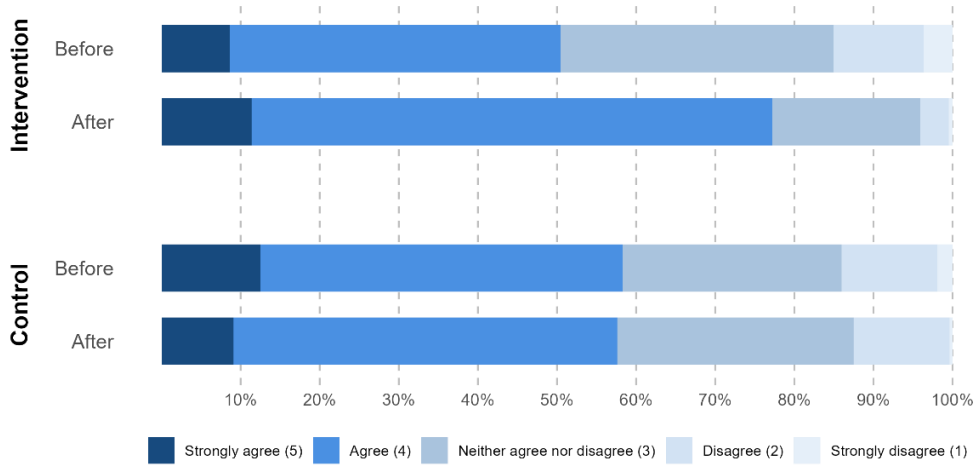

Figure 2: Pre- and post-survey results for cognitive skills.

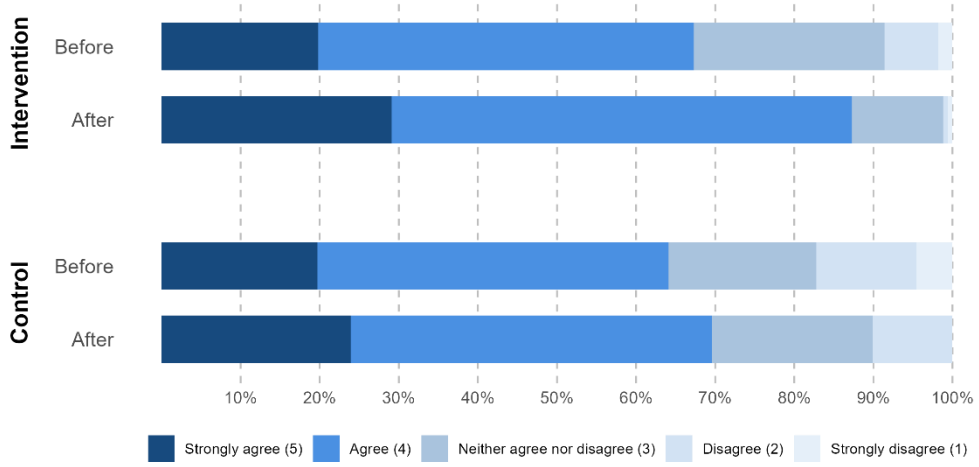

Figure 3: Pre- and post-survey results for listening skills.

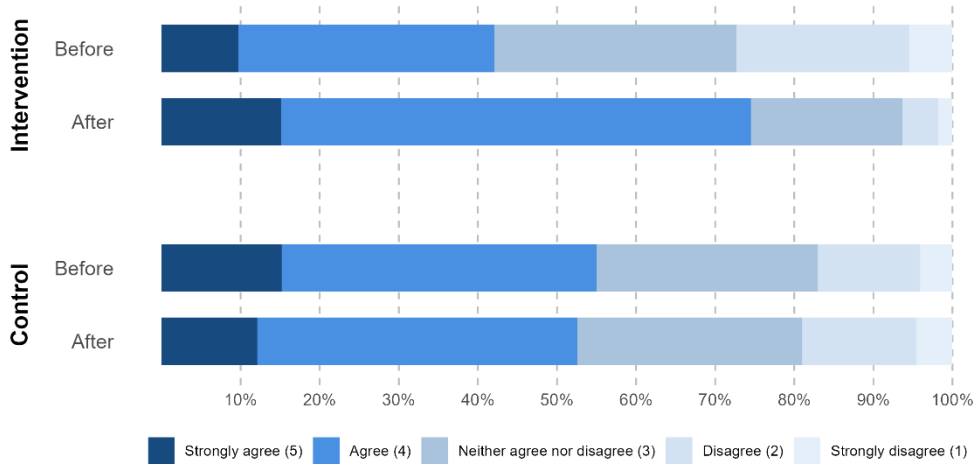

Figure 4: Pre- and post-survey results for speaking skills.

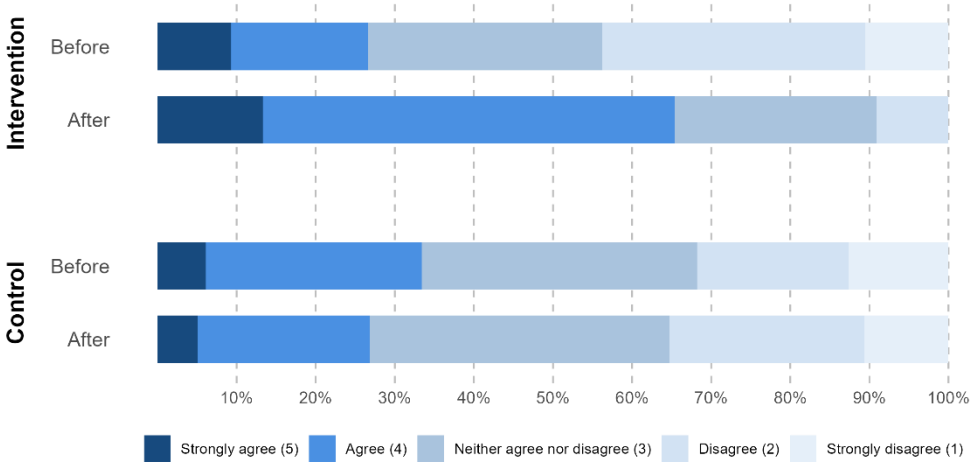

Figure 5: Pre- and post-survey results for presentation skills.

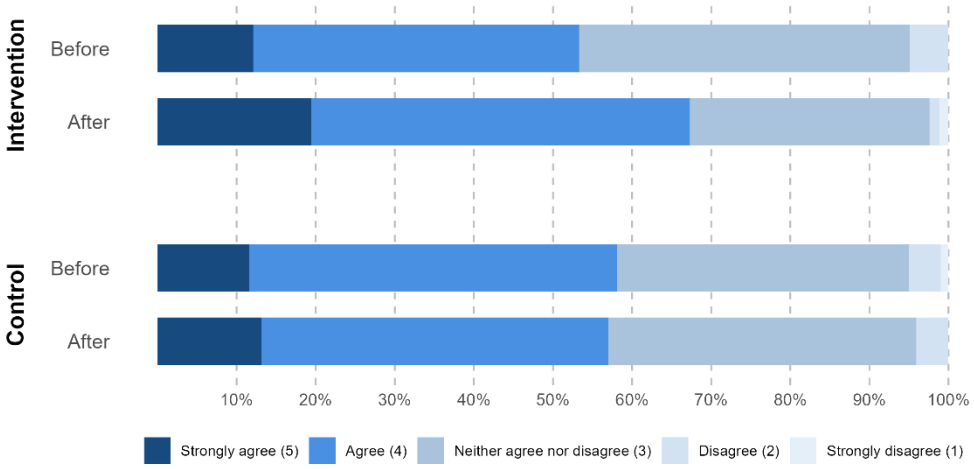

Figure 6: Pre- and post-survey results for self-efficacy (post-16).

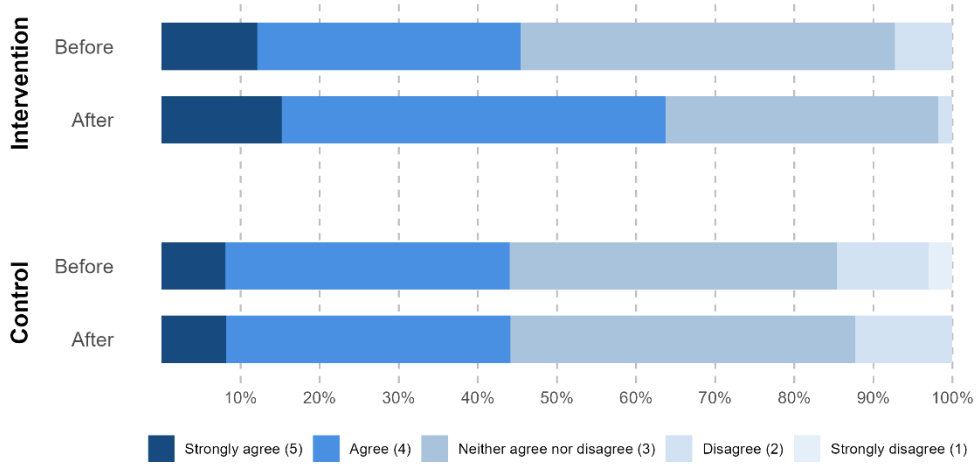

Figure 7: Pre- and post-survey results for self-efficacy (HE).

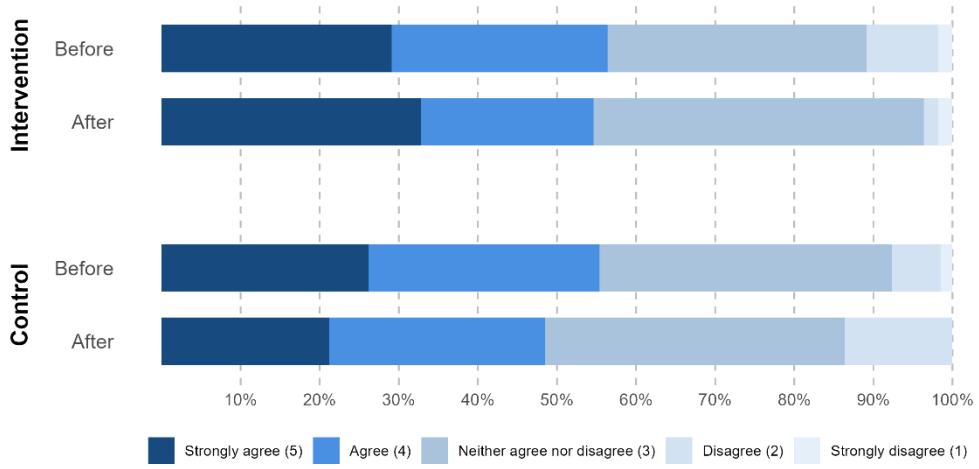

Figure 8: Pre- and post-survey results for HE expectation.

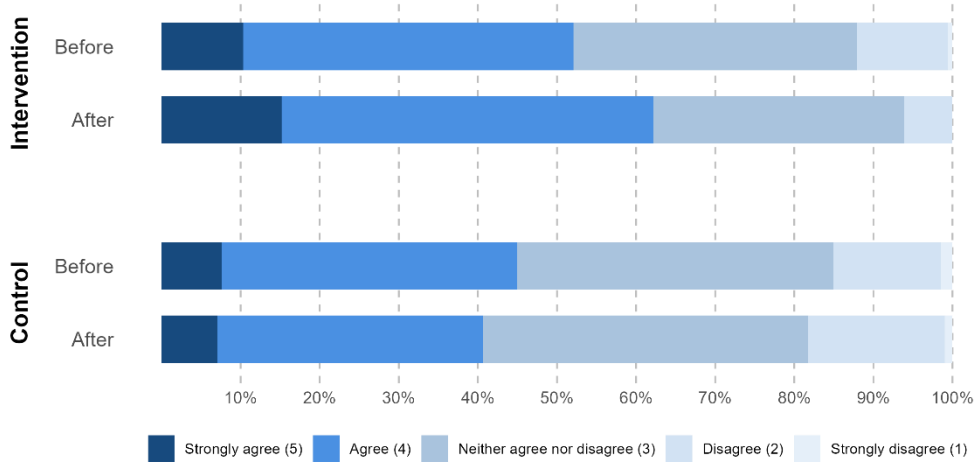

Figure 9: Pre- and post-survey results for HE knowledge.

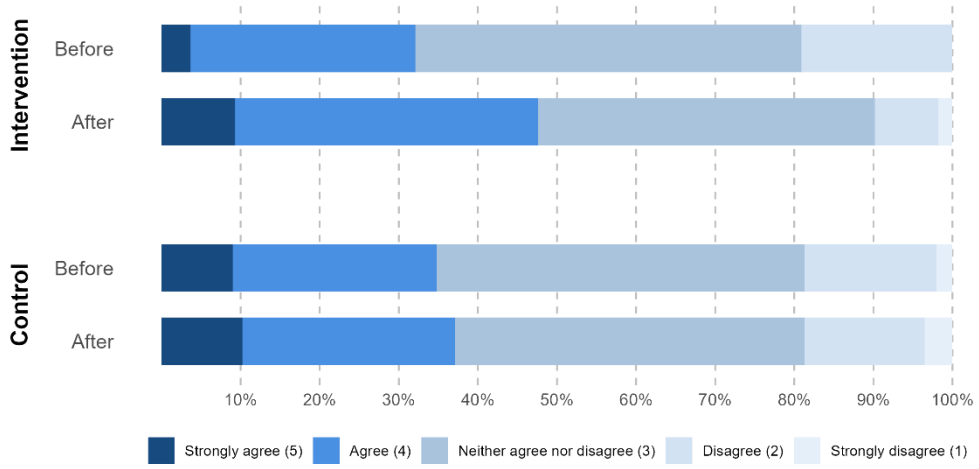

Figure 10: Pre- and post-survey results for sense of belonging.

Differences in impact by programme group

The models used for the overall impact analysis in the previous section examined whether the intervention led to greater changes over time compared to the control group (i.e., differences in improvement between groups). In order to better understand the patterns of chance more precisely, post hoc analyses were conducted to explore changes within each group separately to examine how participants' responses evolved from pre to post within each programme group.

Control group

No meaningful changes were observed in the control group across either assessment scores or survey responses. As such, these results are not discussed in the main body of the report. However, full results, including breakdowns by outcome block and individual questions, as well as assessment data, are provided in Appendix A for reference.

Intervention group

Starting with the **assessment** data, the post hoc within-group analysis revealed that not only did the intervention group scores significantly increased when compared to the control group (see Figure 1 above), but they also showed a significant improvement over time within the group itself ($p < 0.0001$), indicating a meaningful impact of participating in the programme on the learners' oracy skills.

Moving to the **survey** data, while the between-group comparisons indicated that certain outcome categories showed significantly more positive responses in the intervention group compared to the control group, further within-group analyses offer additional insights. These analyses, summarised in Table 3 below, revealed statistically significant improvements from pre- to post-programme across all outcome categories, including those that did not reach significance in the between-group comparisons (see Table 2 above).

Outcome category	Estimate (OR)	p-value	Statistical significance
Cognitive skills	1.27	< 0.001	Significant
Listening skills	0.99	< 0.001	Significant
Speaking skills	1.63	< 0.001	Significant
Presentation skills	2.06	< 0.001	Significant
SE (post-16)	0.90	< 0.001	Significant
SE (HE)	1.09	< 0.001	Significant
HE expectation	0.17	0.63	Not significant
HE knowledge	0.62	0.004	Significant
SoB	1.07	< 0.001	Significant

Table 3: Within-group survey outcome categories in intervention group.

In turn, this pattern suggests that although all outcomes improved within the intervention group, significant between-group differences were primarily observed in areas where the intervention group's improvements were more pronounced, likely due to the explicit nature of the content, as previously discussed. For outcomes such as HE knowledge or self-efficacy, the absence of statistically significant differences between groups may be attributed to greater variability in the responses, or to relatively small effect sizes to detect clear changes. Taken together, these findings also suggest that the intervention may have supported progress across all measured areas, but had a more confident and reliable impact on outcomes that were most explicitly embedded in the programme design and events.

Building on the block-level within-group analyses, which demonstrated statistically significant improvements from pre- to post-programme across almost all outcome categories, post hoc analyses were also conducted at the level of individual questions within each block. This more granular approach allowed for the identification of specific items driving the overall improvements, as well as those where changes were less pronounced. By disentangling these item-level effects, a clearer understanding emerged of which aspects of each outcome category were most responsive to the programme and which may warrant further attention in future iterations (see Recommendations section).

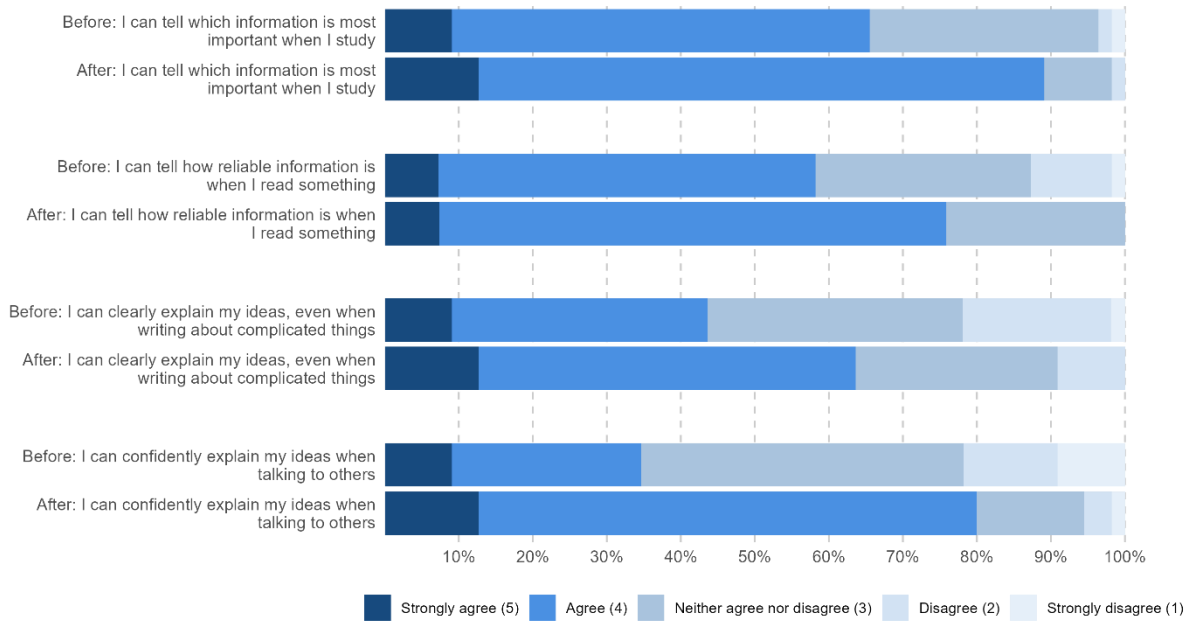

Figure 11: Intervention group pre- and post-survey results for Cognitive skills questions. Post-hoc CLMM pairwise comparisons showed significant increases for all questions ($p = 0.0073$, 0.02, 0.004, < 0.001, respectively).

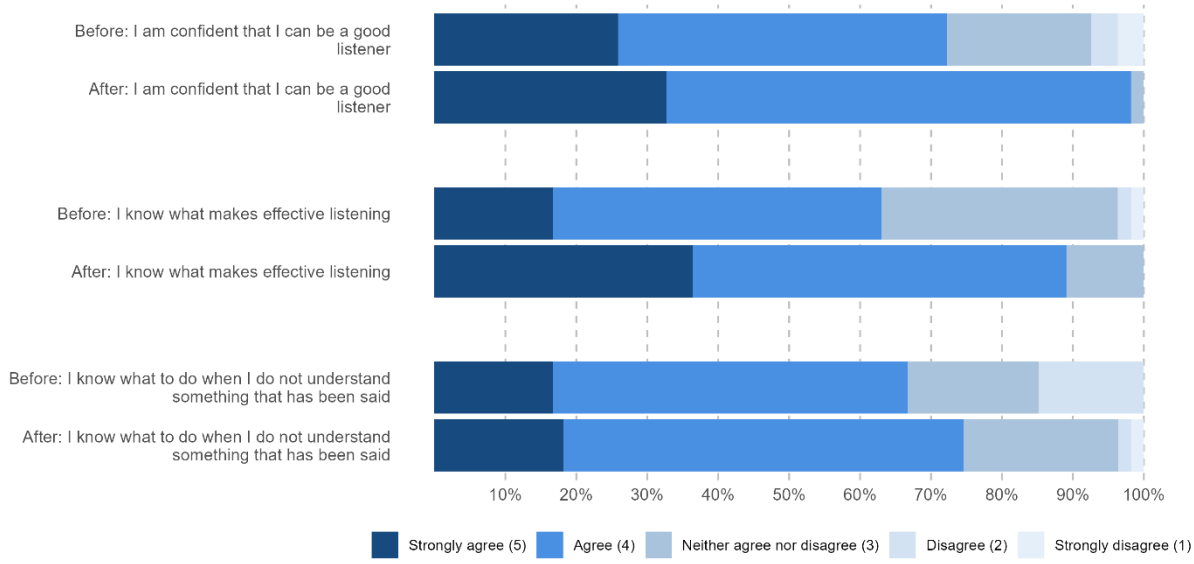

Figure 12: Intervention group pre- and post-survey results for Listening skills questions. Post-hoc CLMM pairwise comparisons showed significant increases for the first and second questions ($p = 0.005$, < 0.001, respectively), while no significant change was observed for the third question ($p = 0.267$).

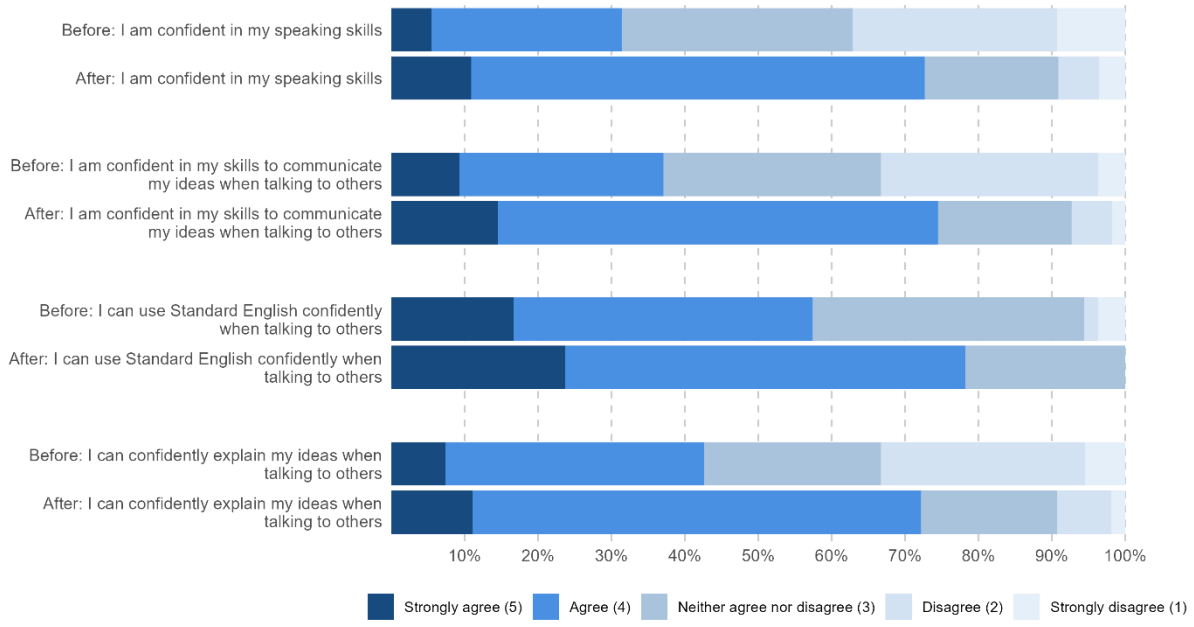

Figure 13: Intervention group pre- and post-survey results for Speaking skills questions. Post-hoc CLMM pairwise comparisons showed significant increases for all questions ($p < 0.001$, < 0.001, 0.005, < 0.001, respectively).

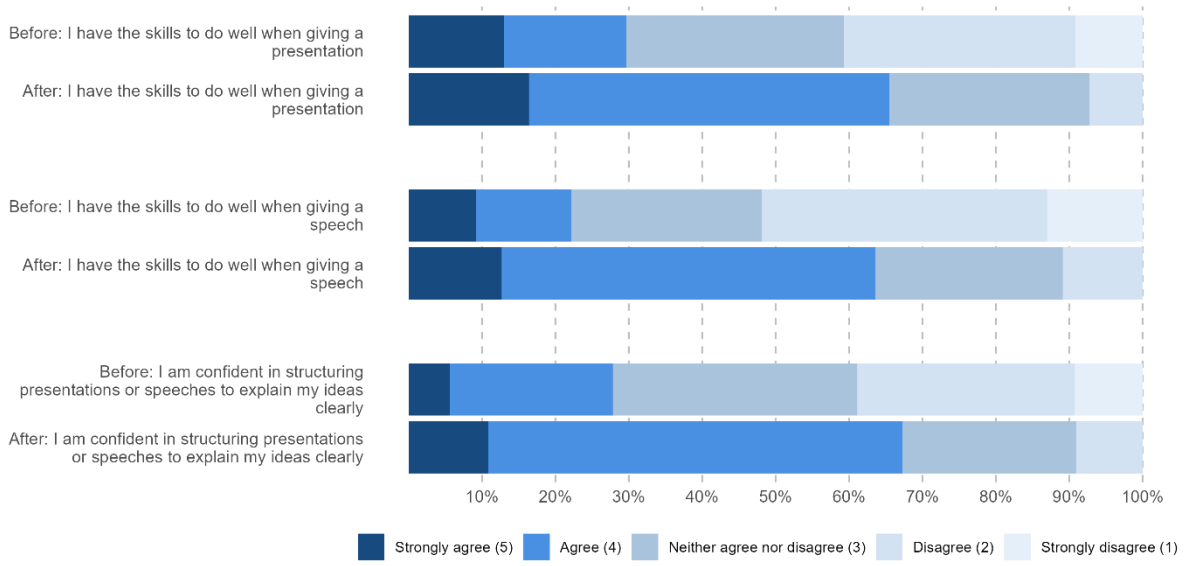

Figure 14: Intervention group pre- and post-survey results for Presentation skills questions. Post-hoc CLMM pairwise comparisons showed significant increases for all questions ($p < 0.001$ for all).

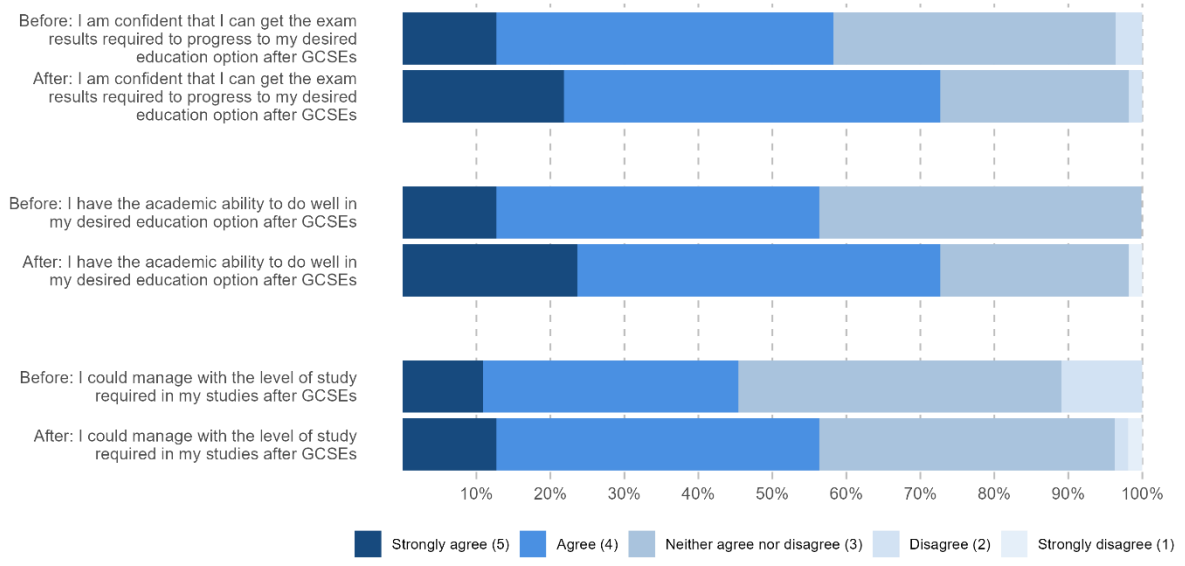

Figure 15: Intervention group pre- and post-survey results for SE (post-16) questions. Post-hoc CLMM pairwise comparisons showed significant increases for the first and second questions ($p = 0.008$, 0.01, respectively), while no significant change was observed for the third question ($p = 0.108$).

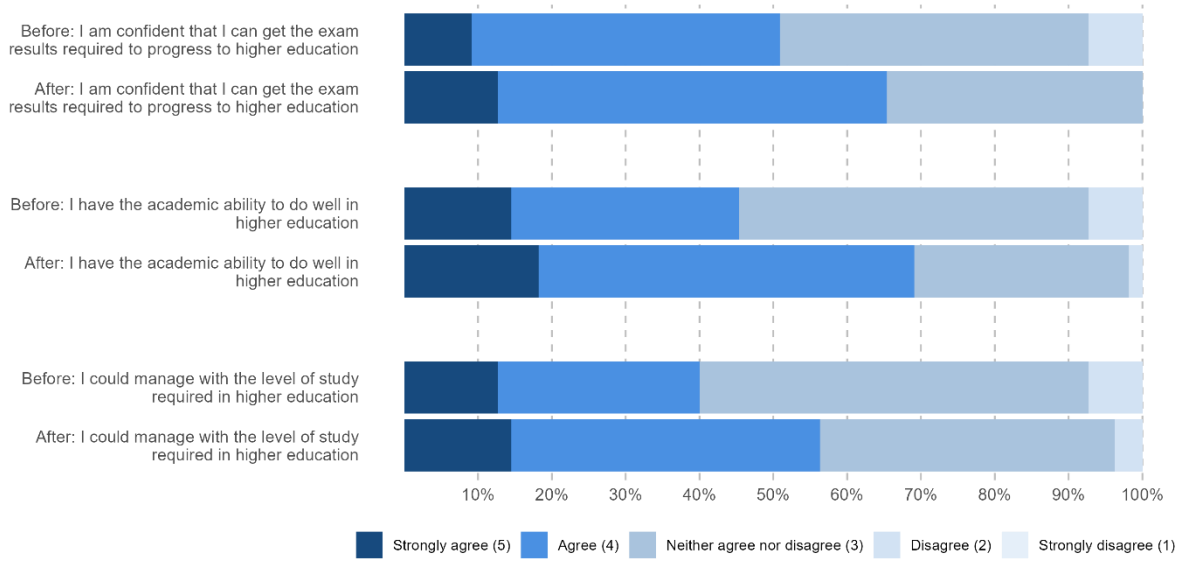

Figure 16: Intervention group pre- and post-survey results for SE (HE) questions. Post-hoc CLMM pairwise comparisons showed significant increases for all questions (p = 0.01, < 0.001, 0.04, respectively).

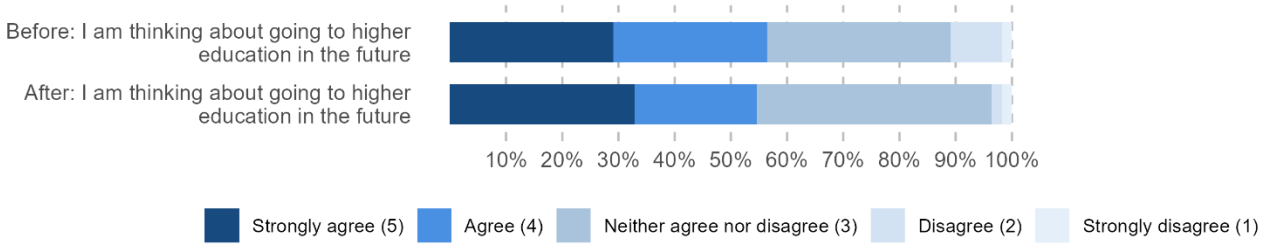

Figure 17: Intervention group pre- and post-survey results for HE expectation question. Post-hoc CLMM comparisons revealed no significant difference between the pre- and post-survey results (p = 0.638).

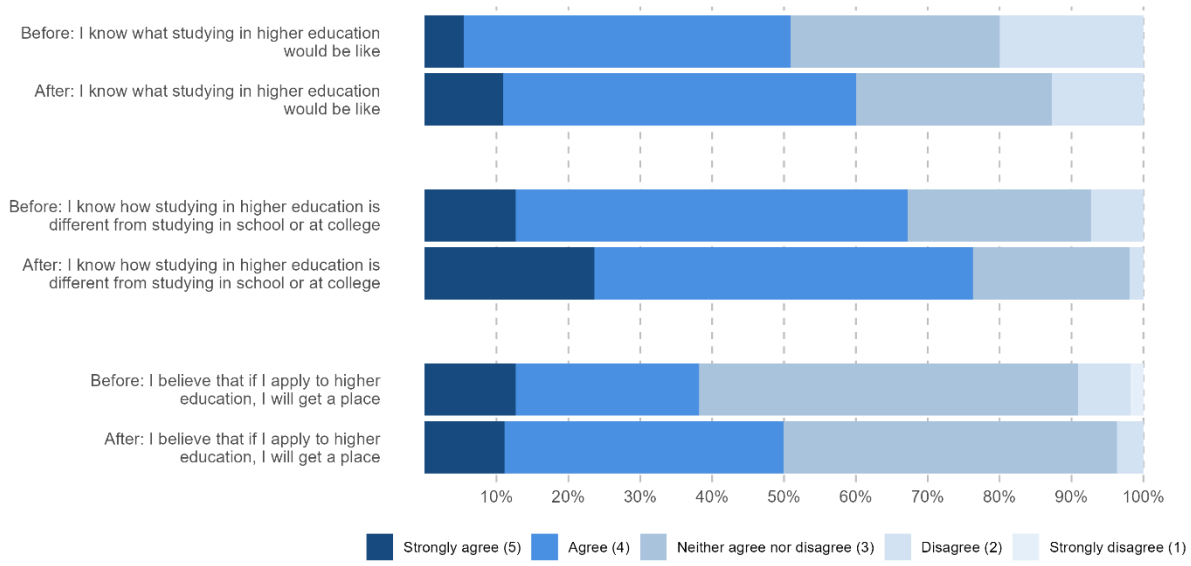

Figure 18: Intervention group pre- and post-survey results for HE knowledge questions. Post-hoc CLMM pairwise comparisons showed a significant increase for the second question (p = 0.02), while no significant change was observed for the first and third questions (p = 0.099, 0.171, respectively).

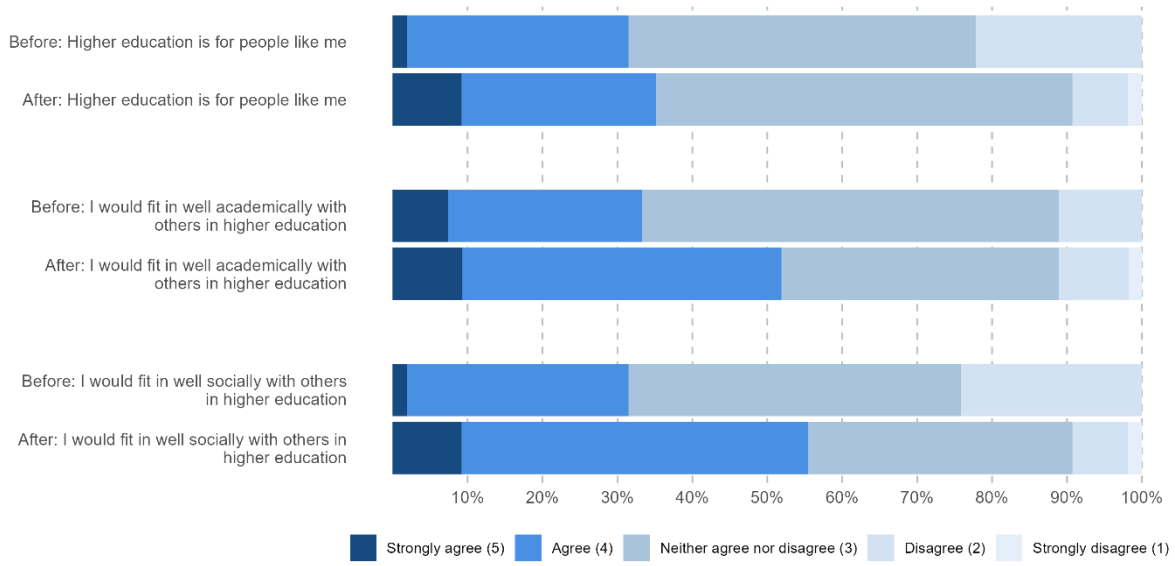

Figure 19: Intervention group pre- and post-survey results for Sense of belonging questions. Post-hoc CLMM pairwise comparisons showed significant increases for the first and third questions (p = 0.02, < 0.001, respectively), while no significant change was observed for the second question (p = 0.06).

The changes observed in the individual question graphs align with the findings reported in the between-group and block-level, within-group analyses. However, the post-hoc analysis of the Sense of belonging questions revealed a notable nuance in the significant development of intervention learners' sense of belonging: the improvement appears to be primarily driven by feelings of seeing HE as a place with people like them (first question) and, most importantly, by a sense that they would fit well *socially* in HE. No significant changes were observed regarding academic fit in HE (see Figure 19 above). This observation is further supported by student feedback, in which learners emphasised campus visits and meeting people there as their key takeaways from the programme (see Table 4 and its discussion below). Taken together, all this contributes to a third key finding:

KEY FINDING 3: Intervention learners reported a significant development and improvement of their cognitive and oracy skills, as well as of sense of belonging (particularly a *social* sense of belonging) after participating in the Raise Your Words programme.

The strong positive outcomes identified in both between-group and within-group analyses regarding students' oracy skills are also further reinforced by qualitative data on the programme's perceived impact, which provide additional insights into the quantitative observations:

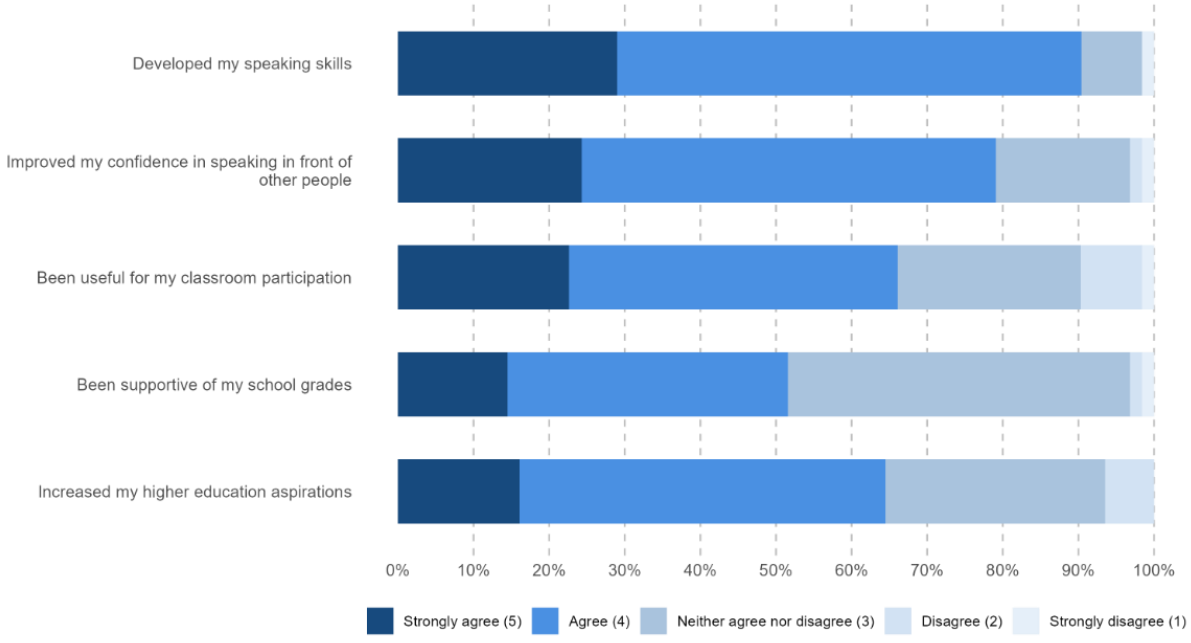

Figure 10: Perceived impact of the Raise Your Words programme.

In line with the quantitative analyses, when students were directly asked about their perceived impact of the programme on their skills, 90.4% of respondents agreed or strongly agreed that the Raise Your Words programme had developed their skills, and

79.1% reported improved confidence in speaking in front of others. This increase in confidence likely contributed to the 66.1% of students reporting that the programme had been helpful for their classroom participation. A similar proportion (66.4%) reported an increase in higher education aspirations. This might appear contradictory given that HE expectation was the only outcome that did not show a statistically significant positive change within the intervention group (see Table 3). However, the positive estimate indicates that there was still an improvement from pre- to post-programme, which aligns with the majority of positive responses in the question on perceived impact. Two possible explanations for the lack of statistical significance might be a smaller effect size despite a positive direction of change, or external factors beyond the programme. Lastly, when asked about the impact of the programme on their school grades, just over 51% of students agreed or strongly agreed – the lowest proportion across the perceived impacts. Despite significant shifts in academic self-efficacy when looking within intervention group, this outcome category was not statistically significant in comparison with the control group. This combination of results might be explained by the fact that noticeable changes in grades are typically a longer-term outcome, and their connection with the skills practised in the programme might not be as obvious for the students, hence leading to a lower perceived link between the programme and academic performance. Future iterations would benefit from incorporating perspectives from the participants' teachers, who may have a broader understanding of students' academic progress and the programme's influence, as well as including longer-term outcome measures such as academic grades to better capture sustained impact (see more in the Recommendations section).

These findings are complemented by the qualitative data, where learners were asked open-ended questions regarding the key takeaways and most significant aspects they gained from the programme, compiled below:

Most useful part of the programme	Percentage of responses[3]
Improved confidence and speaking skills	54.5%
Learnt new, useful resources	38.1%
Trips to campus	30.9%
Developed an interest in speeches and public speaking	16.3%
Met new people	5.5%

Table 4. Summary of topics raised in the intervention learners' responses to open-ended questions.

As shown in Table 4, 54.5% of students identified improvements in confidence and speaking skills as the most useful part of the programme, while 38.1% highlighted learning new, useful resources. Additionally, 30.9% valued the trips to campus, 16.3% developed an interest in speeches and public speaking, and 5.5% noted meeting new people. These responses align with the positive shifts in confidence and oracy skills observed in the quantitative results and offer additional nuance regarding potential unintended outcomes of the programme. These include developing participants' interest in activities where they can apply the acquired skills (e.g., public speaking settings) and fostering a social element through campus trips with other participating schools. The former, in turn, might also serve as a link to the Year 10 oracy programme in the neaco AR offer and its aims of supporting schools and students to engage in events where students apply their oracy skills (e.g., debate competitions or setting up a debate club at their school). Overall, these qualitative reflections emphasise the multifaceted nature of students' experiences and highlight the value of integrating qualitative data alongside quantitative findings to capture the full scope of programme effects – for further discussion, see the Recommendations section.

[3] Please note that 62 out of 66 intervention students decided to answer the open-ended questions.

Lastly, in addition to the qualitative responses above, several students provided brief testimonials reflecting on their participation in the programme. A selection of these, from students across different participating schools and counties, is presented below:

> ***"I found this really fun! It improved my speaking skills and taught me a lot about different universities and how they work."*** – Student at Sewell Park Academy
>
> ***"[My biggest takeaway from the sessions was] to not feel scared of talking in public and how to calm down, to be confident in who I am."*** – Student at Stanground Academy
>
> ***"Thank you so much for organising such an amazing course, I got a big confidence boost and it really helped me to communicate. Would love to participate again."*** – Student at Jane Austen College

Differences in impact by gender

In the Raise Your Words programme, there was an even split between male and female participants in both the intervention and control groups at baseline. In the matched sample used for the overall analysis, the intervention group comprised 33 females and 22 males, and the control group comprised 35 females and 31 males. Assessment scores and survey responses were compared between genders and groups to examine potential differences in programme outcomes.

Regarding **assessment** scores, post hoc pairwise comparisons of estimated marginal means were performed to assess score changes and differences in improvement by gender within each group. Figures 20 and 21 show the pre-post changes by gender within control and intervention groups, respectively:

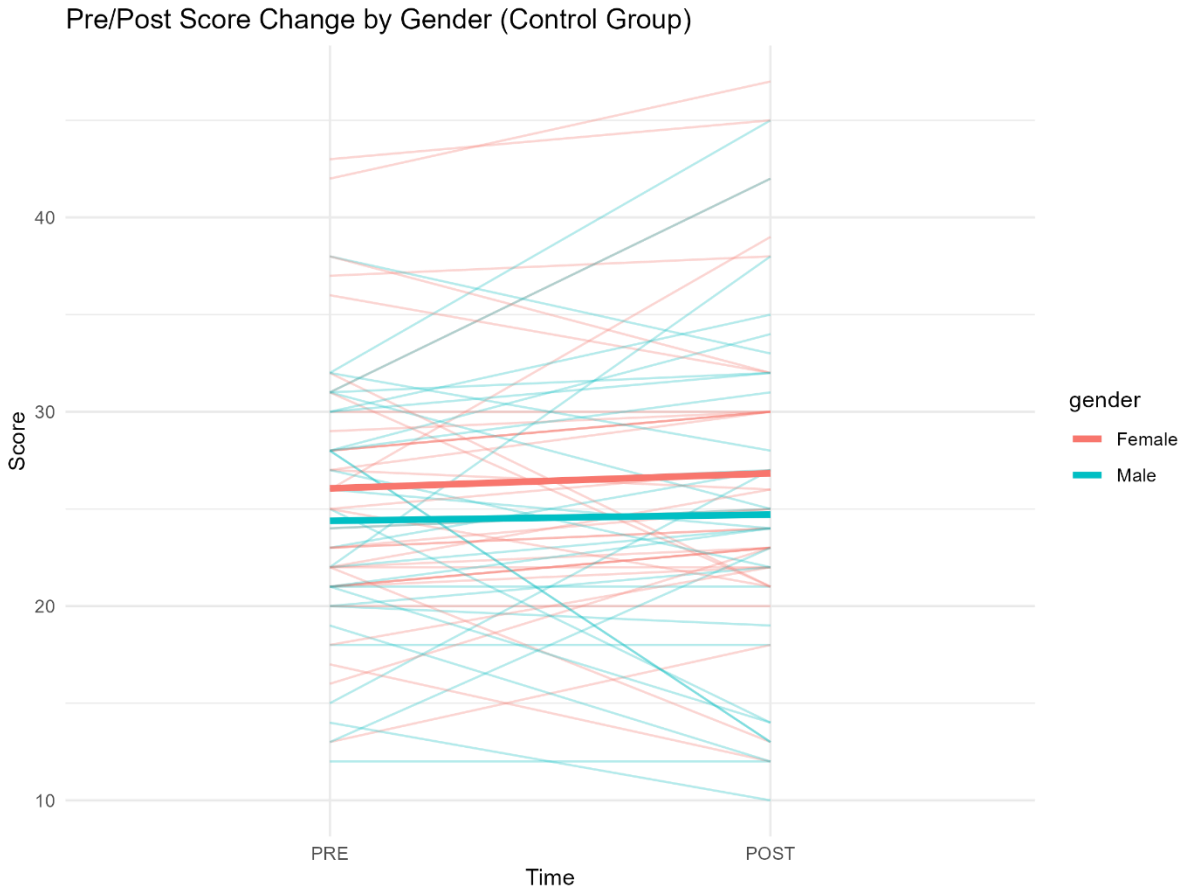

Figure 20. Pre- to post-assessment score change by gender in control group.

In the **control group**, there was no significant change in scores from pre- to post-assessment for either females (p = 0.53) or males (p = 0.81). The difference in score change between females and males was also not significant (p = 0.80), indicating no gender effect in the absence of intervention. However, it is interesting to note that females in the control group started at a (non-significantly) higher baseline score than males, similar to the pattern observed in the intervention group (see below).

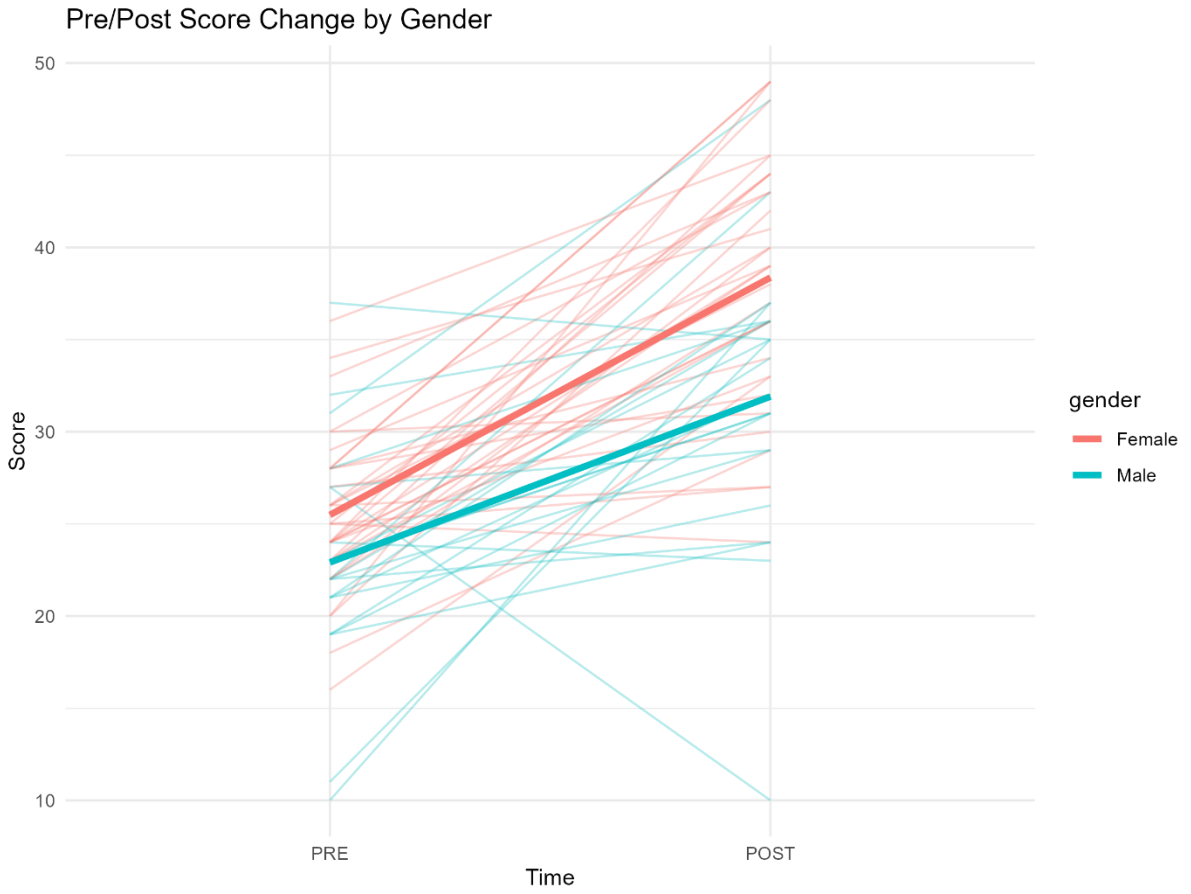

Figure 21. Pre- to post-assessment score change by gender in intervention group.

In the **intervention group**, both females ($p < 0.001$) and males ($p < 0.001$) showed significant score improvements from pre- to post-assessment. The difference in improvement between females and males was marginally non-significant ($p = 0.056$), suggesting a possible trend toward greater improvement in females. However, this trend should be interpreted cautiously, as the relatively small and unequal sample sizes between females and males in the matched intervention responses may limit the statistical power to detect significant gender differences.

With regards to the **survey** data, post hoc pairwise comparisons of estimated marginal means were conducted after the CLMM to explore gender differences within each group. In the **control group**, as discussed in the Differences in impact by programme group section, there was no significant change in survey scores from pre- to post-programme ($p = 0.75$). Neither females ($p = 0.75$) nor males ($p = 0.75$) showed significant improvements. The difference in change between females and males was also not statistically significant ($p = 0.65$), indicating no meaningful change for either gender.

In the **intervention group**, to reiterate, survey scores increased significantly from pre- to post-programme ($p < 0.001$). Both females ($p < 0.001$) and males ($p = 0.0003$) showed significant improvements, with females showing a slightly greater positive change in their pre- to post-programme responses compared to males. This difference in improvement

was not statistically significant (p = 0.25), suggesting that the positive change in survey responses was similar for both genders overall.

Putting these analyses together, they point to the fourth key finding, highlighting an area of further exploration in future iterations of the programme:

KEY FINDING 4: Both female and male intervention learners showed similar significant improvements, although there might be a tentative (currently non-significant) trend toward slightly greater improvement among females.

Recommendations

1. **Strengthen the evaluation by moving towards Type 3 evidence.** Following the positive impact observed in this report, future iterations of the programme should adopt methodologies that enable causal inference, such as quasi-experimental designs or randomised control trials (RCTs). To achieve the latter, it would be key to follow a process of randomisation to allocate students into intervention and control groups and to run calculations to ensure appropriate sample size and statistical power.

2. **Incorporate more qualitative components.** The present results illustrate the nuances that qualitative insights can bring to the discussion of findings, particularly when learners bring up topics and unintended outcomes that might not be targeted by the set quantitative questions. Future evaluations could also explore qualitative methodologies beyond the written medium, open-questions, such as interviews or focus groups, to ensure richer insights into students' experiences during and after the programme.

3. **Consider and incorporate teacher feedback and school grades collection.** Beyond students' self-reports, future evaluations should gather data from teachers on observable changes in classroom participation and performance. Similarly, academic grades and longer-term school outcomes could be integrated into the evaluation framework to provide a more objective measure of impact and to capture the longer-term, ultimate goals of the programme.

4. **Enhance gender analysis.** The current evaluation did not find statistically significant gender differences in outcomes, but a tentative trend pointing towards greater improvements in females was observed. Future evaluations should strengthen the gender analysis to explore gender patterns more systematically.

This could involve larger sample sizes or a more detailed breakdown of the results per question to explore whether differences in gender improvements depend on the question and/or block of outcomes.

5. **Consider and incorporate more IAG elements.** While findings show some positive impact on higher education aspirations, this remained one of the lowest perceived benefits. Future iterations of the programme should embed a clearer, more structured IAG strand. This could include workshops or talks during campus events, mentoring opportunities with student ambassadors or inclusion of an IAG component within the delivery or continuation of the programme.

6. **Retain and expand campus visits.** Student and HEC feedback consistently the value of campus events, which appear to play a key role in fostering a social sense of belonging around HE. These visits should therefore remain a central feature of the programme. Where possible, and in line with Recommendation number 5, they could also be expanded to include more IAG elements to maximise their impact on students' HE expectations, knowledge and aspirations.

7. **Adjust session content based on student and HEC feedback.** Both learners and HECs emphasised the value of having more time to prepare and practise their speeches, while also suggesting a reduction in time allocated to online and video materials. Future iterations of the programme should prioritise interactive, skill-building activities, which were consistently valued as the most beneficial by students and delivery staff.

References

Aubin, G. (2023). *What exactly is explicit instruction?* Education Endowment Foundation Blog. EEF blog: What exactly is explicit instruction? | EEF

The Education Endowment Foundation (2025). *Oral language interventions*. Oral language interventions | EEF

Appendix A: Control group data

Within-group assessments

The post hoc within-group analysis revealed that the control group scores did not show a significant change over time (p = 0.545), indicating that, in the absence of the programme, learners' oracy skills remained largely unchanged.

Within-group surveys

Results by outcome block

The within-group analyses for the control group, summarised in Table A.1 below, revealed that most outcome categories did not show statistically significant changes from pre- to post-programme in the control group. The only exception was listening skills, which showed a modest but significant improvement (p = 0.02), likely attributable to external factors. These findings indicate that, without participation in the programme, learners' responses largely remained stable across cognitive, speaking, presentation, self-efficacy, HE knowledge, expectations, and sense of belonging measures.

Outcome category	Estimate (OR)	p-value	Statistical significance
Cognitive skills	-0.06	0.70	Not significant
Listening skills	0.43	0.02	Significant
Speaking skills	-0.18	0.28	Not significant
Presentation skills	-0.34	0.08	Not significant
SE (post-16)	0.07	0.73	Not significant
SE (HE)	0.16	0.41	Not significant
HE expectation	-0.32	0.32	Not significant
HE knowledge	-0.15	0.42	Not significant
SoB	0.04	0.81	Not significant

Table A.1: Within-group survey outcome categories in control group.

Results for individual questions by outcome block

Building on the block-level within-group analyses, which showed little or no significant change from pre- to post-programme across most outcome categories in the control group, post hoc analyses were also conducted at the level of individual questions within each block. This allowed for a more granular view of the (lack of) changes per question in the control group in the absence of the programme intervention. A visual summary of these analyses is offered in Figures A.1-A.9 below.

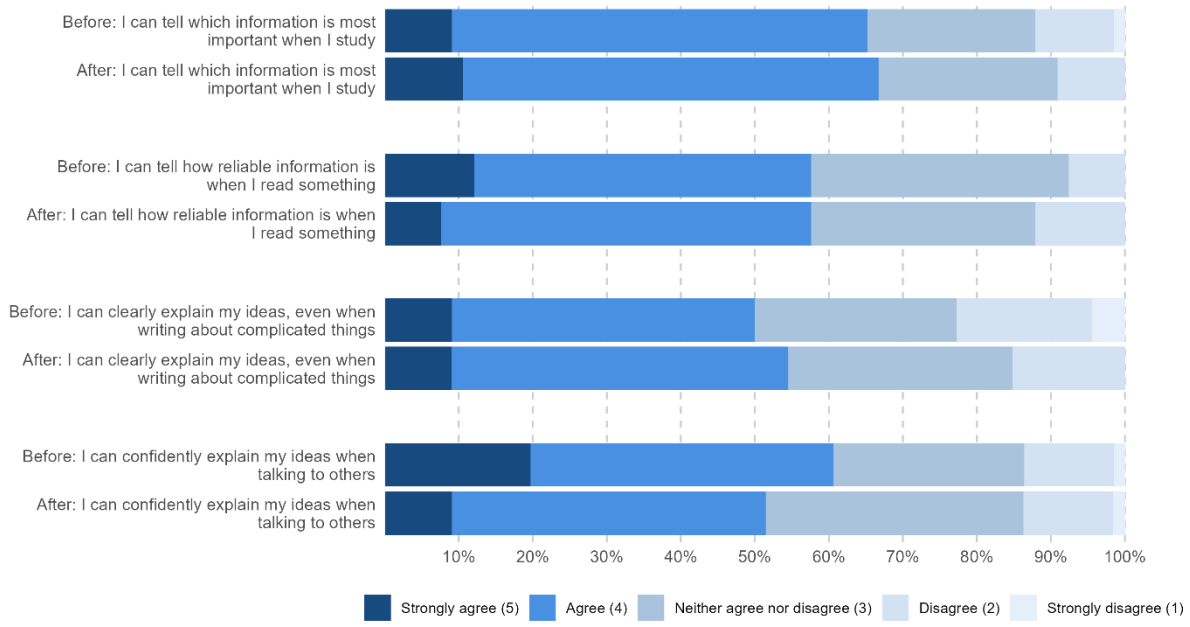

Figure A.1: Control group pre- and post-survey results for Cognitive skills questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.702, 0.521, 0.215, 0.101, respectively).

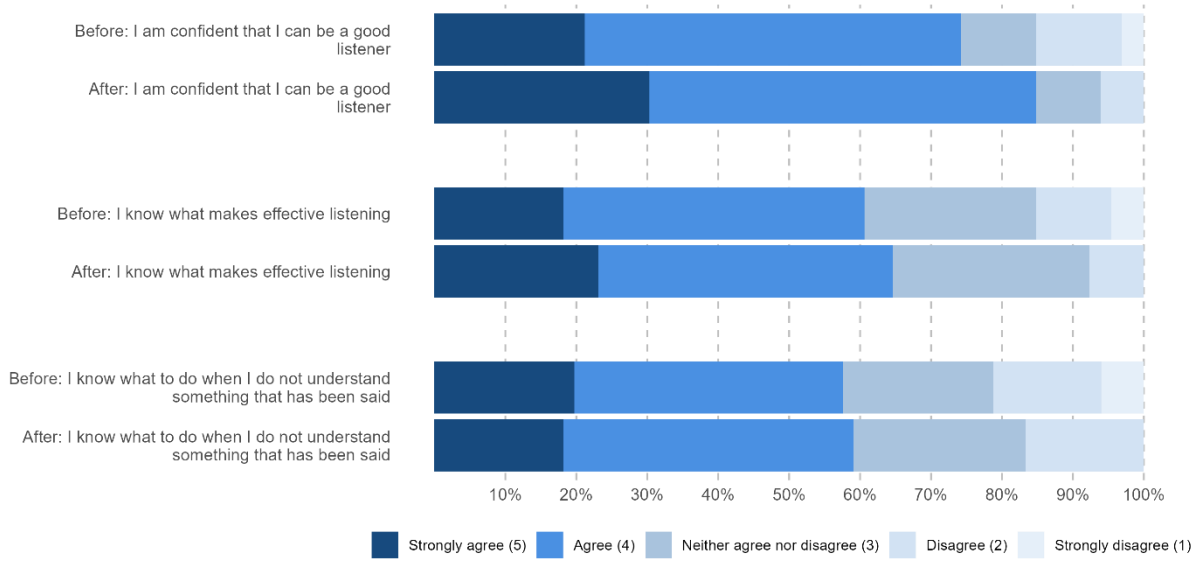

Figure A.2: Control group pre- and post-survey results for Listening skills questions. Post-hoc CLMM pairwise comparisons showed a significant increase[4] for the first question (p = 0.01), while no significant changes were observed for the second and third questions (p = 0.21, 0.659, respectively).

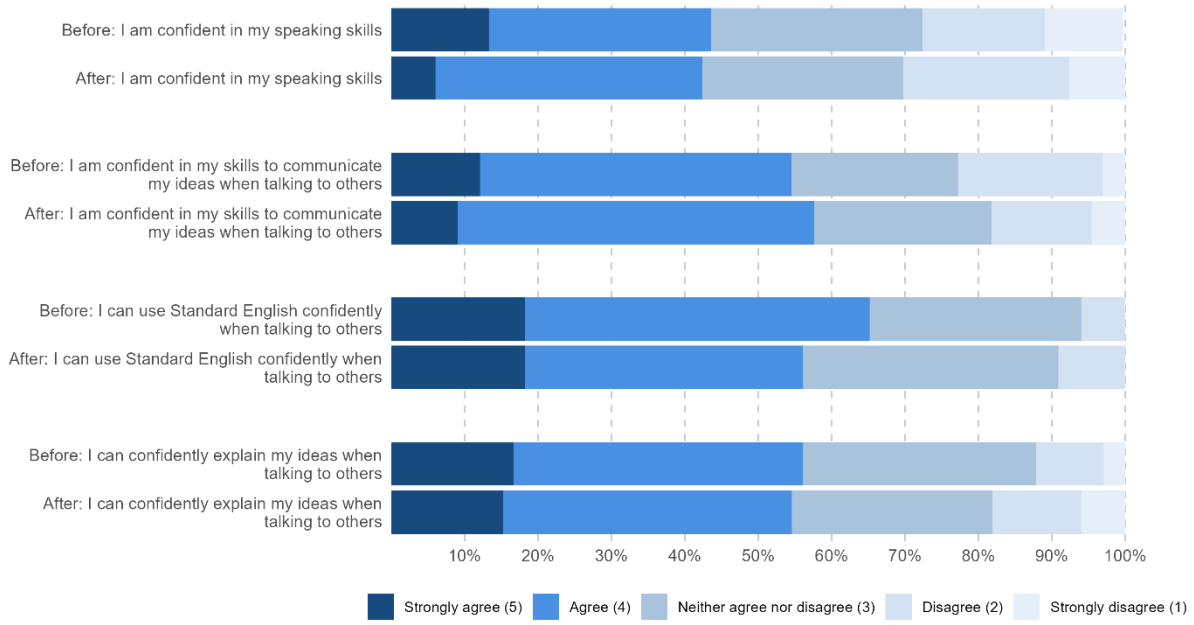

Figure A.3: Control group pre- and post-survey results for Speaking skills questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.486, 0.803, 0.378, 0.447, respectively).

[4] The statistically significant differences observed cannot be interpreted as programme effects and are likely attributable to external factors unrelated to the programme as the control group learners did not participate in the intervention.

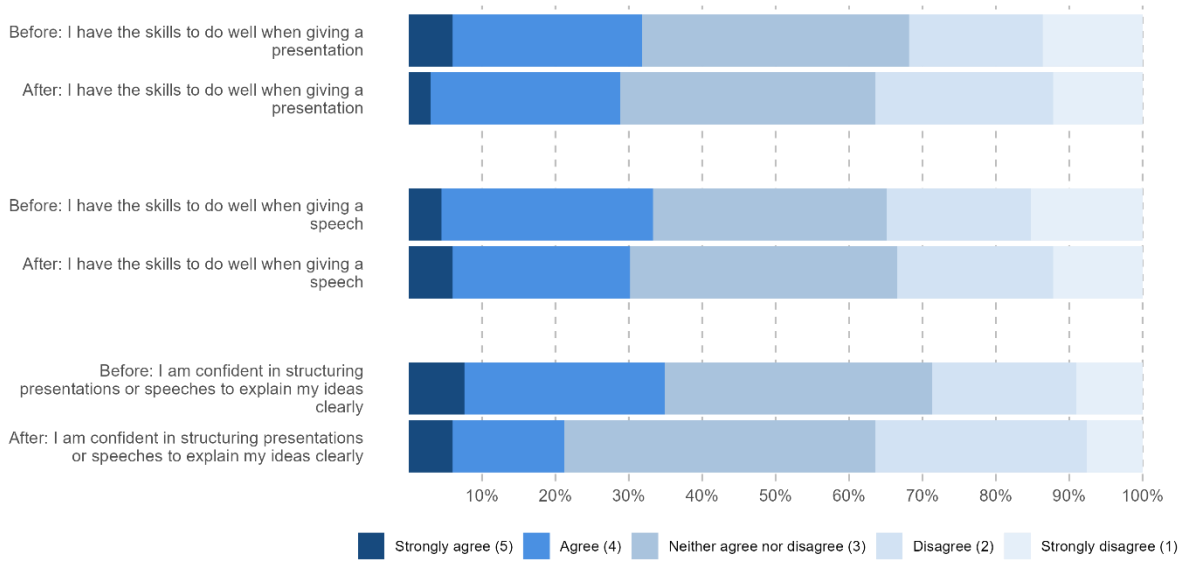

Figure A.4: Control group pre- and post-survey results for Presentation skills questions. Post-hoc CLMM pairwise comparisons showed a significant increase[5] for the third question (p = 0.0345), while no significant changes were observed for the first and second questions (p = 0.345, 0.929, respectively).

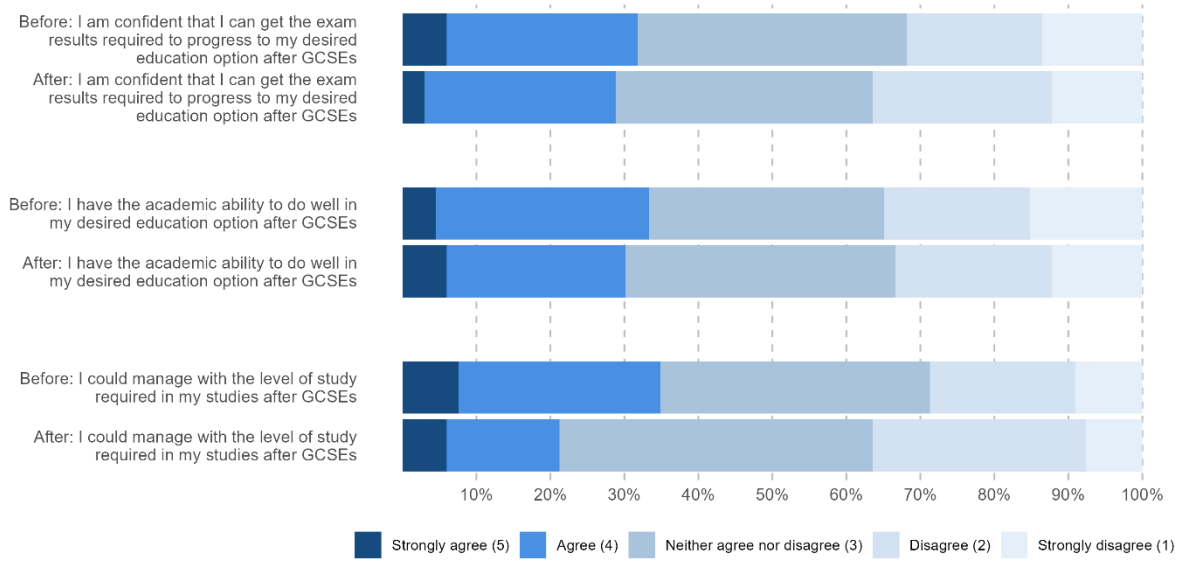

Figure A.5: Control group pre- and post-survey results for SE (post-16) questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.943, 0.630, 0.870, respectively).

[5] See footnote 4.

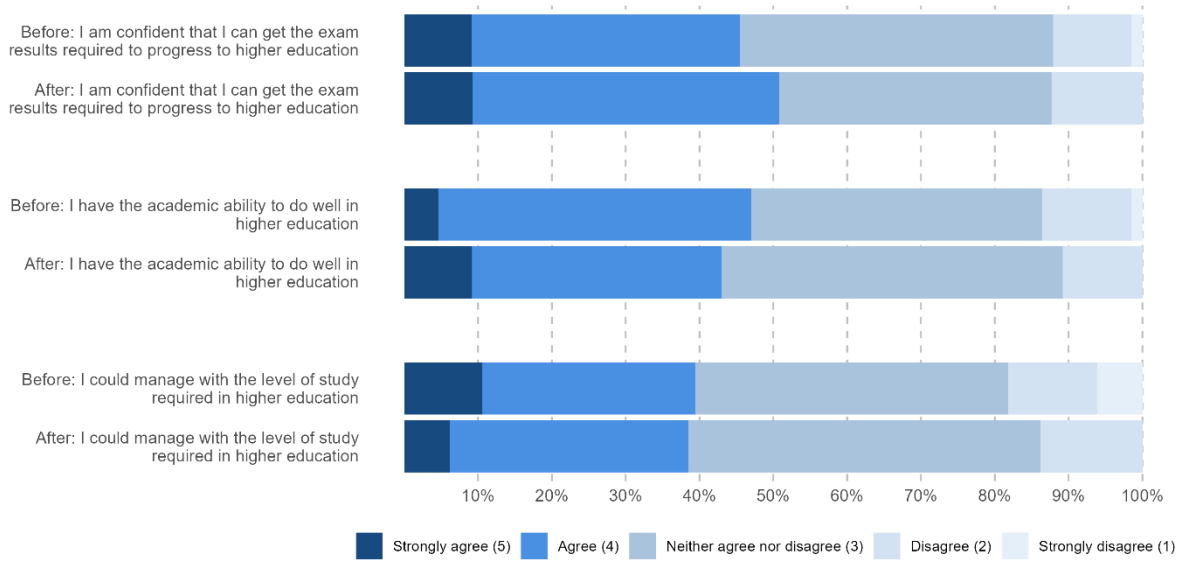

Figure A.6: Control group pre- and post-survey results for SE (HE) questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.634, 0.598, 0.628, respectively).

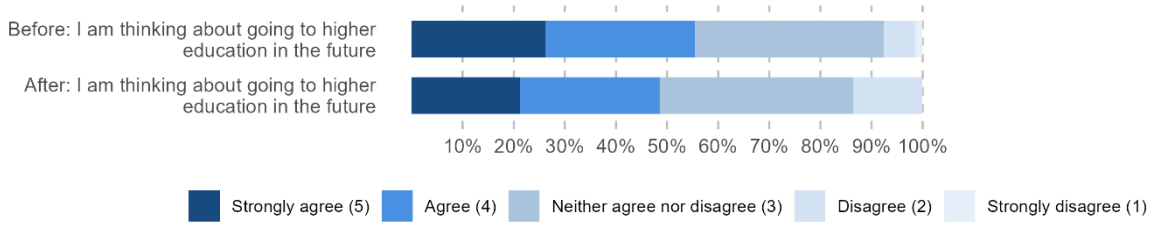

Figure A.7: Control group pre- and post-survey results for HE expectation question. Post-hoc CLMM comparisons revealed no significant difference between the pre- and post-survey results (p = 0.32).

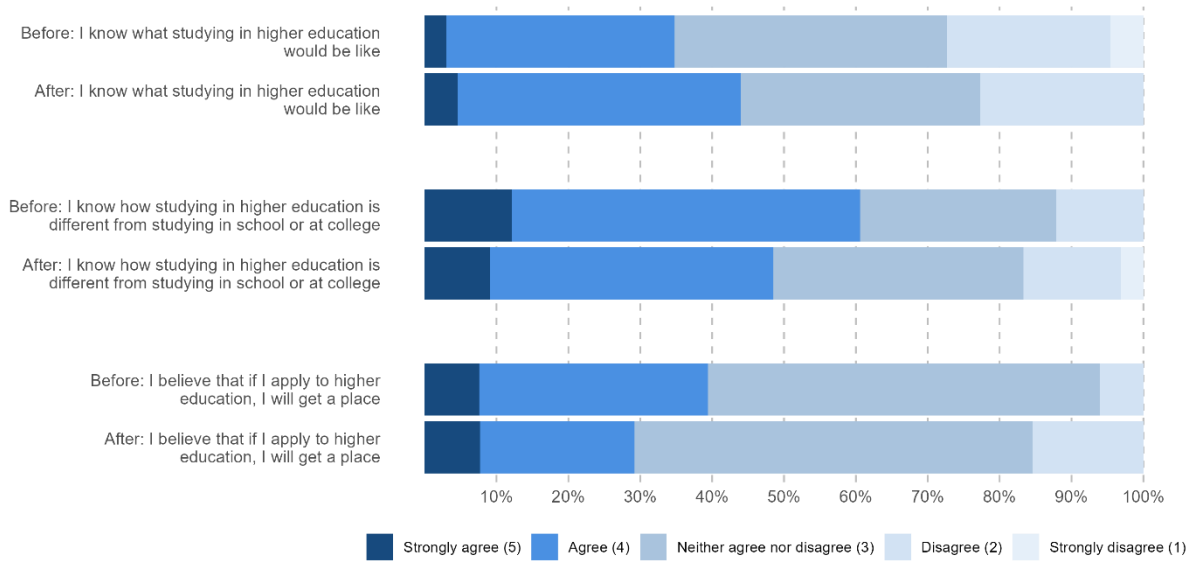

Figure A.8: Control group pre- and post-survey results for HE knowledge questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.08, 0.09, 0.17, respectively).

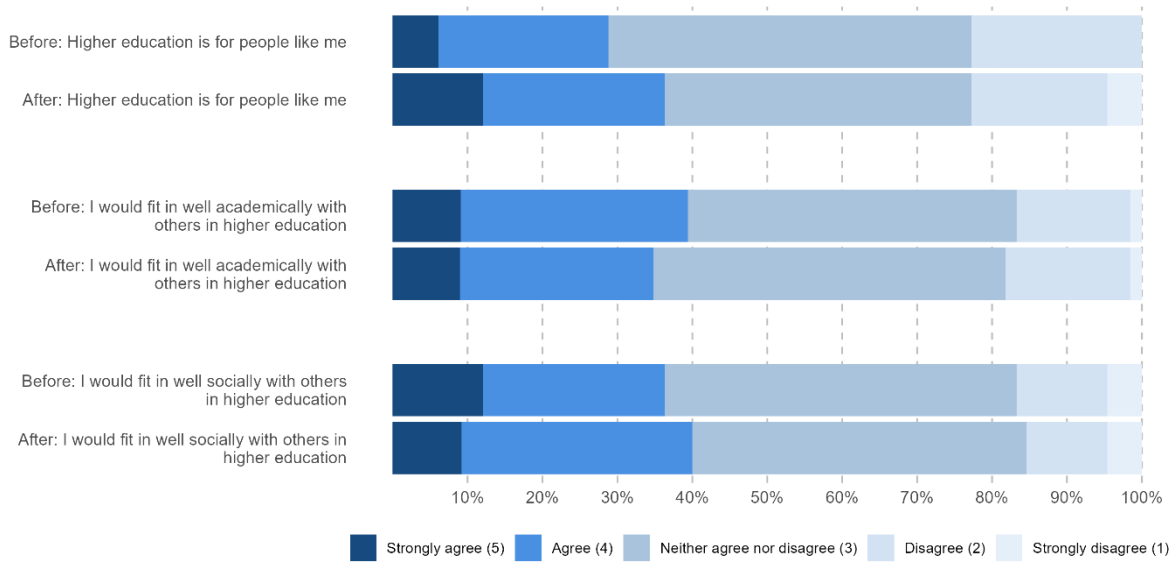

Figure A.9: Control group pre- and post-survey results for Sense of belonging questions. Post-hoc CLMM pairwise comparisons showed no significant changes for any of the questions (p = 0.466, 0.542, 0.784, respectively).