Transforming Access and Student Outcomes in Higher Education

Trial protocol
Institutional Data Use: University of Huddersfield – Score As I Learn (SAIL)

September 2024

Authors:
Sally Andrews, Staffordshire University
Vanessa Dodd, Staffordshire University
Sam Vizcaino-Vickers, Staffordshire University
Sehrish Ghayas, Staffordshire University
Jarek Bryk, University of Huddersfield

QA:
Denise Hawkes, Anglia Ruskin University

TASO Transforming Access and Student Outcomes in Higher Education

VERSION	DATE	REASON FOR REVISION/NOTES
Any changes to the design to be agreed between the implementation partner(s) and the evaluators. Note any agreed changes in the table below.		
5	5/3/24	Clarifications of random effects structure and control variables.
4	28/2/24	Formatting and typographical changes.
3	21/2/24	Second round of QA
2	13/2/24	Revision to address QA issues
1.0 [original]	8/2/24	Original version post QA.
Pre-registration		This design has been pre-registered on the Open Science Framework.[1]

The QA rating system is based on Evaluation Security tool presented in the TASO Monitoring and Evaluation Framework.[2]

[1] https://osf.io/b4xqa/
[2] https://taso.org.uk/evidence/evaluation/

QA	Comments	Rating (out of 5)
Design	Given the data constraints, this design will provide initial evidence of potential causality. A matching model between computing and engineering, given the issues of changes in sector and HEI limiting the ability to do a diff-in diff model between cohorts of engineers, could be used to provide a better match with the control group.	3
Sample size	I am still concerned that we do not know the actual sample size. Once this is known we may expect changes in aspects of the research design and power calculations.	3
Outcome measure	These are the expected outcomes.	4
Attrition	Attrition and continuation are clearly linked. You will need to think carefully about those who appear in the data in one year and then not the next. Given the issues with continuation data, a proxy could be missing data, given this is institutional data the attrition bias is likely to be quite small (compared to survey data).	4
Validity	Still a lot to learn from this, especially as the sample size is unknown and the HEI access to data for the evaluator is a challenge. Let's see what the results give but it would provide a baseline for further analysis.	3
Overall	I would like to thank the evaluation team for responding to my concerns within the constraints of the data access.	3.5

1. Summary

Background

This evaluation design has been developed as part of a project funded by TASO on the use of institutional data to generate causal (Type 3) evidence for interventions designed to increase equality of opportunity within the higher education (HE) sector. Four HE Providers (HEPs) are taking part in the project and a team from Staffordshire University are designing and carrying out the evaluation. Two types of evaluation for each HEP's intervention will be conducted: an impact evaluation and an implementation and process evaluation. This analysis protocol covers the impact evaluation of the Score As I Learn (SAIL) intervention at the University of Huddersfield.

Aims

SAIL is a universal whole-curriculum approach used for students in the Department of Engineering and Technology within the School of Computing and Engineering. SAIL's aims are to support students to engage early and consistently with their course material, by tying engagement to assessment outcome. In doing so, SAIL aims to ultimately support students' continuation, grades, and degree outcomes.

While SAIL's aim is to support all students, differences in attainment were particularly noticeable for students coming to the University of Huddersfield with BTEC or other non-A Level qualifications. This pattern is attributed to different prior educational experiences between BTEC and other students, particularly with respect to self-guided learning principles and assessment practices then employed at the university. SAIL's aims are therefore particularly aimed at supporting these student groups, amongst the wide cohort.

Intervention

SAIL offers students the opportunity to engage in weekly 'low-stakes' summative assessments, each weighted at 3% of the overall module grade. That is, the grades students receive do contribute to their module aggregate, and therefore are credit bearing assessments, but the individual impact of each assessment is nominal.

Weekly assessments are weighted at 3% of the overall module mark. The tasks are typically multiple-choice question quizzes, delivered through the University's Virtual Learning Environment (VLE). Students receive automated feedback, and quiz results are released 24 hours after the passing of the weekly deadline. Students can choose to complete any number of the weekly quizzes, including none at all.

This regular, self-checking exercise also presents an opportunity for academic staff to monitor student and class performance and adapt classroom content based on the outcomes of the whole cohort, where the outcome demonstrates the need to enhance student understanding.

Design

In this study we will apply an ex-post facto quasi-experimental evaluation design to determine whether SAIL participation increases student engagement and student outcomes.

Outcome measures

This study has two primary outcome measures: course engagement and substantive assessment submission habits. We will explore two additional secondary outcome measures: stage grade (students' end of year aggregate grade) and degree award.

Analyses

We will use linear mixed models (LMMs) and model comparison using Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) to address each research question.

2. Background

Table 1. Personnel involved in the project

Organisation	Name	Role and responsibilities
TASO	Dr Rob Summers	Project/Contract Manager
TASO	Luke Arundel	Project Assistant
Staffordshire University	Dr Sally Andrews	Project Lead. Responsible for day-to-day management of the project.
Staffordshire University	Vanessa Dodd	Project Co-lead. Responsible for supporting day-to-day management of the project.
Staffordshire University	Juan Raman Mullor	General project support. Report writing and interpretation.
Staffordshire University	Reagon Alford	Research Assistant. Responsible for data cleaning, analysis, and reporting.
Staffordshire University	Sehrish Ghayas	Research Assistant. Responsible for data cleaning, analysis, and reporting.
University of Huddersfield	Dr Jarek Bryk	Project Lead at University of Huddersfield. Responsible for data curation and distribution, and supporting with exploratory analyses.
University of Huddersfield	Steve Bentley	Strategic Learning Technology Advisor at the University of Huddersfield. Responsible for VLE data collection and curation.

3. Aims

The aim of this evaluation is to determine the effectiveness of the Score as I Learn (SAIL) initiative at the University of Huddersfield. SAIL's major aim is to increase student engagement, and ultimately course and module outcomes by introducing weekly 'low-stakes' summative assessments into modules throughout the student journey. This evaluation is being undertaken to develop the evidence base for the effectiveness of this initiative to support students' engagement and attainment at the University of Huddersfield. The evaluation will meet these aims via robust, inferential statistical techniques so the evaluators can infer causation. In this impact evaluation we will test the following research questions (RQs):

RQ1: Does SAIL impact students' engagement on their course?

> H_0: students on courses with SAIL engage with their course to the same extent as students on non-SAIL courses.

H_1: students on SAIL courses engage with their course to a different extent than those on non-SAIL courses.

RQ2: Does SAIL impact students' degree outcomes on their course?

H_0: there is no difference in degree outcomes between students on SAIL courses compared to non-SAIL courses.

H_1: students on courses with SAIL have different degree outcomes than students on non-SAIL courses.

RQ3: Does SAIL differentially impact non-A Level and international students' grades depending on their qualifications on entry to the university or home/international status?

H_0: there is no difference in grades for non-A Level and international students on SAIL courses compared to non-SAIL courses.

H_1: non-A Level and international students on SAIL courses will be awarded different grades to those on non-SAIL courses.

RQ4: Does engagement with SAIL impact on students' assessment submission habits?

H_0: there is no difference in substantive assessment submission habits between students on SAIL and non-SAIL courses.

H_1: the timing of substantive summative assessment submissions is different for students on SAIL courses relative to students on non-SAIL courses.

RQ5: Does engagement with SAIL impact grades for engineering students?

H_0: there is no difference in the grades of students who engaged in SAIL courses compared to the student who didn't engage in SAIL courses.

H_1: The students on SAIL will have different grades than those who didn't engage with SAIL course.

RQ6: Does participation in SAIL program impact students' continuation from Level 4 to Level 5 and Level 5 to Level 6?

H_0: Participation in SAIL has no impact on continuation from Level 4 to Level 5 or from Level 5 to Level 6?

H_1: Participation in SAIL has an impact on continuation from Level 4 to Level 5 and from Level 5 to Level 6?

4. Intervention

In SAIL, each of the 11 weekly assessments is weighted at a 3% of the module mark. The tasks are typically multiple-choice question quizzes, delivered through the

University's Virtual Learning Environment (VLE). Each week, a quiz related to that week's content (in-person lectures and other materials released through the VLE) is released and students have one week to complete it. They only have one attempt to complete it, but their time for completion is not limited other than with a weekly deadline. Students are able to work collaboratively on the SAIL tasks and are able to leave the assessment and return later. Every module has a question bank, from which the questions are randomly drawn for each student. Through the VLE, automatic feedback and quiz results are released 24 hours after the passing of the weekly deadline. Students can choose to complete any number of the weekly quizzes, including none at all.

Crucially, as the SAIL task submission deadline cannot be extended and resubmission is not allowed even in cases of extenuating circumstances, only the best 8 out of the 11 weekly tasks marks count towards the final component mark. This means that 24% of the total module mark comes from SAIL low-stakes assessments.

This regular, self-checking exercise also presents an opportunity for academic staff to monitor student and class performance and adapt classroom content based on the outcomes of the whole cohort.

5. Design

This study will use a mixed quasi-experimental design.

This study will merge administrative institutional data with localised SAIL engagement data from the School of Computing and Engineering at the University of Huddersfield. Data will be collated from records collected from academic years 2020-21 and 2021-22.

The independent evaluators had no influence on determining the eligibility, group-allocation, selection criteria or collection of data.

6. Outcome measures

Table 2 outlines the primary and secondary measures that will be used within the analyses to address the hypotheses in Section 3. The primary outcome refers to the core aim of the SAIL programme, to increase course engagement. The secondary outcomes refer to those outcomes that are hypothesised to result from increased engagement and are secondary benefits of the SAIL programme.

Table 2. Outcome Measures

Outcome measure	Data to be collected	Point of collection
Course engagement	Mean average of attendance to lectures, seminars and	Real time data which will be collected in relation to specific cut off criteria

	workshops on their undergraduate degree	
Degree classifications	Final grade given to students at the end of their degree (Fail, 3rd, 2:2, 2:1, 1st)	Administrative data collected routinely
Stage grade	Grade at the end of each relevant period of study (i.e. level 4, 5, 6)	Administrative data collected routinely
Assessment submission time	Difference between students' substantive assessment submission time and the original deadline in minutes	Administrative data collected routinely through the VLE

7. Sample selection

This evaluation will use secondary data from students who are current students or graduates of University of Huddersfield between 2020-23. The institution has previously gathered this data for various purposes, potentially utilising it for institutional metrics. Nevertheless, researchers have not previously examined or accessed this specific dataset, making it suitable for pre-registration purposes.

Given these constraints, providing a precise expected sample size is challenging, as researchers are constrained by the data supplied to the project team by the provider.

From conversations with the team at University of Huddersfield, upper bound sample size estimates are outlined in Table 3.

Table 3. Sample size upper bound estimates

Academic year	SAIL involvement	Level 4	Level 5	Level 6	Total
2020-21	SAIL	750	750	750	2,250
2020-21	Non-SAIL	750	750	750	2,250
2021-22	SAIL	750	750	750	2,250

2021-22	Non-SAIL	750	750	750	2,250
Total		3,000	3,000	3,000	9,000

8. Identification strategy

The participants are identified from the Engineering and Technology department and from the Computer Science department in the School of Computing and Engineering at University of Huddersfield. This is because these schools are comparable in demographics (including entry tariff), structure, and the nature of the course. Students on the Engineering course take part in the SAIL initiative, while students on the Computing course do not. These similarities make computing students a good natural comparator group. No other courses were considered as comparators.

To provide additional credence to claims of causality, covariates will be controlled for within the models.

9. Data collection

The evaluation will use secondary data, therefore data *collation* will be achieved using administrative student records (institutional data) between 2020-22 from the University of Huddersfield. No additional primary data will be collected.

Table 4. Data collation

Data item	Timeframe	Collector
Continuation data (y/n)	2020-22	University of Huddersfield
SAIL engagement data (assessments submitted/total number of assessments) ● 6/10 model (2020-21) ● 8/11 model (2021-22)	2020-22	University of Huddersfield
Attendance - weekly aggregate (N sessions)	2020-22	University of Huddersfield
Retention (y/n)	2020-22	University of Huddersfield
Completion (y/n)	2020-22	University of Huddersfield
Attainment (stage grade, degree classification)	2020-22	University of Huddersfield
Assessment submission data (minutes before or after deadline)	2020-22	University of Huddersfield

Data item	Timeframe	Collector
Library use logs and booking of library support sessions – weekly aggregates (N)	2020-22	University of Huddersfield
Brightspace (VLE) use – aggregate average weekly usage (minutes)	2020-22	University of Huddersfield
Module attainment data • Number of summative assessments • Aggregate grade	2020-22	University of Huddersfield
Assessment attainment data • Assessment type (categorical) • Assessment grade (%) • Assessment weighting (%)	2020-22	University of Huddersfield
Student demographic data: • Academic year • Level of study • UCAS tariff points • Ethnicity • Disability • Gender • Age • IMD • Qualifications on entry	2020-22	University of Huddersfield

10. Procedure

Table 5. Timeframe

Timeframe	Action
January January/February February/March	• Clean data • Analyse & interpret data • Report findings including data visualisations

As the evaluation uses secondary data sources, there is no relevant experimental procedure.

11. Power calculations

Analytic power calculations, as typically used for frequentist statistical designs, are not possible for LMM designs. This is because power analysis in linear mixed models depends on certain assumptions of the data structure and its distribution and variability. The variability of fixed and random effects also needs to be included in these estimations. Simulation-based power analyses have been proposed as an alternative method to analytic power calculations for LMM (Kumle, Võ & Draschkow, 2021). It is important to have an idea of (preliminary test) correlation structure of random effects and variance due to the random effects for accuracy of the power analysis. For this reason these calculations will be conducted with simulated data made once the structure of the underlying data is known.

12. Analytical strategy

We are using a comparator course instead of exploring changes in student engagement and outcomes before and after SAIL was introduced. The reason for this is twofold; primarily, institutional data is not available for the period prior to the introduction of SAIL, and secondly, there have been a number of changes in the higher education sector and within the curricula at University of Huddersfield over the past decade that limit the ability to infer that any observed changes in engagement or outcomes would be attributable to the SAIL initiative.

Linear Mixed Model (LMM) analyses will be used to address each research question in this evaluation. In LMM, fixed effects are used to explore the effect of variables of interest on outcome variables of interest. In this case our fixed effects are SAIL experience. However, as this is a quasi-experimental design where students are not randomly allocated to treatment groups, there is non-independence with the fixed effects. Random error is used to account for such unobserved variance that affects certain groups in the data. Differences in academic year, level of study, module structure, lecturer style, student characteristics, etc. will likely influence the outcome variables being explored and will vary between SAIL and non-SAIL courses. Accounting for these variables as random effects means that any resulting observed differences can be attributed to the SAIL course with greater confidence. Accounting for non-independent variance in this way is a method that is not available in other analytical strategies.

Model diagnostics

LMM requires that data meet a set of assumptions. Model diagnostics will therefore be conducted preliminarily to confirm that data meet the criteria for LMM, and power calculations will be conducted to explore whether the obtained sample size is appropriate for potential effect sizes.

The following designs will be used to address each research question:

RQ1: Does SAIL impact students' engagement on their course?

$$Y_{ij} = \beta_0 + \beta_1 * SAIL_{ij} + \beta_2 * X_{ij} + Z_j * y_j + \epsilon_{ij}$$

where:

- Y_{ij} as the engagement score for the i-th student in the j-th course.
- $SAIL_{ij}$ as a binary variable indicating whether the SAIL approach is implemented for the i-th student in the j-th course.
- X_{ij} as a vector of control variables that might influence engagement.
- y_j as a vector of course-specific random intercepts.
- β_0 is the fixed intercept, representing the average engagement score in the control group.
- β_1 is the coefficient for the SAIL variable, representing the average change in engagement due to the SAIL approach.
- β_2 is a vector of coefficients for the control variables.
- Z_2 represents the random effects for each course, accounting for course-specific variations.
- $\in_{ij}$ is the error term.

RQ2: Does SAIL impact students' degree outcomes on their course?

$$Y_{ij} = \beta_0 + \beta_1 * SAIL_{ij} + \beta_2 * X_{ij} + Z_j * y_j + \in_{ij}$$

where:

- Y_{ij} as the degree outcome score for the i-th student in the j-th course.
- $SAIL_{ij}$ as a binary variable indicating whether the SAIL approach is implemented for the i-th student in the j-th course.
- X_{ij} as a vector of control variables (see Table 6) that might influence degree outcomes.
- y_j as a vector of course-specific random intercepts.
- β_0 is the fixed intercept, representing the average degree outcome in the control group.
- β_1 is the coefficient for the SAIL variable, representing the average change in degree outcome due to the SAIL approach.
- β_2 is a vector of coefficients for the control variables.

- Z_2 represents the random effects for each course, accounting for course-specific variations.
- $\in_{ij}$ is the error term.

RQ3: Does SAIL differentially impact BTEC+ and international students' grades, relative to those with A-Level qualifications?

$$Y_{ijk} = \beta_0 + \beta_1 * SAIL_{ijk} + \beta_2 * BTEC_{ijk} + \beta_3 * International_{ijk} + \beta_4 * Alevel_{ijk} +$$
$$\beta_5 * SAIL_{ijk} * BTEC_{ijk} + \beta_6 * SAIL_{ijk} * International_{ijk} + \beta_7 * X_{ij} +$$
$$y_{0k} + y_{1k} * BTEC_{ijk} + y_{2k} + International_{ijk} + \in_{ijk}$$

Here:

- Y_{ijk} is the grade for the i-th student in the j-th qualification group (BTEC+, International, A-Level) in the k-th course.
- $SAIL_{ijk}$, $BTEC_{ijk}$, and $International_{ijk}$ are binary variables indicating whether the SAIL approach is implemented, whether the student has a BTEC+ qualification, and whether the student is an international student, respectively.
- β_0 is the fixed intercept, representing the average grade for A-Level students without SAIL.
- β_1 through β_6 are the coefficients of the fixed effects representing the average impact of SAIL, BTEC+, International, and their interactions on grades.
- X_{ij} is a vector of control variables (see Table 6) that might affect the outcome (students grades).
- β_7 is a vector of coefficients for the control variables.
- y_{0k}, y_{1k}, and y_{2k} are course-specific random effects, capturing variations specific to each course.
- $\in_{ijk}$ is the error term.
- Interaction terms in the model ($\beta_5 * SAIL_{ijk} * BTEC_{ijk}$, $\beta_6 * SAIL_{ijk} * International_{ijk}$) allow you to examine whether the impact of SAIL differs for BTEC+ and international students compared to A-Level students. The random effects y_{1k} and y_{2k} capture course-specific variations for BTEC+ and international students, respectively.

RQ4: Does engagement with SAIL impact on students assessment submission habits?

$$Y_{ij} = \beta_0 + \beta_1 * SAIL_{ij} + \beta_2 * SAIL_{ij} + \beta_3 * \left(Engagement_{ij} * SAIL_{ij}\right) + \beta_4 * X_{ij} + y_{0j} + y_{1j} * Engagement_{ij} + \epsilon_{ij}$$

Here:

- Y_{ij} is the assessment submission habit, defined as the amount of time between submitting the substantive summative assessment and the assessment deadline, for the i-th student in the j-th course.
- $Engagement_{ij}$ representing the level of engagement of the i-th student in the j-th course.
- $SAIL_{ij}$ is a binary variable indicating whether the student engages with the Score as I learn (SAIL) approach in the j-th course.
- X_{ij} is a vector of control variables (see Table 6) that might affect the outcome (assessment submission habit).
- β_0 is the fixed intercept, representing the average assessment submission habit when both engagement and SAIL are zero.
- β_1 and β_2 are the coefficients of the fixed effects representing the average impact of engagement and SAIL, respectively, on assessment submission habit.
- β_3 is the fixed effect representing the interaction between engagement and SAIL, allowing you to assess if the impact of engagement differs when SAIL is implemented.
- β_4 is a vector of coefficients control variables that might affect the outcome (student assessment submission habits).
- y_{0j} and y_{1j} are course-specific random effects, capturing variations specific to each course.
- ϵ_{ijk} is the error term.
- The interaction term ($\beta_3 * \left(Engagement_{ij} * SAIL_{ij}\right)$) allows you to examine whether the impact of engagement on assessment submission habit differs when students are engaged with the SAIL approach.

RQ5: Does engagement with SAIL impact grades for engineering students?

$$Y_{ij} = \beta_0 + \beta_1 * X_1\left(SAILEngagement_{ij}\right) + \ \beta 2 \ * X_{ij} + \epsilon_{ij}$$

- Y_{ij} is the assessment grade for the i-th student in the j-th course.
- X_1 is an independent variable that represents engagement with SAIL Courses.
- X_{ij} as a vector of control variables (see Table 6) that might influence dependent variable
- ϵ_{ij} is the error term.

RQ6: Does participation in SAIL impact students continuation from Level 4 to Level 5 and Level 5 to Level 6?

$$Y_{ij} = \beta_0 + \beta_1 * X_1\left(SAILParticipation_{ij}\right) + \ \beta 2 \ * X_{ij} + \epsilon_{ij}$$

- Y_{ij} is the continuation status for i-th student in the j-th course.
- X_1 is a binary variable that represents if a student participated in a SAIL course or not.
- X_{ij} as a vector of control variables (see Table 6) that might influence dependent variable
- ϵ_{ij} is the error term.

The covariates to be included in the model are shown in Table 6 below.

Table 6. Covariates

Covariate name	Type	Levels
Subject area	Categorical	HESA subject level CAH2 grouping
Entry points	Continuous	UCAS tariff points
Programme mode	Categorical	Full-time, Part-time
Commuter status	Categorical	Commuter, Not Commuter
IMD	Categorical	Quintile 1, Quintile 2, Quintile 3, Quintile 4, Quintile 5
Gender	Categorical	Male, Female, Non-binary, Other

Age	Categorical	Young, Mature
Ethnicity	Categorical	Black/Black British, Asian/Asian British, Mixed or Multiple Ethnicity, White/White British
Disability	Categorical	Disability declared, No disability declared
Care leaver	Categorical	Care leaver, Not care leaver

13. Ethical considerations

Ethical approval has been sought and granted by the University of Huddersfield.

The following ethical considerations have been addressed:

Confidentiality and Privacy:

1. Safeguard the confidentiality and privacy of student data. Implement procedures to protect sensitive information and ensure that individual student identities are not disclosed without explicit consent. Use anonymization or pseudonymization techniques when reporting findings to prevent the identification of individual participants.

Data Security:

2. Implement robust data security measures to protect student data from unauthorized access, disclosure, or loss. Ensure that data storage and transmission comply with relevant data protection regulations. Use secure servers, encryption, and access controls to safeguard the integrity of the data.

Minimisation of Harm:

3. Take steps to minimise any potential harm to students. Be mindful of the potential psychological, emotional, or social impact of the research on participants. If there is a possibility of harm, provide adequate support mechanisms and resources for participants.

14. Risks

Part of evaluation	Risk	Mitigation strategy	Risk owner
Ethical approval	Failure to get ethical approval in time - Delay to University of Huddersfield ethical approval would delay starting on data sharing and analysis	• University of Huddersfield to submit ethics early. • evaluator to adapt timeline to conduct evaluations for University of Huddersfield with ethical approval first, freeing	Jarek Bryk

		up time later for those facing delays with ethical approval.	
Data curation	University of Huddersfield does not agree to share required institutional data with independent evaluator - Limited access to some or all institutional data would impact the robustness of the evaluation	• independent evaluator to lead data sharing agreement with each University of Huddersfield and TASO at the outset of the project. • Research protocols developed based on available data. Independent evaluator document if more relevant institutional data is available but not permitted. • Independent evaluator will work flexibly with University of Huddersfield to develop arrangements that work with University of Huddersfield requirements (e.g., temporary staff account for project members requiring data access negates the need for external data sharing).	Jarek Bryk
Data analysis	Institutional Data accuracy is limited – would impact on robustness of findings	• Independent evaluator to maintain honest dialogue with University of Huddersfield on data accuracy. • Recognising the messiness of real-world data, the independent evaluator will make an informed decision about how to balance depth of findings with robustness of data (using data cleaning and conversations to inform appropriacy).	Independent evaluator

15. Bibliography

Kumle, L., Võ, M.LH. & Draschkow, D. (2021) Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods* **53**, 2528–2543 (2021).