Trial protocol Institutional Data Use: Nottingham Trent University – Black Leadership Programme

September 2024

Authors:
Sally Andrews, Staffordshire University
Reagon Alford, Staffordshire University
Juan Raman Mullor, Staffordshire University
Josh Francis, Staffordshire University
Vanessa Dodd, Staffordshire University

QA:
Denise Hawkes, Anglia Ruskin University

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.		
4	5/3/24	Inclusion of MANOVA diagram in Appendix
3	1/3/24	Formatting and typographical changes.
2	13/2/24	Feedback to QA from team included.
1.0 [*original*]	8/2/24	Original version Post QA
Pre-registration		This design has been pre-registered on Open Science Framework (OSF) registry.[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	I am largely supportive of this design, although we need to be clear that it is going to produce evidence of the success of the BLP conditional on being a Black student. We will therefore not be able to look at whether BLP closed the awarding/graduate outcome gaps, without adding in non-Black students for comparison. The use of the PSM and ITS are valuable in making use of the recruitment process to BLP in the design of the study. The one aspect I would like focused on the design is the impact of BLP and the role of gender, thinking about the intersectionality issues for Black Women in Leadership. I would like to see if the programme works to help Black Women explicitly given the intersectionality issues in the labour market when it comes to graduate outcomes later.	4 I would like to see explicit mention of the BLP impact on Black women in relation to intersectionality. Likewise, the focus on mental health would suggest a focus on Black students with Mental Health needs again intersectionality.
Sample size	This is a modest programme, specially in the start. But there is sufficient sample for the models proposed. A consideration of the role of COVID in this timeline is important and maybe something to think about in interpreting the results. We could make more use of the eligible applicants in terms and wonder if we could run models both for application and for attendee, to explore the sample selection explicitly, is it the programme or the willingness to give it a go that matters?	4 Sufficient sample for model proposed
Outcome measure	I am happy with these outcome measures. We will need to be careful in the interpretation of the outcomes based on attending the library and gym use, the distance to campus of where the student lives could be a big factor in these.	4 Sensible outcome measures for this model
Attrition	As this is largely a one year programme this is likely to be less important. This will need exploring in the descriptive statistics of the paper to confirm this	4 Unlikely to be a concern as a one year programme
Validity	The research design is sufficient to ensure validity. It would be nice to build on the findings to see if it did close the awarding gaps too – as without the comparison to non-black students, we are not showing this explicitly.	4 Design is sufficient

Overall	This is a good design, I would like to see it extended to include the comparison to non-Black students to see if there was an impact on awarding gaps.	4 Good design

Response from Evaluation team re intersectionality: These issues never came up when the Theory of Change was developed by the evaluator (in conjunction with the staff at NTU) hence why they were not explored as part of the trial protocol. We agree that this may well have been overlooked and will suggest to NTU that they consider this issue as they refine the programme in the context of the evaluation.

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 post-entry 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 Black Leadership Programme (BLP) at NTU.

Aims

The programme primarily aims to support black and black heritage students' soft skill development (leadership, resilience, career readiness, and confidence), their sense of belonging and mattering at Nottingham Trent University (NTU) to increase engagement with the wider university and academic experience whilst on their course. This increased engagement enables BLP participants to achieve positive student outcomes.

Intervention

The BLP is a year-long, targeted programme primarily for students in their second year of study. A range of complementary activities scaffold support for targeted students in their first and third year of study. The BLP has three strands of community focused, self-focused and development focused activities.

Design

We will apply an ex-post facto quasi-experimental evaluation approach to this study. We will interrogate secondary administrative data using propensity score matching (PSM) and interrupted time series (ITS) methods to explore the impact of the BLP on student engagement and outcomes.

Outcome measures

There is one primary outcome measure, engagement, categorised across four dimensions:

- structured academic engagement (timetabled attendance)
- structured social engagement (clubs and societies)
- unstructured academic engagement (VLE use, library use)
- unstructured social engagement (NTU gym use)

There are two secondary outcome measures, including degree classification and graduate outcome.

Analyses

The analyses will use MANOVA (Multivariate Analysis of Variance) comparing BLP participants with eligible potential participants using PSM and ITS. The impact of BLP on degree outcomes and graduate outcomes will be tested using path analyses. If a relationship emerges between BLP participation and degree outcome and graduate outcome path analyses can help to understand the relative contribution of the different engagement domains through mediation.

2. Background

Organisation	Name	Role and responsibilities
TASO	Dr Rob Summers	Project/Contract Manager
TASO	Luke Arundel	Project Assistant
Staffordshire University	Dr Sally Andrews	Pedagogic Projects Development Manager
Staffordshire University	Reagon Alford	Research Assistant
Staffordshire University	Joshua Francis	Research Assistant
Staffordshire University	Juan Raman Mullor	Evaluation Officer
NTU	Mike Kerrigan	Head of Research and Insights
NTU	Peter Crowson	Research and Evaluation Coordinator
NTU	Reuel Blair	Collaborative Engagement and Retention Team Manager & BLP Coordinator
NTU	Laura Hope	Research and Data Coordinator

3. Aims

The current research's overall objective is to evaluate whether BLP participation has a positive impact on student engagement across four primary dimensions: structured academic, structured social, unstructured academic, and unstructured social. The BLP consists of a core programme of activities in students' second year of study with complementary activities for students in their first year and third year of study. The evaluation aims and research questions are related to the core programme for students in their second year of study.

As part of this research, we also explore the relationship between BLP participation and degree attainment, end of year attainment and graduate outcomes. We have articulated the following research questions and testable hypothesis:

RQ1: Does BLP participation influence students' engagement at NTU?

H_0: Participation in BLP has no relationship with engagement at NTU.

H_1: BLP participants have significantly different engagement scores across four domains of engagement (structured academic, structured social, unstructured academic, unstructured social), compared to those that did not participate in BLP.

RQ2: Does academic engagement mediate the relationship between BLP participation and degree outcomes?

H_0: Academic engagement has no relationship with BLP participation and degree outcomes.

H_1: BLP participation impacts student engagement which in turn impacts degree award.

RQ3: Does social engagement mediate the relationship between BLP participation and graduate outcomes?

H_0: Social engagement levels do not mediate the relationship between BLP participation and graduate outcomes.

H_1: BLP participation impacts employability and structured social engagement levels which in turn impacts graduate outcomes.

We will test these hypotheses through appropriate inferential statistical analysis of the variables and covariates outlined in the tables below. The way in which the following variables and covariates will be used to meet the research aims and answer the core research questions will be discussed in the sections below.

Table 2: Predictor and outcome variables

Variable name	Type	Levels/created from
BLP Group	Categorical	Attendee, eligible applicant, eligible non-applicant, non-eligible
Structured Academic Engagement	Continuous	Course attendance
Structured social engagement	Continuous	Attendance at extracurricular activities

Unstructured academic engagement	Continuous	Library attendance
Unstructured social engagement	Continuous	NTU gym usage
Final degree classification	Categorical	Good degree outcome, Other degree outcome
Graduate outcomes	Categorical	Progressed to graduate outcome[3] as defined by the Graduate Outcomes Survey Did not progress to a graduate outcome as defined by the Graduate Outcomes Survey

Table 3: Covariates

Variable name	Type	Levels/created from
Age	Continuous/discrete	Age of student at enrolment
Sex	Categorical	Male, female, other
Race and ethnicity	Categorical	Black African, Black Caribbean, Black (mixed heritage), Black Other
Commuter status	Categorical	Commuter, non-commuter
Care leaver	Categorical	Care leaver, non-care leaver
Programme	Categorical/Nominal	Input based (unless provided to participant as a selection list)
Programme mode	Categorical	Full time, Part time

---

[3] A graduate outcome is achieved if a student articulates they are in skilled employment or further study as part of a census taken 15 months after graduation.

Employability engagement	Continuous	Visits to employability services

4. Intervention

The BLP was developed in response to research on the ethnicity degree awarding gap produced by Nottingham Trent Students' Union (NTSU). One of the recommendations focused on increasing leadership opportunities for Black students. The programme's core participants are level 5 students, though undergraduate students at all levels may attend events. There are three types of BLP activities:

1. Community focused
2. Development focused
3. Self-focused

Community focused activities

Community focused activities consist of the BLP launch event, the end of year Celebration and several social events scheduled within the academic year. The launch event is a celebration of black heritage where participants hear from inspirational speakers. Social events are delivered separately for the core second year participants which focus on networking with themed games and quiz nights organised for first year students. A celebration event is held at the end of each delivery year, with participants who successfully complete the programme (by attending a minimum of three mentoring sessions) are awarded a certificate and programme success stories celebrated.

Development focused activities

Development focused activities consist of a range of workshops to support BLP participants to develop leadership, confidence, resilience and employability skills. These activities provide participants with an opportunity to connect with leaders, experts and employers as part of the programme. Workshop sessions for this strand include:
- Employability module and workshop
- GRIT personal development workshops
- Leadership workshops with inspirational leaders and experts

BLP participants also receive a skilled dedicated mentor who offers them support throughout the programme.

Self-focused activities

Self-focused activities are designed to enhance representation and increase a sense of belonging for BLP participants. Activities in this strand include Black studies sessions and BICEP mental health support. Black studies sessions are designed to address gaps in Black representation in formal curricular study. BICEP Mental Health support (available to all BLP participants regardless of year of study) are offered to ensure Black students have a safe space in which to talk about mental health issues directly.

4. Design

We will apply a post-hoc evaluation approach to answer the research questions outlined in Section 2. The data is drawn from student records collected between 2019–2023. This study will use matched administrative data with localised BLP engagement data from academic years 2019-20 to 2023-24.

5. Outcome measures

Table 4 lists the primary and secondary outcome measures identified to test our hypotheses.

Table 4: Outcome measures

Outcome measure	Type	Level
Primary: Structured academic engagement	Continuous	Mean average of attendance to lectures, seminars, and workshops on their undergraduate degree
Primary: Structured social engagement	Continuous	Count of extra-curricular and student union activities attended
Primary: Unstructured academic engagement	Continuous	Total time (in minutes) spent in the library and/or total amount of times the library was accessed
Primary: Unstructured social engagement	Continuous	total count of attendance at NTU gym
Secondary: Degree award	Categorical	Good degree outcome (1st, 2.1), Lower degree outcome (2.2, 3rd)

Secondary: Graduate outcome	Categorical	Progressed to graduate outcome[4] as defined by the Graduate Outcomes Survey, Did not progress to a graduate outcome as defined by the Graduate Outcomes Survey

6. Sample selection

The evaluation will use secondary data from current students or graduates of NTU between 2019-20 and 2023-24. Students should identify as having a Black heritage background as part of BLP eligibility criteria. Participation is also subject to an application process due to interest in the programme and the limited spaces available. The application of PSM will enable the creation of the treatment assignment variable. The groupings are as follows: eligible potential applicants and participants.

A breakdown of the current estimated sample size of eligible and non-eligible NTU students for participation in the BLP since its inception in 2020 is provided below. Data from 2019-20 Is included as it includes level 4 engagement data for those students eligible in 2020-21.

Table 5: Sample population breakdown

Sample group	Programme year				
	2019-20	2020-21	2021-22	2022-23	2023-24
Number of eligible potential applicants	n/a	1,086	1,362	1,570	1,618
Total number of eligible applications	n/a	104	103	132	220
Total number of participants	36	53	103	111	153

7. Identification strategy

---

[4] A graduate outcome is achieved if a student articulates they are in skilled employment or further study as part of a census taken 15 months after graduation.

Sample participants will be identified as either: participants or eligible potential applicants. Participants are those who actively applied for and participated in BLP, while eligible potential applicants are Black and Black heritage students who did not participate in BLP. These may include applicants and non-applicants.

Eligible potential applicants are not an ideal comparator group for two main reasons. Firstly, applicants have a self-selection bias; they have already demonstrated an intrinsic motivation to benefit from the BLP, which non-applicants do not share. However, there are not enough eligible applicants to create a more suitable comparator group.

Secondly, where there were more applicants than places available, students were selected based on loose selection criteria from the BLP team. This included engagement prior to BLP. This means that BLP participants may have lower engagement than eligible potential applicants.

A combination of Propensity Score Matching (PSM) and Interrupted Time Series (ITS) design will be used to control for confounds across the five research questions. Specifically, PSM will facilitate a comparator group to be generated from students with similar Level 4 engagement characteristics. ITS will be used for the comparison of data pre-and post-BLP, specifically engagement at Level 4 relative to engagement at Level 5 and Level 6. This means that changes in student engagement can be observed within-student, which offers greater potential to attribute causality to the BLP.

8. Data collection

Data will be obtained from NTU's administrative records on students between 2019–23. No data will be collected by the researchers at any point.

9. Procedure

A high-level project timeline can be found in Table 6.

Table 6: Project timeline

Timeline	Action
October 2023-January 2024	- Set up data sharing process and agreement - Conduct enhanced theory of change workshop - Achieve ethics approval - Complete draft enhanced theory of change - Complete Trial Protocol
February 2024 – March 2024	Analyse data and deliver final report

10. Power calculations

As prior research has not established the common effect size of interventions like this, certain assumptions must be made; this includes the effect, which will be tested three times to determine the required sample size to observe small, medium and large effects. According to Cohen (1988), this equates to f^2 = .02, .15 and .35, respectively. This covers the first and primary research question. This will be analysed using a 2 x 3 MANOVA (multivariate analysis of variance) with 4 outcome variables.

As mentioned, the overall power analysis will be based on a multivariate regression model which will address the first research question (see section 12 for more details). This decision is motivated by it being the evaluation's primary analysis.

The power analysis was run via the G*Power software, with the settings as follows:

Test family: "F tests"

Statistical test: "MANOVA: Global effects"

Approximation (F-Transformation): "Pillai Trace"

Type of power analysis: "A priori: Compute required sample size – given α, power and effect size"

Effect size f^2(V) = .02, .15, .35

α err prob = 0.05

Power (1 – β err prob) = 0.80

Number of groups = 12

Response variables = 4

As the researchers have no control or influence over the total sample population, nor the allocation to the intervention or comparator groups, only the total required sample will be reported instead. The Critical F and Actual Power will also be reported. The table also gives values assuming a 20% loss (due to missing data) to the minimal required sample, in the event that the sample provided is less than the sample size desired by the evaluators. This is calculated using the same parameters as above, with the "Post hoc: Compute achieved power" type of power analysis option selected and 80% of the minimum required sample size for each given effect size (f^2).

Table 7: Power calculation

f²	Minimum required sample size	Critical F-value	Pillai Trace (V)	True Power (1 – β err probability)
0.02 (small)	270	1.38	0.08	0.81
20% sample loss		1.58	.08	0.68
0.15 (medium)	42	1.64	0.52	0.84
20% sample loss		1.68	0.52	0.63
0.35 (large)	24	1.71	1.04	0.92
20% sample loss		1.79	1.04	0.74

The researchers have no control over the final sample size, nor any ability to collect more data, as it is historical institutional data provided by relevant data providers. A post-hoc power analysis will also be conducted to accurately determine the power once data are curated.

11. Analytical strategy

Three analyses will be conducted, a primary analysis and two secondary analyses. As mentioned in the previous section, the foremost analysis will be a 2 (BLP participation; participant, eligible non-participant) x 3 (level of study; level 4, level 5, level 6) MANOVA with four outcome (*response)* variables measuring engagement (structured academic; unstructured academic; structured social; unstructured social).

Level of study is a within-subjects factor, while BLP participation is a between-subjects factor. The 2 x 3 mixed-design MANOVA was chosen as it provides the ability to infer the effect of BLP attendance by comparing the trajectory of engagement for BLP participants relative to non-BLP participants. This is possible as the BLP begins at level 5, enabling level 4 engagement to be used as a comparator reference with non-participants. As this explores the trajectory of individual students' engagement, it avoids issues of confounding variance.

An illustration of the model can be found in Appendix 1: 'MANOVA Diagram'. This model will be used to address the first research question.

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i + \beta_3 X_i + \epsilon$$

- Y_i is a matrix of the four outcome (response) variables (structured academic; unstructured academic; structured social; unstructured social).
- β_0 is the intercept
- β_i are matrices of regression coefficients.
- X_i are matrices of the two predictor variables (BLP participation; accepted applicant, non-accepted applicant and level of study; level 4, level 5, level 6).
- ϵ Is a matrix of residuals.

Two path-analyses will be run to answer the remaining two research questions. These path analyses will allow for a nuanced and robust analysis of whether BLP participation impacts Final degree award and graduate outcomes, whilst accounting for other mediating variables. On the surface these may seem to play a similar role to ANCOVA, but path analyses can be used to measure the effect of the primary predictor on the covariates, and then their effect on the outcome variable, as well as the primary effect of the predictor variable on the outcome.

Whilst ANCOVA is specifically designed for comparing group means while accounting for the effects of covariates, path analysis is suitable for modelling complex relationships and understanding paths within a broader theoretical framework.

The first path analysis (see Figure 1) will have four variables: BLP attendance, structured academic engagement, unstructured academic engagement, and final degree award. If a relationship is found between BLP attendance and final degree award, then interaction (mediation) effects will be explored for structured and unstructured academic engagement. This model will have a path going directly from BLP attendance to the final degree award, a path from the BLP attendance to structured academic engagement and unstructured academic engagement, which each their own path towards degree outcome.

$$Y_i = \beta 0_i + \beta 1_i + \beta 2_i + \beta 3_i + \epsilon$$

- Y_i is Final Degree Award
- $\beta 0_i$ is the intercept
- $\beta 1_i$ is BLP Participation
- $\beta 2_i$ is Structured Academic Engagement
- $\beta 3_i$ is Unstructured Academic Engagement, and
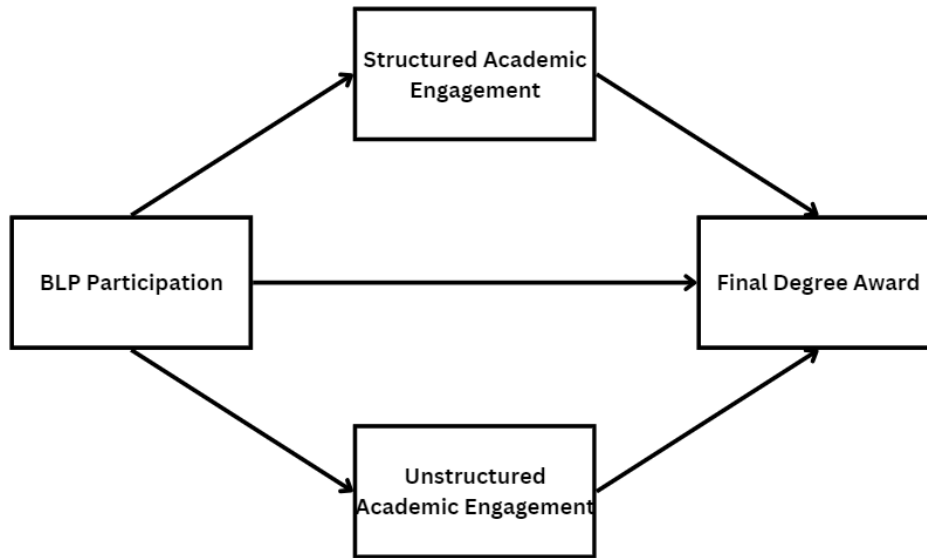- ϵ is a matrix of residuals.

Figure 1: Path analysis structure of BLP attendance to final degree award

The final analysis will be similar to the previous path-analysis model. However, the mediators will be changed to structured social engagement and employability engagement (Figure 2), and the outcome will be changed to graduate outcome.

$$Y_i = \beta 0_i + \beta 1_i + \beta 2_i + \beta 3_i + e$$

- Y_i is Graduate Outcome
- $\beta 0_i$ is the intercept
- $\beta 1_i$ is BLP Participation
- $\beta 2_i$ is Structured Social Engagement
- $\beta 3_i$ is Employability Engagement, and
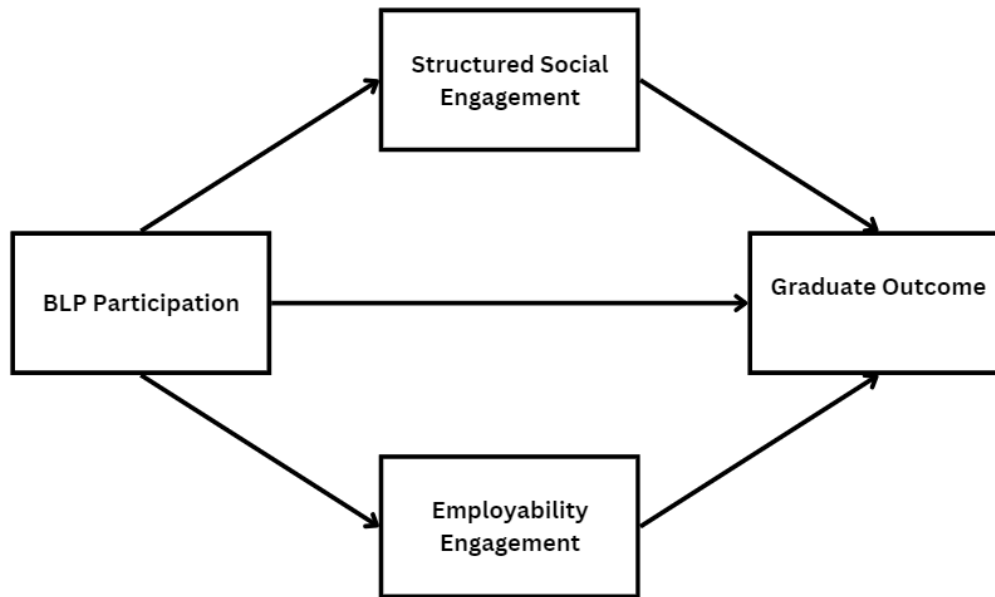- e is a matrix of residuals.

Figure 2: Path analysis structure of BLP attendance to graduate outcomes

12. Multiple Comparisons

Due to the need for multiple tests to be conducted upon the same dataset in order to address the multiple hypotheses adjusting for these comparisons will be necessary. Conducting multiple testing upon related or connected data increases the likelihood of rejecting the null hypothesis incorrectly (increased Type I error/false discovery rate). We will not need to adjust during the MANOVA analysis, as the error rate is adjusted as part of the analysis, but will need to adjust for the multiple ANOVAs conducted and path analysis. As such we will utilise the Benjamini and Hochberg method for controlling false discoveries. The procedure to do this correction is the following:

1. Perform all required statistical test required, and record the p-value for each.
2. Arrange the recorded p-values in ascending order, assigning a rank to each (1 for the smallest p-value, 2 for the one immediately larger and so on).
3. Utilise the following equation to find each p-values Benjamini-Hochberg critical value:

BH Critical Value = $(i/m) * Q$

- i = the rank of p-value
- m = total number of p-values recorded
- Q = the chosen false discovery rate (commonly 5% in line with 0.05)

17

4. Compare the *p*-values recorded against the critical value, note the largest *p*-value below the critical value and take every *p*-value below this to be significant.

13. Ethical Considerations

This project has received ethical approval from Nottingham Trent University's ethics committee. The following ethical considerations are key to the research and an approved ethics submission will be submitted to provide further detail:

Confidentiality and Privacy: We will safeguard the confidentiality and privacy of student data in line with GDPR (2016) regulation. In addition, the providers' privacy notice informs students that their administrative data may be used for research and evaluation purposes. We have implemented procedures to protect sensitive information and ensure that individual student identities are not disclosed without explicit consent. Data owners developed robust anonymisation protocols prior to disseminating data to evaluators. These protocols prevent the identification of individual participants when conducting analyses and reporting findings.

Data Security: Data owners and evaluators have implemented robust data security measures to protect student data from unauthorised access, disclosure, or loss. Data will be shared using secure servers, encrypted data files, and two factor authentication access controls to safeguard the integrity of the data.

Minimisation of Harm: We have taken steps to minimise any potential harm to students through the procedures outline above. This research will be undertaken using large scale secondary datasets which reduces the probability of identification. We will not report descriptive statistics on control or covariate data where counts are considered low (n<15) and will aggregate data where necessary. For example, it may be appropriate to report on ethnicity using the aggregate groupings Black, Asian, mixed ethnicities and white rather than disaggregating this data into more granular groupings.

14. Risks

Part of evaluation	Risk	Mitigation strategy	Risk owner
Ethical approval	Failure to get ethical approval in time - Delay to NTU ethical approval would delay starting on data sharing and analysis	• NTU to submit ethics early • evaluator to adapt timeline to conduct evaluations for NTU with ethical approval first, freeing up time later for those facing delays with ethical approval	Mike Kerrigan

| Data curation	NTU 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 NTU 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)	Mike Kerrigan
Data analysis	Institutional Data accuracy is limited – would impact on robustness of findings	• Independent evaluator to maintain honest dialogue with NTU 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)	Juan Raman Mullor

15. Appendix 1: MANOVA diagram
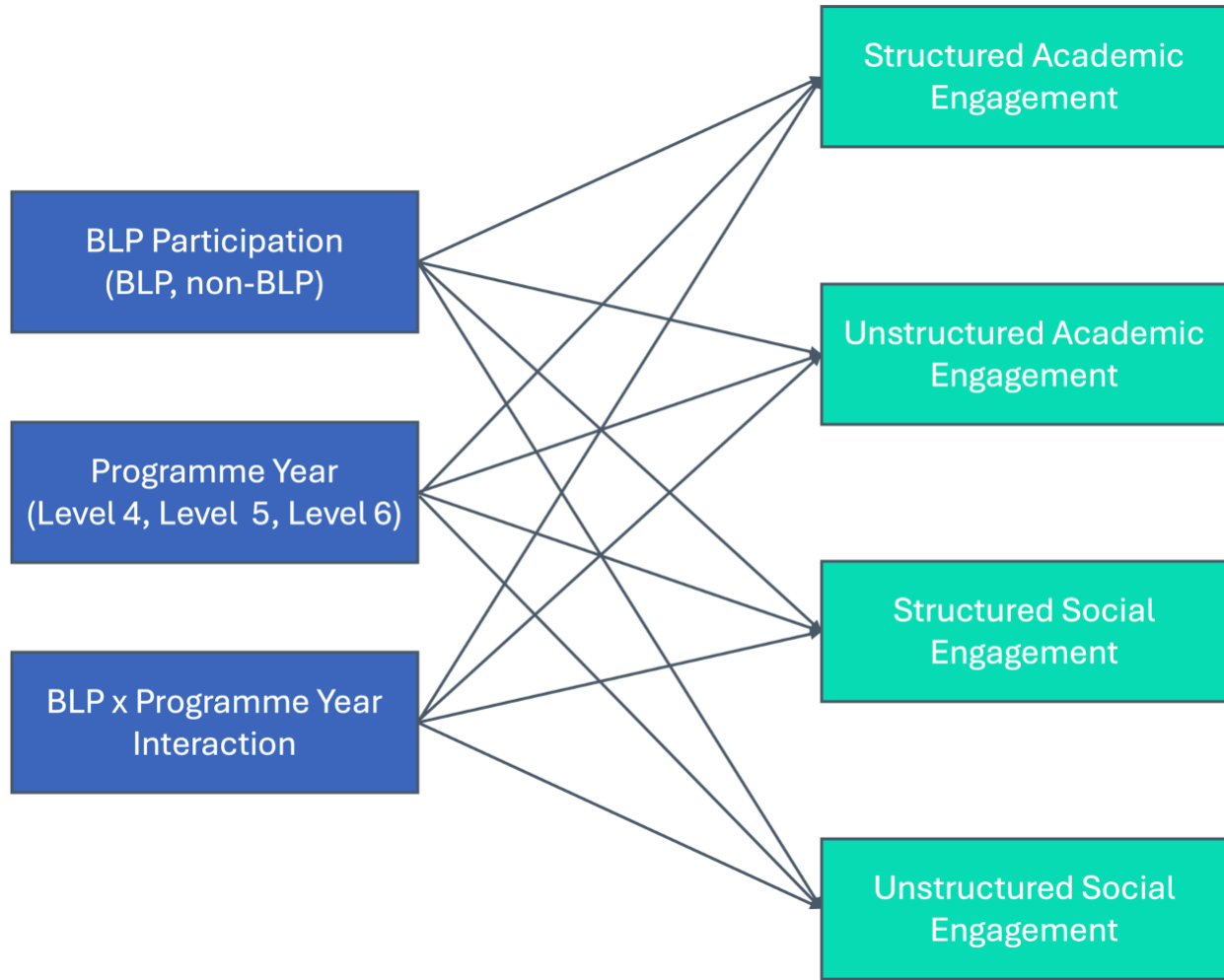
Figure A1: MANOVA diagram illustrating the model for the 2x3 Mixed design MANOVA.