

Pre-analysis plan: Strengthening Students' Resilience

Note: this document will be pre-registered prior to trial launch and then published once the trial report is released. Therefore it will be circulated to relevant partners for comment before finalisation.

Policy problem, trial aims and research question

The Try, Test and Learn Fund supports innovative projects that aim to boost employment among groups of people at high risk of long-term welfare dependence. One of the groups identified by the Fund is young students from disadvantaged backgrounds who are at risk of dropping out of their post-secondary studies. The Strengthening Students' Resilience program seeks to support these students to stay connected to their education, complete their study and attain employment.

Our research question is whether a mobile application that teaches students how to buffer themselves against the stressors of university life through their social group identities can improve students' completion of tertiary studies and academic performance. A secondary question is whether the app improves re-enrolment rates in the subsequent semester.

Outcome measures

The primary outcome measure for academic performance is the average subject marks received for units studied in semester 1, 2020. Subject marks are expressed as a score out of 100.

The primary outcome measure for completion of tertiary studies is the completion rate of study units that students are enrolled in beyond the census date in the 2020 autumn semester (31 March for WSU, 20 March for UON).

We define 'completion' of an academic unit as any instance where students successfully complete a course. Non-completion is defined as any instance where students withdraw from a course after the census period, achieve a 'failing' grade for a particular course, or have not completed requirements to pass. The full protocol for constructing the completion outcome variable is included at Appendix A.

As a sensitivity analysis, we will use a slightly different measure of completion based on all courses in which students have enrolled for the trial semester. This will allow us to investigate differences in course selection prior to census that may be affected by the app.

Outcomes will be analysed at the “study unit” level (treating students as clusters) — as opposed to the student level.

Inclusion criteria

The trial will not be available to all students at participating universities. To be eligible, a student must: be an undergraduate, be enrolled at a participating university, and have their student number on the whitelist (for validation of student IDs by the app).

The Intervention

The treatment group will receive access to an app called *Grok*, which involves students maintaining a zen garden that is a visual representation of their social connections, both within and beyond university. The app is designed to shift student’s mindset over the course of around 16 weeks to develop better means for coping, reacting to academic stressors.

Students must complete different social, wellbeing as well as academic exercises each week in order to maintain different elements of the garden. Each element of this garden is directly linked to an outcome designed to minimise negative attributions about one’s performance at university.

The app will:

- Educate students on the importance of social groups for their sense of belonging and their overall health and wellbeing at university;
- Encourage students to identify their most important and current social groups;
- Provide students with opportunities for reflection and affirmation of these social group identities throughout the semester where they are relevant;
- Help students to foster new identities that may improve their identification with and thus, retention at university in the long term; and,
- Provide students with study tips.

The ‘control’ group will receive information about the key student support services available to students at their respective universities.

Hypotheses

The hypotheses for this trial are:

H1. Proportion of completed study units: treatment group > control group.

H2. Average subject mark received: treatment group > control group.

For both hypotheses we have chosen a one-sided test because it is unlikely that the app intervention will reduce the completion rate of study units, or reduce average subject marks.

Sample size and power calculations

We present power calculations for the two primary outcomes below. Power calculations were conducted in R Studio (R version 3.6.0, R Studio version 1.2.1335). Calculations are based on marks for study units undertaken by our population of interest at Western Sydney University in semester 2, 2018.

Calculations were conducted at the study unit level, meaning the study load of each student represents a 'cluster' of units that is taken into account when calculating the required sample size. Cluster sizes vary in line with the number of study units a given student undertakes in a semester. We estimate that the average study load is 2.99 study units, with a standard deviation of 1.02 and an intra-cluster correlation coefficient of 0.46 for the completion outcome and 0.68 for the marks outcome.

Subject marks

For the 'performance' outcome measure, our power calculations suggest that we need 1,842 participants to detect an increase of around 2 subject marks (a Cohen's *d* of 0.1) at an 80% level of statistical power (treatment proportion = 0.5, alpha = 0.05, one-tailed test).

Completion

For the 'completion' outcome measure, our power calculations suggest that 6,602 participants will detect an increase in the completion rate from approximately 80 per cent to approximately 82 per cent at an 80% level of statistical power (treatment proportion = 0.5, alpha = 0.05, one-tailed).

Trial design

The intervention will be evaluated through a multisite two-arm randomised controlled trial (RCT) in two public universities. The target population for the trial is undergraduate students studying at the University of Newcastle (UON) and Western Sydney University (WSU) during semester 1, 2020. Students from all campuses at both universities will be invited to participate, though marketing and outreach activities will focus primarily on three UON campuses (Callaghan, Ourimbah and NewSpace – a total of over 20,000 students) and four WSU campuses (Parramatta, Parramatta CBD, Penrith and Campbelltown – a total of around 16,000 students).

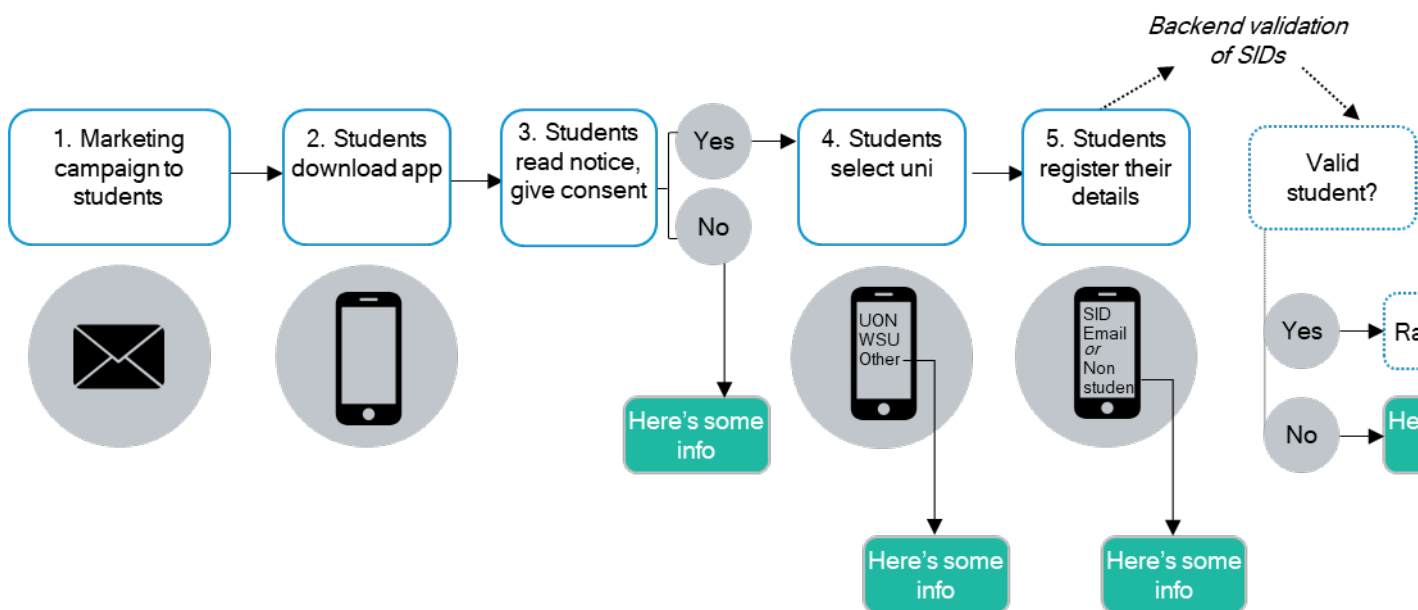
The experimental sample will consist of undergraduate students who register their interest in participating in the research. All eligible participants from UON and WSU

will be emailed an invitation to participate in the trial, which will direct any interested students to the app.

Our goal is to recruit at least 3,000 students at each university. This will give sufficient power to detect plausible effects for academic completion within each university population.

Once registered, participants will be randomly assigned “on-the-spot” to the treatment group or control group with a 50% chance of being assigned to treatment. Randomisation will take place at the student-level, and will occur when an eligible student successfully downloads the app, consents to participation in the trial, and enters a valid student identifier and email address. The randomization will be implemented in the app’s Application Programming Interface (API) as part of a students’ registration and onboarding process (see figure 1).

Figure 1. Student registration and onboarding process



In addition to the results of the RCT, we will also analyse app usage data in a descriptive fashion to determine how students are using the app and to identify areas for improving the app.

Trial threats

The most significant threat to this trial relates to the recruitment and adherence of participating students. To maximise student recruitment for the trial, we will deliver an extensive marketing campaign supporting the launch of the app, along with incentives for students who sign up to the research. Incentives will include free merchandise (e.g. stationery that is university-branded and/or app-branded) for downloading the app. Students who are allocated to treatment and control groups

will also enter a prize draw to win up to five cash vouchers valued at up to \$500, per university. Small cash vouchers will also be offered for survey completion and focus group participation.

We have worked with our app developers to carefully design *Grok* to maximise adherence among app users. *Grok* will include 'gamified' elements to encourage students to continue to use the app, including push-notifications to remind students to check the app, and reward systems and collectible items linked to completing the *Grok's* social, wellbeing and study-related exercises.

This trial is unblinded, meaning participating students will be aware that they are participating in pilot-testing a new app designed to improve their university experience, and will be randomly allocated to either a 'user' group or a 'non-user' group.

The risk of spillovers between app users and non-users participating in the trial is relatively low. Some app activities that build social connections at university and engagement in learning may have positive effects on students that associate with the app user, but these effects are unlikely to provide enough stimulus to affect academic performance or completion rates of students in the 'non-user' group, or the broader university population.

There is a small risk that some student might find a way to subvert the randomisation. Student who are randomised to control may enter more student IDs in an effort to be randomised into treatment, or use a friend's ID to access the app. The validation of students IDs prior to randomisation will help prevent these issues.

Missing data strategy

Covariates

Missing data are not expected for covariates as universities are required to collect this information for the national datasets. In the event that some data are missing, we will use missingness dummies for binary variables and a grand mean imputation for continuous covariates.

Completion outcome

For the academic completion outcome the risk of missing is low. The universities assign grades to every unit at the end of the semester. Each of these will be recoded into 'complete' or 'incomplete' in accordance with the business rules for each university. To minimise the number of pending grades, we will delay the data extraction until after all supplementary examinations have been completed and graded. Any remaining pending grades will be coded as 'incomplete' in line with business rules.

Achievement outcome

For the academic achievement outcome, there is a greater risk of missing data. In Situation 1, the course is graded as Pass/Fail and students do not receive a mark. These data are technically censored. This type of missing data will be independent of treatment status and while it may lead to reduced power it will not introduce bias.

In Situation 2, a student may fail but the lecturer does not record a mark (or decides to record marks for 'near misses' but not low fails). In this instance, had the student not failed a mark would have been recorded. Given the app is designed to improve marks, there could be a relationship between treatment status and missingness.

In either case, we will assess missingness and try to ascertain if there is a relationship with treatment status. If there is less than five per cent missing, we will conduct a complete case analysis. If more than five per cent of students have missing data, we will still report a complete case analysis, however, we will run additional analyses to assess the type of missingness and may use techniques such as bounds and multiple imputation to inform our interpretation.

Method of analysis

For both primary outcome measures, we will estimate treatment effects using Ordinary Least Squares (OLS) regression with standard errors clustered at the student level using the following specification:

$$Y_{ij} = \alpha + \beta T_i + \gamma X_i + X_i * T_i + v_i + \omega_{ij}$$

Where Y_{ij} is either a binary variable indicating whether student i completes unit j , or the raw mark received by student i for unit j during semester 1, 2020 (out of a possible 100 marks). T indicates whether student i is allocated to the treatment or the control group, X_i indicates the set of de-meant baseline covariates, $X_i * T_i$ is the interaction of these covariates with the treatment indicator, v_i is a student-level error term and ω_{ij} is the study unit-level error term.

The following baseline covariates will be included in our primary analyses: student age, gender, full-time/part-time status, university, and socioeconomic status.

Given our completion outcome is binary, we will conduct logistic regression as a robustness check for our primary OLS specification and report average marginal effects. As a second robustness check, we will conduct student-level analyses, as opposed to study-unit-level analysis.

If our hypothesis tests fail to reject the hypothesis that the app has no effect, we will conduct an equivalence test with a smallest effect size of interest (SESOI) as determined with university partners.

Randomisation checks

We will conduct a post-trial balance check in which treatment status will be regressed on all pre-treatment covariates and assess the model for joint orthogonality. A p-value < 0.01 will prompt a review of the random assignment procedure and possible data-handling mistakes. If the review finds no errors, we will report the imbalance test and proceed on the assumption that the imbalance is due to chance.

Interpretation of results

We conduct statistical tests both for public reporting purposes and to inform future government decision making. For public reporting, we will be guided by the effect size and by the conventional threshold for 'statistical significance' used in the social sciences ($p < 0.05$).

For decisions about whether to promote and improve the app in future, however, we will make a judgement about the likelihood of an effect (of a certain size) from the app and weigh this against the costs of maintaining and promoting the app. In practice, this could mean we recommend continuing with the app if the possible effect size is large even if the p-value is larger than 0.05. Alternatively, if the effect size is small (eg, less than one per cent), we may not recommend using the app even if that effect is 'statistically significant'.

Primary subgroup analyses

University: We will test whether there is a difference in effect between the two universities. Our hypothesis is that there will be no difference in effect so we will use a two-sided test.

Student disadvantage: We will test whether there is a difference in effect by socio-economic status (SES). We will use the binary low-SES indicator in the university administrative data. This represents the bottom quartile of Index of Education and Occupation (IEO), one of the indices within the Socio-Economic Indexes for Areas (SEIFA). Our hypothesis is that the intervention will be more effective for low-SES students as compared with other students. We will use a one-sided test for this hypothesis. We will conduct a secondary analysis by using the IEO to create SES terciles and then testing for effects in each tercile and for differences between terciles.

Secondary and exploratory analyses

Subgroup analyses

We will conduct secondary subgroup analyses by domestic/international status, full-time/part-time status, attendance mode (e.g. on-campus or distance) and 'first-in-family' status.

‘Dose response’ analysis

We will conduct dose-response analyses for our two primary outcome measures based on the duration of a student’s active app usage during the semester. We will analyse dose using three approaches: number of weeks with at least one completed activity, mean weekly number of activities completed, and number of reflection exercises completed.

Secondary outcomes

Secondary outcome measures for the trial relate to: withdrawal rates (as opposed to non-completion rates), student re-enrolments in semester 2, and students’ subjective well-being and identification with university life.

Withdrawal rates are described in the ‘Primary outcome variables’ section above. The re-enrolment rate is measured as the group attrition rate of students from semester 1, 2020 to semester 2, 2020. Administrative data on student retention will be available from census dates in Semester 2, 2020.

Measures of subjective wellbeing will be compiled through a survey launched at the end of the semester’s teaching period, questions are currently being finalised, but will be two to three validated scales.

Qualitative Research

In addition to the RCT, we will undertake qualitative research through surveys, focus groups and/or interviews with staff and students discussing what prompted them to download and use the app, and their experiences using it. Research involving university staff will involve discussions on whether they observed differences in the level of engagement among students during the semester. This research will be conducted at both UON and WSU in the 2020 mid-year break.

Pre-analysis plan commitments

Although trial launched prior to the publication of this plan, no data has been received by BETA and no analysis has been undertaken. We will be transparent about, and provide justification for, any deviations from this plan.

Appendix A: Protocols for the construction of outcome variables

Binary variable: complete/ incomplete

To construct the outcome variable 'complete', we will use administrative datasets from WSU and UON and recode course outcomes into a binary variable using the following protocol:

Table 1: Conversion of grades to binary completion

WSU	UON
High Distinction, Distinction, Credit, Pass grade	High Distinction, Distinction, Credit, Pass grade
F (Fail) – student has shown unsatisfactory performance.	UP (Ungraded Pass) – student has met level requirements to pass the course. Applies for courses for which only a pass or fail is available.
CF (Compulsory Fail) – student failed a threshold assessment component of the unit other than the practicum, but has achieved a total mark of 50 or more for the unit.	FF (Fail) – student has failed to satisfactorily achieve learning outcomes. If all compulsory course components are not completed, the mark will be zero.
FNS (Fail Non Submission) – student has not officially withdrawn from the unit and has not completed one or more of the mandatory assessment requirements for the unit.	
<i>Ungraded assessments/units</i>	
S (Satisfactory) – student has met defined criteria within specified time.	
U (Unsatisfactory) – student has not met defined criteria within specified time.	
PF (Practicum Fail) – student has failed to satisfy criteria for practicum assessment and will not progress further in the unit (and course if the unit is compulsory). The student will not be	

WSU	UON
allowed to re-enrol in the unit or its equivalent and therefore will be unable to meet the requirements of the course.	
W (Withdrawn without academic penalty) – student provides evidence of serious illness or misadventure after census date.	
Y (Continuing unit) – unit covers more than one teaching session and final assessment has not been made.	
<i>Temporary grades (finalised before the following census date)</i>	
I (Incomplete) – assessment tasks incomplete due to illness or misadventure.	
J (Deferred Exam) – due to illness or misadventure, with formal approval granted for the student to sit an alternative to the formal scheduled exam.	
N (Result Pending) – grade still to be finalised.	
R (Re-assessable Fail) – allows for a further assignment or other work, and must be converted to Pass or Fail.	
<i>Administrative grades (no marks are given)</i>	
E (Fail – Discontinued) – student has withdrawn from unit after census date without authorisation.	<p>I (Incomplete) – not all information is available to calculate a mark at this stage. Reasons might be student needs to provide an assessment requirement or lecturer has yet to finish marking.</p> <p>I results are automatically converted to FF after 90 days.</p>

WSU	UON
K (Academic Credit – Specified) – academic credit for specified core or elective units.	S (Special Consideration) – result is pending the outcome of supplementary exam. S results are automatically converted to FF after 180 days, if result is not amended.
L (Academic Credit – Unspecified) – academic credit for non-specified core or elective units.	WW (Withdrawn No Penalty) – student withdraws from a course by census date, or has permission to withdraw without penalty.
Z (Aegrotat Pass) – awarded on compassionate grounds because of inability to complete unit.	NA (Grade Not Applicable) – used only for some components of Research Higher Degrees and multi-term sequence courses.
	- (Grade or score not applicable)
	EX (Credit Awarded) – credit has been approved and added to student’s record.
	CF (Component Fail) – student has failed one or more compulsory components of a course. CF grade appears on internal transcripts; FF grade will appear on external transcripts.

Note: Green indicates the grade will be coded as complete, red indicates the grade will be coded as incomplete