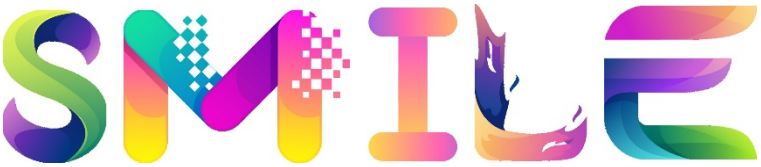

SMILE: **S**upporting **M**ental Health in Young People: **I**ntegrated Methodology for c**L**inical d**E**cisions and evidence-based interventions – Proof-of-concept Study

Pilot PROTOCOL

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No°101080923.

Contents

1. Summary

Adolescence is a sensitive period marked by increased exposure to a range of mental disorders, the incidence of which has been rising in recent decades. Young people often face difficulties in actively seeking help. On one hand, they might seek support within the perceived safety of social media, preferring to share their feelings in an anonymous online environment. On the other hand, the technological environment can negatively affect the psychological well-being of vulnerable young people.

Serious games are interventions where gaming elements are integral and primary methods for achieving mental health goals. Gamification, in turn, involves using game elements in non-game contexts and can be found in digital interventions. Both approaches have recently shown promise in enhancing digital mental health technologies and increasing adherence.

The European project SMILE (Supporting Mental Health in Young People: Integrated Methodology for Clinical Decisions and Evidence-based Interventions) aims to promote resilience in young people by providing a gamified platform with digital cognitive behavioural interventions. These interventions aim to increase skills like cognitive flexibility, self-efficacy, critical thinking, self-regulation, and self-confidence, and by this prevent symptoms of depression and anxiety.

This study is part of a European consortium within the Horizon Europe program and will be conducted in seven countries—referred to as pilot sites: Cyprus, Germany, Italy, Poland, Slovenia, Spain, and the UK. The Protocol includes two phases: a proof-of-concept pilot study and key stakeholder interviews.

First, we will test the acceptability, feasibility, efficacy and preliminary effectiveness of the SMILE intervention in a proof-of-concept study using a cluster randomised multi-site multi-arm adaptive trial design. The participants will be adolescents and young adults, aged 10-24.

Then, we will conduct a process evaluation using semi-structured interviews with a small number of adolescents and stakeholders, including, parents, school professionals, and clinicians.

Study results will be analysed at both the country level and at trial level.

2. Scientific background

2.1. Serious games and gamification for the prevention of mental disorders in adolescents

The World Health Organization (WHO) defines mental health as “a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community. It is an integral component of health and well-being that underpins our individual and collective abilities to make decisions, build relationships and shape the world we live in” [1]. This definition highlights that mental health is more than just the absence of a disorder; it is a major global health concern, with an estimated 125 million Disability-Adjusted Life Years (DALYs) worldwide in 2019 [2].

Adolescence is widely recognized as a sensitive period for the development of mental disorders, whose incidence has been increasing in recent decades [3]. The WHO specifically emphasizes the poor mental health among young people, who often struggle to access mental health services and avoid seeking professional help due to fear of stigmatization. Consequently, these disorders often remain undetected and untreated until later in life.

Depression is highly prevalent among young people worldwide, and suicide is the fourth leading cause of death among European adolescents [4].

Established prevention recommendations for improving adolescent mental health include connecting with other people, self-help strategies after self-harm, accessibility of care, and a supportive school environment [5]. In the wake of the COVID-19 pandemic, which imposed contact restrictions, online interventions for preventive measures have been increasingly developed [6]. A recent meta-analysis demonstrated that these interventions can significantly prevent increasing depression scores; however, more evidence is needed regarding their effectiveness in addressing anxiety and stress [7].

Serious games are interventions where gaming elements are an integral and primary method for achieving goals related to mental health [8]. Gamification, in turn, is a technology used in digital interventions that incorporates game elements in non-game contexts. Both approaches have recently emerged as promising methods to enhance digital mental health technologies and increase adherence [9]. Serious games for adolescents include mini-games, adventure worlds, and social simulations, often addressing anxiety and/or depressive symptoms. However games targeting all age ranges are currently still missing [10]. As such, the SMILE project strives to develop a serious game helping young people to develop a range of meta-skills and cognitive competencies to resiliently cope with psychological distress related to day-to-day stressors, ultimately increasing their well-being.

The European project SMILE (Supporting Mental Health in Young People: Integrated Methodology for Clinical Decisions and Evidence-Based Interventions) aims to promote resilience in young people by providing a gamified platform that incorporates elements of digital cognitive behavioural interventions. This platform is designed to achieve change in mental health and well-being by enhancing key psychological factors, such as cognitive flexibility, self-efficacy, critical thinking, self-regulation, and self-confidence. Within the application, algorithms will provide a gamified environment focused on mental health, including virtual self-assessments and self-monitoring, all integrated into a mobile application. Additionally, SMILE will measure specific gaming behaviours, known as “digital biomarkers”, described in this Protocol as in-game measures in subsequent sections. These measures have been previously used to assess, for example, social anxiety by collecting proximity measures and movement patterns in interaction with an avatar [11], or motor and balance skills as biomarkers for physical well-being in exergames [12]. Figure 1 gives an overview of the core components of the SMILE concept.

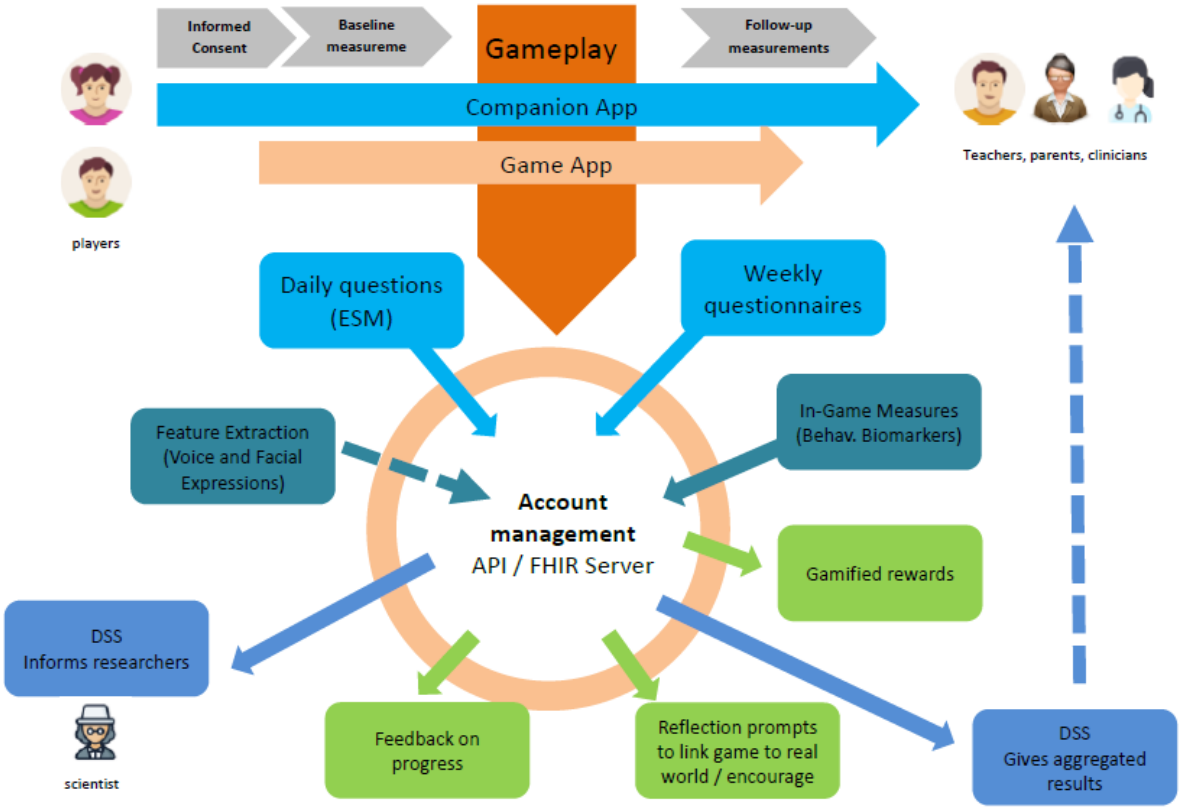

Figure 1. *Core components of the SMILE concept.*

The SMILE approach places special emphasis on co-creation, defined as "the participation of users/consumers in the development process (…) significantly improving the quality, usability, and social acceptance of new solutions" [13]. Our previous studies within this project included focus groups and workshops organized with adolescents and young adults in seven participating European countries (ages 10-14, 15-19, and 20-24), as well as the Living Labs testing phase, where a first version of the game was tested by a small number of representatives from each age group (see a comparable approach with Living Labs in [14]). The current stage of the project described in this Protocol will consist a proof-of-concept study to test the feasibility, acceptability, efficacy and preliminary effectiveness of the SMILE tools. The stage will conclude with stakeholder interviews aimed at facilitating future implementation.

2.2. The game concept

The game's storyline centres around the player moving to the cyberpunk-style city of Hopetown with their sister, Seraphina. In Hopetown, people's automatic thoughts manifest as small, robot-like creatures. As the player adjusts to their new home, Hopetown is attacked by a cybernetic dragon, who captures Seraphina. As the player embarks on a quest to save their sister and the city, they are guided by key characters to learn different techniques and competencies based on psychological principles and Cognitive Behavioural Therapy (CBT) techniques. The journey takes the player through diverse environments, each module designed to teach and reinforce different aspects of CBT, such as identifying cognitive distortions, evaluating evidence, and challenging core beliefs. By engaging with in-game challenges that mirror real-life psychological struggles, players learn to reframe unhelpful thinking patterns, build resilience, and develop healthier mental habits. Figure 2 shows an example of planned gameplay.

***Figure 2**. Example of planned gameplay.*

The game will consist of a range of scenarios, divided into five modules, with goal-oriented and story-driven game activities. Each scenario will include compelling dialogues with non-player characters (NPCs), context-aware challenges, all to provide the player with an interactive and immersive experience. By prioritising positive user experience, the SMILE serious game strives to improve the intervention acceptance, adherence and consequently effectiveness [e.g., 15]. This game will serve as an engaging and innovative medium for the delivery of psychological intervention elements, including psychoeducation about negative automatic thoughts, cognitive restructuring, self-regulation training, etc.

In Module 1 ("Hopetown"), the player's goal is to uncover why the city was attacked and how to save their sister. As the player is exploring the game world, they witness the Dragon Wielder (the villain of the game) hack the city's network to take control of citizens' negative automatic thoughts. The player then encounters a series of challenges left behind by the Dragon Wielder, designed to teach them how to distinguish between thoughts, feelings, and behaviours, and how they influence one another in line with the CBT "hot-cross bun model" (Figure 3) [16]. They will also learn about automatic thoughts and how positive and negative automatic thoughts produce different feelings, and behavioural outcomes. Finally, they will explore different patterns of thinking – thinking errors – which influence automatic thoughts.

***Figure 3.** The CBT hot cross bun model, adapted from Fenn and Byrne [16].*

In Module 2 ("Starfall District"), the player navigates Starfall District, a once vibrant green space now corrupted by the Dragon Wielder's influence. The player witnesses some NPCs struggling with their negative automatic thoughts that are under the Dragon Wielder's control, and they

must learn some new tools to help them. To do so, the player will explore the relationships between thoughts, feelings and behaviours in more depth: they will learn how negative automatic thoughts, reinforced by individual thinking errors, can create a negative cycle [17]. They will also be introduced to two techniques of challenging this negative cycle – cognitive restructuring and self-regulation (paced deep breathing) – which could strengthen the player's resilience when coping with negative automatic thoughts in their day-to-day life. After battling the Dragon Wielder's cybernetic dragons, the player obtains a cryptex (a portable vault that will only open its contents for the owner), which provides the next objective for Module 3: to find the owner in the hopes the cryptex contains some useful information about Seraphina's whereabouts.

In Module 3 ("The Technicians' Quarters"), the player infiltrates the high-security quarters of the Dragon Wielder's technician. They must avoid guards and gather information about the technician's core beliefs, to access the cryptex. After correctly identifying the Technician's core belief, the player can read the cryptex's contents. However, before they get a chance, the player is discovered and thrust into a fight, flight, or hide scenario. To escape successfully, the player must choose fight and overcome their negative thoughts using the CBT-based tools they have learned so far. After escaping, the player can read the cryptex's contents which contains a clue about how to defeat the Dragon Wielder, left behind by Seraphina.

Module 4 ("Library") sends the player on a quest to find information about how to defeat the Dragon Wielder, after gaining a clue from Seraphina. The player explores the city's vast digital library, with the objective to crucial information about how to defeat the Dragon Wielder. Critical thinking and thought challenging are essential as the player evaluates evidence to test the validity of information – a skill which can enhance identifying evidence for and against negative automatic thoughts as a part of cognitive restructuring [18]. By critically analysing information, the player gains insight into the Dragon Wielder's backstory, learning more about his childhood and core beliefs. Armed with this knowledge, the player can now defeat the Dragon Wielder and save their sister, but they must escape the library to do so. While avoiding guards and hacking locks, the player encounters some information about the Dragon Wielder's whereabouts. By evaluating the validity of the evidence, they can decide on the Dragon Wielder's location, needed for Module 5.

The player's goal in Module 5 ("The Data Vault") is to rescue their sister and stop the Dragon Wielder's plan to take over Hopetown. The final boss battle consists of overcoming peer pressure and utilising all techniques learned thus far to de-corrupt the dragon before battling the Dragon Wielder. After doing so, the player attends a party held in the centre of the city, to celebrate their win over the Dragon Wielder. The player is forced to give a speech, with the opportunity to practice assertive communication. It will build on the previously acquired competencies by focusing on pre- and post-event processing and challenging fear of negative evaluation through cognitive restructuring, reducing avoidance and increasing social confidence through behaviour modification, reducing the harmful effects of social influence through assertiveness training [19]. At the end of the game, with Seraphina free and the Dragon Wielder defeated, Hopetown recovers. The robot-like creatures that once terrorised the city are now peaceful, reflecting the player's newfound inner peace.

2.3. Game variants: feedback vs no feedback

In SMILE we have incorporated a feedback element as an intrinsic motivator for engagement and learning and to facilitate greater applicability of the skills learned in the game to real world scenarios and experiences. This feedback (presented graphically) will be delivered to only half of the sample (i.e., half of the clusters as outlined in Section 4 of this Protocol) in order to evaluate its effectiveness. Feedback is based on Nicolson's [23] "Recipe for meaningful game

engagement". Feedback will be generated across time points from a triad of measures from the in-game metrics, experience sampling methods (ESM) and self-report measures. The player will be delivered feedback following completion of the module and will be invited to click on specific time-points of interest. The system will generate some brief information (e.g., "your mood appeared to be a little more positive than earlier in the day"), linking contextual and game-based metrics with daily mood and well-being ratings. The clusters of participants who do not receive the feedback will still participate in the game, ESM and reflective exercises, but will not be able to visualise and reflect on their data.

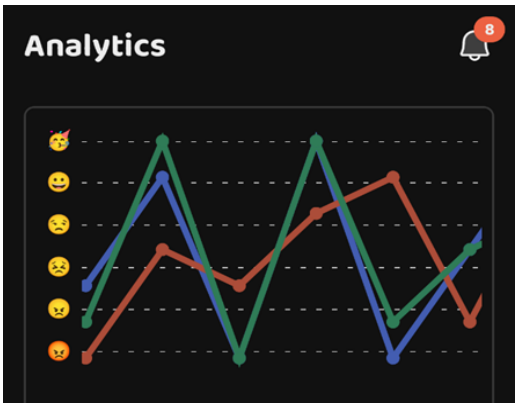

Figure 4. *Prototype of the feedback provided by the SMILE companion app.*

The SMILE companion app will deliver feedback in graphical form collated from in-game measures, ESM measures and reflective exercise measures including simple daily averages and trends over time as well as basic association between activity, context and well-being (see Figure 4). The player can select one of these measures to see how it relates to the activity in the game or the real world and how it relates to other measures.

3. Aims of the study

This study comprises two phases: a proof-of-concept study using a cluster randomised multi-site multi-arm adaptive trial design and key stakeholder interviews.

3.1. Proof-of-concept study

The overarching aim of the proof-of-concept study is to investigate the acceptability, feasibility, and preliminary efficacy of the gamified platform designed to enhance the well-being of adolescents and young adults, in a pragmatic cluster randomised multi-site multi-arm adaptive trial. In the pilot study, we will test 6 modes of intervention including a control period; (a) neutral baseline with a period of experience sampling (ESM), (b) Game Modules 1-3, and (c) Game Modules 4-5, each with and without feedback.

To achieve our overarching aim, this study will have two objectives:

1. To investigate feasibility and acceptability; i) Reach (i.e., participation), ii) Preliminary Effectiveness (defined as the interaction of efficacy × use), operationalized as greater improvement of psychological distress post-test, in the experimental conditions compared with control period as primary outcome, and iii) Engagement with the gamified intervention platform. This will provide the basis for assessing the public health impact of the use and engagement with the gamified platform.
2. To identify contexts, psychological processes, and mechanisms of change, and how these are associated with outcomes of the intervention in participants.

To meet objective 1, we will test the following hypotheses:

Primary hypothesis: It is hypothesized that, compared with the control period, anxiety and depression scores (primary outcomes) will be lower in the experimental conditions (i.e., with and without feedback) post-test.

Secondary hypotheses 1: It is hypothesized that, compared with the control period, the secondary outcomes (well-being, resilience, emotion regulation and self-efficacy) will be higher and social anxiety lower in the experimental conditions (i.e., with and without feedback) post-test.

Secondary hypothesis 2: It is hypothesized that feedback will be positively associated with the feasibility, acceptability, and preliminary effectiveness of the SMILE tools.

To meet objective 2, we will test the following hypotheses:

Secondary hypothesis 3: Compared with the control period and the experimental condition without feedback, well-being and resilience will be higher in the experimental condition with feedback post-test.

We will also conduct exploratory analysis to map digital biomarkers from in-game metrics and diaries to subjective anxiety and mood scores collated from in-game, ESM and reflective exercises from the Companion App.

3.2. Key stakeholder interviews

The aim of this part is to understand how participants leverage the SMILE intervention to support their mental health and well-being and evaluate the process of implementing the intervention within their day-to-day contexts using a realist evaluation framework [29] to identify in vivo configurations of contexts, processes, and mechanisms of change, and how these are associated with outcomes of implementation and intervention. The goal is to 1) collect feedback from the young people who participated in the proof-of-concept study and their parents, and 2) discuss potential implementation of SMILE tools with school professionals, and clinicians.

Specifically, we aim to:

1. Gain deeper insight into how, when, and where adolescents used SMILE tools and whether they found it useful and interesting (interviews with young people, their parents and school professionals).
2. Identify barriers and facilitators to the future use and implementation of SMILE tools (interviews with parents, school professionals, and clinicians).

4. Study design

4.1. Proof-of-concept study

This proof-of-concept study utilizes a cluster randomised multi-site multi-arm adaptive trial design. Participants will be allocated to clusters, which will then be randomly assigned to one of six sequences across three conditions – each delivered with or without feedback, (a) neutral baseline with a period of experience sampling (ESM), (b) Game Modules 1-3, and (c) Game Modules 4-5.

The adaptive trial design will be based on an initial sample size per site based on the assumption that each step in the sequence will be completed. We will conduct 2 stages of interim analyses, calculating the effects of each intervention modality, including control. Based on these initial effects we will adjust sample size per cluster/site, and make decisions to amend or discontinue relevant intervention elements. Each site will receive randomisation of groups/clusters to assign groups to one of the six sequences. Each interim analysis will

calculate initial effects of the primary outcomes and detect the key predictors of these effects. Sample size and randomisation sequences will be adjusted accordingly.

At each interim analyses point we will also be able to adjust the active intervention elements based on the model parameters; if elements have consistently been accessed or uses, and whether or not indicators related to active intervention elements contribute to interim effect sizes. See Section 11.1 for pre-defined criteria for adjustment.

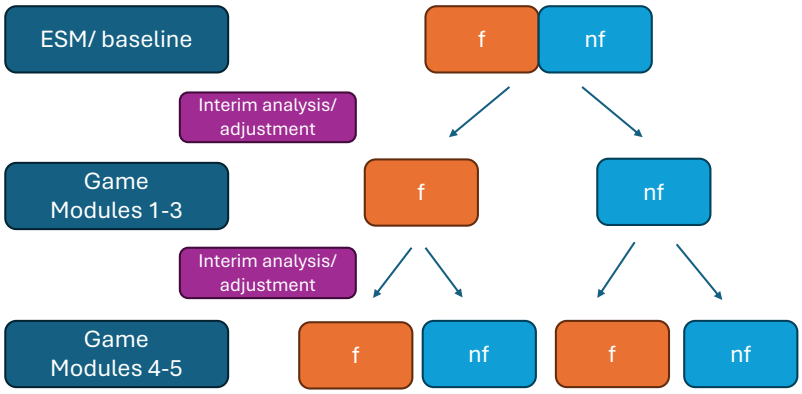

***Figure 5a**. Proof-of-concept study design*

	FEEDBACK													
	1	2	3	4	INTERIM ANALYSIS	9	10	11	12	INTERIM ANALYSIS	17	18	19	20
ESM/Baseline														
Game Module 1-3														
Game Module 4-5														
	NO FEEDBACK													
	5	6	7	8	INTERIM ANALYSIS	13	14	15	16	INTERIM ANALYSIS	21	22	23	24
ESM/Baseline														
Game Module 1-3														
Game Module 4-5														

Figure 5b: *Randomized Cluster Distribution*

This design will be combined with a process evaluation, which will apply a mixed-methods approach using realist evaluation. It reflects the optimum design for investigating all aspects of the proof-of-concept study (i.e., acceptability, feasibility, reach, effectiveness), with a particular focus on deriving estimates of effectiveness with low risk of bias (e.g., avoiding contamination between experimental and control periods), while at the same time allowing for an in-depth process evaluation as well as optimal generalizability of study findings, given the naturalistic field setting.

The intervention will be rolled out over a period of 12 months, in sequences of 2-week intervals. Eight clusters will be randomized into each sequence, amounting to a total of 24 clusters across seven pilot sites. Randomization will be conducted by a statistician who is not involved in conducting the trial. Each intervention period will last 6 weeks.

4.2. Key stakeholder interviews

Qualitative material will be collected from semi-structured in-depth interviews with adolescents, parents, school professionals, and clinicians after adolescents have completed the intervention period. Interviews will be conducted at each pilot site, either in person or online.

5. Randomization plan

Proof-of-concept study employs a cluster randomised multi-site multi-arm adaptive trial design, meaning that randomization occurs at the cluster level and involves assigning each cluster to either feedback or no feedback variant within each condition (i.e., neutral, game modules 1-3, games modules 4-5; see Figure 5). Randomization will be conducted in blocks (using block randomisation) at the consortium level using REDCap software by a statistician who is not involved in the study. A total of 24 clusters (8 clusters per sequence) will be randomized into two conditions, feedback and no feedback; these will be allocated pragmatically between sites and adjusted in size and distribution if indicated by the interim analyses. There is no randomization at the individual level.

Clusters will be blinded as to which of the two interventions they will receive, i.e., with or without feedback.

6. Study schedule and time required

6.1. Proof-of-concept study

The study will span 12 months, scheduled to run from September 2025 to September 2026. Every module is planned to last no longer than 20 minutes. With 5 modules in total and diaries, ESM items and questionnaires to be answered, we expect the participants to spent maximum 4.5 hours for the study, spread over several weeks.

6.2. Key stakeholder interviews

Interviews can be conducted at any point during the proof-of-concept study, but only after the participants of a given sequence have completed the game. Each interview will last 30-60 minutes.

7. Sample

In sum the expected number of participants includes:

- N=1438 for the proof-of-concept study (considering a Conservative 45% attrition rate)
- N=210 for the key stakeholder interviews

7.1. Proof-of-concept study

The following assumptions are made for the sample size calculation:

- 3 sequences
- 4 clusters per sequence (as the calculator can only handle one intervention, the sample size has to be doubled afterwards)
- Cross-over after 2 weeks
- 2 interim analyses
- Alpha = 0.05

- Power = 0.8
- Intracluster Correlation = 0.01 (from Girls on the Go! Study, that used schools as cluster)
- Individual Autocorrelation = 0.6 (Girls on the Go! Study had a value of 0.65, but I would suggest a more conservative estimate)
- We assume our outcome scales to be continuous
- Effect size: 0.2 (mean difference of 0.2 with standard deviation of 1)
- Drop-out Rate: 50%

With these assumptions, we would need 24 participants per arm per site to ensure a power of 80% at a significance level of 5%, meaning 992 participants overall [32]. To account for the dropout which was found to fall between 24% and 83% [32, 33], we would have to recruit 1438 participants for the study.

However, since this is an adaptive trial design, we need to account for the interim analyses to adjust sample size for subsequent steps based on the actual interim effects. This approach will enable us to gather sufficient data to inform potential subsequent definitive RCTs and to refine the intervention as necessary.

7.2. Key stakeholder interviews

Semi-structured interviews will be conducted with adolescents aged 10-24 years old who completed the SMILE intervention as part of the proof-of-concept study (N=10), as well as parents/guardians (N=10), school professionals (N=5) and clinicians (N=5). In total. we intend to interview N=30 per site, amounting to 210 participants in total.

8. Inclusion and exclusion criteria

8.1. Inclusion criteria

8.1.2. Proof-of-concept study

Participants will be adolescents and young adults aged 10-24 that meet the following inclusion criteria:

- Participants under 16 require consent from a parent or guardian, as well as their own agreement to participate. Informed Consent Forms and their digital equivalents will be tailored to be age-appropriate, ensuring both guardians and participants can make informed decisions. If either the parent or the participant does not agree, the participant will not be included in the study.
- Must complete measures of primary outcome.
- Must have normal or corrected to normal vision.

For those who do not own/have access to a smartphone or digital device, we will provide a phone handset to facilitate access to the app and will cover the cost for all data network charges.

8.1.3. Stakeholder interviews

The only inclusion criterion for parents/legal guardians is to have a child in care who has already completed their participation in the active part of the intervention (i.e., has finished playing the game). For school professionals and clinicians working in the field of mental health, the sole inclusion criterion is working with adolescents and young adults aged between 10 and 24. For adolescents, the only inclusion criterion is that they have completed playing the game.

8.2. Exclusion criteria

8.2.2. Proof-of-concept study

Participants will not be eligible if they are unable to consent to participation or if they have a current confirmed diagnosis or treatment episode for:

- severe mental disorder (schizophrenia, bipolar disorder, severe depression),
- substance use disorder,
- epileptic disorders,
- gaming addiction.

8.2.2. Stakeholder interviews

None.

9. Recruitment

9.1. Proof-of-concept study

Recruitment will commence as soon as SMILE tools are ready for testing, approximately in May 2025 or July/August 2025, depending on regional or site specific summer holidays. Potential participants will be recruited purposively by the researchers. Recruitment and sampling procedures will be documented in detail for reporting and transparency.

Prior to commencing the study but after consenting to taking part, participants will be offered a short onboarding session, either individually or in a group. During the session they will receive information and a brief demonstration on how to install, set up and use the SMILE apps.

9.2. Key stakeholder interviews

We will conduct a detailed process evaluation provide in-depth insight into the implementation and associated processes of the SMILE intervention and to establish what works, for whom, in what circumstances, in what respects, to what extent, and why. The quantitative data collected will describe the specific outcome patterns and investigate quantitative metrics of the implementation/intervention fidelity. However, this will not in itself explain underlying processes that generate these patterns. We will use a mixed-method approach that combines the strengths of quantitative metrics and qualitative user and stakeholder experience data to produce a coherent and plausible explanation. We will attend to individual person-, system- and context-based factors that influence the effective use, and implementation of SMILE within existing contexts of support across different settings. Specifically, the process evaluation will use a mixed-methods approach and take a realist evaluation approach. This implies that configurations of contextual factors, mechanisms of implementation, and outcomes of the implementation and intervention are explored across all levels of agents within the intervention and its implementation (i.e., individual participants, parents, educators and clinicians, and socio-economic and contextual factors that may impact their intentionality, behaviour and decision making at different stages of the intervention). This will also allow us to examine how participants and stakeholders appropriate SMILE to serve their particular needs. Stakeholder interviews will be conducted with legal guardians of the adolescents who joined the proof-of-concept study, school professionals, and clinicians.

When recruiting young people to the proof-of-concept study, we will additionally ask them to state whether they are interested in participating in the interviews after the game is completed. We will also extend this invitation again when they complete playing the game. We will then contact randomly chosen 10 participants per site and invite them to the individual interviews.

When recruiting young people to the proof-of-concept study, we will additionally invite their parents/ legal guardians to take part in interviews once their children have completed playing the game. To this end we will distribute a separate Participant Information Sheet and Informed Consent Form. We will then contact randomly chosen 10 parents per site and invite them to the individual interviews.

Teachers who facilitated access to schools/ colleges during the proof-of-concept study will be invited to participate in the interviews, as well as extend this invitation to their colleagues at the participating schools/ colleges. To this end we will distribute a separate Participant Information Sheet and Informed Consent Form. We will then contact randomly chosen 5 school professionals per site and invite them to the individual interviews.

Where appropriate we will also choose 5 Child and Adolescent Mental Health professionals from participating local services to take part in the interviews. We will use social media and traditional media to advertise the study among clinicians working with young people aged 10-24. We will also use the researchers' professional networks and word of mouth. People who express their interest in participating will be approached by the research staff, who will provide them with an oral and/or written explanation of the study. They will need to confirm eligibility and sign an Informed Consent Form or its digital equivalent prior to participating.

10. Measures

10.1. Proof-of-concept study

There are seven types of data that we will collect in this phase. First, majority of the measures of primary and secondary outcomes and potential confounders will be collected at the following time points: baseline (T0; two weeks before the intervention), right before the intervention (T1) during the intervention (T2), and post-test (T3; after completion of the intervention). Additionally, measures of depression and anxiety (i.e., primary outcomes) will be distributed on a weekly basis, alternating between a depression scale one week and an anxiety scale the next. Some secondary outcomes will be collected only after specific modules of the game. These measures are detailed in Appendix 1.

Second, participants will be invited to respond to a short daily survey using ESM. Third, there will be in-game measures related to the current activity in the game. These measures are detailed in Table 2. Fourth, we will collect vocal and facial data via weekly diary recordings. Fifth, we will assess the feasibility and acceptability of the game through measures other than participants' responses, which are described in detail below. Finally, we will collect user data that are restricted to the SMILE apps: a game app, and a SMILE companion app.

All measurements described below will be collected by the SMILE end-user applications: SMILE game app and SMILE companion app. These apps do not store data, they only collect data according to the specifications and send them to SMILE Application Programming Interface for storage and processing. The data is stored on a server that is using the Fast Healthcare Interoperability Resources standard. Once the data is stored, the access to the data is guarded by SAPL (Simplified Attribute-based Policy Language) that is part of the SMILE project and solution. The rules for viewing the data are made based on discussions within consortium to ensure that all parties have access only to the necessary data. The data visualization tools and analytics are available for consortium partners (based on the established rules) using the SMILE DSS (Decisions Support System). All SMILE tools are using Keycloak for federation, strong authentication and user management.

10.1.1. Primary outcomes

Primary outcomes in this proof-of-concept study are depression and anxiety. Detailed description of the scales can be found in Appendix 1. Completing these scales is mandatory for the participants.

10.1.2. Secondary outcomes

Secondary outcomes in this proof-of-concept study are: well-being, resilience, privacy concerns, emotion regulation, self-efficacy, social anxiety, and two scales related to the usage of the SMILE apps. Detailed description of the scales can be found in Appendix 1. Completing these scales is optional for the participants.

In addition to self-reported measures, some secondary outcomes will be evaluated based on ESM data.

10.1.3. Confounders

Potential confounders in this study relate to privacy concerns. Items measuring this construct can be found in Appendix 1 and completing them is optional.

10.1.4. In-game measures

Examples of in-game measures ("digital markers") are detailed in Table 2. These measures are subject to change based on feedback from Living Labs testing phase, during which target users will play the game and provide their input.

***Table 2**. In-game measures ("digital markers") to be collected in the proof-of-concept study.*

Category	**Measure**
Behavioural markers	Time on task
	Approach/avoidance count
	Speed of approach/avoidance
Cognitive performance	Memory: Remembering instructions
	Decision making: Deciding which route or course of action to take under a limited time (decision making)
	Speed of decision making
	Time interacting with NPCs and artefacts in the game
	Selection of positive/ negative/neutral/self-referential/other-referential thoughts
Motor performance	Number of clicks, taps, or other interactions, and movement patterns within the game
Social biomarkers	Negative self-referential descriptions (from menu selection)
	Interaction with NPCs (dwell time)
	Proximity to NPCs
	Speed of approach to NPCs

	NPC selection
	Drop-off Points: Where users typically stop playing or lose interest
Affect	Mood/anxiety visual analogue scale
	Brief recording of facial expression in limited exercises
Engagement	Duration: Average length of time spent per session/per level
	Number of attempts at level
	Frequency: How often the user engages with the game
	Interaction with objects (frequency/duration)
Problem Solving	Correct/incorrect responses - puzzles
	Number of attempts to correctly solve puzzle (related to identifying maladaptive thoughts)
	Number of items (e.g., creatures) captured

10.1.5. Diary recordings

Participants will be asked to record short videos every week (responding to approximately 3-4 questions each week) and slightly longer videos at baseline and post-test (approximately 8 questions). Questions before and after the intervention will be general, and will include topics such as self-perception, personality traits, cognitive patterns, potential symptoms of anxiety and depression, emotional well-being including mood and energy, coping mechanisms, and self-esteem. Weekly videos will be in their content tied to the modules of the game, i.e., they will invite the participant to reflect on the specific module and their experience with its content.

Examples of the diary prompts include:

- Can you tell me about yourself so I can get an idea of who you are?
- How would you describe your overall feelings and energy levels recently? Have you noticed any changes affecting your daily life, work, family, or friends?
- How does playing this game make you feel, and why?

After the recording is complete, the video is encrypted and sent to the server, where the video and audio are separated. Neither the video nor the audio are seen by humans at any point; they remain encrypted until processed. Encryption ensures that if anyone attempts to open it, they would only see noisy audio and pictures, not the actual video and audio recordings.

Diary recordings will be used to extract digital markers that have been suggested by the previous literature as indicative of anxiety and depression. This is an exploratory part that will enable the consortium to test the associations between validated self-report assessments of anxiety and depression and extracted digital markers with the overreaching aim of development of screening explainable AI tool, potentially applicable in young population to detect signs of mental distress as soon as possible.

All processing of the recordings is done by AI. From the video, AI extracts features called Action Units and uses them to calculate metrics such as gaze, energy of head movements, facial expressions, etc. From the audio, AI analyses the signal to estimate energy, MEL frequency, and other acoustic parameters. The audio also goes through speech recognition for text feature extraction to identify mentions of symptoms such as fever and tiredness. The

use of positive, negative, and neutral words/sentences, as well as the number of words in sentences, are counted. Specific terms related to depression are counted, but the overall semantics are not analysed. The transcriptions are not stored nor available to human observers.

10.1.6. Feasibility and acceptability outcomes

Feasibility and acceptability will be assessed with the following measures: (1) cluster recruitment yield, defined as the proportion of clusters included to the total number of clusters invited; (2) uptake, defined as the proportion of participants who completed the minimum dose of the intervention to all participants included; (3) engagement with the SMILE serious game based on in-game measures (see Table 2); (4) adherence, measured as the average proportion of completed game modules to the total number of modules; (5) dropout, defined as the proportion of participants who disengaged with the study at any point between baseline and post-test assessments to all participants; (6) adverse event analysis (see Section 13.1.); (7) User Experience Questionnaire (UEQ-S) assessed at post-test (see Table 1).

10.1.7. User data

User data collected will be restricted to the apps used in this study (i.e., game app and companion app). These are e.g., frequency of the app usage, drop-off points (where users typically stop using the app), length of app usage per session, when the sessions begins and ends, whether it is completed with or without interruptions, completion rate, how often do users engage with the game, average length of time spent on session.

10.1.8. Use of data for research purposes

The data collected in the study will be anonymised and then used to develop a decision support system (DSS) for researchers in order to carry out an exploratory analysis to 1) recognize and map key factors associated with mood changes, 2) develop transparent algorithms to visualize patterns in mood-change prediction, and 3) support end-user self-monitoring and self-assessment in serious-games. The DSS will be developed within a GDPR compliant software environment.

Participants' personal information will be used to create synthetic data – artificial information that statistically resembles real data without containing any actual personal details. This process helps the research team develop and test systems while completely protecting individuals' privacy. The synthetic data maintains the patterns and relationships found in the original information but cannot be traced back to any specific person. This approach allows researchers to gain valuable insights while ensuring the highest standards of confidentiality and data protection. Additionally, this methodology helps decrease bias in the resulting algorithms, creating more equitable systems that better represent and serve all populations. Since the information is completely anonymous, this data can be shared among different researchers, even outside the SMILE consortium, without exposing the participants, the owners of the original data, to any risks of re-identification or identification exposure.

The digital diary recordings will be processed through a multilingual named entity recognition system (NER) to extract health-significant entities. The output will comprise structured annotations of three primary categories: PROBLEM (e.g., symptoms, conditions, diseases), TEST (e.g., diagnostic procedures, assessments), and TREATMENT (e.g., medications, therapeutic interventions) entities. This will enable automated identification and classification of health-related expressions within participants' natural language diary entries, as additional quantifiable linguistic markers that can be aligned with standardized instruments, such as PHQ-9, GAD-7.

Example:

Diary input: "After a restless night with little sleep, I felt extremely tired and anxious all day, which made it hard to concentrate at work, so I tried some deep breathing exercises that my therapist recommended, and later took the medication my doctor prescribed."

The NER system would identify and categorize:

PROBLEM: "restless night," "little sleep," "tired," "anxious," "hard to concentrate"

TEST: none in this example

TREATMENT: "deep breathing exercises," "medication"

10.2. Key stakeholder interviews

Two kinds of qualitative interviews are planned after completion of a respective cluster: a) interviews with young who played the game and b) stakeholder interviews.

Qualitative interviews with adolescents of all ages and groups (feedback vs. no feedback) will be organized by following a semi-structured interview guide regarding their experiences of the SMILE game and the companion app. In line with the realist evaluation approach [29], these interviews will focus on what works, for whom, in what circumstances, in what respects, to what extent, and why (i.e., context, mechanisms and outcomes).

Stakeholder interviews will take the form of semi-structured interviews with parents, school professionals, and clinicians. The interviews will cover the following topics:

- What are the main concerns of children and young people?
- What are their resources, what helps them cope?
- What is the feedback from parents and school professionals on the game?
- What barriers do respondents identify in implementing SMILE tools at home and at school?
- What facilitators do respondents identify in implementing SMILE tools at home and at school?

11. Statistical design

11.1. Proof-of-concept study

The hypotheses regarding the preliminary effectiveness of the SMILE interventions in relation to the outcome measures will be verified with Linear Mixed Effects Models built for each primary and secondary outcome. In the case of the outcomes measured with age-specific scales, raw response scores will be transformed to z-scores or percentages to ensure comparability. The models assessing the linear time-on-treatment effect will consider fixed effects of: time variables (calendar time and exposure time), condition (game with feedback, game without feedback, and baseline control), as well as an interaction of exposure time and condition. The random effect structure will include three levels: country, cluster and individual. Interaction between cluster and calendar time as well as individual and calendar time will also be included as random effects to account for variance over time.

Analyses will be conducted on the minimum dose intention-to-treat sample that is those participants who completed baseline assessments and at least one game module. As this is an exploratory study, p-values are interpreted in a descriptive way.

Interim analyses will determine the actual effects of primary and secondary outcomes at the end of each interim sequence. Based on the empirical effect we will be able to adjust number and size of clusters needed to meet the ultimate target sample in order to reliably answer our

research questions. In addition to adaptations to sample size, we will also consider measures that do not produce reliable variables (floor or ceiling effects, lack or variance or extreme skewness) and examine the validity of our primary outcome measures. Each interim analysis will also evaluate the level of intervention and participants' engagement with individual intervention components and the level of feedback provided in the randomised clusters, which will be considered in each of the following intervention sequences (e.g. additional game modules or modes of interactions, level and detail of feedback provided).

Pre-defined criteria for adjustment:

1. Primary outcomes – if little variance is detected in the primary outcomes, that is depression and anxiety (e.g., if >80–90% of participants score in the top or bottom 10% of the scale) we will promote well-being to primary outcomes. This will be associated with the sample size recalculation focused on the expected size effect for the new primary outcome.
2. Sample size adjustment
3. Adaptation of intervention elements

11.2. Key stakeholder interviews

The interviews of the process evaluation will be semi-structured and take a realist evaluation approach [29] to establish what works, for whom, in what circumstances, in what respects, to what extent, and why. This implies that configurations of contextual factors, mechanisms of implementation, and outcomes of the intervention are explored across all levels of agents within the intervention and its implementation (i.e., individual participants, cluster characteristics, and socio-economic and contextual factors that may impact their intentionality, behaviour and decision making at different stages of the intervention). This will allow us to identify key aspects of successful and effective implementation of the intervention in different settings from a participatory perspective.

Qualitative material will be collected from semi-structured in-depth interviews with adolescents, parents, school professionals, and clinicians after adolescents have tested the SMILE tools. This sample will include approximately N = 40 per pilot site.

Quantitative material of this stage will include demographics that will be analysed descriptively. The interviews will be recorded and transcribed. The qualitative evaluation will be carried out with a suitable program, e.g., MAXQDA [30]. We will apply thematic analysis using an inductive approach [31]. Thematic analysis is an established method in qualitative research that employs five main steps:

1. Familiarizing with the data (transcribing, reading the data, noting down initial ideas)
2. Generating initial codes (coding of interesting features of the data in a systematic fashion, collating data relevant to each code)
3. Searching for themes (collating codes into potential themes, gathering all data relevant to each potential theme)
4. Reviewing themes (checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis
5. Defining and naming themes (ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme [31].

This analysis will be done by each pilot site and will result in themes that in turn can be used to gain a better understanding of the needs, experiences and best practices that have been discussed. Further results will be used for a cross-national analysis. An English report will be

developed, describing the findings with the aim to be used as further contribution for the interpretation of the project outcomes and future perspectives.

12. Benefits

12.1. Benefits for the research community

We hope that this study will provide us with an evidence base about the acceptability and feasibility of the SMILE game for the three age groups young, middle and late adolescents. We expect data that supports preliminary effectiveness of the game in helping young people to cope with stressors and improving their resilience and mental well-being. By including stakeholders who will be parents, teachers and clinicians, we expect to get a deeper understanding of the context of our results. We moreover hope to gain a deeper understanding of the interlocking of mental states and digital gaming behaviour measured by behavioural biomarkers. Finally, as our study is a proof-of-concept design, we aim to test the methodology that could be used in the future for an effectiveness trial.

12.2. Benefits for the participants

All participants have the opportunity to gain insights into current technological advancements in the field of mental health for the benefit of adolescents in Europe.

13. Potential risks

The SMILE game aims to build resources and competencies without requiring players to explore or disclose their personal negative experiences if they do not wish to do so (this includes the diary recordings which will not be obligatory). However, playing the game might make participants aware of the challenges in their lives. The game is designed to suggest coping methods for some challenges but not all. Participants (and their legal guardians in the case of minors) will be informed in advance about the available help in their respective countries. This information will be provided in leaflets, Participant Information Sheets and/or Informed Consent Forms. Moreover, helplines and other available resources will be included in the automatic responder in the project-dedicated email – whenever a participant or their legal guardian emails the research team, they will receive this information.

13.1. Adverse Events Monitoring

13.1.1. Key definitions

Adverse Event (AE)[1] = any untoward medical occurrence, unintended disease or injury, or untoward clinical signs (including abnormal laboratory findings) in participants, users or other persons, in the context of a clinical investigation, whether or not related to the investigational medical device.

Serious Adverse Event (SAE) = adverse event that leads to any of the following:

- death,
- serious deterioration in the health of the participant, that resulted in any of the following:
    - a life-threatening illness or injury, or
    - permanent impairment of a body structure or a body function, or
    - hospitalisation or prolongation of patient hospitalization, or

[1] This definition includes events related that are anticipated as well as unanticipated events. This definition includes events occurring in the context of a clinical investigation related to the investigational device, the comparator or the procedures involved.

  - medical or surgical intervention to prevent life-threatening illness or injury or injury or permanent impairment to a body structure or a body function
  - chronic disease
- foetal distress, foetal death or a congenital physical or mental impairment or birth defect.

13.1.2. Handling of AE

Adverse events will be reported during the complete participation in the study. This also includes adverse events that occur after the participant's termination of the study and if a causal relationship to the intervention or study procedures is assumed.

All adverse events observed during any period of a clinical trial are to be recorded, treated by the investigator as needed, and followed up. This also includes adverse events that occur after the completion of the clinical investigation and are still related to the intervention or investigation procedures.

Non-serious adverse events/adverse device effects will also be documented in the eCRF (electronic Case Report Form; Appendix 2) and listed, verified during the routine monitoring visits.

All adverse events/adverse device effects will be listed, statistically analysed, and verified and evaluated regarding the causal relationship to the investigation by the sponsor.

13.1.3. Handling and reporting of SAE

All serious adverse events that are ongoing after the completion of the clinical investigation will be followed up until a final assessment can be made. Serious adverse events are also documented (in the eCRF; Appendix 2) and verified during routine monitoring visits.

The investigator must report all serious adverse events/adverse device effects to the sponsor immediately.

After receipt, the SAE report will be checked for correctness, completeness, and plausibility by the coordinator of the investigation. If necessary, further information will be requested from the investigation site. The report will then be forwarded to the coordinating investigator for review, classification and assessment.

In the scope of the SAE assessment, a comparison with the investigation-specific risk analysis will be performed. If the reported SAE is assessed or not assessed according to its severity regarding the risk analysis, corrective and preventive action is determined. This will be reported to the sponsor. The suspension of the investigation will be taken into consideration until the corresponding action will have completely been taken.

13.1.4. Procedures and timelines for SAE reporting

The site coordinators have to report SAEs to the competent authority immediately (without undue delay) if a causal relationship between the SAE and the intervention or study procedures performed as part of the trial or other conditions of the trial conduct cannot be excluded.

13.1.5. Documentation of AE and SAE

All adverse events and adverse effects (serious or non-serious) will be documented in the adverse event report pages of the eCRF (Appendix 2).

13.1.6. Assessment of causality

The causality of AEs will be first assessed by the principal investigator at the site at with an AE took place and then confirmed by the chief investigator. The relationship or association of the AE to the investigated product will be characterized using definitions in Table 3.

Table 3. *Causality definitions for AE and SAE reporting*

Causality category	Definition
Definitely not	The relationship to the intervention or research procedures can be excluded when: the event is not a known[2] side effect of the category the intervention belongs to or of similar interventions and procedures. the event has no temporal relationship with the use of the intervention or the procedures. the serious event does not follow a known response pattern to the intervention (if the response pattern is previously known) and is biologically implausible. the discontinuation of intervention or the reduction of the level of activation/exposure - when clinically feasible - and reintroduction of its use (or increase of the level of activation/exposure), do not impact on the serious event. the event involves a body-site, or an organ not expected to be affected by the intervention or procedure. the serious event can be attributed to another cause (e.g., an underlying or concurrent illness/ clinical condition, an effect of another intervention, drug, treatment, or other risk factors). In order to establish the non-relatedness, not all the criteria listed above might be met at the same time, depending on the type of intervention/procedures and the serious event.
Probably not	The relationship with the intervention seems not relevant and/or the event can be reasonably explained by another cause, but additional information may be obtained.
Possibly	The relationship with the intervention is weak but cannot be ruled out completely. Alternative causes are also possible (e.g., an underlying or concurrent illness/ clinical condition or/and an effect of another intervention, drug, or treatment). Cases where relatedness cannot be assessed, or no information has been obtained should also be classified as possible
Probably	The relationship with the intervention seems relevant and/or the event cannot reasonably be explained by another cause, but additional information may be obtained.

[2] When the event is not a known side effect of the category the intervention belongs to or of similar interventions and procedures, generally is considered “not related”. Yet, the unexpected effect shall not be excluded from evaluation and reporting.

Definitely	The serious event is associated with the intervention or with procedures beyond reasonable doubt when: the event is a known side effect of the category the intervention belongs to or of similar interventions and procedures. the event has a temporal relationship with intervention or procedures. the serious event follows a known response pattern to the intervention (if the response pattern is previously known). the discontinuation of the intervention (or reduction of the level of activation/exposure) and reintroduction of its use (or increase of the level of activation/exposure), impact on the serious event (when clinically feasible). other possible causes (e.g., an underlying or concurrent illness/ clinical condition or/and an effect of another intervention, drug, or treatment) have been adequately ruled out. In order to establish the relatedness, not all the criteria listed above might be met at the same time, depending on the type of intervention/procedures and the serious event.

13.1.7. Sponsor`s reporting of SAE and device deficiencies

Any SAE that has a causal relationship with the intervention or trial processes, or where such causal relationship is reasonably possible. These “reportable events” must be reported within the following timelines:

- A reportable event which results in imminent risk of death, serious injury, or serious illness that requires prompt remedial action for other participants must be reported immediately, but not later than 2 calendar days after awareness to the sponsor of a new reportable event or of new information in relation with an already reported event.
- Any other reportable event or a new finding/ update so it must be reported immediately, but not later than 7 calendar days following the date of awareness by the sponsor of the new reportable event or of new information in relation with an already reported event.

Every side adverse event occurrence will be documented in the eCRF (Appendix 2). Other intervention or trial processes related complications must be reported by mail or email within 7 calendar days.

14. Withdrawal and termination of participation

14.1. Stopping rules for individual participants

With regard to the safety and selection of participants, all required efforts have to be undertaken to ensure that the participant completes the investigation.

Participants may be excluded at any time from the investigation. Specific reasons for excluding a patient may be:

- Voluntary termination by the participant who is free to terminate his/her participation in the study at any time, without prejudice to further treatment.
- Serious violations of the Protocol identified by the investigator and/or sponsor.

- Unauthorized admission, i.e., the participant does not meet the required inclusion/exclusion criteria for the study.
- Any adverse event or illness that, in the opinion of the chief investigator, may prevent further participation in the study.

14.2. Withdrawal and replacement of a participant

Participants can leave the study at any time for any reason if they wish to do so without any consequences. The investigator can decide to withdraw a participant from the study for urgent medical reasons. Participants will be informed that their participation is voluntary and that they can withdraw from the study at any time without having to provide a reason. Participants will be given a choice upon their study withdrawal whether for their data collected up to that point to be retained in the study or if they wish for their data to be removed from the study datasets. A removal of already collected data will only be possible before data has been anonymised. Participants will be informed about the process of pseudonymisation and anonymisation before they agree to participate. To ensure that participants are aware that they can withdraw from the study at any time for any reason without any consequences, this is described in detail in the Participant Information Sheet and the Informed Consent Sheet. When participants sign the informed consent sheet, they agree that they are aware of the voluntary nature of the study. Participants who stop the study should always be asked for the reason for their interruption and the occurrence of adverse events. Adverse events must be tracked.

The number of withdrawals and the reasons for withdrawal will be documented and reported in the dissemination of the study. It will be examined whether these participants differ in any characteristics from those who did not withdraw from the study. Participants can withdraw from the study at any time for any reason if they wish to do so without any consequences. The data of participants, who choose to withdraw from the study, will continue to be used or anonymised if participants require their data to be anonymised depending on national legislation.

Upon participant's request their identifiable data will be removed.

There will be no replacement of participants once the trial starts.

15. Privacy Policy

Storage and analysis will follow the data protection regulations of the provisions of the Data Protection Act 2018 as well as the EU General Data Protection Regulation (GDPR). The data important for the study will be transferred and stored in pseudonymised form, evaluated and, if necessary, passed on to the partnering institutions based on the Data Processing Agreement that will be prepared and signed within the SMILE consortium, who are: Universitats Klinikum Heidelberg (Germany), FTK - Forschungsinstitut für Telekommunikation und Kooperation e. V. (Germany), RDIUP (France), Nurogames GmbH (Germany), Heriot-Watt University (UK), IRCCS-Azienda Ospedaliera-Universitaria Bologna (Italy), Universita di Bologna (Italy), University of Maribor (Slovenia), Municipality of Maribor (Slovenia), C.I.P. Citizens in Power (Cyprus), NVISION Systems and Technologies, S.L. (Spain), WIZ Development & Services SRL (Romania), SWPS University (Poland), Fundación INTRAS (Spain), and The University of Edinburgh (UK).

The study management will take all reasonable steps to ensure the protection of the data is in accordance with European Union data protection standards.

The data is secured against unauthorized access. Decryption is only carried out by the study management. As soon as it is possible for the research or statistical purpose, the personal data will be anonymised. After completion of the study, the decryption list will be destroyed so that none of the data can be traced back to the person (anonymised). The evaluation of the

data is carried out in pseudonymised form. When results are published, no reference to the persons will be possible. Anonymised data will be stored for up to 10 years from the project completion.

16. Legal and ethical aspects

The investigation is conducted in accordance with the Declaration of Helsinki in its current version. The study will not commence before the approval from the Ethics Committee is obtained. Each trial partner is responsible for obtaining own approval.

In this study minors will be invited to take part in the Living Labs testing phase and the proof-of-concept study. Following chapters 28 – 30 of the Declaration of Helsinki the participation of adolescents is a necessary requirement for successfully fulfilling the main target criteria of this study, as the perspective of children and adolescents is a key issue for understanding their needs. Adolescent participants, who are not yet at a legal age, will be informed and invited after parents or those with parental authority have given their consent (§ 28). Children and adolescents will be informed in a way that is adequate and understandable for their respective age group. In case of approval by parents or those with parental authority but rejection by the adolescents, the rejection will be respected (§ 29). The study won't be conducted without informed consent of the adolescents, their parents or those with parental authority (§ 30). In the final phase, that is semi-structured interviews, only adults will participate. They will also be informed about the benefits and risks of participation and will be asked to sign respective Informed Consent Forms or complete their digital equivalents.

Participation is voluntary. The consent can be withdrawn by the participant at any time, without giving reasons. The study attendees will be informed in writing prior to the start of the study about the nature and scope of the planned investigation, the possible benefits and any risks. Their consent is collected either digitally or physically. In case of withdrawal from the study, any (data) material already obtained will be destroyed, or the participant will be asked whether they agree to the evaluation of the material.

No commuting accident insurance was taken out. Participants are not insured for study-related trips to or from the test centre. They will be informed of this in advance.

This study is part of the project SMILE which is funded by the European Union Horizon Europe Research and Innovation Programme, Grant Agreement No 101080923.

There are no conflicts of interest.

17. References

1. Mental health fact sheet [https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response]

2. Arias D, Saxena S, Verguet S: Quantifying the global burden of mental disorders and their economic value. eClinicalMedicine 2022, 54:101675.

3. Mental health of adolescents [https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health]

4. Brás M, Elias P, Cunha FF, Martins C, Nunes C, Carmo C: Vulnerability to Suicide Ideation: Comparative Study between Adolescents with and without Psychosocial Risk. Healthcare (Basel, Switzerland) 2023, 11(19).

5. Looijmans M, van Bergen D, Popma A, van Eijk N, Mérelle S, van Veen S, Hawton K, Gilissen R: The self-perceived needs of adolescents with suicidal behaviour: a scoping review. European child & adolescent psychiatry 2023.

6. Pillai R, Patel I, Balkrishnan R: Risk factors and strategies for prevention of depression, anxiety and eating disorders among adolescents during the coronavirus disease 2019 (COVID-19) pandemic: a review. Global health journal (Amsterdam, Netherlands) 2023, 7(2):78-86.

7. Noh D, Kim H: Effectiveness of Online Interventions for the Universal and Selective Prevention of Mental Health Problems Among Adolescents: a Systematic Review and Meta-Analysis. Prevention science : the official journal of the Society for Prevention Research 2023, 24(2):353-364.

8. Fleming TM, Cheek C, Merry SN, Thabrew H, Bridgman H, Stasiak K, Shepherd M, Perry Y, Hetrick SJRdPyPC: Serious games for the treatment or prevention of depression: A systematic review. 2014, 19(3).

9. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, Lucassen M, Lau HM, Pollmuller B, Hetrick S: Serious Games and Gamification for Mental Health: Current Status and Promising Directions. Frontiers in psychiatry 2016, 7:215.

10. Martinez K, Menéndez-Menéndez MI, Bustillo A: Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: Systematic Review. JMIR Serious Games 2021, 9(4):e30482.

11. Dechant MJ, Frommel J, Mandryk RL: The Development of Explicit and Implicit Game-Based Digital Behavioral Markers for the Assessment of Social Anxiety. Frontiers in psychology 2021, 12:760850.

12. Petsani D, Konstantinidis E, Katsouli AM, Zilidou V, Dias SB, Hadjileontiadis L, Bamidis P: Digital Biomarkers for Well-being Through Exergame Interactions: Exploratory Study. JMIR Serious Games 2022, 10(3):e34768.

13. Laurisz N, Ćwiklicki M, Żabiński M, Canestrino R, Magliocca P: Co-Creation in Health 4.0 as a New Solution for a New Era. Healthcare (Basel, Switzerland) 2023, 11(3).

14. Santonen T, Petsani D, Julin M, Garschall M, Kropf J, Van der Auwera V, Bernaerts S, Losada R, Almeida R, Garatea J et al: Cocreating a Harmonized Living Lab for Big Data-Driven Hybrid Persona Development: Protocol for Cocreating, Testing, and Seeking Consensus. JMIR research protocols 2022, 11(1):e34567.

15. Litvin S, Saunders R, Maier MA, Lüttke S: Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: A randomized controlled trial. PloS one 2020, 15(9):e0237220.

16. Fenn K, Byrne M: The key principles of cognitive behavioural therapy. InnovAiT 2013, 6(9):579-585.

17. Stallard P: Think good-feel good A cognitive behavior therapy workbook for children and young people. Cognitive behaviour therapy 2003, 32(1):47-48.

18. Facione PAJIa: Critical thinking: What it is and why it counts. 2011, 1(1):1-23.

19. Scaini S, Belotti R, Ogliari A, Battaglia M: A comprehensive meta-analysis of cognitive-behavioral interventions for social anxiety disorder in children and adolescents. Journal of anxiety disorders 2016, 42:105-112.

20. Dechant MJ, Frommel J, Mandryk R: Assessing Social Anxiety Through Digital Biomarkers Embedded in a Gaming Task. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Yokohama, Japan: Association for Computing Machinery; 2021: Article 544.

21. Wols A, Lichtwarck-Aschoff A, Schoneveld EA, Granic I: In-Game Play Behaviours during an Applied Video Game for Anxiety Prevention Predict Successful Intervention Outcomes. Journal of Psychopathology and Behavioral Assessment 2018, 40(4):655-668.

22. Tsionas A, Lazaridis A, Vlahavas IJPotG: Serious game development for the diagnosis of major depressive disorder cases using machine learning methods. 2020, 2020:2-4.

23. Nicholson SJGie, business: A recipe for meaningful gamification. 2015:1-20.

24. Brooke J: SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, Weedermeester A, McClelland II (Eds.): Usability evaluation of industry. Taylor & Francis. 1996: 189-194.

25. Gao M, Kortum P, Oswald FL: Multi-language toolkit for the system usability scale. International Journal of Human–Computer Interaction 2020, 36(20): 1883-1901.

26. Laugwitz B, Schrepp M, Held T: Construction and evaluation of a user experience questionnaire. In: Holzinger A (Ed.): USAB 2008, LNCS 5298: 63-76.

27. Schrepp M, Hinderks A, Thomaschewski J: Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In: Marcus A (Ed.): Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science 2008, 8517: 383-392.

28. Baauw E, Markopoulous P: A comparison of think-aloud and post-task interview for usability testing with children. Proceedings of the 2004 conference on Interaction design and children: building a community 2004: 115-116.

29. Pawson R, Tilley NJEftscAh: An introduction to scientific realist evaluation. 1997, 1997:405-418.

30. Rädiker S, Kuckartz U: Analyse qualitativer Daten mit MAXQDA: Springer; 2019.

31. Braun V, Clarke V: Using thematic analysis in psychology. Qualitative Research in Psychology 2006, 3:77-101.

32. Supplementary materials: Sample size calculation for multiple period cluster randomized trial designs - Examples for stepped-wedge cluster randomized trials (s) [https://rpubs.com/derek6561/swcrtexample]

33. Linardon J, Rosato J, Messer M: Break Binge Eating: Reach, engagement, and user profile of an Internet-based psychoeducational and self-help platform for eating disorders. International Journal of Eating Disorders 2020, 53(10):1719-1728.

18. Appendices

Appendix 1: Proof-of-concept study outcome measures

Table 1. Self-report measures used in the SMILE proof-of-concept study.

Variable*	Outcome*	Questionnaire	Age group	Validation/Source
Anxiety	Primary	Penn-State Worry Questionnaire for Children (PSWQ-C)	10-17	[1]
Anxiety	Primary	Generalized Anxiety Disorder Scale (GAD-7)	18-24	[2]
Depression	Primary	Patient Health Questionnaire for Adolescents (PHQ-A)	10-17	[3, 4]
Depression	Primary	Patient Health Questionnaire (PHQ-9)	18-24	[5]
Well-being	Secondary	Warwick-Edinburgh Mental well-being scale (WEMWBS)	10-24	[6, 7]
Resilience	Secondary	Child and Youth Resilience Measure (CYRM-12)	10-17	[8]
Resilience	Secondary	Brief Resilience Scale (BRS)	18-24	[9]
Privacy concerns*	Exploratory	Online privacy concerns (adapted to game)**	10-24	[10]
Acceptance SMILE Game	Exploratory	Unified Theory of Acceptance and Use of Technology Scale (UTAUT2)***	10-24	[11]
Acceptance SMILE Companion app	Exploratory	Unified Theory of Acceptance and Use of Technology Scale (UTAUT2)***	10-24	[11]
User experience SMILE Game	Exploratory	User Experience Questionnaire short version (UEQ_S)	15-24	[12]
User experience SMILE Companion app	Exploratory	User Experience Questionnaire short version (UEQ_S)	15-24	[12]

* Variables and outcomes are subject to change based on the results of interim analyses.
** Originally used by Dhir et al. (2017) [10] in relation to social media.
*** Originally used to Venkatesh et al. (2012) [11] in relation to mobile Internet.
here “mobile Internet” was replaced with “the SMILE Game” and “the SMILE Companion App”.

Table 2. Frequency of self-report measures across the 6-week intervention period.

	T0		T1	T2				T3
	Sequence 1			Sequence 2		Sequence 3		Post-test
	Week 1	Week 2		Week 3	Week 4	Week 5	Week 6	
Anxiety	x			x		x		
Depression		x			x		x	
Well-being	x			x		x		
Resilience	x			x		x		
Privacy concerns	x							
Acceptance (SMILE Game)								x
Acceptance (SMILE Companion App)								x
User experience (SMILE Game)								x
User experience (SMILE Companion App)								x

Validation/Source:

1. Chorpita BF, Tracey SA, Brown TA, Collica TJ, Barlow DH: **Assessment of worry in children and adolescents: An Adaptation of the Penn State Worry Questionnaire**. *Behaviour research and therapy* 1997, **35**(6):569-581.

2. Spitzer RL, Kroenke K, Williams JB, Löwe B: **A brief measure for assessing generalized anxiety disorder: the GAD-7**. *Archives of internal medicine* 2006, **166**(10):1092-1097.

3. Cortez AB, Wilkins J, Handler E, Lerner MA, Burchette R, Wissow LS: **Multistage Adolescent Depression Screening: A Comparison of 11-Year-Olds to 12-Year-Olds**. *The Permanente journal* 2021, **25**.

4. Johnson JG, Harris ES, Spitzer RL, Williams JB: **The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients**. *J Adolesc Health* 2002, **30**(3):196-204.

5. Kroenke K, Spitzer RL, Williams JB: **The PHQ-9: validity of a brief depression severity measure**. *J Gen Intern Med* 2001, **16**(9):606-613.

6. Melendez-Torres GJ, Hewitt G, Hallingberg B, Anthony R, Collishaw S, Hall J, Murphy S, Moore G: **Measurement invariance properties and external construct validity of the short Warwick-Edinburgh mental wellbeing scale in a large national sample of secondary school students in Wales**. *Health and Quality of Life Outcomes* 2019, **17**(1):139.

7. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, Parkinson J, Secker J, Stewart-Brown S: **The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation**. *Health Qual Life Outcomes* 2007, **5**:63.

8. Liebenberg L, Ungar M, LeBlanc JC: **The CYRM-12: a brief measure of resilience**. *Canadian journal of public health = Revue canadienne de sante publique* 2013, **104**(2):e131-135.

9. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J: **The brief resilience scale: assessing the ability to bounce back**. *International journal of behavioral medicine* 2008, **15**(3):194-200.

10. Dhir A, Torsheim T, Pallesen S, Andreassen CS: **Do Online Privacy Concerns Predict Selfie Behavior among Adolescents, Young Adults and Adults?** *Frontiers in psychology* 2017, **8**:815.

11. Venkatesh V, Thong, JYL, & Xu, X: **Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology**. *MIS Quarterly* 2012, *36*(1), 157–178.

12. Schrepp M, Hinderks A, Thomaschewski J: **Design and evaluation of a short version of the user experience questionnaire (UEQ-S)**. *International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108* 2017.

Appendix 2: Adverse Events reporting form

Adverse Event Report Form

Study Title: SMILE Proof-of-Concept Study **Chief Investigator:** Prof. Matthias Schwannauer

Study Information:

Participating site:	
Event reported by:	
Email address of person completing form:	
Participant ID:	

Adverse event:

Date of report:	/ /
Type of report:	Initial ☐ Follow up ☐
Unique adverse event reference:	
Adverse event summary (Written description):	
Start date of adverse event:	/ /
Start time of adverse event:	: Time unknown ☐
Date PI aware of adverse event:	/ /
Time PI aware of adverse event:	: Time unknown ☐
Resolution date:	/ /
Resolution time:	: Time unknown ☐
Participant outcome:	☐ Resolved ☐ Resolved with sequelae ☐ Ongoing ☐ Death ☐ Unknown
Action taken with participant: If discontinued, please complete a withdrawal form.	☐ Continued with study ☐ Discontinued study
	If discontinued, was it due to AE? ☐ Yes ☐ No

Seriousness classification:

Serious adverse event?	Yes ☐	No ☐
	If **Yes**, report to Sponsor within 7 days and complete seriousness criteria section below If **No**, do not complete Seriousness Criteria section below	
An AE is "serious" if it results in the following: (a) Death; (b) Life-threatening illness or injury; (c) Disability or incapacity (including permanent impairment of a body structure or a body function); (d) Medical or surgical intervention to prevent the above; (e) Requires hospitalisation or extends hospitalisation; (f) Foetal distress or death; or (g) Otherwise medically significant.		

Seriousness criteria:

Please tick at least one box and provide detailed additional information where required. Complete only if answered "Yes" in the Seriousness Classification section.										
☐ Death										
Date of death:	D	D	/	M	M	/	Y	Y	Y	Y
Cause of death:										
☐ Life threatening illness or injury										
☐ Disability or incapacity										
☐ Medical or surgical intervention to prevent life-threatening illness, injury, disability, or incapacity										
☐ Initial hospitalisation or prolonged admission										
Initial hospitalisation ☐ OR Prolonged admission ☐										
Physical Health ☐ OR Mental Health ☐										
Admission date:			/			/				
Discharge date:	D	D	/	M	M	/	Y	Y	Y	Y
☐ Led to foetal distress or death										
☐ Other										
Please specify:										

Causality, severity, and expectedness:

Relationship to app (software-related)	Definitely not ☐	Probably not ☐	Possibly ☐	Probably ☐	Definitely ☐
Relationship to hardware (handset, network)	Definitely not ☐	Probably not ☐	Possibly ☐	Probably ☐	Definitely ☐
Relationship to other trial procedure:	Definitely not ☐	Probably not ☐	Possibly ☐	Probably ☐	Definitely ☐
Severity grade:	Mild ☐	Moderate ☐	Severe ☐		
Expectedness:	Anticipated ☐	Unanticipated ☐			

Event narrative:

Narrative (e.g. background and context, onset of symptoms, treatment, medications, outcome, reason for causality assessment):
Relevant Medical History (e.g. please state any relevant pre-existing conditions):

Details of team discussion:
To include date of discussion, who present and decision.

Guidance from DMEC:

Follow-up (if not resolved at time of completing form):

Signatures.										
PI Name:										
PI signature:										
Date:			/			/				
CI Name:										
CI signature:										
Date:			/			/				