FEBRUARY 2024

Alena is a safe and effective digital therapeutic for social anxiety

Clinical trial report

Prepared in 2024 by

alena

Acknowledgements

This report was prepared by Dr Jessica McFadyen, Alex Long, and Dr Mandana Ahamdi.

The clinical trials presented in this report were designed and implemented by Dr Tayla McCloud, Alex Long, Dr Mona Garvert, and Dr Sofie Meyer. Data analysis was conducted by Dr Jessica McFadyen.

Contributions to the creation and design of the *Alena* digital therapy program were made by Sandra Sobanska, Dr Sofie Meyer, Alex Long, Dr Stuart Linke, Dr Quentin Huys, Blanca Guardia, Dr Gwydion Williams, Dr Jessica McFadyen, Dr Mona Garvert, Sebastian Traeneus, and Omar Reid.

Clinical supervision throughout the trial was provided by Dr Quentin Huys and Dr Stuart Linke.

This report and the clinical trials it contains were funded and conducted by Aya Technologies Ltd.

Contents

Summary

Aim

Social anxiety significantly impacts individuals' work, personal relationships, and overall ability to engage in life opportunities, yet seeking face-to-face therapy can be daunting and inaccessible for those affected. Alena aims to address this gap by providing a convenient, digital solution. The primary objective of this research was to gather robust scientific evidence supporting the effectiveness of Alena, a digital cognitive-behavioural therapy (CBT) app, in alleviating symptoms of social anxiety disorder.

Methods

This research comprised two randomised controlled trials (RCTs):

1. **RCT #1 (2022)**: A pilot study involving 102 female participants aged between 18 and 35.
2. **RCT #2 (2023)**: A more extensive study with 267 participants (64% females, 36% males), aged between 18 and 75.

Both studies divided participants into two groups: 1) an intervention group receiving access to the digital therapy app, and 2) a waitlist control group. RCT #1 provided access for 4 weeks, while RCT #2 extended this to 8 weeks. Participants completed weekly surveys assessing their social anxiety symptoms using the Social Phobia Inventory (SPIN) and Work and Social Adjustment Scale (WSAS) and were monitored for safety. Follow-up surveys were conducted after the intervention periods (2-week follow-up for RCT #1 and 4-week follow-up for RCT #2).

Results

The intervention groups in both RCTs experienced a threefold greater reduction in social anxiety symptoms compared to the waitlist group, as measured by the SPIN. Participants also reported improvements in work, home management, and social and private leisure activities, as gauged by the Work and Social Adjustment Scale (WSAS). There were no significant safety issues reported with the app's use. Participants found the app helpful, user-friendly, and were likely to recommend it. The intervention group showed a high median completion rate of 84%-91% for the therapy program, despite not being incentivised to do so. Overall, the consistency of outcomes across both studies strengthens the overall reliability of this research.

Introduction

As highly social creatures, we are highly susceptible to experiencing social anxiety at some point in our lives. Social anxiety is broadly defined as a fear of being negatively evaluated by others. It likely evolved as a way to enhance group cohesion, encourage risk avoidance, and facilitate cooperation, thereby increasing the individual's chances of reproductive success and survival[1]. In many ways, these advantages still apply in today's world.

Social anxiety exists on a continuum, affecting almost everyone to some degree. For some people, however, social anxiety can increase to an unmanageable level and start to harm our quality of life[2]. Significant anxiety, self-consciousness, and embarrassment in everyday interactions can impair our ability to make and sustain personal relationships[3], and is a risk factor for mental health complications such as depression[4] and substance abuse[5]. Avoidance of social situations can significantly interfere with work and school life and can lead to social isolation, missed opportunities, and loneliness[6].

It is estimated that 36% of the global population have symptoms of social anxiety that exceed a clinical threshold, and yet half of these people do not perceive themselves as having social anxiety[7]. This highlights not only the extremely high prevalence of social anxiety, but also a possibility that it can impact our lives in subtle yet detrimental ways. Even those who might not be considering seeking professional help might therefore benefit from effective treatment for social anxiety.

Social anxiety is commonly treated with talking therapies, such as cognitive-behavioural therapy (CBT)[8]. CBT is highly effective in treating social anxiety disorder[9] and is the treatment recommended by the National Institute for Health and Care Excellence (NICE). The NICE guidelines endorse a CBT program centred on the Clark & Wells cognitive model of social anxiety disorder[10], which emphasises the role of cognitive processes and behavioural patterns in the development and maintenance of social anxiety. It identifies key elements such as negative thoughts about the self in social situations, attentional biases towards perceived threats in social interactions, and safety behaviours that individuals employ to mitigate anxiety but which paradoxically maintain it (e.g., while one might avoid eye contact to alleviate anxiety, this can actually reinforce feelings of social disconnection and awkwardness). By addressing these components, CBT under the Clark & Wells model empowers individuals to challenge and change their thought patterns, ultimately altering their emotional and behavioural responses to social situations. This model of CBT is implemented by the National Health Service (NHS) Talking Therapies[11] in the UK and achieves improvement rates of 67.1% and recovery rates of 36.4%[12] for social anxiety.

Despite effective treatment being available, it is estimated that only 4-20% of people with social anxiety ever seek treatment[13]. This may be somewhat unsurprising, given that interacting with a stranger in therapy is likely highly aversive to someone with social anxiety. This paradox underscores a critical need for an accessible, alternative solution — one that caters to a wide spectrum of social anxiety symptoms, from mild to extreme. The development of a digital therapy solution presents a promising avenue, with the potential to make a profound impact on society by offering discreet, immediate, and personalised support. Digital CBT has been shown to be as effective in treating social anxiety disorder as traditional face-to-face therapy[14,15], and thus offers a convenient treatment pathway.

Building on the proven efficacy of CBT and the Clark & Wells model, our company sought to transpose this standard treatment for social anxiety into a modern, digital format. We developed an interactive mobile application called *Alena for social anxiety*, with a focus on an appealing visual design and a highly usable interface. By combining evidence-based therapy with a user-friendly platform, we envisioned an innovative approach to managing social anxiety.

Our goal was to assess whether *Alena* effectively improves symptoms of social anxiety, making a significant stride in mental health treatment accessibility and convenience. In this report, we present the results of two clinical trials that test this hypothesis, exploring the impact of our digital therapy app on individuals with social anxiety. Our findings offer insights into the future of mental health treatment, highlighting the potential of digital solutions in reaching and assisting a broader segment of the population.

Methods

Our approach

We conducted two separate Randomised Controlled Trials (RCTs) - an initial smaller pilot study (RCT #1), followed by a larger and more inclusive trial (RCT #2). The primary goal of these trials was to evaluate the safety and effectiveness of the Alena app. In each trial, participants first underwent a screening process, followed by a baseline assessment one week later. They were then assigned to either a four-week (RCT #1) or eight-week (RCT #2) period where they either had access to the Alena app (**intervention group**) or were placed on a waitlist (**control group**). After the intervention or waitlist period, we conducted follow-up assessment to evaluate the impact of the app.

All of this was conducted entirely online using Prolific, a platform for study recruitment and data collection. The study was approved by the Reading Independent Ethics Committee (study reference: AYSATOL).

Participants

Recruitment

For both trials, we recruited initial pools of participants (350 for RCT #1 and 1,282 for RCT #2) to complete an online screening questionnaire so that we could select participants who experienced at least moderate levels of social anxiety (see Eligibility criteria below). In RCT #1, our pilot study, we specifically recruited women aged between 18 and 35, as research indicates this group is more prone to social anxiety[16] and more likely to adhere to therapy[17]. The broader trial, RCT #2, expanded these criteria to include a wider age range and all biological sexes, allowing us to evaluate the app's effectiveness across a more diverse population.

We collected information on demographics, lifestyle habits, mental health history, and access to technology. Participants provided informed consent and received £1 for their involvement in the screening process.

Eligibility

Participants were screened according to the following criteria:

1. **Social anxiety severity**: We used the Social Phobia Inventory (SPIN)[18] to measure social anxiety levels. To qualify for our study, participants needed to score 30 or higher, indicating a moderate to severe level of social anxiety.
2. **Alcohol use**: We used the Alcohol Use Disorders Identification Test for Consumption (AUDIT-C)[19] to identify any potential risk of alcohol dependence, with participants required to have less than a severe risk.
3. **Prior use of Alena app**: Participants who had previously used the Alena app were excluded to ensure that all participants had similar levels of exposure to the treatment at the study's start.
4. **Recreational drug usage**: We asked questions about recreational drug usage, with participants required to have minimal to no use to qualify.
5. **Recent changes in mental health medication**: If participants were on mental health-related medication, they needed to have been on a stable dose for at least the past 8 weeks.
6. **Technology access**: For RCT #1, participants needed access to an iPhone with an internet connection, as the Alena app was initially developed for the iOS platform. In RCT #2, this requirement was expanded to include both iOS and Android devices, allowing a broader range of participants.

Through these criteria, we aimed to create a study group that was representative of the target population for the Alena app.

Procedure

Once participants passed the eligibility criteria, they were randomly placed into one of two groups with a 1:1 ratio: the intervention group, which would use the Alena app, or the waitlist control group. The study began with an initial survey to establish a baseline understanding of each participant's social anxiety levels, using the SPIN, and the impact of this anxiety on their daily functioning, assessed with the Work and Social Adjustment Scale (WSAS)[20]. We also gathered information on demographics, expectations about the Alena app, and previous experiences with mental health apps.

Following this initial survey, participants were notified about their group assignment. Those in the intervention group received detailed instructions on how to download and access the Alena app. Those in the control group were informed that they would receive access to the Alena app in several weeks' time (4 weeks in RCT #1 and 8 weeks in RCT #2).

During the course of the intervention/waitlist period - four weeks for RCT #1 and eight weeks for RCT #2 - participants were asked to fill out weekly surveys. These surveys repeated the SPIN and WSAS measures to track changes in their social anxiety and its effects over time. For those with access to the Alena app, additional questions regarding

their usage of the app were included. To encourage consistent participation, we offered a compensation of £5 per week for completing these surveys. Additionally, reminder emails were sent to the intervention group to encourage (but not incentivize) regular engagement with the app.

After the designated intervention or waitlist period, access to the Alena app was revoked for the initial intervention group and granted to the waitlist group. To assess the lasting impact of the Alena program, we conducted a follow-up survey several weeks after the end of the intervention/waitlist period - two weeks later for RCT #1 and four weeks later for RCT #2. This survey aimed to evaluate any sustained clinical improvements and gather final feedback from participants.

Intervention

The Alena app represents a digital adaptation of CBT specifically tailored for social anxiety, based on the Clark & Wells cognitive model of social phobia[10]. This program was carefully designed in line with the University College London (UCL) CBT competencies framework, a standard in psychological therapy practices, and developed under the expert guidance of a clinical psychologist and a clinical psychiatrist.

Program structure and content

The Alena app's program was organised into five distinct modules, each targeting a key mechanism of social anxiety disorder:

1. **Introduction:** Serves as an introductory overview, setting the stage for the program and providing insight into the drivers of social anxiety symptoms
2. **Beliefs:** Focuses on conditional beliefs about oneself and others
3. **Attention:** Concentrates on self-awareness and self-focus during social interactions
4. **Avoidance:** Deals with safety behaviours and avoidance patterns
5. **Rumination:** Addresses the tendency to overthink or analyse social interactions after they occur

Each module contained various elements, including psychoeducational audio lessons and practical worksheets that help participants to challenge their thought and behaviour patterns in a guided manner. In RCT #2, we introduced game-like assessments to engage participants further and assess their cognitive and behavioural patterns related to social anxiety. These assessments, lasting between 5 to 15 minutes, were positioned at the start of

each module, and completion was required to unlock the rest of the exercises within that module.

Program pacing

The exercises, each taking between 1 to 8 minutes to complete, were thoughtfully designed to fit into the user's daily routine. The app encouraged participants to repeat exercises if needed and to extend their learning outside the app through real-life exposure experiments, supported by in-app exercises that assisted participants with planning and reflecting on these experiments.

To optimise the learning curve and ensure a structured progression through the program, the availability of modules was controlled. In RCT #1, modules were sequentially unlocked each week, while in RCT #2, all modules were accessible from the start, but participants were advised to complete one module every two weeks. To complete all recommended content in the app, participants would have needed to spend between 10 and 20 minutes on the app per week.

Outcome measures

To comprehensively evaluate the safety and efficacy of the Alena app, our study utilised a set of outcome measures divided into primary and secondary categories.

Primary outcome measures

Efficacy was measured with the SPIN[18], a 17-item self-rating scale designed to measure the severity of social anxiety disorder. The scale is rated over the past week and includes items assessing the spectrum of fear, avoidance, and physiological symptoms. A SPIN score ranges between 0 and 68, with a score of 19 distinguishing between individuals with social phobia and controls and a reduction of 10 points or more indicating significant improvement.

Safety was monitored through items in the weekly surveys that asked participants to report any new serious health effects experienced in the past week. Any reported effects were reviewed by a psychiatrist who determined whether the effect matched criteria for a "Serious Adverse Event", as defined by the ISO 14155 (A:14)[21].

> A Serious Adverse Event (SAE) is any untoward medical occurrence, unintended disease or injury, or any untoward clinical signs (including an abnormal laboratory finding) in subjects, users or other persons, whether or not related to the investigational medical device, that results in death, is life-threatening, requires (or prolongs existing) hospitalisation, results in persistent or significant

disability or incapacity, causes a congenital anomaly or birth defect, or induces any other condition that may jeopardise the participant or require medical or surgical intervention to prevent one of the listed outcomes.

Secondary outcome measures

Daily functioning was measured with the WSAS, which provides insights into how social anxiety affected participants' ability to do daily tasks in various domains, such as work, social activities, and home management. This scale helped us understand the broader impact of the Alena app beyond the reduction of anxiety symptoms.

Acceptability was measured by asking participants to rate their experience with the Alena app in terms of acceptability, ease of use, helpfulness, and overall satisfaction. These subjective measures offered insight into the app's practicality in a real-world setting.

Therapy adherence is a critical factor in the effectiveness of any digital therapy program, and so we measured adherence by tracking how consistently participants engaged with the Alena app.

Results

Below, we report the findings from both the pilot RCT (#1) and the larger RCT (#2). Details of the statistical analysis behind each finding can be found in the **Statistics** section.

Participants

Retention

Eligible participants from the screening studies (45.14% in RCT #1 and 27.22% in RCT #2) were randomly allocated to either the intervention or waitlist control groups and then invited to participate in the baseline assessment. Recruitment to the baseline assessment was capped once the target sample size was reached (N = 50 per group for RCT #1, and N = 125 per group for RCT #2). Note that in RCT #1, a technical error meant that two extra participants allocated to the intervention group completed the baseline assessment and were thus included in the remainder of the study (hence, N = 52 in the intervention group and N = 50 in the control group). In RCT #2, one participant in the intervention group was excluded from the study due to reporting they no longer had access to a smartphone with Internet access in the baseline assessment (hence, N = 124 in the intervention group and N = 124 in the control group).

After the baseline assessment, participants completed weekly surveys throughout the intervention/waitlist period, and then a final follow-up assessment (see **Table 1**). In RCT #1, 88.22% of participants retained from baseline to follow-up. Retention was similar for both groups. In RCT #2, 90.36% retained from baseline to follow-up, with better retention seen for the waitlist control group (see Statistics section for details).

RCT #1

			Intervention/Waitlist				***Follow-up***
Group	**Screened**	**Baseline**	**Week 1**	**Week 2**	**Week 3**	**Week 4**	**Week 6**
Intervention	350	52	50	51	50	47	45
Waitlist		50	50	50	50	45	44
Total		102	100	101	100	92	89

Number of participants who completed weekly surveys at each stage of RCT #1.

RCT #2

			Intervention/Waitlist								***Follow-up***
Group	**Screened**	**Baseline**	**Wk 1**	**Wk 2**	**Wk 3**	**Wk 4**	**Wk 5**	**Wk 6**	**Wk 7**	**Wk 8**	**Week 12**
Intervention	1282	124	106	109	107	107	105	103	106	105	104

Waitlist		125	123	125	124	124	124	124	121	124	121
Total		249	229	234	231	231	229	227	229	225	225

Table 1. Number of participants who completed weekly surveys at each stage of RCT #2.

Baseline characteristics

In both RCTs, participants were randomly allocated to the intervention and waitlist groups. To ensure that the groups were balanced, we statistically compared a variety of factors such as age, baseline SPIN scores, expectations for the Alena app, alcohol use, educational background, ethnicity, education, and employment (see Statistics for a descriptive table).

The groups in RCT #1 were equivalent on all measures except for age, where the intervention group were older on average (mean age = 29 years) compared to the waitlist group (mean age = 27 years old). The intervention group were also slightly more likely to have had therapy for social anxiety before, or used apps for their mental health. The groups in RCT #2 were perfectly balanced on all baseline characteristics.

Primary outcomes

Efficacy

Our primary clinical outcome was the change in SPIN scores, indicating the severity of social anxiety symptoms. In both trials, all participants started with a median SPIN score of 43 (severe). We tracked how SPIN changed throughout the intervention/waiting period, and then if there were any further changes at follow-up (see **Figure 1**).

RCT #1

In the pilot RCT #1, participants with access to Alena saw a significantly greater reduction in SPIN (9.8 points on average) compared to the waitlist control group (4.1 points on average) throughout the 4-week intervention period. Additionally:

- **Significant improvement:** 51% percent of the intervention group saw a significant (≥ 10 point reduction) improvement in social anxiety, compared to only 22% of the control group.
- **Recovery rates**: 19.15% of participants in the intervention group reached subclinical levels of social anxiety symptoms (SPIN ≤ 19) by the end of the 4-week intervention, compared to 6.67% of the waitlist group. At the 2-week follow-up, 17.78% had reliably recovered, compared to 9.09% of the waitlist group.
- **Lasting effects**: At follow-up, SPIN in the intervention group remained stable, while the waitlist group saw a further reduction of 2.7 points on average (note that, during

this time, the waitlist group had received access to the *Alena* app and 8% were using it).

RCT #2

We observed the same effects in RCT #2, which had a longer intervention period. Here, the intervention group reduced by 12.8 points on average, while the waitlist control group only reduced by 7.5 points on average. Additionally:

- **Significant improvement**: 61% percent of the intervention group saw a significant improvement in social anxiety, compared to only 36% of the control group.
- **Recovery rates**: 21.9% of participants in the intervention group reached subclinical levels of social anxiety symptoms by the end of the 8-week intervention (compared to 10.48% in the waitlist group). At the 4-week follow-up, participants in the intervention group were 2.7 times more likely to have recovered (26.92%) than waitlist participants (11.57%).
- **Last effects**: At follow-up, SPIN in the intervention group reduced by a further 2.8 points on average, while the waitlist group remained stable (zero participants in either group could use the *Alena* app during this time).

Overall, these results suggest that having access to the *Alena* app significantly reduces social anxiety symptoms. Both RCTs show that this improvement persists over time, suggesting a lasting impact of the *Alena* app. In both RCTs, participants in the intervention group were approximately 3 times more likely to improve by 10 points or more (indicating a significant change) than those in the control group.

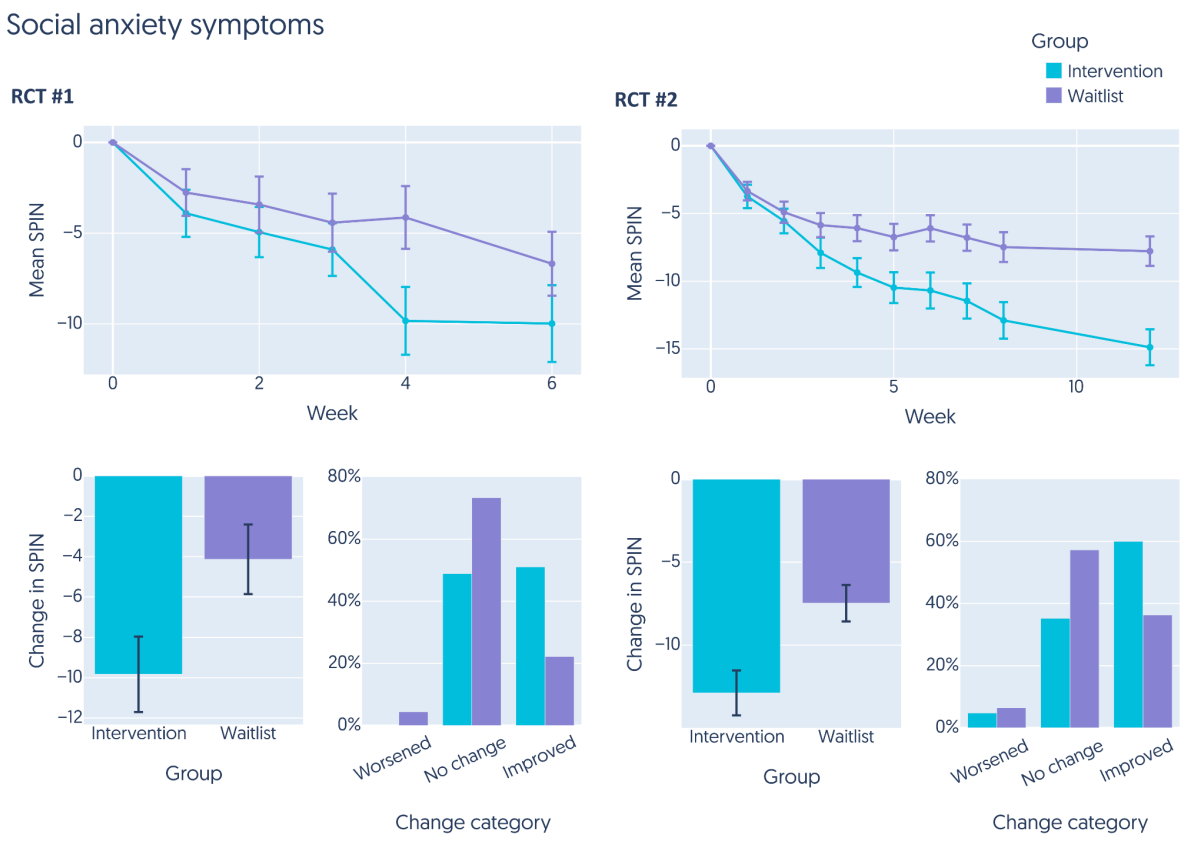

Figure 1. Improvement in social anxiety symptoms over time. Results are shown for RCT #1 (left) and RCT #2 (right). The mean SPIN score for participants in the intervention group (blue) and waitlist control group (purple) was tracked in each week of the study, with the first week (week 0) being the baseline assessment and the last week being the follow-up assessment. All weeks in between were the intervention/waitlist period. Error bars represent the standard error of the mean. The total change in SPIN from week 0 to the final week of the intervention/waitlist period is shown in the left-most bar plot per RCT. The right-most bar plot in each RCT shows the proportion of participants in each group whose change in SPIN indicated a worsening of symptoms (increase in SPIN of 10 points or more), an improvement in symptoms (decrease in SPIN of 10 points or more), or no change.

Safety

An important aim of the study was to evaluate whether using the *Alena* app was associated with any increased risk of serious adverse health effects. Throughout the intervention and waitlist periods, we monitored participant safety through weekly self-reports (see **Figure 2**). Fortunately, very few negative health effects were reported and none were classified as a Serious Adverse Event (see definition in Primary outcomes section). In fact, participants with access to Alena tended to report fewer negative health effects than participants in the waitlist group (but note this difference was not significant).

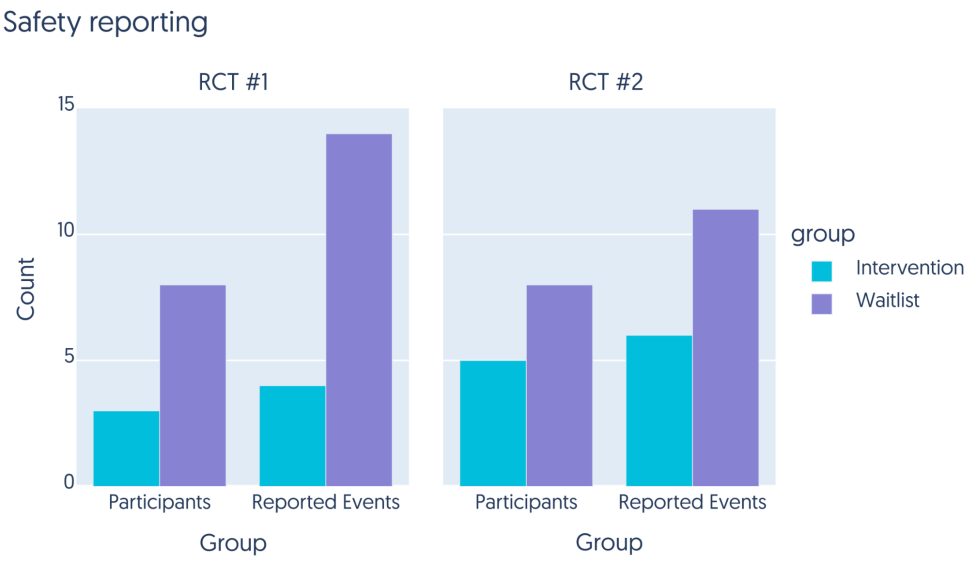

Figure 2. Incidence of adverse health effects. In both RCT #1 (left) and RCT #2 (right), we counted the number of participants reporting serious adverse health effects throughout the intervention/waitlist period, and also the total number of reports made summed across participants, in both the intervention (blue) and waitlist control (purple) groups.

Secondary outcomes

Daily functioning

In addition to clinical efficacy and safety, we also evaluated the impact of *Alena* on daily functioning, as measured by the WSAS. At baseline, all participants reported significant functional impairment due to their social anxiety (the median WSAS score across all participants was 19; see **Figure 3**).

In RCT #1, the intervention group saw significantly greater improvement in daily functioning than the control group. WSAS scores in the intervention group were reduced by 4.5 points on average from baseline to the end of the intervention, compared to a reduction of only 2 points in the control group. At the 2-week follow-up, WSAS scores slightly improved for the waitlist group (1 point on average) and slightly worsened for the intervention group (1 point on average).

In RCT #2, both the intervention group and waitlist control group saw a similar improvement in daily functioning through the 8-week intervention/waitlist period, with an average reduction in WSAS of 4 points in the intervention group and 3 points in the control group. At the 4-week follow-up, WSAS scores reduced to a greater degree for participants in the intervention group (5.2 points on average) than in the waitlist group (3.3 points on average).

Overall, these results underscore the potential of the Alena app not only in reducing social anxiety symptoms but also in enhancing the overall quality of life and daily functioning of individuals.

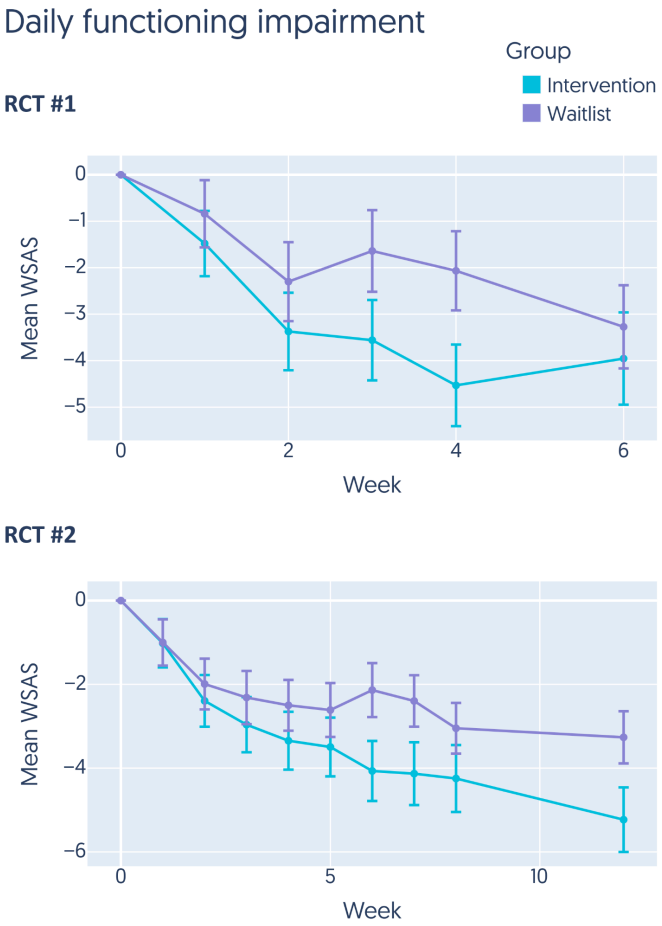

Figure 3. Improvement in daily functioning over time. We used the Work and Social Adjustment Scale (WSAS) to track daily functioning impairment in both RCT #1 (top) and RCT #2 (bottom). The y-axis represents the mean WSAS score across participants in either the intervention (blue) or waitlist control (purple) group each week (x-axis). Error bars represent standard error of the mean. Week 0 is the baseline assessment, and the final week is the follow-up assessment. All weeks in between constitute the intervention/waitlist period.

Acceptability

A high level of acceptability is critical in the context of digital therapies, as user satisfaction and perceived utility are key drivers for continued engagement and adherence to the program. To gauge this, we collected subjective ratings from the intervention group on various aspects of their experience using the app each week. These aspects included overall satisfaction with the app, its perceived helpfulness, the ease of use, and the likelihood of recommending the app to others.

The feedback from participants in both RCT #1 and RCT #2 consistently reflected high levels of acceptability (see **Figure 4**). Participants rated the app highly across all measures, with median ratings reaching 4 out of 5 for satisfaction, helpfulness, and likelihood of recommendation, and median ratings reaching the maximum 5 out of 5 for ease of use. Overall, these findings suggest that the *Alena* app was highly acceptable to participants.

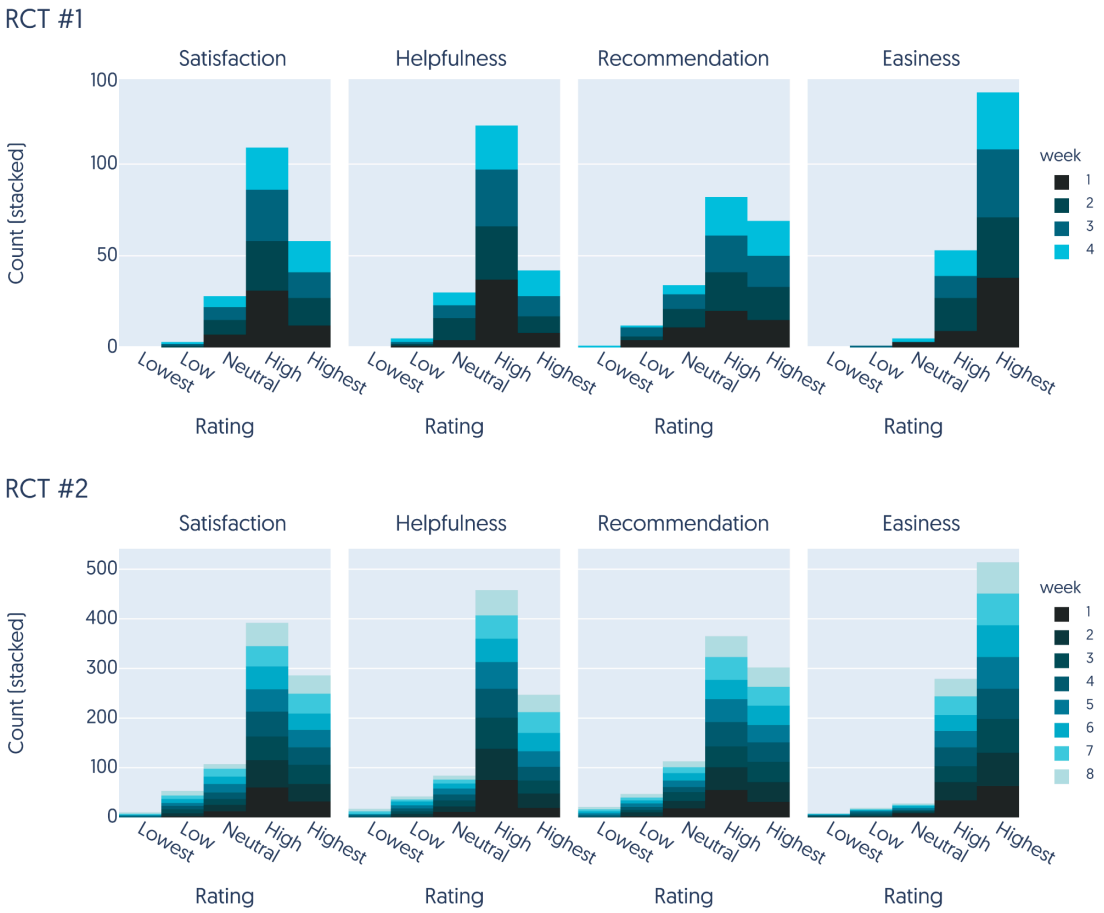

Figure 4. Acceptability ratings of the *Alena* app. We measured acceptability in four categories: how satisfied participants were with the app, how helpful they found the app, how likely they were to recommend the app, and how easy the app was to use. Response ranged from 1 (lowest) to 5 (highest). Measures were taken each week (see legend for colour scale) in both RCT #1 (top) and RCT #2 (bottom).

Therapy adherence

Throughout the intervention period, we tracked how well participants adhered to Alena's therapy program, monitoring the number of audio lessons listened to and interactive worksheets finished by each participant.

The data revealed encouraging levels of adherence to the app's program (see **Figure 5**), despite participants not being incentivised to do so (they were only compensated for the time required to complete the weekly surveys). In RCT #1, participants in the intervention

group showed a commendable median completion rate of 90.91% (mean = 76.92%, SD = 29.17%). For RCT #2, which featured a longer treatment program, the median completion rate was slightly lower but still robust at 84.21% (mean = 70.88%, SD = 27.22%).

We delved deeper to understand the relationship between adherence to the therapy program and clinical outcomes. We found that the total completion rate of therapy content (measured as the number of exercises completed out of the total number of available exercises) did not significantly influence the reduction in social anxiety symptoms, as measured by the SPIN, in either RCT #1 or RCT #2. However, in RCT #2 (but not RCT #1), a higher completion rate correlated with a more substantial improvement in daily functioning, as indicated by the WSAS. Thus, while higher adherence to the digital therapy did not consistently relate to a reduction in social anxiety symptoms, it did relate to improvements in daily functioning in the longer trial. This suggests that sustained interaction with *Alena* could enhance overall well-being.

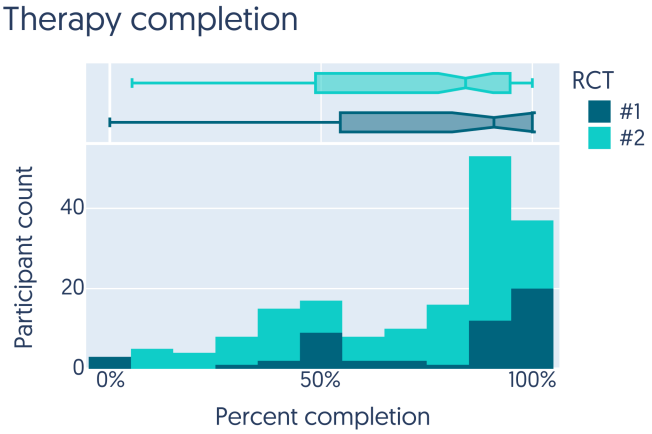

Figure 5. Therapy completion rates. Histogram shows the proportion of exercises completed across participants in the intervention group for RCT #1 (dark blue) and RCT #2 (green). Box-and-whisker plot (top) shows the distribution of therapy completion rates across participants in each RCT, with the median at the notch, the 25th to 75th percentiles represented by the box (i.e., the "interquartile range"), and the "whiskers" of the plot representing each box boundary ± 1.5 × the interquartile range.

Discussion

Our study unveils exciting insights into the effectiveness and safety of the *Alena* app, a digital therapy tool, in reducing social anxiety. Key findings indicate a significant reduction in social anxiety symptoms among users, as well as improvements in their daily functioning, compared to participants in a waitlist control group. The *Alena* app was highly accepted by users, as evidenced by their positive feedback on its usability, helpfulness, and overall satisfaction. The high therapy adherence rates observed indicate that users were not only willing to engage with the app but also maintained their engagement over time. Furthermore, safety monitoring revealed no significant health risks associated with the use of the Alena app, underscoring its safety profile as a mental health intervention. These aspects combined – effectiveness, acceptability, adherence, and safety – highlight the Alena app's potential as a comprehensive, user-friendly, and safe digital solution for reducing social anxiety.

Key Findings

51-65%	3×	84-91%	0	4.75/5
of users significantly reduce social anxiety with *Alena*	more likely to see significant improvement with *Alena*	adherence to *Alena's* therapy program (median completion)	health risks associated with *Alena*	ratings across acceptability measures

The *Alena* app achieves comparable outcomes to NHS Talking Therapies. In data published in 2022, the NHS observed improvement rates of 67.1% and recovery rates of 36.4%[12] for social phobia disorder treated with CBT. This is similar to our own observed improvement rates of 65% and recovery rates of 27% in RCT #2. The NHS also reports an average reduction in WSAS of 5.8 points, which is similar to the 5.2 point reduction we observed in RCT #2. Thus, *Alena* offers an effective alternative or complement to existing treatment options.

Importantly, *Alena's* efficacy either matches[22,23] or surpasses[24] that observed by previous digital interventions for social anxiety, although interventions including support from a human therapist can show enhanced effects[25]. *Alena's* success as a standalone tool is likely partially attributable to the high acceptability of the app to users, afforded by the appealing visual design, easy-to-use interface, and bite-sized therapeutic content. This combination of factors not only makes the *Alena* app appealing to users but also encourages consistent engagement, as evidenced by the high level of adherence to the

therapy program we observed in both RCTs (users completed 84-91% of material in the app).

As we look toward the future of digital mental health solutions, the format of the *Alena* app presents a promising template that could potentially be adapted to a broader spectrum of psychological conditions. *Alena's* success in treating social anxiety disorder provides a strong foundation but careful research and development are needed to adapt it effectively for other conditions, or potentially for transdiagnostic factors that are common across various disorders, such as self-confidence issues in both social anxiety and depression, or pervasive worry seen in both clinical anxiety and non-clinical populations.

Future research should focus on directly comparing digital interventions like Alena with their face-to-face therapy counterparts. Such studies would provide valuable insights into the relative strengths and limitations of digital therapy, helping to refine these tools and better integrate them into mainstream mental health care. This direct comparison would also aid in identifying specific patient profiles that may benefit more from digital or traditional therapy modalities.

In summary, the Alena app exemplifies the significant potential of digital therapy for social anxiety. It adapts a gold-standard model of CBT into a format that is not only effective but also near-infinitely scalable and accessible. This study demonstrates that high-quality therapy can be delivered through a digital medium, reaching more people who need it, at their convenience and comfort.

FAQs

What is the *Alena* app?

Alena for social anxiety is an app that provides a structured cognitive-behavioural therapy (CBT) program to reduce social anxiety. It is entirely digital, meaning there is no human in the loop. The program is split into separate modules that each target a specific driver of social anxiety.

What is a Randomised Clinical Trial (RCT)?

A randomised controlled trial (RCT) is a study where people are randomly assigned to two different groups: one where they receive a new treatment, or a "control" group that does *not* receive a new treatment. Because people are *randomly assigned*, there should not be any systematic differences between the groups on things like people's background or their personal characteristics. This means that any difference between groups in, say, symptom reduction is most likely due to people having access to the new treatment or not.

Why did you run two RCTs?

The first RCT was a pilot study, which means it was a smaller study (102 people in total) to quickly test whether the *Alena* app had a positive effect on social anxiety. After seeing significant results in this pilot study, we ran a second study (249 people in total) to replicate the effects in a larger sample of people.

Why did the first RCT only recruit women aged between 18 and 35?

Social anxiety reaches its peak between the ages of 18 and 35, and is more likely to affect women. Therefore, it was important to us that *Alena* was effective in women aged between 18 and 35, as these are the people most likely to seek help. This is why our pilot study focused on this group first, before we conducted the second RCT which included a larger, wider range of people.

Why was the control group put on a waitlist for Alena?

Sometimes, just knowing that you're going to receive help can reduce symptoms of mental health conditions. By having one group on a waitlist to receive *Alena* (the control group) and the other group with access to *Alena* (the treatment group), we could see how having access to Alena improves symptoms above and beyond just the knowledge that you will have mental health support. In addition, for ethical reasons, we wanted to provide the same benefit to both groups, meaning that both were incentivised similarly.

Why wasn't there an "active" control group?

An “active” control group receives an alternative treatment or intervention, as opposed to the primary treatment being tested. While we could have compared *Alena* with an alternative treatment, such as a meditation app or a general wellness app, to more rigorously assess the benefit of *Alena* over alternative interventions, we chose a waitlist control group for the following reasons:

1. In our own user research, we discovered that the vast majority of individuals in the UK who suffer from social anxiety do not routinely use a particular type of intervention, making it difficult for us to discern what would be a suitable and representative alternative.
2. The purpose of these clinical trials was to establish a baseline level of efficacy, and a waitlist control group serves as a clear baseline against which to measure the app's effects. This is a common first, fundamental step in the development of a new treatment.

Why did the control group get better over time?

Part of the reason could be because of what we describe above (i.e., that just knowing you will receive support can alleviate mental health symptoms). There is, however, also an effect known as “regression to the mean”. Regression to the mean is a statistical concept that explains how extreme situations tend to become less extreme over time, just by chance. Imagine a group of people who are feeling particularly socially anxious at the start of a study. They are selected for the study precisely because their symptoms are severe. However, over time, their condition is likely to improve somewhat, just by chance, as their symptoms naturally fluctuate and return closer to their average level of health. This improvement would happen even without any treatment. This is why it's important to have a control group in studies, to distinguish between real treatment effects and natural fluctuations like regression to the mean.

Does it matter how much you use the *Alena* app to see benefits?

No, in our study, we did not see a significant effect of how much someone used *Alena* on how much their social anxiety reduced over time. There are several potential explanations for this. For instance, *how* people used the app might have been more important than *how much* people used the app. We did not directly manipulate the “dosage” of the app (rather, participants could complete as much or as little of the app’s therapy program as they liked), but this could be the focus of a future study.

Is the *Alena* app safe to use?

Yes, the *Alena app for social anxiety* is not associated with any increased risk of adverse health effects. In fact, in both clinical trials, participants with access to the app tended to experience fewer adverse health effects than participants in the control group.

Statistics

We used custom Python code to preprocess and analyse the data, and used R to conduct fixed and mixed effects modelling.

Participants

Retention

To see if participant retention changed over time, and whether this depended on the group participants were in, we conducted the following linear regression:

Formula: `number of participants ~ group × week`

RCT	Effect	Estimate (β)	Standard error	t value	p value
RCT #1	*Intercept*	*0.512*	*0.931*	*54.999*	*< 0.001 ****
	Group	1.000	1.317	0.759	0.469
	Week	-1.143	0.281	-4.071	0.004 **
	Group × Week	< 0.001	0.397	< 0.001	1.000
RCT #2	*Intercept*	*124.806*	*1.990*	*62.708*	*< 0.001 ****
	Group	-12.357	2.815	-4.390	< 0.001 ***
	Week	-0.272	0.337	-0.807	0.432
	Group × Week	-0.738	0.477	-1.547	0.141

Baseline characteristics

	RCT #1			RCT #2		
Characteristic	**Intervention**	**Waitlist**	**BF_{01}**	**Intervention**	**Waitlist**	**BF_{01}**
Age (years) - *mean (SD)*	29.12 (4.07)	27.46 (4.61)	0.933	39.37 (10.53)	38.15 (10.84)	4.910
SPIN - *mean (SD)*	43.81 (9.14)	43.28 (7.59)	4.575	44.54 (8.35)	43.96 (9.34)	6.348
WSAS - *mean (SD)*			4.087			7.192
Expectations for Alena - *mean (SD)*	2.37 (0.69)	2.28 (0.73)	4.055	2.27 (0.7)	2.24 (0.72)	6.730
Ethnicity			37.312			> 100
White - *N (%)*	43 (82.69%)	43 (86.0%)		103 (83.06%)	111 (88.8%)	
Black - *N (%)*	2 (3.85%)	2 (4.0%)		6 (4.84%)	1 (0.8%)	
Asian - *N (%)*	1 (1.92%)	3 (6.0%)		6 (4.84%)	8 (6.4%)	
Other - *N (%)*	0 (0.0%)	0 (0.0%)		3 (2.42%)	2 (1.6%)	
Mixed/Multiple - *N (%)*	6 (11.54%)	2 (4.0%)		6 (4.84%)	3 (2.4%)	
Employment			> 100			> 100
Full-time - *N (%)*	34 (65.38%)	32 (64.0%)		66 (53.23%)	62 (49.6%)	
Part-time - *N (%)*	8 (15.38%)	8 (16.0%)		24 (19.35%)	27 (21.6%)	
Student - *N (%)*	5 (9.62%)	7 (14.0%)		7 (5.65%)	5 (4.0%)	
Retired - *N (%)*	0 (0.0%)	0 (0.0%)		2 (1.61%)	2 (1.6%)	
Unemployed - *N (%)*	2 (3.85%)	2 (4.0%)		13 (10.48%)	15 (12.0%)	
Unable to work - *N (%)*	1 (1.92%)	1 (2.0%)		6 (4.84%)	5 (4.0%)	
Temporarily not working - *N (%)*	2 (3.85%)	0 (0.0%)		6 (4.84%)	9 (7.2%)	
Education			14.507			> 100
No qualifications - *N (%)*	0 (0.0%)	0 (0.0%)		1 (0.81%)	1 (0.8%)	
GCSE or equivalent - *N (%)*	2 (3.85%)	2 (4.0%)		16 (12.9%)	15 (12.0%)	
A-level or equivalent - *N (%)*	10 (19.23%)	17 (34.0%)		18 (14.52%)	30 (24.0%)	
Apprenticeship, higher-education diploma or equivalent - *N (%)*	4 (7.69%)	4 (8.0%)		14 (11.29%)	8 (6.4%)	
Bachelor's degree or equivalent - *N (%)*	36 (69.23%)	27 (54.0%)		49 (39.52%)	54 (43.2%)	
Postgraduate degree or equivalent - *N (%)*	-	-		24 (19.35%)	14 (0.0%)	
PhD or equivalent - *N (%)*	-	-		2 (1.61%)	3 (2.4%)	
Alcohol use - *mean (SD)*	2.71 (1.71)	2.48 (1.74)	3.898	2.4 (1.97)	2.02 (2.16)	2.685
Any drug use - *N (%)*	5 (9.62%)	2 (4.0%)	4.524	3 (2.42%)	6 (4.8%)	10.461
Ever had therapy for social anxiety - *N (%)*	44 (84.62%)	35 (70.0%)	1.069	53 (42.74%)	50 (40.0%)	5.860
On medication - *N (%)*	12 (23.08%)	9 (18.0%)	4.180	27 (21.77%)	27 (21.6%)	7.688
Used apps for mental health before - *N (%)*	28 (53.85%)	22 (44.0%)	2.534	38 (30.65%)	42 (33.6%)	6.009

The table above describes the baseline characteristics of each group in each RCT, with the group mean and standard deviation (SD) shown for continuous variables (e.g., age) and the number of participants (N) and group percentages shown for categorical (e.g., education) or binary (e.g., any drug use) variables. We conducted Bayesian analyses in JASP (0.18.3) to assess evidence for a null hypothesis that both groups were the same (BF_{01}). If $BF_{01} \geq 3$, this indicates evidence for the null hypothesis, whereas a value < 1 indicates

evidence for the alternative hypothesis (that the groups are different). A value between 1 and 3 indicates insufficient evidence for either hypothesis. For continuous variables, we implemented Bayesian independent samples t-tests[26], and for categorical or binary variables, we implemented Bayesian contingency tables using an independent multinomial sampling method (groups fixed)[27].

Primary outcomes

Efficacy

Independent samples t-tests comparing groups, using the Benjamini-Hochberg method of False Discovery Rate (FDR) correction for multiple comparisons:

RCT	Comparison	Intervention group	Waitlist group	t value	df	p value	FDR-corrected p value
1	Δ SPIN from baseline to end of intervention	M = -9.83 SD = 12.80	M = -4.13 SD = 11.59	-2.23	90	0.028 *	0.037 *
	Δ SPIN from end of intervention to follow-up	M = 0.05 SD = 6.74	M = -2.71 SD = 6.10	1.97	83	0.052	0.052
2	Δ SPIN from baseline to end of intervention	M = -12.89 SD = 13.87	M = -7.48 SD = 12.24	-3.13	227	0.002 **	0.008 **
	Δ SPIN from end of intervention to follow-up	M = -2.39 SD = 6.15	M = -0.29 SD = 6.41	-2.48	227	0.014 *	0.028 *

Δ = change, M = mean, SD = standard deviation

Linear mixed-effects regression analysis on the change in SPIN over the intervention period (including baseline but not including follow-up), modulated by group (intervention vs waitlist) and controlling for age, sex (RCT #2 only), and the plateau effect of SPIN over time ($week^2$):

Formula: `SPIN ~ group × week + week² + age + sex + (1|participant)`

RCT	Effect	Estimate (β)	Standard error	t value	p value
RCT #1	*Intercept*	*38.16*	*1.469*	*25.976*	*< 0.001 ****
	Group	2.002	2.059	0.972	0.333
	Week	-3.149	0.398	-7.922	< 0.001 ***
	$Week^2$	0.406	0.336	1.207	0.228
	Age	1.408	1.03	1.367	0.175
	Group × Week	1.691	0.566	2.99	0.003
RCT #2	*Intercept*	*35.908*	*1.116*	*32.879*	*< 0.001 ****
	Group	36.423	1.108	1.558	0.121
	Week	2.095	1.345	-18.956	< 0.001 ***
	$Week^2$	-3.801	0.201	7.059	< 0.001 ***
	Age	1.092	0.155	0.043	0.966
	Sex	0.029	0.677	-1.948	0.053
	Group × Week	-2.73	1.402	6.894	< 0.001 ***

Variance Inflation Factor (VIF) < 2.5.

We conducted chi-square contingency tests to compare recovery rates (SPIN ≤ 19) between groups at the end of intervention and at follow-up:

RCT	Time point	Recovery	Intervention group	Waitlist group	X^2	p value
1	End of intervention (week 4)	Recovered	9	3	2.153	0.142
		Not recovered	38	42		
	Follow-up (week 6)	Recovered	8	4	0.791	0.374
		Not recovered	37	40		
2	End of intervention (week 8)	Recovered	23	13	4.769	0.029 *
		Not recovered	82	111		
	Follow-up (week 12)	Recovered	28	14	7.701	0.006 **
		Not recovered	76	107		

Safety

We conducted chi-square contingency tests to compare groups (intervention vs waitlist) on either:

- The number of **participants** *reporting* serious health effects vs the number who *didn't*
- The total number of **reports** *made* for serious health effects vs the total number of reports that *weren't* made: `(number of participants × number of weeks) - total number of reports made`

RCT	Category	Absence/Presence	Intervention group	Waitlist group	χ^2	p value
1	Number of participants	Made a report	3	8	1.812	0.178
		Didn't make a report	49	42		
	Number of reported events	Reported events	4	14	5.086	0.024 *
		Missed reports	204	186		
2	Number of participants	Made a report	5	8	0.308	0.579
		Didn't make a report	119	117		
	Number of reported events	Reported events	6	11	0.917	0.338
		Missed reports	986	989		

Secondary outcomes

Daily functioning

Independent samples t-tests comparing groups, using the Benjamini-Hochberg method of False Discovery Rate (FDR) correction for multiple comparisons:

RCT	Comparison	Intervention group	Waitlist group	t value	df	p value	FDR-corrected p value
1	Δ WSAS from baseline to end of intervention	M = -4.53 SD = 6.02	M = -2.07 SD = 5.71	-2.01	90	0.047 *	0.073
	Δ WSAS from end of intervention to follow-up	M = 1.05 SD = 5.06	M = -1.24 SD = 2.77	2.57	82	0.012 *	0.048 *
2	Δ WSAS from baseline to end of intervention	M = -4.25 SD = 8.19	M = -3.05 SD = 6.74	-1.22	227	0.225	0.225
	Δ WSAS from end of intervention to follow-up	M = -1.04 SD = 4.36	M = 0.08 SD = 4.22	-1.93	220	0.055	0.073

Δ = change, M = mean, SD = standard deviation

Linear mixed-effects regression analysis on the change in SPIN over the intervention period (including baseline but not including follow-up), modulated by group (intervention vs waitlist) and controlling for age, sex (RCT #2 only), and the plateau effect of SPIN over time ($week^2$):

Formula: `WSAS ~ group × week + week² + age + sex + (1|participant)`

RCT	Effect	Estimate (β)	Standard error	t value	p value
RCT #1	*Intercept*	*15.653*	*1.064*	*14.717*	*< 0.001 ****
	Group	2.093	1.504	1.391	0.167
	Week	-1.572	0.237	-6.629	< 0.001 ***
	$Week^2$	0.418	0.2	2.088	0.037 *
	Age	-0.014	0.752	-0.018	0.985
	Group × Week	0.896	0.337	2.657	0.008 **
RCT #2	*Intercept*	*15.827*	*0.838*	*18.882*	*< 0.001 ****
	Group	0.956	1.023	0.935	0.351
	Week	-1.327	0.122	-10.862	< 0.001 ***
	$Week^2$	0.444	0.094	4.72	< 0.001 ***
	Age	0.497	0.515	0.965	0.336
	Sex	-0.315	1.065	-0.296	0.768
	Group × Week	0.597	0.166	3.587	< 0.001 ***

Variance Inflation Factor (VIF) < 2.5.

Therapy adherence

Linear mixed-effects regression analysis on the change in SPIN *or* WSAS in the **intervention group only** over the intervention period (including baseline but not including follow-up), modulated by therapy completion rate and controlling for age, sex (RCT #2 only), and the plateau effect of SPIN over time (week^2):

Formula: `SPIN ~ completion × week + week² + age + sex + (1|participant)`

RCT	Effect	Estimate (β)	Standard error	t value	p value
RCT #1	*Intercept*	*38.769*	*1.48*	*26.187*	*< 0.001 ****
	Completion rate	0.487	1.499	0.325	0.747
	Week	-3.169	0.422	-7.513	< 0.001 ***
	Week^2	0.042	0.501	0.085	0.933
	Age	2.047	1.491	1.373	0.176
	Completion rate × Week	0.171	0.427	0.399	0.690
RCT #2	*Intercept*	*35.879*	*1.218*	*29.457*	*< 0.001 ****
	Completion rate	-0.32	0.98	-0.326	0.745
	Week	-3.795	0.21	-18.06	< 0.001 ***
	Week^2	1.174	0.237	4.956	< 0.001 ***
	Age	1.039	0.955	1.088	0.279
	Sex	-1.487	1.98	-0.751	0.454
	Completion rate × Week	-0.021	0.214	-0.1	0.921

Variance Inflation Factor (VIF) < 2.5.

Formula: `WSAS ~ completion × week + week² + age + sex + (1|participant)`

RCT	Effect	Estimate (β)	Standard error	t value	p value
RCT #1	*Intercept*	*15.683*	*1.041*	*15.064*	*< 0.001 ****
	Completion rate	-0.661	1.073	-0.616	0.541
	Week	-1.569	0.249	-6.31	< 0.001 ***
	Week^2	0.393	0.295	1.333	0.184
	Age	0.822	1.069	0.769	0.446
	Completion rate × Week	-0.168	0.252	-0.667	0.506
RCT #2	*Intercept*	*15.505*	*0.944*	*16.432*	*< 0.001 ****
	Completion rate	-1.51	0.765	-1.974	0.051
	Week	-1.281	0.125	-10.285	< 0.001 ***
	Week^2	0.503	0.14	3.583	< 0.001 ***
	Age	1.507	0.746	2.02	0.046 *
	Sex	0.622	1.546	0.402	0.688
	Completion rate × Week	-0.443	0.127	-3.501	< 0.001 ***

Variance Inflation Factor (VIF) < 2.5.

References

1. Bulley, A., Miloyan, B., Brilot, B., Gullo, M. J. & Suddendorf, T. An evolutionary perspective on the co-occurrence of social anxiety disorder and alcohol use disorder. *J. Affect. Disord.* **196**, 62–70 (2016).
2. Dryman, M. T., Gardner, S., Weeks, J. W. & Heimberg, R. G. Social anxiety disorder and quality of life: How fears of negative and positive evaluation relate to specific domains of life satisfaction. *J. Anxiety Disord.* **38**, 1–8 (2016).
3. Tonge, N. A. *et al.* Interpersonal problems in social anxiety disorder across different relational contexts. *J. Anxiety Disord.* **75**, 102275 (2020).
4. Stein, M. B. *et al.* Social anxiety disorder and the risk of depression: a prospective community study of adolescents and young adults. *Arch. Gen. Psychiatry* **58**, 251–256 (2001).
5. Lemyre, A., Gauthier-Légaré, A. & Bélanger, R. E. Shyness, social anxiety, social anxiety disorder, and substance use among normative adolescent populations: A systematic review. *Am. J. Drug Alcohol Abuse* **45**, 230–247 (2019).
6. Croot, G. *All by Myself: Dealing with Anxiety and Lonliness*. (Amazon Digital Services LLC - Kdp, 2023).
7. Jefferies, P. & Ungar, M. Social anxiety in young people: A prevalence study in seven countries. *PLoS One* **15**, e0239133 (2020).
8. Rodebaugh, T. L., Holaway, R. M. & Heimberg, R. G. The treatment of social anxiety disorder. *Clin. Psychol. Rev.* **24**, 883–908 (2004).
9. Yoshinaga, N. *et al.* Cognitive Behavioral Therapy for Patients with Social Anxiety Disorder Who Remain Symptomatic following Antidepressant Treatment: A Randomized, Assessor-Blinded, Controlled Trial. *Psychother. Psychosom.* **85**, 208–217 (2016).
10. Clark, D. M. & Wells, A. A cognitive model of social phobia. in *Social phobia: Diagnosis, assessment, and treatment (pp* (ed. Heimberg, R. G.) vol. 435 69–93 (The Guilford Press, xii, 1995).
11. Social anxiety disorder: recognition, assessment and treatment (Clinical guideline [CG159]). https://www.nice.org.uk/guidance/cg159 (2013).
12. Psychological Therapies, Annual report on the use of IAPT services, 2020-21. *NHS Digital* https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2020-21.
13. Patel, T. A., Schubert, F. T., Hom, M. A. & Cougle, J. R. Correlates of treatment seeking in individuals with social anxiety disorder: Findings from a nationally representative sample. *J. Anxiety Disord.* **91**, 102616 (2022).
14. Williams, A. D., O'Moore, K., Mason, E. & Andrews, G. The effectiveness of internet cognitive behaviour therapy (iCBT) for social anxiety disorder across two routine practice pathways. *Internet Interventions* **1**, 225–229 (2014).
15. Pauley, D., Cuijpers, P., Papola, D., Miguel, C. & Karyotaki, E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. *Psychol. Med.* **53**, 567–579 (2023).
16. Asher, M., Asnaani, A. & Aderka, I. M. Gender differences in social anxiety disorder: A review. *Clin. Psychol. Rev.* **56**, 1–12 (2017).
17. Grubbs, K. M. *et al.* The role of gender in moderating treatment outcome in collaborative care for

anxiety. *Psychiatr. Serv.* **66**, 265–271 (2015).
18. Connor, K. M. *et al.* Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *Br. J. Psychiatry* **176**, 379–386 (2000).
19. Saunders, J. B., Aasland, O. G., Babor, T. F., de La Fuente, J. R. & Grant, M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction* **88**, 791–804 (1993).
20. Mundt, J. C., Marks, I. M., Shear, M. K. & Greist, J. H. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br. J. Psychiatry* **180**, 461–464 (2002).
21. Clinical investigation of medical devices for human subjects [ISO 14155:2020]. *ISO* https://www.iso.org/standard/71690.html (2022).
22. Tulbure, B. T. *et al.* Internet-delivered cognitive-behavioral therapy for social anxiety disorder in Romania: a randomized controlled trial. *PLoS One* **10**, e0123997 (2015).
23. Nauphal, M., Swetlitz, C., Smith, L. & Rosellini, A. J. A Preliminary Examination of the Acceptability, Feasibility, and Effectiveness of a Telehealth Cognitive-Behavioral Therapy Group for Social Anxiety Disorder. *Cogn. Behav. Pract.* **28**, 730–742 (2021).
24. Powell, J. *et al.* Effectiveness and Cost-Effectiveness of a Self-Guided Internet Intervention for Social Anxiety Symptoms in a General Population Sample: Randomized Controlled Trial. *J. Med. Internet Res.* **22**, e16804 (2020).
25. Clark, D. M. *et al.* More than doubling the clinical benefit of each hour of therapist time: a randomised controlled trial of internet cognitive therapy for social anxiety disorder. *Psychol. Med.* **53**, 5022–5032 (2023).
26. Gronau, Q. F., Ly, A. & Wagenmakers, E.-J. Informed Bayesian t-Tests. *Am. Stat.* **74**, 137–143 (2020).
27. Gûnel, E. & Dickey, J. Bayes factors for independence in contingency tables. *Biometrika* **61**, 545–557 (1974).