RESEARCH ARTICLE

Creating authentic assessment in mathematics

Ria Symonds, School of Mathematical Sciences, University of Nottingham, Nottingham, UK.
Email: ria.symonds@nottingham.ac.uk
Lisa Mott, School of Mathematical Sciences, University of Nottingham, Nottingham, UK.
Email: lisa.mott@nottigham.ac.uk

Abstract

Assessment of students' mathematics knowledge within higher education (HE) has normally taken a very traditional approach. Closed-book assessments have long been the favoured mode of assessment (Iannone & Simpson, 2011) which often requires students to recall facts, formulae, and methods. One could argue that this type of assessment is limited in its ability to effectively assess how well a student's ability to authentically use mathematics has developed. Due to the recent pandemic, many institutions were forced to rethink their assessment methods so that they could be delivered online and remotely. As such, there has been a renewed sense of need for more 'authentic' assessments for mathematics-based programmes.

In this paper, we will discuss our journey of creating more authentic assessments for apprentices enrolled on a new Data Science Degree Apprenticeship, particularly in mathematics/statistics. We will compare two years of delivery of the course; the first year of delivery which comprised of traditional assessment methods (coursework/exam) and the second year of delivery that used more authentic assessment methods. We will discuss the pros and cons of each model by reflecting on our practice and drawing on apprentices' feedback.

Keywords: authentic, traditional, assessment, mathematics, statistics, apprentices.

1. Introduction

Degree apprenticeships (DAs) form an extension to higher education level of the partly publicly funded British apprenticeship system and involve a period (typically between two and five years) of employment and workplace training alongside a part-time bachelor's or master's degree (predominantly taught remotely). The case for the development of a new degree apprenticeship in data science at the University of Nottingham was approved in 2019. Owing to the pandemic, the DA finally launched in September 2021, recruiting 30 apprentices from companies such as Experian, Ford, Toyota, Western Power (National Grid), and Rolls Royce. The original programme design was modelled quite closely on a more traditional in-person BSc in Data Science that had run up to 2018, but during 2021/22 it became apparent that this model needed to be considerably refined to take account of the needs of employers and apprentices. Part of this redesign was to create more authentic assessment opportunities to equip apprentices for the workplace by ensuring that learning and assessment was integrated into their activities. In this article 'authentic assessment' refers to assessment of learning that is conducted through 'real world' tasks requiring students to demonstrate their knowledge and skills in meaningful contexts' (Swaffield, 2011).

DAs are not yet widely offered within the UK and so, little is known about best practice in their design and delivery. In the 2022/23 academic year 46,790 apprentices were enrolled onto degree apprenticeships in England, which only accounts for 13.9% of new apprenticeships in England. However, the number of apprentices undertaking DAs has been increasing every year since 2017/18 when only 10,870 apprentices started DAs which was only 2.9% of apprentice starts in England (Department for Education, 2023). Utilising existing knowledge and experience from the programme

team, the curriculum was refreshed for 2022/23 delivery so that the assessment was decoupled from the teaching. In addition, in response to employer and apprentice feedback, the assessment was made more authentic for the learner to avoid compartmentalising knowledge which can lead to a lack of understanding and an inability to draw upon knowledge and skills when faced with unfamiliar problems.

Since the curriculum refresh, the programme is split into 'teaching blocks' and 'assessment blocks', the latter of which assesses a range of different learning objectives from the taught content. 'Traditional' style assessments (such as closed book/open book exams) were replaced with a variety of 'authentic' assessments. Examples of authentic assessments used on the programme include:

- Portfolios
- Posters
- Presentations
- Statistical Reports
- Business Case Report

This research study investigates the effect of the change of assessment style on performance outcome. We investigate whether there was a marked difference in performance of the 2021/22 cohort (whose first year on the programme was assessed using the traditional style of assessment and their second year was assessed using authentic assessment). It also investigates apprentices' attitudes towards both modes of assessment from two intakes (2021/22 and 2022/23) and seeks to discover whether one is more favourable over the other.

2. Rationale and relevant literature

Assessment is a critical component of higher education, serving as a tool for gauging student learning and performance. Traditional assessment methods, such as examinations, have long been the standard in evaluating students' understanding of mathematical concepts. However, recent pedagogical shifts and advances in educational research have prompted a revaluation of these methods. Authentic assessments focus not on how much students remember but on how they can reflect what has been learnt in a new environment (Gibson & Shaw, 2011). Authentic assessment is a form of assessment which involves students conducting 'real world' tasks in meaningful contexts (Swaffield, 2011).

Lam (2013) has shown that when students are engaged in authentic tasks that mimic real-world applications of mathematics, their understanding and motivation increase significantly. Authentic assessments, by nature, bridge the gap between theory and practice, making mathematical concepts more accessible and meaningful.

A driving factor for changing assessment practices in this programme was to ensure apprentices are equipped with the necessary skills for the workplace. Traditional assessment methods, such as examinations, often have a focus on the ability to memorise and recall facts and formulae. On the contrary, it can be argued that authentic assessment methods in mathematics challenge students to think critically and apply mathematical concepts to solve complex, non-routine problems. This shift from memorisation to application has been shown to enhance problem-solving and critical thinking skills (Brookhart, 2010; Pellegrino et al., 2001) and hence these skills are more transferable to the workplace. Niss (1998) advocates this, suggesting that assessment of a broader range of skills can be attained through report writing, projects, investigations, and/or oral examinations. Implementing a range of assessment approaches allows students multiple opportunities to utilise feedback and demonstrate their learning.

A further benefit of authentic assessments is that they are deemed more inclusive, allowing for a broader range of learning styles and abilities (Tai et al, 2023). For example, traditional exams may inadvertently disadvantage students with test anxiety or different learning preferences. Authentic assessments can reduce test anxiety whilst providing an opportunity for students to showcase their mathematical understanding.

3. Methodology

The purpose of this research was to investigate whether there was a marked difference in the performance of apprentices completing mathematics authentic assessments versus traditional assessments in the DA. We also sought to understand apprentices' perceptions of the two different assessment styles in mathematics. We address the following research questions.

1. Did the change in assessment style affect the apprentices' results?
2. What were the apprentices' opinions of the change in assessment style?

The cohort of apprentices is small, so we cannot make statistically significant conclusions from our analysis.

3.1 Assessment analysis

The 2021/22 cohort of apprentices completed traditional assessment for the statistics aspects of the DA during their first year on programme (prior to curriculum and assessment changes in 2022). During their second year on programme, the same apprentices were assessed using authentic assessment methods in statistics. This cohort was unique since they were the only cohort who were exposed to both assessment styles on the DA. Therefore, we wanted to compare their module results for the statistics modules in both year 1 and year 2.

We also compared the year 1 results from the Probability and Statistics assessment in 2021/22 (traditional assessment style) with the year 1 results from the Probability and Statistics assessment in 2022/23 (authentic assessment style). This was to investigate whether there was a difference in the performance of the two different cohorts who were exposed to the two different assessment styles for Year 1 on programme.

For both analyses the mean, median, minimum, maximum, lower quantile, and upper quantile were compared to see how these differed between assessment types.

3.2 Survey analysis

Three surveys were sent to all apprentices enrolled on the DA at the end of the 2022/23 academic year. Survey 1 and survey 2 were sent to the 2021/22 cohort to seek their opinion of the assessments in both year 1 (traditional assessment) and year 2 (authentic assessment). The response rates for surveys 1 and 2 were 50% (8/16) and 87.5% (14/16) respectively.

Survey 3 was sent to the 2022/23 cohort to seek their opinion of the assessments in year 1. 86% (12/14) of the apprentices completed the survey.

Each survey used a Likert scale to collate responses and apprentices were encouraged to give reasons for their responses. They were also encouraged to give any other constructive comments regarding the assessment.

Quantitative results from the responses for the questions from all three surveys were analysed graphically.

4. Results

We now present the results of the analyses. We will address our first research question "*Did the change in assessment style affect the apprentices' results?*" by comparing raw marks of apprentices from the two cohorts.

4.1 Quantitative Results

We first compare the performance of the 2021/22 cohort in their first and second year of the DA. The results of the cohort are summarised in figure 1. The boxplot indicates that there is a slight improvement in the performance from year 1 to year 2 with the change from traditional to authentic assessment. This change is not statistically significant, and we are working with small sample sizes.

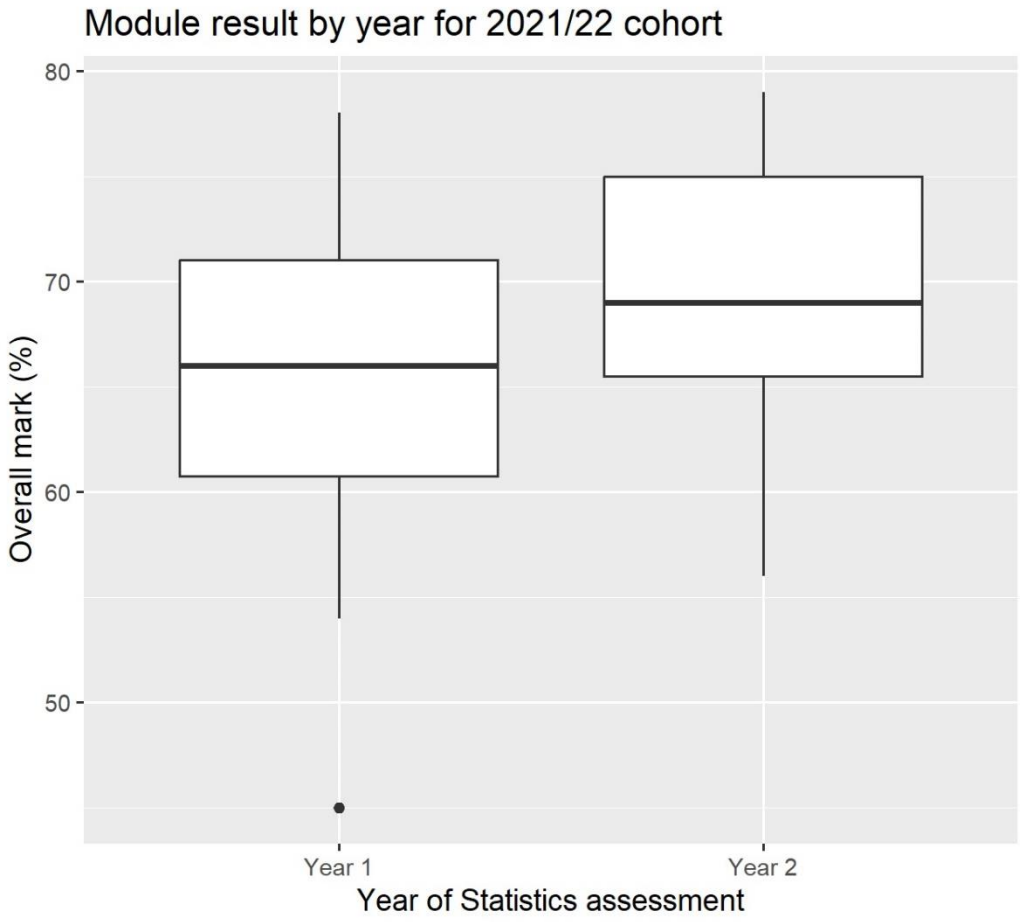

Figure 1. A boxplot that compares the Statistics marks of the 2021/22 cohort in their first and second years of the degree apprenticeship.

The analysis indicates that the change to authentic assessment has not improved assessment grades (but neither has it impaired apprentices' performance). Although most apprentices' marks in year 2 were within 5% of the mark they achieved in year 1, further investigation reveals 4 exceptions.

- "Apprentice 1"'s mark decreased by 7% in year 2.
- "Apprentice 2"'s mark increased by 6% in year 2.
- "Apprentice 3"'s mark increased by 7% in year 2.
- "Apprentice 4"'s mark increased by 18% in year 2.

We note that the performance of “Apprentice 4” was statistically significant. It could be that the traditional form of assessment in year 1 was not suitable as a way of assessing this apprentice’s ability or it could be the case that the apprentice had a better understanding of Statistics in year 2.

To further investigate the effect of authentic assessment on performance we compared the first-year marks of the 2021/22 cohort (assessed by traditional assessment) with the first-year marks of the 2022/23 cohort (assessed by authentic assessment). The results are shown in figure 2.

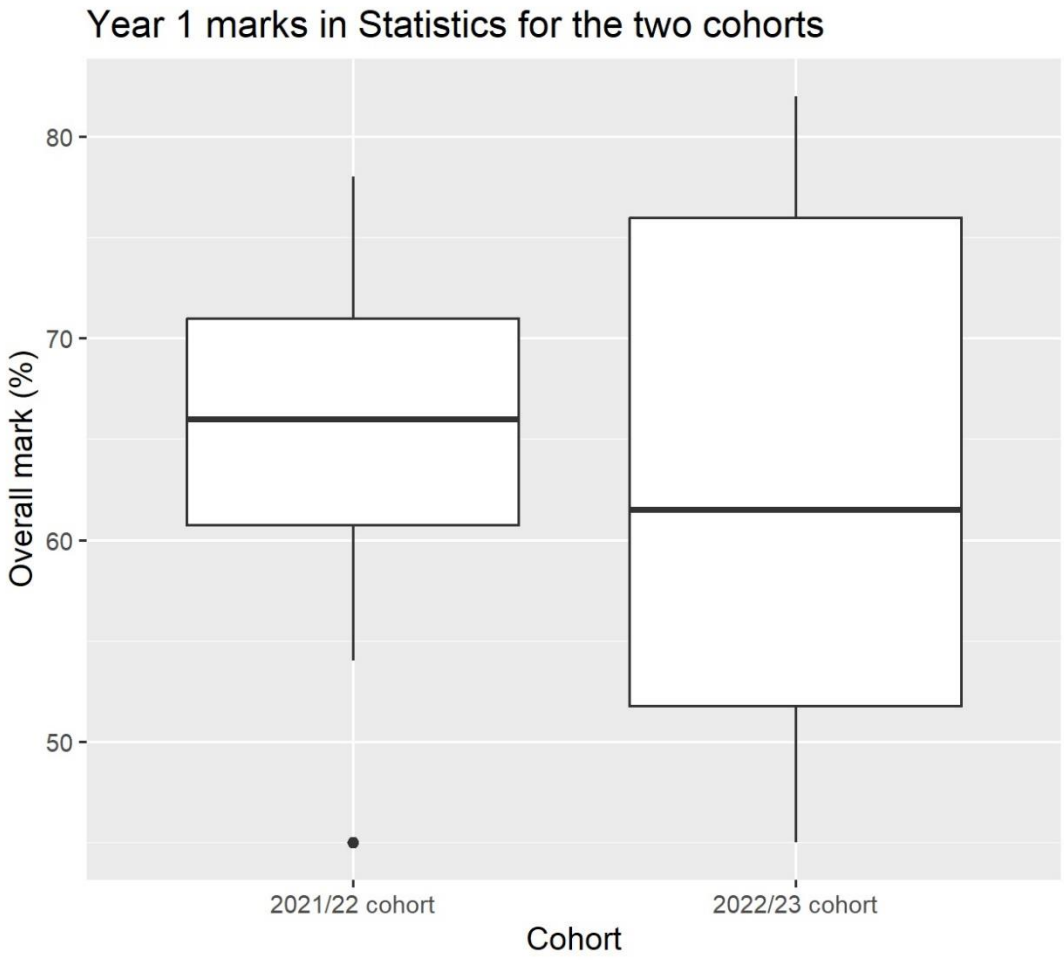

Figure 2. A comparison of the year 1 Statistics marks in 2021/22 and the year 1 Statistics marks in 2022/23.

The results indicate that the marks are more varied for the 2022/23 cohort than the 2021/22 cohort. The median and mean marks are similar for the two cohorts. It is perhaps difficult to draw conclusions from this analysis since the two cohorts comprise of different backgrounds and prior skills (for example the 2021/22 cohort comprised of many more mature learners than the 2022/23 cohort which consisted of mainly school leavers who has just completed their A-levels). However, since there is no significant difference in the average performance of both cohorts, we could assume that the use of authentic assessment is a suitable method of assessing learning objectives comparable to that of traditional assessments.

4.2 Qualitative Results

We now seek to investigate the second research question "*What were the apprentices' opinions of the change in assessment style?*"

In Survey 1, the 2021/22 cohort were asked to indicate which assessment style they preferred. 100% of respondents stated a preference towards the year 2 authentic assessment commenting that they appreciated the workplace application that authentic assessment lends itself to. Some of these comments can be seen below:

> *"I feel like this approach to learning is more representative of real-world application, as opposed to under exam conditions."*
>
> *"The assessments were designed to use the skills learned practically in problem scenarios. This can then be simulated at a workplace."*
>
> *"The year 2 assessment were more relatable to the workplace and therefore more appropriate to an apprenticeship."*

The graph in figure 3 summarises the responses from the same cohort 2021/22, indicating whether they found the individual assessments (comprising of a group presentation, business case report, statistical report, and R portfolio) in year 2 useful for embedding their knowledge.

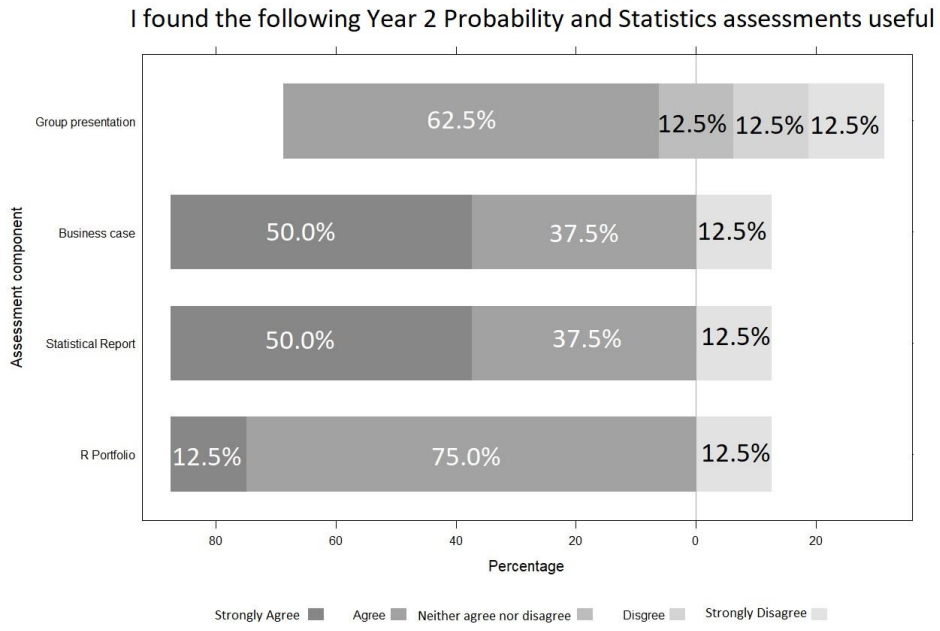

Figure 3. A Likert plot showing the responses of the 2021/22 cohort of apprentices. They were asked whether they agree or disagree with the statement "*I found the following Year 2 Probability and Statistics assessments useful*" for the four different Statistics assessments they completed in year 2.

Most apprentices selected 'agree' or 'strongly agree' in their responses to how useful they found each of the assessments in year 2. Interestingly, the group presentation was not as well received in comparison to the other assessments (25% of respondents selected either 'disagree' or 'strongly disagree'). In addition, it appears that some apprentices did not favour being assessed as a group.

This is perhaps not surprising since the literature suggests that group work can be a frustrating experience for some students (Hall & Buzwell, 2012).

> *"I found the content of the group presentation useful but did not like being assessed as a group."*

> *"Regarding the group exercise. I found it to be a bit like individual pieces of work we just stitched together, so there was no real element of group work. It could be that this was specific to our team though."*

In general, other comments from apprentices suggested that some of the assessments were too time-consuming and that apprentices would have benefitted from additional support in seeking out relevant data required for the assessments. Overall, the analysis suggests that the authentic assessment was well received, and any constructive feedback was linked to the logistical side of implementing the assessment (rather than the assessment itself).

One surprising observation was related to an individual response who has selected 'strongly disagree' for all four assessments, indicating that this apprentice did not like the new form of authentic assessment. However, analysis of their accompanying comments could suggest that perhaps they selected the wrong response on the scale.

> *"I think that all the assessments presented a good learning experience and a chance to apply and demonstrate our understanding of the probability and statistical methods."*

In survey 2, the same cohort (2021/22) were asked whether the different assessment blocks for the whole programme provided a positive experience. We can see in figure 4 that 78.6% of apprentices either agreed or strongly agreed that the Probability and Statistics assessment provided a positive learning experience. In fact, all assessment on the DA course were well received. The only exception was for the 'Synoptic Assessment'. Interestingly, this was the only assessment block that had a traditional assessment style component in the form of a multiple-choice test. Overall the apprentices' responses suggests that the second year was a positive learning experience for the majority of the cohort.

In the final survey, survey 3, opinions of the 2022/23 cohort (who completed the first year using authentic assessment only) were sought, as can be seen in figure 5. 75% of apprentices either agreed or strongly agreed that the Probability and Statistics assessment provided a positive learning experience. Again, similar to the analysis above, all assessment on the DA course were well received by this cohort with an exception for the 'Synoptic Assessment' (which contained a traditional assessment component in the form of a Multiple-Choice Test). Further analysis of the comments received suggest that some apprentices (two respondents) would have preferred to have been assessed via a test rather than a portfolio for statistics/probability; *"Multiple choice for stats not just portfolios".* It is perhaps worth noting, that this cohort of apprentices were mostly school leavers who have only ever been assessed using exams and have limited experience of working in industry.

Selected comments from survey 3 (from the 2022/23 cohort) indicate that apprentices found the authentic assessment in the DA useful for applying skills at work.

> *"I had fun with all of these and made the most of the courses as they helped with work and vice versa."*

> *"I was able to apply or [sic] my skills on this course"*

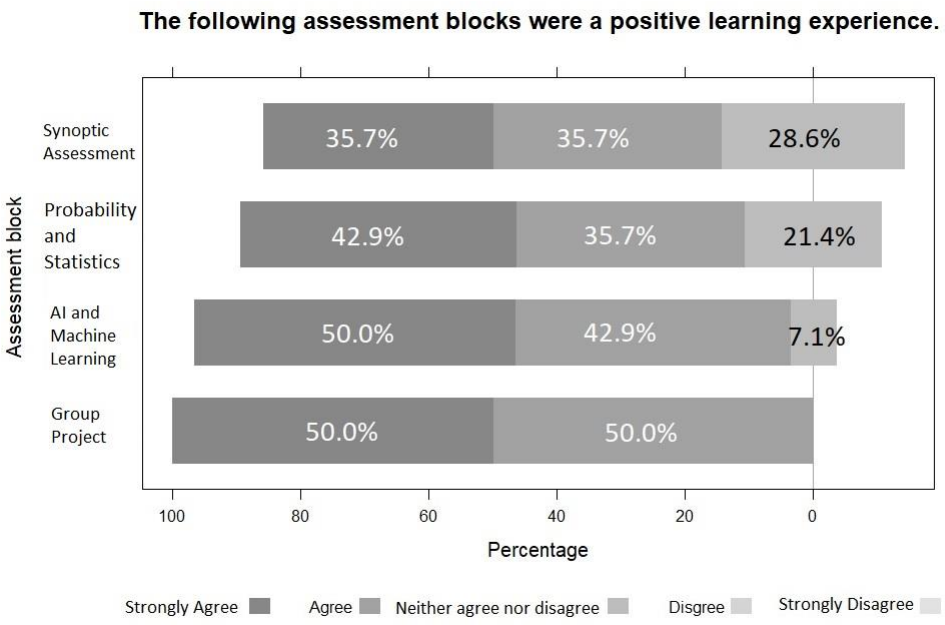

Figure 4. A Likert plot showing the responses of the 2021/22 cohort of apprentices. They were asked whether they agree or disagree with the statement "*The following assessment blocks were a positive learning experience*" for the four different assessment blocks in year 2 of the DA.

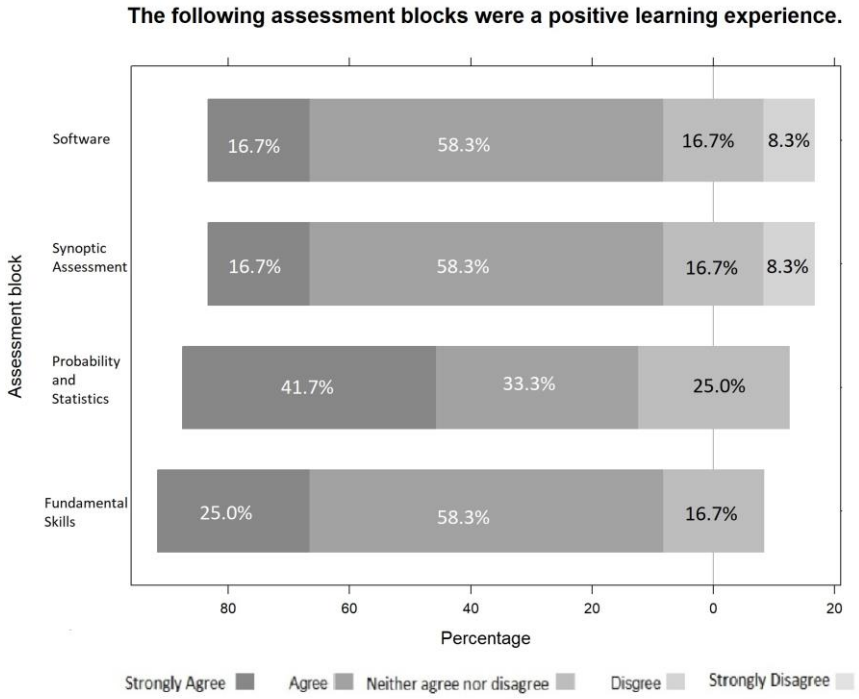

Figure 5. A Likert plot showing the responses of the 2022/23 cohort of apprentices. They were asked whether they agree or disagree with the statement "*The following assessment blocks were a positive learning experience*" for the four different assessment blocks in year 1 of the DA.

Constructive comments were related to the logistics of carrying out the assessment, such as timing of assessments in the academic year.

> *"Would be better if the assessments were spread out over year (not too many in exam season)."*

It is worth noting, that although assessment for the DA is more spread out over the whole year than a traditional degree (final assessment was due mid-July) we are still constrained by some university processes during the summer period for marking/progression of apprentices in terms of the degree qualification.

5. Discussion

Our analysis suggests that using authentic assessment was a positive experience for both cohorts of apprentices on the DA. Analysis of the qualitative data suggests that the use of assessments which mimic workplace tasks (such as writing statistical reports and giving presentations) made the DA content more accessible and meaningful to their individual organisations. For example, at the end of the academic year, both cohorts of apprentices were able to showcase work-based projects they had contributed to which utilised assessment and taught methods from the course. Whilst none of the comments were necessarily 'negative' there were some detailed constructive feedback with regards to the logistics of implementing authentic assessment. For example, spreading the workload more evenly in the academic year or replacing some of the more similar style assessments (business case report, statistics report) with an alternate authentic assessment such as on oral viva.

One consideration is whether the use of authentic assessments is disadvantageous to apprentices who excel in exams. However, one of the main aims of the DA is to prepare apprentices for industry by mimicking workplace tasks and, therefore, authentic assessment is more relevant and useful in this context. In addition, more apprentices seem to favour the authentic assessment over the traditional assessment (such as the multiple-choice test that is used in the Synoptic Assessment block for both cohorts).

The DA commenced in 2021/22 and we currently have three cohorts of apprentices. Recruitment numbers are significantly growing (annual growth of 300% from 2022/23 to 2023/24) which will undoubtedly involve more work to mark which adds to staff workload. We may need to implement automated processes or utilise more peer marking to assess apprentices effectively whilst also keeping the balance of authentic assessment. Further research could investigate whether the use of authentic assessment is also helping to develop other transferable skills in the apprentices' roles.

6. References

Brookhart, S. M. (2010). *How to assess higher order thinking skills in your classroom*. Alexandria, VA: ASCD.

Gibson, K. and Shaw, C.M. (2011). Assessment of Active Learning. *Oxford Research Encyclopedia of International Studies*. https://doi.org/10.1093/acrefore/9780190846626.013.120

Hall, D. and Buzwell, S. (2012). The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution. *Active Learning in Higher Education*, 14, 37-49. https://doi.org/10.1177/1469787412467123

Iannone, P. and Simpson, A. (2011). The summative assessment diet: how we assess in mathematics degrees. *Teaching Mathematics and its Applications: An International Journal of the IMA*, 30(4), pp.186-196, https://doi.org/10.1093/teamat/hrr017

Lam, B.H. (2013). Authentic Learning. Available at: https://www.eduhk.hk/aclass/Theories/AuthenticLearning_28June.pdf [Accessed 8 November 2023].

Niss, M.A. (1998). Dimensions of geometry and assessment. In C. Mammana and V. Villani (Eds.), *Perspective on the Teaching of Geometry for the 21st Century: An ICMI Study*, pp.263-274. Springer Netherlands.

Pellegrino J.W., Chudowsky N. and Glaser R. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), pp.433-449. https://doi.org/10.1080/0969594X.2011.582838

Tai, J., Ajjawi, R., Bearman, M., Boud, D., Dawson, P. and Jorre de St Jorre, T. (2023). Assessment for inclusion: rethinking contemporary strategies in assessment design, *Higher Education Research & Development*, 42(2), pp.483-497. https://doi.org/10.1080/07294360.2022.2057451