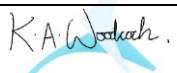
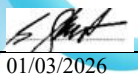
Statistical Analysis Plan (SAP):
Study Title

Three cluster randomised controlled trials in Portugal, Czechia and the United Kingdom to test the effectiveness and cost-effectiveness of an augmented social play intervention in schools to support real-world connection and sense of belonging (ASP#1 trial)

Version: 1.0

Date: 01 March 2026

Person(s) contributing to the analysis plan	
Name(s) and position(s)	
Authorisation	
Position	**Chief or principal investigator**
Name	Kate Woodcock
Signature	
Date	25/02/2026
Position	**Senior trial statistician**
Name	**Beth Stuart**
Signature	
Date	01/03/2026

.

Not to be used without prior permission by the PCTU

1. Administrative Information

1.1 Trial information:

The ASP trial has ISRCTN registration number ISRCTN11613092. This SAP was written in conjunction with protocol version 0.2 (date 27/02/2025). This SAP will be amended and updates listed in section 1.2, if impacted by any subsequent amendments to the protocol.

1.2 SAP revision history

Protocol version*	Updated SAP version no.	Section number changed	List/description of changes from previous version/protocol	Justification for revision†	Author of change	Date

1.3 Members of the writing committee

Beth Stuart (QMUL, PCTU) wrote the Statistical Analysis Plan with input from the chief investigator and study team. Beth Stuart will be responsible for writing the computer code to implement the procedures described in the SAP and implementing the strategy at the point of analysis. If decisions are required during the course of the analysis they may be discussed with a statistician within the PCTU, independent of the trial

1.4 Timing of SAP revisions in relation to unblinding of data/results

Version 1.0 of the SAP was written whilst the statistician and trial team had no access to unblinded trial data or to interim trial results.

The trial statistician will perform the analysis unblinded. Any potential deviations from the agreed sap can be discussed with the independent oversight committees, or a statistician independent of the trial, to ensure decisions are not influenced by the data or emerging results. The primary analysis will be independently checked against the finalised SAP to ensure it has been carried out in accordance with it. “

1.5 Timing of statistical analysis

The statistical analysis is conducted once the SAP has been signed off, the last participant has completed the last follow-up, the data has been cleaned and the database locked.

1.6 Remit of SAP

The purpose of this document is to provide details of the statistical analyses and presentation of the quantitative results of the ASP Trial including the primary objective, secondary objectives and exploratory objectives as described in the protocol version 0.2, 27/02/2025. The following aspects of analysis work described in the protocol are not covered by this SAP: the economic evaluations and the process evaluations.

Subsequent papers of a more exploratory nature will not be bound by this SAP but will be expected to follow the broad principles laid down in it. Any deviations from the SAP, or exploratory, post hoc or unplanned analyses will be clearly identified in the respective study analysis report."

1.7 Public availability of the SAP

The SAP will be made available on the trial registration webpage.

1.8 Statistical software

The analysis will be carried out in Stata v.18 or above.

2. Introduction

2.1 Background and rationale

Study objectives	The primary aim of the trials is to assess the effectiveness of the ASP#1 programme (called *Lina*) in improving children's sense of belonging within a classroom environment in Portugal, Czechia and the United Kingdom
Study design	Cluster randomised, parallel group controlled, school-based trial in three countries.
Setting	Three separate trials will be conducted concurrently in schools in Portugal (PT), Czechia (CZ) and the United Kingdom (UK).
Participants	Schools will be included as 'trial schools' if they: - are a school in Portugal, Czechia or the UK who have classes of students primarily aged 12-13 - agree to support the delivery of the ASP#1 intervention (*Lina*) as part of routine school activities in the 2025/26 academic year - agree to the process of randomisation, with the understanding that classes which are randomly allocated to the control group and will have to wait until the 2026/27 school year to implement the ASP#1 intervention Individual young people will be included in the trial if: - they are 12-13 years old or in the appropriate class year group (as per ethical approvals in each country) - they are part of a class with similarly aged students - assent to the collection of personal data at key timepoints through the trial (baseline, 6-10 weeks post baseline, and 6 months baseline)
Interventions	Control group classrooms will receive routine lessons based on the established local curriculum as normal. The intervention under investigation (ASP#1, called *Lina*) is a newly developed software designed to enhance students' sense of social belonging through a blended approach of interactive gameplay facilitated by smartphones in a classroom setting. All necessary equipment, including smartphones, routers, headphones, and markers will be provided by the trial team to the class. The intervention is primarily user-driven but will be guided by the classroom teacher.
Primary outcome measure(s)	The primary outcome will be sense of belonging/ class climate as measured by the Delaware School Climate Survey- Student (DSCS-S) at 6-10 weeks post intervention

2.2 Objectives

The primary objective of the trials is to assess the effectiveness of the ASP#1 programme (called *Lina*) in improving children's sense of belonging within a classroom environment in Portugal, Czechia and the United Kingdom.

Secondary objectives of the study are:

- To test the impact of *Lina* on improving young people's mental health and wellbeing (more broadly and, specifically, in terms of depression and social anxiety)
- To test the impact of *Lina* on other secondary outcomes including social awareness, loneliness and young people's stigmatising beliefs.
- To collect the costs of delivering the intervention in comparison to routine school costs, and conduct a comprehensive economic evaluation to establish whether *Lina* presents good value for money. (To be detailed in a separate Health Economics Analysis Plan)

3. Study methods

3.1 Trial design

The design of the study is a cluster randomised, parallel group controlled, school-based trial in three countries. Three separate trials will be conducted concurrently in Portugal (PT), Czechia (CZ) and the United Kingdom (UK). Overall, these three parallel trials will employ the same design, although there will be some differences to accommodate local contexts- where these differences are planned, they will be highlighted in the protocol.

The intervention provided will replace usual lessons within the school term for a period of six lessons and will be facilitated by a teacher. Outcome data will be collected at baseline and immediately following intervention completion (6-10 weeks post-baseline), and at 6 months follow up.

Trial data will be analysed separately in the first instance (i.e., country specific analysis) but will be later pooled to look for overall patterns in the data, particularly for the primary outcome (i.e., class belonging). Pooled data analysis will seek to determine if trial findings are replicated across different European national contexts, and to provide a larger, more varied sample for analysis of effectiveness.

3.2 Planned interim analyses

There are no formal interim analyses planned. The trial oversight committee will undertake confidential review of accumulating data, and will advise if, in their view, the trial has indicated 'proof beyond reasonable doubt' that the intervention is clearly indicated or contraindicated in terms of efficacy or safety."

3.3 Randomisation procedure

The trial will employ a cluster randomised design, with class being the unit of randomisation. Whole classes will be randomised to either the *Lina* or the control arm.

Randomisation will be stratified by country, and then blocked.

Randomisation will occur only when: schools have agreed to be part of the trial and appropriate classes that meet the inclusion criteria have been identified. The allocation ratio will be 1:1.

3.4 Sample size calculation

Based on an assumed small to medium effect size (Cohen's d = 0.4), a conservative ICC of 0.1, which takes into account the clustering at the class and school level, and a design effect of 2.7, 40 classes per country will provide 90% power (alpha = 0.05).

The analysable trial sample assumes an average cluster size of 18, (a total of 720 participants per country). To allow for an anticipated drop-out rate of 20%, an average cluster size of 22 will be recruited at baseline (i.e., average 22 adolescents per classroom), thus giving a total sample size of 880 per country. The cluster size of 22 at baseline takes into account the variability in class sizes within and across each country, and expected number of participants who opt out of the study.

3.5 Outcome definitions

	Name of outcome measure	Measuring?	Timpoints	Scoring
Primary Outcome	Delaware School Climate Survey-Student (DSCS-S)	Sense of belonging / class climate	Baseline, post intervention, 6 months	23 items with 5 subscales: student-student relationships, teacher-students relationships, school safety, fairness of rules, school liking. All items / items within a subscale are summed to a form total score/ subscale scores.[1]
Secondary outcome	WHO-5	General wellbeing	Baseline, post intervention, 6 months	Items transformed to a score 0-100 with lower scores indicating worse well-being
	PHQ-8	Depression severity	Baseline, post intervention, 6 months	The total score can range from 0–24 with a score of ≥10 indicative of clinical-level depression.
	Social Anxiety Scale for Adolescents (SAS-A)	Levels of social anxiety	Baseline, post intervention, 6 months	18-item questionnaire with three subscales: Fear of Negative Evaluation, Social Avoidance and Distress-New and Social Avoidance and Distress-General. Total score ranges from 0-90 with a higher score indicating a higher level of social anxiety. *

[1] Since *Lina* aims to improve *class* climate, items referring to the whole "school" will be changed to "class". Thus, the school safety subscale will refer to class safety; and the school liking subscale to class liking.

	UCLA Loneliness Scale for Children	Level of loneliness	Baseline, post intervention, 6 months	4 item scale summed to a total score
	SSIS Social Emotional Learning Scale-brief version. Social Awareness subscale	Social awareness competency	Baseline, post intervention, 6 months	4 item scale summed to a total score
	Peer Mental Health Stigmatision Scale-Revised (PMHSS-R) Stigma Awareness subscale	Stigma	Baseline, post intervention, 6 months	5 Likert items summed to a total score
	EQ-5D-Y-3L	Youth general quality of life (as part of economic evaluation)	Baseline, post intervention, 6 months	To be detailed in the HEAP
	CHU-9D	Measure of health and a tool to calculate QALYs for adolescents.	Baseline, post intervention, 6 months	To be detailed in the HEAP

**Note that the baseline SAS-A in Portugal was administered with incorrect labels for Q14 and Q18. These item responses will be ignored and the resulting value scaled up to be consistent with a score out of 90.*

3.6 Timing of outcome assessments

Baseline measures will be collected at least one week before the first session of ASP#1 (*Lina*). To allow enough time to collect data from all students in the class, data can be collected up to three weeks before delivery of the first episode of *Lina*.

Post intervention measures will be collected between 6 and 10 weeks from the date of baseline assessments (allowing a 2 week time window). Six month follow up measures will be collected within four weeks (either side) of the date of six whole months since baseline assessments. We will include all data collected in the primary analysis, even if outside the allocated time window, acknowledging the challenges of data collection in the school setting.

4. Statistical Principles

4.1 General analysis principles

Each analysis of primary or secondary outcome will adjust for the corresponding baseline value of that outcome. Due to the clustered nature of the trial, all analyses will include a random effect that allows for the clustering of children within classes. For the analyses of the primary and secondary outcomes the following information will be presented:

- The number of participants included in each analysis, by treatment arm
- A summary statistic of the outcome (e.g. number (%)) by treatment arm
- The estimated treatment effect
- A 95% confidence interval for the estimated treatment effect
- A two-sided p-value

No formal adjustment for multiple testing will be made. However, the number of secondary outcomes will be noted when reporting results.

The analysis of the primary outcome will be conducted on all participants and in line with the intention to treat principle, missing outcome data will be imputed using a multiple imputation model detailed in section 6.2.

Once country level data has been analysed, then data from across the three countries will be pooled and analysed as one dataset. This will follow analysis principles described above but with further level of clustering at the country level (children nested within classes nested within countries).

Descriptive statistics will be reported to 1 decimal place as number and percentage for categorical variables, and mean and standard deviation for continuous variables, or median and interquartile range for variables with a skewed distribution. The groups will be labelled Usual Education and Intervention

4.2 Methods and assumptions for dealing with data anomalies

Outliers: Extreme values in continuous outcomes will be identified using boxplots and z-scores. Sensitivity analyses excluding confirmed outliers will be conducted if necessary.

Any logically inconsistent data (e.g., contradictory responses) will be recorded in an appendix to the statistical analysis report together with rules for handling these inconsistencies.

It is possible that some schools may not be able to collect their 6 month follow up data within the 2025-6 school year as planned. In this case, the data will be collected early and for the purposes of the primary analysis will be treated as the 6 month follow up data for those participants. We will report the summary statistics for this group separately in an Appendix.

5. Trial population

5.1 Participant flow

.

Participant flow through the trial will be summarised by a CONSORT flow diagram for cluster randomised trials. This will include the numbers: screened, consented and randomised clusters. It will further include clusters and individual participants withdrawn and include in the analysis of the primary outcome at both follow-up time-points. See Appendix for a draft CONSORT diagram. There will be one CONSORT per country as these are three separate trials.

5.2 Withdrawals

Participants may withdraw before or during any data collection session. Any withdrawals will be detailed in the CONSORT diagram.

5.3 Baseline characteristics

Demographic and primary/secondary outcomes at baseline will be summarised by trial arm. Demographic information will include all items collected on the baseline demographics form with age presented as mean (SD) and the remaining categorical variables presented as n (%). Outcomes will be summarised using the mean (SD) for continuous measures and n (%) for binary or categorical measures. Any count outcomes will be summarised using median (lower quartile, upper quartile).

6. Analysis

6.1 Analysis of primary and secondary outcomes

	Name of outcome measure	Model
Primary Outcome	Delaware School Climate Survey-Student (DSCS-S) at 6-10 weeks	Generalised linear mixed model controlling for baseline score and with random effect for class
Secondary outcome	WHO-5	Generalised linear mixed model controlling for baseline score and with random effect for class
	PHQ-8	Generalised linear mixed model controlling for baseline score and with random effect for class
	Social Anxiety Scale for Adolescents (SAS-A)	Generalised linear mixed model controlling for baseline score and with random effect for class

Not to be used without prior permission by the PCTU

	UCLA Loneliness Scale for Children	Generalised linear mixed model controlling for baseline score and with random effect for class
	SSIS Social Emotional Learning Scale-brief version. Social Awareness subscale	Generalised linear mixed model controlling for baseline score and with random effect for class
	Peer Mental Health Stigmatision Scale-Revised (PMHSS-R) Stigma Awareness subscale	Generalised linear mixed model controlling for baseline score and with random effect for class
	EQ-5D-Y-3L	To be detailed in the HEAP
	CHU-9D	To be detailed in the HEAP

Assumptions for linear mixed models will be checked using QQ plots for normality and residual vs fitted value plots for linearity and homoscedasticity. If linear modelling assumptions are not met, another appropriate parametric distribution will be sought and a suitable generalised linear mixed model will be used. Where an instrument has subscales of interest, the same modelling approach will be taken for the subscales as was used for the total score for that outcome measure.

The analysis which pools these outcomes across participating countries will use the same modelling approach above but will also include a random effect for country, allowing for children nested within classes nested within countries.

6.2 Missing data

The number (percentage) of missing data on the demographic questionnaire and all clinical outcomes at baseline and follow-up time points will be summarised, and any prior, trial-based, knowledge around reasons for missing data described and discussed.

The missing data mechanism will be assumed to be Missing at Random (MAR), provided some strong predictors are present for each outcome and no prior knowledge exists that would suggest violation of this assumption i.e. that observations are missing not at random (MNAR).

Multiple imputation by chained equations using the MI package in Stata will be used; the number of datasets generated (at least 5) will reflect the percentage of missing data present.

The imputation model will include variables identified to be predictors of the missing data mechanism; auxiliary variables correlated with outcome; and all variables contained in the substantive model. For continuous outcomes, truncated regression will be used to ensure imputed data do not extend beyond the range of the questionnaire.

Individual analyses on each imputed dataset will be combined using Rubin's rules.

6.3 Subgroup analyses

The primary analysis will be repeated to include an interaction term between trial group and gender to explore any differential treatment effect. We will also explore whether there is an interaction term with those in the upper quartile for the SAS-A at baseline, indicating the group with the highest levels of pre-existing social anxiety.

The presence of an interaction will be tested using a likelihood ratio test comparing the subgroup analysis model, including the interaction term, and the primary analysis model, not including the interaction term. The test result will be considered significant at the 5% level. Within each category of participant, we will report summary statistics of the outcome by treatment arm, a treatment effect and 95% confidence interval. A p-value for the interaction test will also be reported."

6.4 Sensitivity analyses

As a sensitivity analysis around the MI imputation model, we will carry out a complete case analysis of the primary outcome with the significant predictors for missingness added as covariates.

6.5 Adherence

We will undertake an analysis to explore whether the effect of *Lina* on sense of belonging (Delaware Class Climate scale) is affected by whether a student engages with **a sufficient** number of *Lina* sessions to be considered to have received the intervention. This effect might exist for the post-intervention and/or the long term follow up.

A sufficient number of Lina sessions = At least 4 sessions, including the last 2 *and* either session 1 or session 2.

Rather than the randomised group, the population will be divided into those who did not receive *Lina* + those who did not receive a sufficient number of sessions compared to those who did receive a sufficient number of *Lina* sessions.

We will also include a sensitivity analysis that allows for actual allocation delivered, rather than randomised allocation, in case of any cluster misallocation.

6.6 Safety analysis

SAEs will be summarised descriptively by randomised arm. No SAEs are considered expected for the purposes of this trial. Occurrence of SAEs among all randomised participants will be analysed using logistic regression.

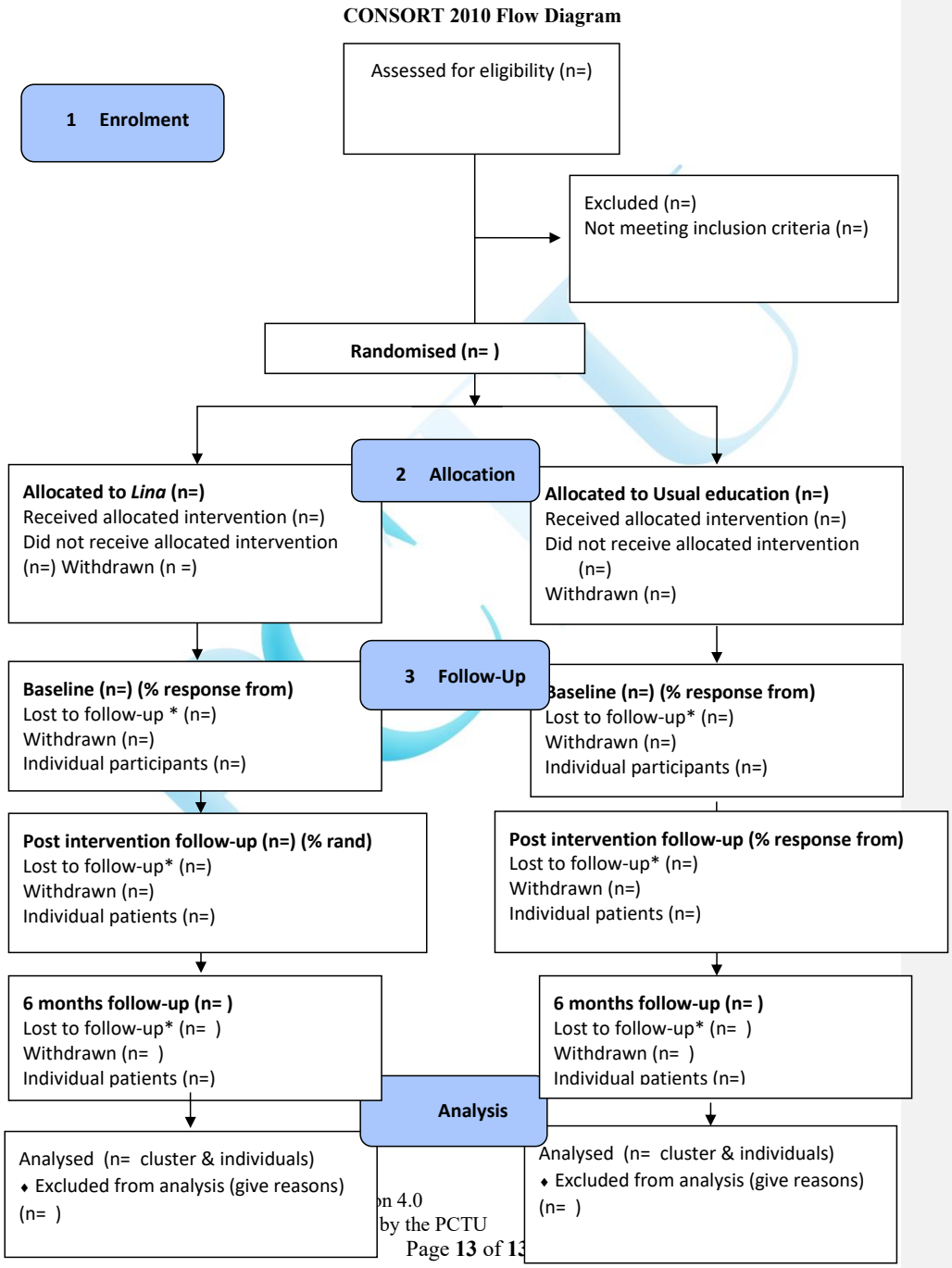
7. References

Provide a list of references.

Commented [BS1]: To add for each outcome measure

8. Appendices

Example dummy CONSORT diagram



on 4.0
by the PCTU