GAMES FOR HEALTH JOURNAL: Research, Development, and Clinical Applications
Volume 5, Number 1, 2016

DOI: 10.1089/g4h.2015.0033

Original Article

Feasibility and Efficacy of an mHealth Game for Managing Anxiety: "Flowy" Randomized Controlled Pilot Trial and Design Evaluation

Quynh Pham, MSc,[1,2] Yasmin Khatib, PhD,[1] Stephen Stansfeld, MD, PhD,[1] Simon Fox,[2] and Tobias Green, LLB[2]

Abstract

Objective: Meeting the complex needs of patients with chronic common mental health disorders (CMHDs) may be the greatest challenge facing organized medical practice. On the basis of a well-established and proven theoretical foundation for controlled respiration as a behavioral intervention for CMHDs, as well as preliminary evidence that gamification can improve health outcomes through increasing patient engagement, this randomized controlled pilot study evaluated the feasibility and clinical efficacy of a mobile health game called ''Flowy'' (www.flowygame.com) that digitally delivered breathing retraining exercises for anxiety, panic, and hyperventilation symptom management.

Materials and Methods: We designed an unblinded, Web-based, parallel-group randomized controlled trial focusing on feasibility, clinical efficacy, and design proof of concept. In the intervention condition ($n = 31$), participants received free access to ''Flowy'' for 4 weeks. In the control condition ($n = 32$), participants were placed on a waitlist for 4 weeks before being offered free access to ''Flowy.'' Online measurements using psychological self-report questionnaires were made at 2 and 4 weeks post-baseline.

Results: At trial conclusion, participants found ''Flowy'' acceptable as an anxiety management intervention. ''Flowy'' engaged participants sufficiently to endorse proactive gameplay. Intent-to-treat analysis revealed a reduction in anxiety, panic, and self-report hyperventilation scores in both trial arms, with the intervention arm experiencing greater quality of life. Participants perceived ''Flowy'' as a fun and useful intervention, proactively used ''Flowy'' as part of their care, and would recommend ''Flowy'' to family and friends.

Conclusions: Our results suggest that a digital delivery of breathing retraining exercises through a mobile health game can manage anxiety, panic, and hyperventilation symptoms associated with CMHDs.

Introduction

Background

Common mental health disorders (CMHDs) account for a quarter of the global burden of disease, defined as premature death combined with years lived with disability.[1] They include depression, generalized anxiety disorder, panic disorder, obsessive–compulsive disorder, posttraumatic stress disorder, and social anxiety disorder[2] and are classified as a chronic disease by the World Health Organization alongside asthma, cancer, and diabetes.[3] Worldwide, 272 million people suffer from anxiety disorders alone.[3] The Mental Health Foundation reports one in five Britons feel anxious a lot or all the time; of these, over 6 million will experience panic attacks, and of this group, 1.7 million will suffer from panic disorder.[4]

Meeting the complex needs of patients with chronic CMHDs may be the greatest challenge facing organized medical practice. Despite recent global commitments to addressing mental disorders,[5] national and international responses have been slow, inadequate, and fragmented. This has resulted in exorbitant costs to the healthcare system; 15 percent of all United

[1]Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom.
[2]Playlab London, London, United Kingdom.
This study was registered with a clinical trials registry under trial identifier number ISRCTN78356215.

Kingdom National Health Service (NHS) expenditure on chronic disease is linked to poor mental health and well-being, totalling nearly £10 billion in England each year.[6] CMHDs disrupt existing acute healthcare structures built to be reactive to disease—responding only when the patient is sick—but not proactive and focused on keeping the patient healthy.

Characterized by a long duration and slow progression, anxiety disorders are an extreme manifestation of an otherwise functional and appropriate human response to situational stressors. For those with the most severe forms, their lives can become completely dominated and debilitated by their condition. The epidemiology of anxiety disorders notably postulates respiratory abnormalities as both a cyclic cause and effect for panic attacks.[7–12] As a result controlled respiratory interventions by which a patient exerts voluntary control over his or her own breathing are hypothesized to alleviate anxiety symptoms. Breathing retraining methodologies have been successfully validated in numerous outcome studies for effecting anxiogenic health outcomes.[12–17] They are endorsed by the NHS and have been integrated into the national Improving Access to Psychological Therapies (IAPT) program, where patients are encouraged to practice controlled breathing for 2–3 minutes, three to four times daily, to cope with panic and anxiety.[18]

Despite their effectiveness and proven ability to optimize anxiety symptom management, current methods of breathing retraining delivery are limited in both availability and access, making them difficult to integrate into anxiety self-management regimens in their current form. Breathing retraining exercises are primarily taught in services, which is problematic given that nearly 85 percent of people with mental disorders are not receiving treatment.[19] Moreover, although regulated breathing techniques are easy to learn, they are also easy to perform incorrectly. Mitigating the risk of people seeking misleading psychoeducation on breathing retraining to manage their condition, as well as respecting that intensive self-management is challenging for everyone with anxiety, is critical when planning novel delivery channels.

A need therefore exists for using integrative healthcare models and designing proactive and feasible interventions for CMHDs with patient-level integration at all stages of research and design, from conceptualization to delivery. The patient-centered model of care is grounded in the belief that people are more than their disorders.[20] Patients should be viewed as individuals with unique experiences, needs, and aspirations and seen in the context of their daily lives, as part of a family and a community. By focusing on people rather than their disease and designing holistic interventions that allow them to actively engage in and self-manage their condition, patients are placed at the center of the healthcare system where they belong.

Mobile health innovation

Advances in technology and particularly mobile digital information and communication technology have enabled a rapid proliferation of mobile health (mHealth) applications (apps), many targeted at health. The developers of these include commercial enterprises, third-sector organizations, and NHS Trusts and staff, as well as patients themselves, with the majority of apps being sold via commercial marketplaces. As of January 2013, 75 percent of the United Kingdom population own a smartphone, with one in five having downloaded at least one mHealth app[21]; it is projected that by 2018 there will be 3.4 billion smartphone users, of whom 50 percent will have downloaded at least one mHealth app.[22] With this level of ubiquitous penetration, mHealth apps are well positioned to improve access and treatment adherence to psychological and behavioral interventions by enabling services to be delivered more flexibly and tailored to individual patient needs. They offer the potential to transform mental health delivery through widening access to information and offering sustained adherence support through real-time symptom monitoring.

The NHS Confederation's Mental Health Network recognizes that mHealth apps represent real opportunities to engage and empower patients with mental health problems, but also cautions that they must have a strong evidence base. The Mental Health Network recommends that patients be actively involved throughout the development process to ensure the trust of the wider public.[23] Additionally, policy makers, practitioners, and providers of mental health services must also be active partners in the development, evaluation, and implementation of new technology to capitalize on opportunities offered by mHealth apps to not just improve efficiency, but also to transform the very nature of mental healthcare.

Employing a user-centered design process—an evidence-based iterative approach endorsed by the WHO and informed by a thorough understanding of a specific end-user group—in creating mHealth apps demonstrates a responsible and appropriate prioritization toward patient-centered care.[24] User-centered design has been used in the design and development of apps for diabetes, heart disease, and asthma to endorse sustained behavioral change and improved self-care.[25] Interventions that are developed or redesigned using this patient-centered ideology can achieve greater health outcomes and deliver measurable and significant health benefits.[26]

Games for mHealth

A subset of mHealth apps being leveraged to improve self-management behaviors is mHealth games. They employ gamification, defined as the implementation of the most common and enjoyable mechanics of videogames in non–videogame contexts.[27] Common gamification mechanics include badges, leaderboards, points, levels, challenges, and quests.[28] Recent research has endorsed mHealth games as positive mechanisms for health behavior change due to their ability to engage and motivate players toward effective disease management.[29] Numerous mHealth games have been clinically evaluated and proven to effectively manage chronic diseases such as diabetes,[26,30] cancer,[31–33] cystic fibrosis,[34–36] and asthma.[37–39] They generated positive health outcomes, with all studies achieving statistically significant results, and were readily adopted by their target patient group. Current games range from exergames to increase both knowledge and motivation of exercise behavior,[40–43] role-playing games that foster greater understanding of sensitive topics like mental health,[44–47] race,[48] and sexuality,[49–52] and virtual rehabilitation games to help patients during recovery.[53–57] It is clear that games are an appropriate and effective medium for addressing numerous health conditions that might otherwise be ill-addressed through conventional care.

By 2015, 50 percent of industries will gamify their innovation processes, although 80 percent of gamified applications are expected to fail because of a lack of research around suitable game design, as well as poor rationale or design of gamification mechanics.[28] The academic and clinical staff responsible for the majority of mHealth games currently on the market are not game designers and consequently develop products that miss the most essential mechanism of engagement in games—the fun.[58] The solution seems to then be that mHealth games should be responsibly designed and developed by a multidisciplinary team of clinicians, researchers, and game designers, as well as clinically evaluated to confirm how much more effective they are in changing behavioral and health outcomes than conventional approaches. mHealth games are capable of teaching new forms of thought and behavior in an enjoyable way. This learning potential has been left largely untapped in the mental health arena, and addressing this gap may lead to a novel validated approach in mental health intervention and self-management.

mHealth evaluation

As the field of mHealth advances and more players enter the arena of digital intervention development for CMHDs, concerns have arisen over the number of apps available for download that have not been adequately tested for efficacy. In the current app marketplace, new health apps are often accompanied by substantial hype and promise but are not supported by evidence.[59–61] Several studies have highlighted apps that could compromise patient safety and are potentially dangerous.[62–65] To mitigate against this, both clinical and risk evaluation must be done to ensure that novel health technologies fulfill basic qualifications: they are evidence-based, provide measurable value to the user, and are risk-adversive and user-centered in design and development.[66] Precedence exists for safe and effective apps for CMHDs, but they are scarce and make up less than 1 percent of commercially available apps.[67]

While it is tempting to shift the blame for this discrepancy onto the app developers themselves and cite a lack of responsibility and best practice on their behalf, an equally feasible explanation is that few evaluation methodologies for mHealth interventions exist. The traditional approach to behavioral intervention evaluation is through a randomized controlled trial (RCT), the gold standard for assessing the effect of an intervention post-development.[68] Current evaluation methodologies for behavioral interventions typically involve development, pilot testing, evaluation in at least two RCTs, and implementation studies.[69,70] However, this process can take up to 17 years from initial research to full implementation, making it fundamentally incompatible with the rapid pace of mHealth; the technology being evaluated may well be obsolete before the trial is completed.[71–73] Traditional evaluation processes are therefore incongruent with the fundamental principles of continuously evolving mHealth applications, which rely on iterative cycles of development to identify design weaknesses and inefficiencies. Time spans of traditional trials also see publications citing effects of technology that has long since been surpassed and therefore often provide limited value to helping inform clinical decisions.

In acknowledgement of the shortcomings stemming from traditional evaluation processes as well as advances in mobile devices that now incorporate sensors and can collect and store both physiological and self-report health data, new mHealth evaluation methodologies must be conceived and adopted to achieve the potential for ''real-time research'' and render conventional research designs more efficient.[74–77] The potential for mobile devices to collect research data under a safe, secure framework has largely remained untapped. Rigorous research is needed to examine both the potential and challenges of using mobile technologies to improve health outcomes, as well as to establish the reliability and validity of the mHealth evaluation methodology used to generate these critical data. The results from this pilot study will therefore lead to further research supporting innovative evaluation methodologies for mHealth interventions.

Study aims

On the basis of a well-established and proven theoretical foundation for controlled respiration as a behavioral intervention for CMHDs and preliminary evidence that gamification can improve health outcomes through increasing patient engagement, we developed an mHealth game called ''Flowy.'' ''Flowy'' is the first mHealth app to operationalize breathing retraining exercises for gameplay. We designed ''Flowy'' to bridge intervention accessibility gaps, reduce the economic burden of chronic mental illness, endorse personalized patient-centered care, and engage users to understand and manage their own condition in a fun and meaningful way.

This randomized controlled pilot study evaluated the acceptability of ''Flowy'' as an mHealth game that digitally delivers breathing retraining exercises for anxiety, panic, and hyperventilation symptom management. We engaged service users with CMHDs, their families, and their care providers in the design, development, and pilot evaluation of ''Flowy.'' We designed this study to establish clinical efficacy in reducing symptom severity alongside improving quality of life. Furthermore, we assessed both user engagement and usability as key study metrics. We hypothesized that ''Flowy'' would be an acceptable, clinically effective, engaging, and useful anxiety management intervention.

Objectives

This study had two main assessment objectives: (1) feasibility of a breathing retraining game called ''Flowy'' as an anxiety management intervention and (2) clinical efficacy of ''Flowy'' to reduce anxiety symptomatology.

The long-term objective of this feasibility and efficacy pilot study was to trial concepts and methods needed to inform a subsequent full-scale clinical evaluation, which will assess whether ''Flowy'' can objectively reduce clinical symptoms in a vulnerable clinical population. We also trialed our study design to assess whether it was an appropriate evaluation methodology for mHealth interventions, which will inform the continued development of evaluation frameworks in the field of mHealth.

Primary hypotheses

We hypothesized that we would achieve participant acceptability of ''Flowy'' as an anxiety intervention. This was assessed by the following criteria: (1) the number of people

who responded to the recruitment call to trial an app that managed their anxiety and completed the study eligibility screener; (2) the response rate for participants in the trial throughout the 4-week assessment period; (3) the number of participants in the intervention condition proactively using ''Flowy'' at least once for over 1 minute of play during the 4-week assessment, as assessed by log data; and (4) the number of participants in the intervention condition who engaged in sustained usage of ''Flowy,'' as assessed by log data.

Secondary hypotheses

Our secondary hypotheses were that participants in the intervention group would experience reductions in anxiety, panic, and hyperventilation, and increase in quality of life at Week 4 assessment compared with baseline ($P < 0.05$). This was defined as a decrease in score on the Generalized Anxiety Disorder Scale (GAD-7), Panic Disorder Severity Scale-Self Report (PDSS-SR), Nijmegan, and Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (QLES-Q-SF) psychological self-report questionnaires. We hypothesized that in comparison with the control condition, participants in the intervention group would experience greater reductions in anxiety, hyperventilation, and panic attack symptoms, as well as increased quality of life at the Week 4 assessment. This was defined as a greater decrease in score on the GAD-7, PDSS-SR, Nijmegan, and QLES-Q-SF self-report psychological questionnaires. Finally, we hypothesized that participants in the intervention group would perceive ''Flowy'' as a useful intervention, defined by a mean rating of ≥ 3 (''neither agree nor disagree'') on a 1–5-point Likert scale usability questionnaire following study completion. Further detailed information on these outcome measures can be found in Materials and Methods.

Materials and Methods

Overview

The study methodology was reviewed according to the Consolidated Standards of Reporting Trials (CONSORT) standards for clinical trials in eHealth[78] and reported according to the CONSORT-eHealth (version 1.6.1) checklist.[79] The study methodology was further reviewed and advised by a research design advisor at the National Institute of Health Research Design Services.

Trial design

We designed an unblinded, Web-based, parallel-group RCT focusing on feasibility, clinical efficacy, and design proof of concept. This was an exploratory study with an iterative design to assess and improve study methodology, game design, and usability. Initial contact with potential participants was made online; trial candidates ($n = 510$) self-generated a login username and password and completed an eligibility screener that required participants to meet inclusion criteria as well as show moderate signs of anxiety, defined as a score of ≥ 16 on the Anxiety Sensitivity Index-3 (ASI-3), ≥ 8 on the Overall Anxiety Severity and Impairment Scale (OASIS), and ≥ 6 on the GAD-7. Eligible participants ($n = 63$) subsequently received a follow-up e-mail to digitally sign the study information sheet and consent form. They then completed the online baseline assessment and were randomized to one of two study conditions following assessment. In the intervention condition ($n = 31$), participants received free access to ''Flowy'' for 4 weeks. In the control condition ($n = 32$), participants were placed on a waitlist for 4 weeks before being offered free access to ''Flowy.'' Online measurements using self-report questionnaires were made at 2 and 4 weeks post-baseline. Participants in both conditions had unrestricted access to professional help. Figure 1 shows the study CONSORT flowchart.

Clinical measures and data collection

Table 1 provides an overview of all clinical measures used alongside their data collection schedule. Table 2 describes all measures with their number of items, minimum–maximum scores, clinical cutoff scores, and psychometric reliability/construct validity (Cronbach's alpha). Participants were asked to provide information on their age, gender, education, employment, ethnicity, and religion/spirituality at baseline assessment (Table 3). Study measures were obtained using six clinical self-report questionnaires. Anxiety symptoms were assessed using three questionnaires: the GAD-7 item scale,[80] the OASIS,[81] and the ASI-3 item scale.[82] Panic symptoms were assessed using the PDSS-SR.[83] Hyperventilation symptoms were assessed using the Nijmegen Questionnaire.[84] Quality of life, enjoyment, and satisfaction were assessed using the QLES-Q-SF.[85] Participant views on using information technology for health were assessed using the eHealth Literacy Scale (eHEALS).[86] Participant perception of ''Flowy'' was measured with a 21-question usability questionnaire alongside open-ended qualitative feedback and user stories around their experience using ''Flowy'' (Table 4).

Outcomes

The primary study outcome was the acceptability of ''Flowy'' as an anxiety management intervention. This was assessed by the following:

1. The number of people who responded to the recruitment call to trial an app that managed their anxiety and completed the study eligibility screener
2. Trial adherence throughout the 4-week assessment period as measured by completion of weekly assessments
3. The number of participants in the intervention condition proactively using ''Flowy,'' defined internally as at least once for over 1 minute of play during the 4-week assessment
4. The number of participants in the intervention condition who engaged in sustained usage of ''Flowy,'' defined internally as at least 5 minutes of uninterrupted play and based on NHS IAPT guidelines for suggested durations of controlled breathing exercises.[18]

Secondary outcomes included the following:

1. A reduction in anxiety, panic, and hyperventilation symptoms and improvement in quality of life reported by participants in the intervention group compared with the control group from baseline to Week 4
2. A reduction in anxiety, panic, and hyperventilation symptoms and improvement in quality of life reported by participants in the intervention group from baseline to Week 4

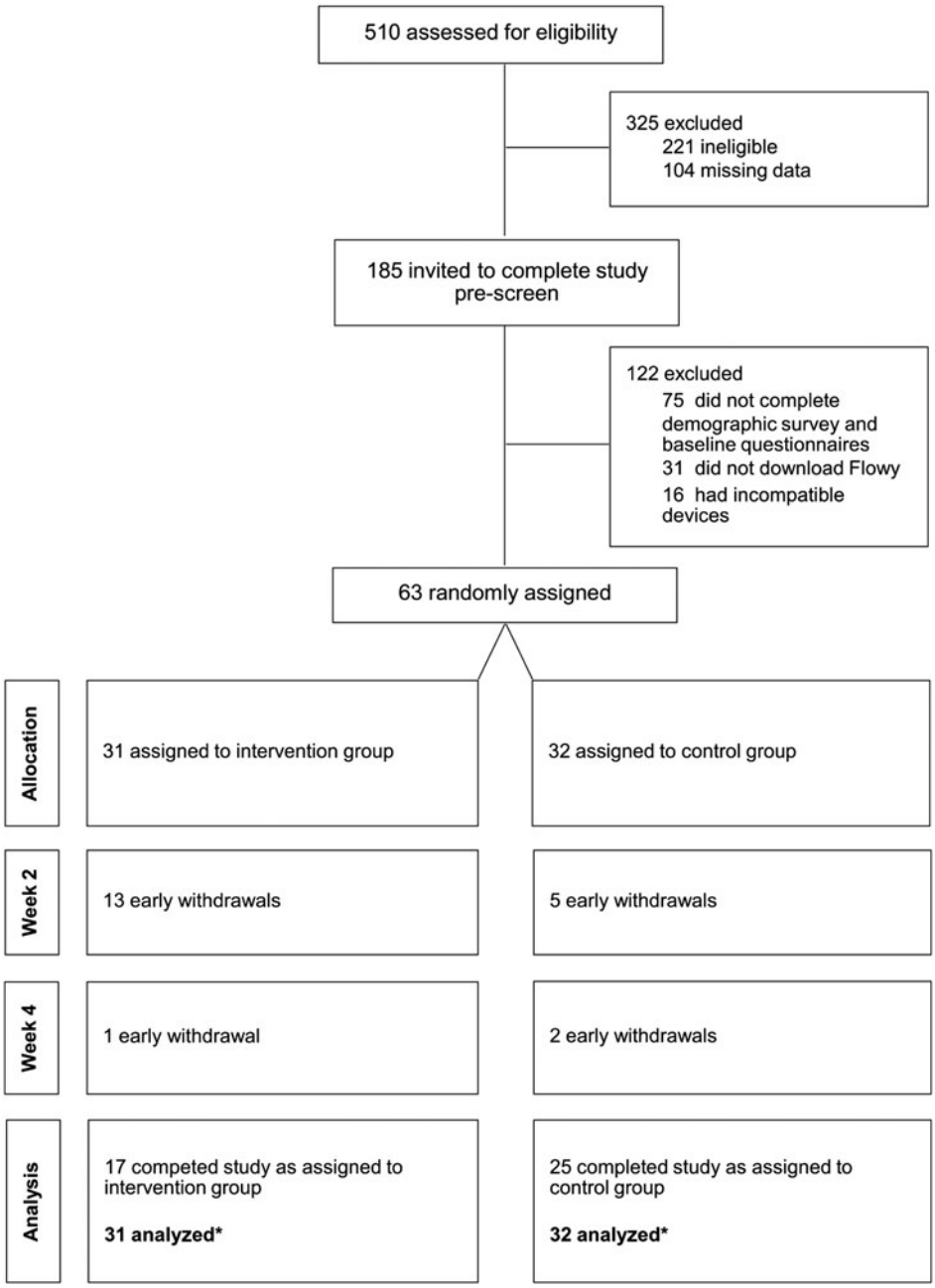

FIG. 1. CONSORT study flow diagram. *Last observation carried forward for full data analysis.

TABLE 1. OVERVIEW OF CLINICAL INSTRUMENTS USED AND DATA COLLECTION SCHEDULE

Variable	*Questionnaire*	*Eligibility screener*	*Baseline*	*Week 2*	*Week 4*
Electronic health literacy	eHEALS	X	—	—	—
Anxiety	ASI-3	X	X	X	X
	GAD-7	X	X	X	X
	OASIS	X	X	X	X
Hyperventilation	Nijmegan	—	X	X	X
Panic attacks	PDSS-SR	—	X	X	X
Quality of life	Q-LES-Q-SF	—	X	X	X

ASI-3, Anxiety Sensitivity Index-3; eHEALS, eHealth Literacy Scale; GAD-7, Generalized Anxiety Disorder Scale; OASIS, Overall Anxiety Severity and Impairment Scale; PDSS-SR, Panic Disorder Severity Scale-Self Report; Q-LES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form.

Table 2. Description of Psychometric Measures Evaluating Primary and Secondary Outcomes

Measure	*Type*	*Description*	*Alpha*
eHealth Literacy Scale (eHEALS)	Self-report	A measure of electronic health resource literacy, defined as a consumer's combined knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems. On this eight-item scale, responses are rated on a scale from 1 to 5, with higher scores indicating greater eHealth literacy.	0.93
Generalized Anxiety Disorder Scale (GAD-7)	Self-report	A brief seven-item anxiety scale to assess symptom severity and identify probable cases of generalized anxiety disorder in clinical practice and research. Responses are rated on a scale from 0 to 3, with a total score of ≥ 6 representing moderate anxiety, ≥ 11 moderately severe anxiety, and ≥ 15 severe anxiety.	0.91
Overall Anxiety Severity and Impairment Scale (OASIS)	Self-report	A measure of the severity and impairment associated with anxiety, as assessed by a brief, continuous five-item scale that can be used across anxiety disorders, with multiple anxiety disorders, and with subthreshold anxiety symptoms. Responses are rated on a scale from 0 to 4, and individual items are added for a total severity score. A cutoff score of ≥ 8 is used to determine an anxiety diagnosis.	0.89
Anxiety Sensitivity Index-3 (ASI-3)	Self-report	An 18-item self-report multidimensional measure of the three domains of anxiety sensitivity (physical, social, and cognitive). Respondents indicate their agreement with each item, from ''very little'' (0) to ''very much'' (4), with total scores ranging from 0 to 72. A cutoff score of ≥ 16 was used to establish moderate anxiety.	0.93
Nijmegan Questionnaire	Self-report	A 16-item measure of hyperventilation and dysfunctional breathlessness, suitable as a screening tool for early detection and also as an aid in diagnosis and therapy planning. Responses are rated on a scale from 0 to 4, with a total score of ≥ 23 out of 64 suggesting a positive diagnosis of hyperventilation.	0.92
Panic Disorder Severity Scale(PDSS-SR)	Self-report	A questionnaire developed for assessing the severity and detecting possible symptoms of panic disorder, measured using seven items, each rated on a 5-point scale ranging from 0 to 4. The overall assessment is made by a total score ranging from 0 to 28, and a score of ≥ 9 suggests the need for a formal diagnostic assessment.	0.91
Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF)	Self-report	A self-report questionnaire with 14 items evaluating overall life enjoyment and satisfaction with physical health, mood, work, social and family relationships, daily functioning, and overall well-being. Responses are scored on a 5-item scale with possible scores ranging from 14 to 70, where high scores indicate better enjoyment and satisfaction with life.	0.86

3. Participant perception of ''Flowy'' as a useful intervention
4. Assessing outcome measure scores in pilot study participants to establish parameters of measurable change in anxiety, hyperventilation, and panic attack symptoms for a subsequent full-powered, full-scale clinical evaluation
5. Study design evaluation to inform sample size calculation and methodology of a subsequent full-powered, full-scale clinical evaluation.

Participant recruitment

Participants were recruited online through extensive media coverage, including a BBC profile, social media press, newsletters, blog posts, Google AdSense, and a Call for Participants landing page. Local businesses were contacted for permission to display recruitment posters on their premises. University campuses were targeted for recruitment. The primary study advertisement included a sample of questions from the eligibility screener to optimize targeting appropriate participants who experienced anxiety. The advertisement emphasized the need to visit the eligibility screener Web site to determine eligibility. It also included a link to the main ''Flowy'' Web site, which contained complete information on the intervention, clinical explanation of breathing retraining exercises, and illustrations of the game mechanism (Supplementary Fig. S1; Supplementary Data are available online at www.liebertonline.com/g4h).

Participants interested in determining their eligibility accessed the eligibility screener Web site (Supplementary Fig. S2) and self-generated a login username with their e-mail

Table 3. Baseline Participant Demographic Characteristics

Characteristic	*Participants [n (%)]*			P
	Intervention (n = 31)	*Control (n = 32)*	*Total population (n = 63)*	
Age (years)				0.69
18–24	7 (22.6)	10 (31.3)	17 (27.0)	
25–34	17 (54.8)	15 (46.9)	32 (50.8)	
35–44	4 (12.9)	6 (18.8)	10 (15.9)	
45–54	1 (3.2)	1 (3.1)	2 (3.2)	
55–64	1 (3.2)	0 (0)	1 (1.6)	
65 and older	1 (3.2)	0 (0)	1 (1.6)	
Gender				0.27
Male	18 (58.1)	13 (40.6)	31 (49.2)	
Female	13 (41.9)	18 (56.3)	31 (49.2)	
Other	0 (0)	1 (3.1)	1 (1.6)	
Country of residence				0.27
Argentina	3 (9.7)	0 (0)	3 (4.8)	
Brazil	2 (6.5)	0 (0)	2 (3.2)	
Canada	3 (9.7)	5 (15.6)	8 (12.7)	
Cyprus	0 (0)	1 (3.1)	1 (1.6)	
Germany	0 (0)	1 (3.1)	1 (1.6)	
Israel	1 (3.2)	0 (0)	1 (1.6)	
Italy	0 (0)	1 (3.1)	1 (1.6)	
Malaysia	0 (0)	1 (3.1)	1 (1.6)	
United Kingdom	12 (38.7)	10 (31.3)	22 (34.9)	
United States	10 (32.3)	13 (40.6)	23 (36.5)	
Education				0.82
High school diploma	7 (22.6)	10 (31.3)	17 (27.0)	
Bachelor degree	14 (45.2)	11 (34.4)	25 (39.7)	
Master degree	6 (19.4)	6 (18.8)	12 (19.0)	
Professional degree	1 (3.2)	3 (9.4)	4 (6.3)	
Doctorate degree	1 (3.2)	1 (3.1)	2 (3.2)	
Not applicable	2 (6.5)	1 (3.1)	3 (4.8)	
Employment status				0.41
Employed	25 (80.6)	23 (71.9)	48 (76.2)	
Unemployed	6 (19.4)	9 (28.1)	15 (23.8)	
Ethnicity				
Asian	3 (10.0)	3 (10.0)	6 (9.5)	0.73
White	22 (73.3)	23 (76.7)	45 (71.4)	
European	3 (10.0)	1 (3.3)	4 (6.3)	
Hispanic	2 (6.7)	2 (6.7)	4 (6.3)	
Mixed	0 (0)	1 (3.3)	1 (1.6)	
Missing	1 (3.3)	2 (6.7)	3 (4.8)	
Religion/spirituality				0.37
Yes	11 (35.5)	8 (25.0)	19 (30.2)	
No	20 (64.5)	24 (75.0)	44 (69.8)	
Mobile platform				0.91
iOS	14 (45.2)	17 (54.8)	28 (44.4)	
Android	14 (43.8)	18 (56.3)	35 (55.6)	
Mobile gameplay history				0.35
Yes	21 (67.7)	25 (78.1)	46 (73.0)	
No	10 (32.3)	7 (21.9)	17 (27.0)	

address and a unique password. Following completion of the screener, their data were exported to a password-protected Excel® (Microsoft, Redmond, WA) database. The e-mail provided was checked for multiple registrations, and participants with missing data were excluded. Participants were notified of their study eligibility by e-mail and directed to the online study information sheet and consent form. Those not eligible for the study were thanked for their time and given the opportunity to sign up and receive a code to download ''Flowy'' in January 2015. Participants were not compensated for taking part in the study but were allowed to download ''Flowy'' and keep the game following study completion. Minors were not asked to participate in research.

Inclusion/exclusion criteria

Study candidates from the general population were eligible for the study if they met the following criteria: (1) males

Table 4. Usability Questions to Assess Participation Perception of ''Flowy''

Number	*Question*	*Score*
1	''Flowy'' is useful.	4.17 (0.38)
2	''Flowy'' makes me more productive.	3.22 (0.55)
3	''Flowy'' gives me more control over the activities in my life.	3.56 (0.62)
4	''Flowy'' makes the things I want to accomplish easier to get done.	3.06 (0.87)
5	''Flowy'' saves me time when I use it.	2.94 (0.64)
6	''Flowy'' meets my needs.	3.39 (0.98)
7	''Flowy'' does everything I would expect it to do.	3.22 (0.94)
8	''Flowy'' is easy to use.	4.17 (1.15)
9	''Flowy'' is simple to use.	4.39 (0.85)
10	Both occasional and regular users would like ''Flowy.''	3.94 (1.11)
11	I can use ''Flowy'' successfully every time.	3.89 (0.96)
12	I learned to use ''Flowy'' quickly.	4.44 (1.04)
13	I learned to use ''Flowy'' easily.	4.61 (0.61)
14	I easily remember how to use ''Flowy.''	4.67 (0.49)
15	I quickly became skillful with ''Flowy.''	3.83 (1.25)
16	I am satisfied with ''Flowy.''	3.94 (0.80)
17	I would recommend ''Flowy'' to a friend.	4.17 (0.99)
18	''Flowy'' is fun to use.	3.89 (0.68)
19	''Flowy'' works the way I want it to work.	3.28 (0.90)
20	I feel I need to have ''Flowy.''	3.06 (0.87)
21	''Flowy'' is wonderful.	3.94 (0.80)

Data are mean (standard deviation) values.

and females at least 18 years of age; (2) screen GAD-7 score of ≥ 6; (3) screen OASIS score of ≥ 8; (4) screen ASI-3 score of ≥ 16; (5) comorbid disorders including major depression, dysthymia, panic disorder, agoraphobia, social anxiety disorder, generalized anxiety disorder, simple phobia, obsessive–compulsive disorder, and somatization disorder were accepted as long as moderate symptoms of anxiety were observed (these broad inclusion criteria maximized generalizability of findings); (6) concurrent use of antidepressants, anxiolytics, hypnotics, and herbal products with psychoactive substances was acceptable provided there was no change in medication type and dose after randomization; (7) provision of written informed consent; and (8) able to comply with the study protocol (e.g., able to download ''Flowy'' onto a mobile device, able to complete assessments).

Study candidates were ineligible if they met any of the following criteria: (1) lack of mobile device or incompatible operating system (Android™ [Google, Mountaine View, CA] and iOS [Apple, Cupertino, CA] compatibility required); (2) lack of access to Internet; (3) any coexisting medical conditions that could have altered the clinical presentation of hyperventilation such as chronic severe asthma or chronic obstructive pulmonary disease; (4) lifetime history of bipolar disorder or psychosis; (5) presence of any Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) Axis I disorder, excluding those listed in the inclusion criteria, that were likely to have interfered with the patient's ability to participate in the study, as judged by the investigator; (6) presence of any DSM-IV Axis II disorder that were likely to have interfered with the patient's ability to participate in the study, as judged by the investigator; or (7) participants who were acutely suicidal to the degree that precautions against suicide were needed, or a history of suicide attempt in the past 5 years.

Sample size

Given that the primary study outcome was the acceptability of ''Flowy'' as an anxiety management intervention, to be assessed by recruitment and response rates, power calculations were not calculated *a priori*. The results from this pilot study will be used to inform the required sample size for a full clinical evaluation. There was no limit on the number of eligible participants during the 4-week recruitment period, and recruitment strategies were exhausted to include as many participants as possible.

Randomization and allocation

Randomization and allocation were done at the individual level using a computerized randomization list to protect against bias. Candidates who met inclusion criteria were randomly allocated to the intervention group (''Flowy'') or to the control (waitlist) group after signing the information sheet and consent form. Participants were equally allocated 1:1 to intervention and control conditions for optimal analysis and maximized statistical power. Both investigator and participants were unblinded to allocation.

Conception

''Flowy'' (www.flowygame.com) was initially conceived as a digital delivery of breathing retraining. ''Flowy'' was designed by start-up venture Playlab London with a team of game developers and animators. Playlab's founder had lived experience of panic and anxiety disorders. The process of recovery helped him to develop insight into chronic mental health conditions and inspired him to investigate how condition symptoms might be ameliorated.

In January 2013 Playlab began conceptual work, informally testing paper and code prototypes of a simple breath-controlled game experience. Early research and development included periodic formative meetings to review each stage, correct problems, and make decisions for the following stages in an iterative process. Qualitative data were collected from various stakeholder groups and thematically analyzed: mental health practitioners and clinicians provided data on feasibility, people supporting those with lived experience of panic and anxiety disorders provided data on desirability and value, and people with lived experience of panic and anxiety disorders provided data on current behaviors concerning panic, anxiety, panic attacks, and technology.

User-centered design

A series of test media was constructed and shown to people with lived experience of panic and anxiety disorders to assess desirable esthetics, themes, and mechanics and their

FIG. 2. Screenshots of ''Flowy'' minigames. (Color graphics available at www.liebertonline.com/g4h)

effect on the user with particular regard to feelings of anxiety, trustworthiness, and ''medical'' framing. Following from this initial research, prototypes were developed, and several rounds of usability tests were performed with a varied group of participants to evaluate functionality, interface, and robustness. Throughout the project, expert advisors contributed to thinking on game design, usability, user experience design, clinical accuracy, and efficacy. ''Flowy'' was designed to endorse empowerment and motivation in managing anxiety by potentially making this process fun. The design team therefore aimed for ''Flowy'' to be visually appealing, easy to navigate, and intuitive to use and to ultimately provide a meaningful user experience. They followed the user-centered design process of development and engaged in usability testing throughout the project lifecycle. This user-centered design process is endorsed by the World Health Organization, which advises user evaluation at every step of mHealth consumer application development to ensure sustained user engagement for effective outcomes.[24] From this user-centered design process, the design team obtained objective feedback that was crucial to producing sustained behavioral changes, as well as key to end-user adoption.

Intervention

''Flowy'' is a mHealth game that engages users in a series of minigames where they use breathing retraining exercises and perform diaphragmatic breathing to alleviate anxiety. Diaphragmatic breathing is the deepest of all breaths, with the most air inhaled. During diaphragmatic breathing, the lowest levels of the lungs are inflated; the lower third of the lungs contains the greatest amount of blood when standing vertically, resulting in more oxygenated blood per breathing cycle. The minigames in ''Flowy'' range in theme, from sailing a boat down a river to flying balloons into the sky (Fig. 2 gives a screenshot of various minigames). Users touch the screen with their finger as they inhale and remove their finger from the screen as they exhale to control the gaming mechanics. A breathing indicator visually represents a full breath; users see a circle expanding as they inhale and contracting as they exhale. This indicator provides a visual guideline of a breathing retraining exercise and also simplifies the cognitive assessment of what constitutes a full diaphragmatic breath. The goal of each minigame is to correctly follow the breathing indicator while advancing in the game narrative; users progress through levels and achieve goals by breathing correctly and staying calm. ''Flowy'' also includes an in-game interactive tutorial on how to properly perform diaphragmatic breathing as guided by both an NHS IAPT protocol[18] and evidence-based research protocols from literature.[12,87] An overview of the game characteristics can be found in Table 5. Usage guidelines were not enforced to allow for assessment of naturalistic usage patterns determined by the patient.

Control

Participants in the control group were placed on a waitlist for 4 weeks. To prevent high attrition rates, participants received a weekly newsletter with curated content on breathing retraining exercises, ''Flowy'' game development, mindfulness meditation, and similar content relevant to ''Flowy.'' Additionally, they received reminder e-mails through online marketing tool MailChimp to complete study assessments. After 4 weeks, participants received a free code to download ''Flowy.''

Table 5. Characteristics of a Videogame for Health (''Flowy'')

Characteristic	*Details*
Health topic(s)	Anxiety disorders, panic attacks
Target age group(s)	All ages
Other targeted group characteristics	Well-being, mindfulness, meditation
Short description of game idea	''Flowy'' is a collection of minigames that digitally deliver breathing retraining exercises.
Target player(s) (check one):	■ Individual ☐ Dyad ☐ Small group ☐ MMOG ☐ Other: _____
Guiding knowledge or behavior change theory(ies), models, or conceptual framework(s)	Capability-Motivation-Opportunity Behavior Model Mechanics-Dynamics-Esthetics Framework Self-Determination Theory
Intended health behavior changes	Reduced hyperventilation Increased activation Increased self-management Increased knowledge and education
Knowledge element(s) to be learned	Breathing retraining
Behavior change procedure(s) (taken from Michie inventory) or therapeutic procedure(s) used	(From the BCT Taxonomy version 1) 2.2 Feedback on behavior 2.7 Feedback on outcome(s) of behavior 4.1 Instruction on how to perform the behavior 6.1 Demonstration of the behavior 8.1 Behavioral practice and rehearsal 8.2 Behavior substitution 8.3 Habit formation 8.4 Habit reversal
Clinical or parental support needed? (Please specify)	No
Data shared with parent or clinician	☐ Yes ■ No
Type of game (check all that apply):	☐ Active ☐ Action ☐ Adventure ☐ Role-playing ☐ Simulation ☐ Strategy ☐ Sports ■ Casual ■ Educational ☐ Other: _____
Story (if any)	
Synopsis (including story arc)	(1) Blow bubbles from a straw (2) Sail a boat down a river (3) Fly balloons into the sky
How the story relates to targeted behavior change	Minigame narratives revolve around air and flow to encourage calm breathing.
Game components	
Player's game goal/objective(s)	Player progresses through levels and achieves goals by staying calm and breathing correctly.
Rules	Follow the breathing indicator: touch the screen to inhale, release to exhale.
Game mechanic(s)	Reward feedback Goals Narrative
Procedures to generalize or transfer what's learned in the game to outside the game	NA
Virtual environment	
Setting (describe)	See Figure 2 for screenshots of the minigames.
Avatar	
Characteristics	NA
Abilities	NA
Game platform(s) needed to play the game: (check all that apply):	■ Smartphone ■ Tablet ☐ Kinect ☐ Xbox ☐ Wii ☐ PlayStation ☐ Computer ☐ Handheld device ☐ Other: _____
Sensors used	No
Estimated play time	Approximately 10 minutes per play session

BCT, Behavior Change Taxonomy; NA, not applicable.

Statistical analyses

Statistical analyses were conducted using IBM (Armonk, NY) SPSS Statistics version 22. All participants who were randomized were included in data analysis in accordance with the intention-to-treat-principle. Chi-squared tests were used for discrete variables to compare control and intervention groups to determine whether they differed at baseline on demographic characteristics. Effect sizes were calculated based on means and odds ratios (Cohen's d[88]).

Primary analysis on engagement, defined as proactive and sustained usage, used frequency and duration measures from gameplay logs to assess outcomes. Secondary analysis on clinical efficacy included a repeated-measures 2×2 analysis of variance (ANOVA) for four clinical outcome variables: (1) anxiety, (2) panic, (3) self-report hyperventilation, and (4) quality of life, with time (baseline versus Week 4) as a within-subjects factor and condition (intervention versus control) as a between-subjects measure. Results were confirmed with a gain score ANOVA with condition as the between-subjects factor alongside an analysis of covariance using pretreatment scores as the covariate. The effect of time on treatment was evaluated using repeated-measures ANOVA at three time points (baseline, Week 2, and Week 4). P levels were corrected for non-sphericity using the Greenhouse–Geisser epsilon when necessary. We tested for linear trends because we hypothesized a decrease in GAD, PDSS, and Nijmegan scores and an increase in QLES-Q-SF scores across treatment weeks. Significance levels for linear trends across all measures were Bonferroni-corrected. Statistical significance was considered at $P < 0.05$ unless otherwise specified. All test results reported are two-tailed.

Ethical approval

Institutional board approval for the study protocol was obtained from the Queen Mary University of London Research Ethics Committee with reference number QMREC2014/20. The study protocol, interventions, participant information, and informed consent procedure were approved and found to be in accordance with all applicable regulations. All participants received and digitally signed a study information sheet and consent form. All participants were given the opportunity to download and keep "Flowy" on their mobile devices following study completion.

Results

Demographics

Demographic characteristics of the control and intervention groups did not differ significantly at baseline (Table 3). Participants were predominantly young adults (18–34 years old; 51 percent) from developed, English-speaking countries (Canada, the United Kingdom, and the United States; 84 percent), highly educated (68 percent university-educated or above), employed (76 percent), and white (71 percent). As part of the eligibility pre-screening, candidates completed the eHEALS questionnaire, which assessed their electronic health literacy. The final study population had high electronic health literacy (mean = 3.58, standard deviation [SD] = 0.22).

Intervention acceptability

In total, 510 candidates were screened for study eligibility from April 2014 to June 2014 (Fig. 1). One hundred eighty-five participants met preliminary eligibility criteria. Of these, 106 did not meet pre-study requirements (baseline assessment, registering a mobile device, providing consent). Sixteen participants were subsequently excluded from the study because of hardware compatibility issues, leaving a final study population of 63. At study completion, 67 percent of participants remained engaged and completed terminal study assessments; 75 percent of control and 58 percent of intervention group participants completed all study assessments.

In total, 63 participants were randomly allocated to either the control ($n = 32$) or intervention ($n = 31$) group of the study (Fig. 1). Both groups experienced attrition during the 4-week study period (45 percent of intervention participants and 25 percent of control participants), with the majority of attrition occurring after Week 2 study assessments (42 percent of intervention participants and 16 percent of control participants). There was a significant association between mobile platform and attrition, with more participants leaving the study if they had downloaded "Flowy" on an iOS platform (64.3 percent) compared with an Android platform (23.5 percent) [$\chi^2(1) = 5.28$, $P = 0.022$]. Participants did not differ on any other baseline characteristics.

Log data were recovered from all intervention group participants who played "Flowy" (26/31, 84 percent). Eighty-four percent of intervention group participants proactively played "Flowy" at least once for a minimum of 1 minute during the study. The number of participants who played "Flowy" per week peaked at baseline (20/31, 65 percent) and was lowest at Week 4 (8/31, 26 percent) (Fig. 3, top). Log data measured a mean weekly sessional play duration of 10 minutes during the 4-week study period, increasing weekly and peaking at 15 minutes in Week 4 (Fig. 3, middle) and a mean weekly total play duration of 20 minutes (Fig. 3, bottom).

Clinical efficacy

Mean scores and SDs for primary and secondary outcome measures for the intent-to-treat sample appear in Table 6. There were no group differences for any of the baseline measures. A group-effects multivariate ANOVA showed no significant difference for psychopathology measures (anxiety, panic, and hyperventilation) based on treatment condition ($F_{3,\ 59} = 2.059$, $P = 0.115$).

At the study conclusion, participants in both groups reported improved scores on anxiety, panic, hyperventilation, and quality of life psychological self-report questionnaires. Univariate gain score and 2×2 treatment by condition ANOVAs showed an increase in quality of life (Δmean = 4.570; $F_{1,\ 61} = 4.845$, $P = 0.034$) and a reduction in anxiety (Δmean = 0.911; $F_{1,\ 61} = 1.049$, $P = 0.310$), panic (Δmean = 0.108; $F_{1,\ 61} = 0.012$, $P = 0.914$), and hyperventilation (Δmean = 0.523; $F_{1,\ 61} = 0.055$, P = 0.816) scores (Fig. 4) in the intervention group playing "Flowy" compared with the waitlist control group from baseline to Week 4. In particular, the intervention group hyperventilation scores decreased meaningfully. A score of ≥ 23 out of 64 on the Nijmegan questionnaire suggests a positive clinical diagnosis of hyperventilation; intervention group participants went from a mean score above the threshold score of 23 to below it and

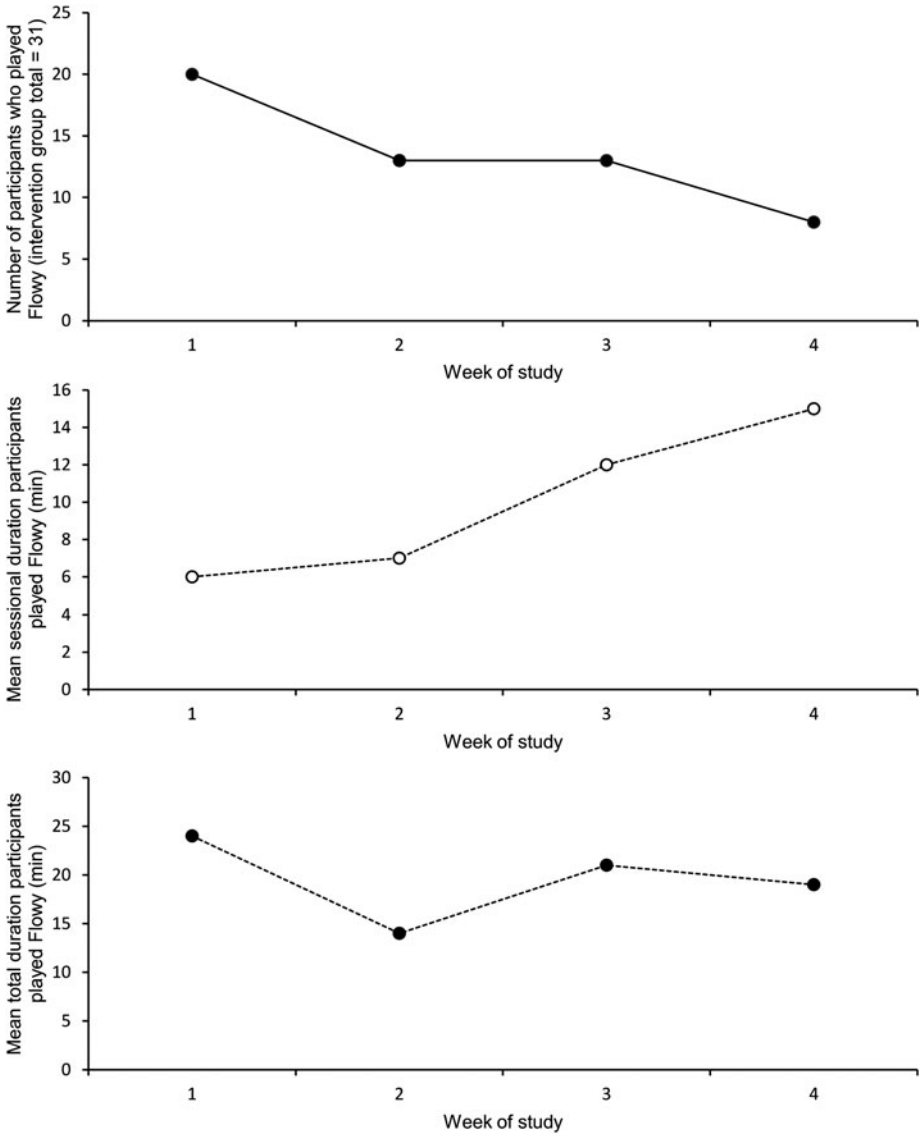

FIG. 3. From baseline to Week 4, **(top panel)** number of intervention group participants who played ''Flowy'' weekly, **(middle panel)** mean weekly play session duration, and **(bottom panel)** mean weekly total play duration.

TABLE 6. GROUP SCORES OF INTERVENTION AND CONTROL GROUP PSYCHOMETRIC QUESTIONNAIRES PRE- AND POST-EVALUATION AND SIGNIFICANCE OF GAIN SCORE ANALYSIS OF VARIANCE, 2×2 ANALYSIS OF VARIANCE, AND ANALYSIS OF COVARIANCE GROUP EFFECTS ANALYSES

	Intervention		*Control*						
Outcome	*Pre*	*Post*	*Pre*	*Post*	*Analysis*	df	F	η^2	P
Anxiety (GAD-7)					GS	1	1.049	0.017	0.310
Mean (SD)	11.55 (5.05)	9.39 (5.21)	10.66 (4.63)	9.53 (4.79)	2×2	1	1.049	0.017	0.310
n	31	32	31	32	ANCOVA	1	0.638	0.011	0.427
Panic (PDSS-SR)					GS	1	0.012	0.000	0.914
Mean (SD)	16.90 (5.24)	15.35 (5.76)	15.78 (5.37)	14.13 (6.57)	2×2	1	0.012	0.000	0.914
n	31	32	31	32	ANCOVA	1	0.045	0.001	0.832
Hyperventilation (Nijmegan)					GS	1	0.055	0.001	0.816
Mean (SD)	24.45 (9.79)	21.74 (11.27)	26.91 (11.01)	24.72 (12.75)	2×2	1	0.055	0.001	0.816
n	31	32	31	32	ANCOVA	1	0.204	0.003	0.653
Quality of Life (Q-LES-Q-SF)					GS	1	4.845	0.108	0.034
Mean (SD)	32.11 (7.32)	35.89 (7.71)	32.71 (6.98)	31.92 (9.17)	2×2	1	4.845	0.108	0.034
n	31	32	31	32	ANCOVA	1	4.698	0.108	0.036

ANCOVA, analysis of covariance; GAD-7, Generalized Anxiety Disorder Scale; GS, gain score; PDSS-SR, Panic Disorder Severity Scale-Self Report; Q-LES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form; SD, standard deviation.

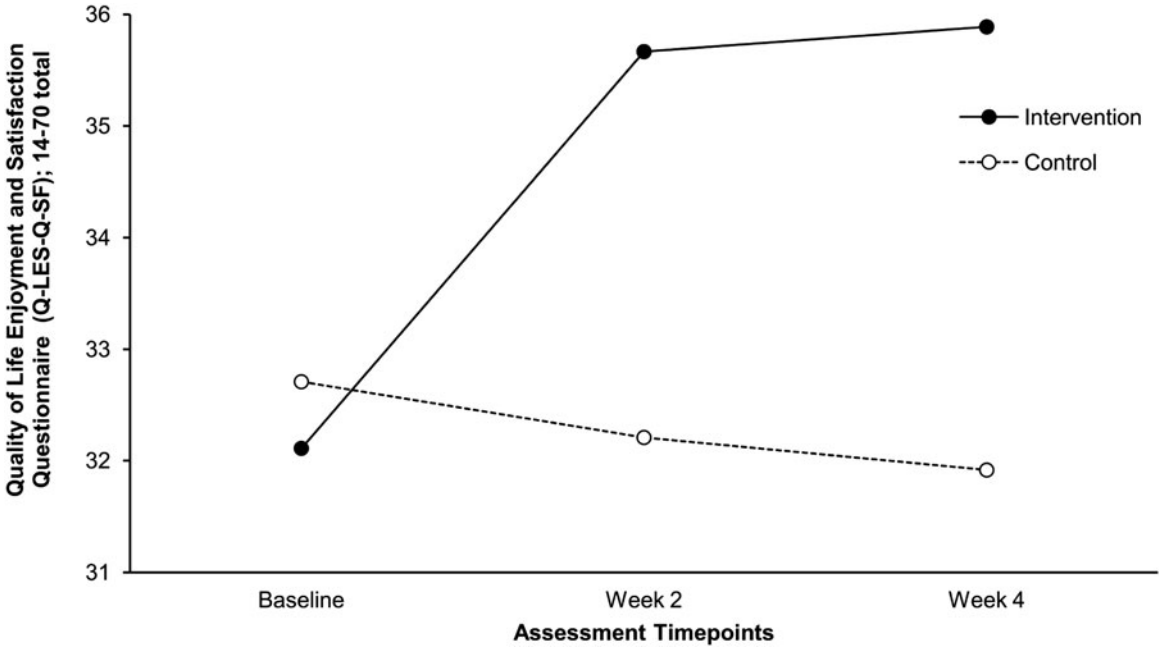

FIG. 4. Quality of Life, Enjoyment and Satisfaction Questionnaire Short Form scores for the intervention (solid circle, solid line) and control (open circle, dashed line) groups at study assessment time points (baseline, Week 2, and Week 4).

outside the threshold zone (baseline mean, 24.45; to Week 4 mean, 21.74). A repeated-measures ANOVA evaluating time as a main effect on treatment showed a decrease in panic and hyperventilation ($P = 0.011$ and 0.016, respectively) and an increase in quality of life ($P = 0.005$) scores in intervention participants from baseline to Week 4 (Table 7). Effect size differences between intervention and control groups post-treatment were small ($d < 0.3$) for anxiety, panic, and hyperventilation measures and moderate ($d = 0.393$) for quality of life measures.

Usefulness and usability

One hundred percent of participants in the intervention group found ''Flowy'' to be a useful intervention (Fig. 5 and Table 4). They also all remembered how to use ''Flowy'' after initial play, confirming that ''Flowy'' can be learned and used regularly with ease. Eighty-nine percent would recommend ''Flowy'' to family and friends. Participants rated the usability of ''Flowy'' highly, with 92 percent saying ''Flowy'' was easy and simple to use. Eighty-three percent of participants found ''Flowy'' fun to use. They did not strongly identify ''Flowy'' as making them more productive, making the things they want to accomplish easier to get done, or saving them time when they use it. In addressing the concern that ''Flowy'' could be used as a safety behavior and avoid the cognitive root of their panic attacks, 72 percent of participants did not identify with feeling like they needed to have ''Flowy.'' Qualitative feedback from participants endorsed the concept of a novel breath-controlled game for managing anxiety and panic attacks:

Table 7. Intervention Group Psychometric Questionnaire Scores at All Evaluations and Significance of Repeated-Measures Analysis of Variance Time Effect Analysis

	Intervention						
Outcome	*Baseline*	*Week 2*	*Week 4*	df	F	η^2	P
Anxiety (GAD-7)							
Mean (SD)	11.55 (5.05)	9.81 (4.60)	9.39 (5.21)	2	0.531	0.018	0.591
n	31	32	31				
Panic (PDSS-SR)							
Mean (SD)	16.90 (5.24)	15.42 (5.42)	15.35 (5.78)	2	4.858	0.139	0.011
n	31	32	31				
Hyperventilation (Nijmegan)							
Mean (SD)	24.45 (9.79)	22.45 (10.61)	21.74 (11.27)	2	4.405	0.128	0.016
n	31	32	31				
Quality of life (Q-LES-Q-SF)							
Mean (SD)	32.11 (7.32)	35.67 (5.08)	35.89 (7.71)	2	6.266	0.269	0.005
n	31	19	18				

GAD-7, Generalized Anxiety Disorder Scale; PDSS-SR, Panic Disorder Severity Scale-Self Report; Q-LES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form; SD, standard deviation.

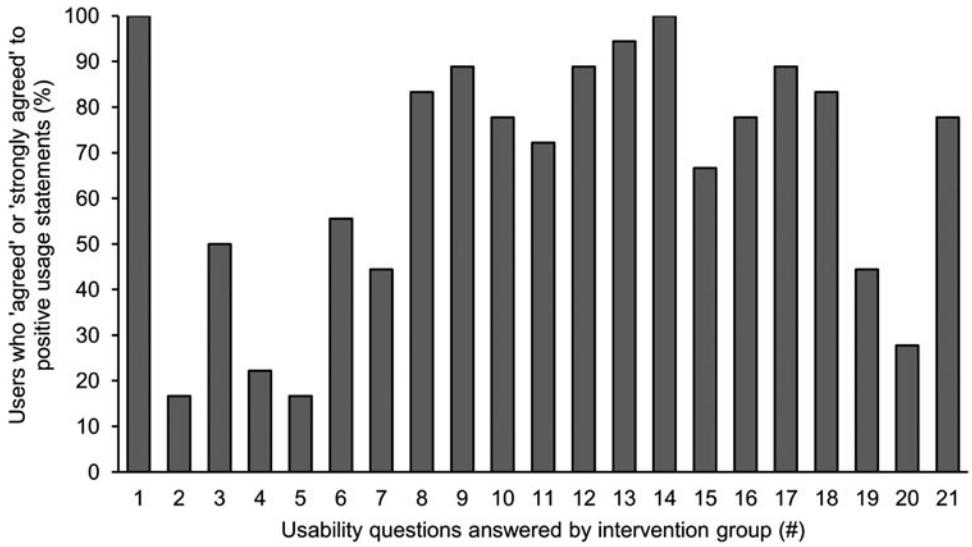

FIG. 5. Frequency of intervention group participants who ''agreed'' (4 of 5) or ''strongly agreed'' (5 of 5) with positive usage statements on a usability questionnaire.

> I just wanted to stress how important I think the application (and moreover the approach) could be in helping people like myself deal with anxiety. For me—it's changed my life completely—and I know this is a story you must hear over and over again.
> I think it's such an amazing idea and there is nothing else like this around. I feel like panic attacks are swept under the rug too much and it feels amazing that someone is finally addressing it. Once the app is ready, I will be promoting it to everyone I know who has anxiety!!!
> The game has helped me on more than one occasion to prevent a big panic attack.
> Since first using ''Flowy'' I have been feeling a lot better. I do still use the app, and I am a lot happier having this game with me, it sometimes feels like a secret weapon!

Discussion

With an initial aim of creating an mHealth game for anxiety management that could digitally deliver health outcomes in a fun and engaging way, ''Flowy'' was consequently conceived and piloted to measure usefulness, desirability, and clinical impact. Preliminary results confirm a demand exists for such an intervention, with a large population perceiving the concept of a breathing regulation game as both viable and desirable. Participants accepted ''Flowy'' as an anxiety management intervention, engaged sufficiently with ''Flowy'' to proactively reach for it when needed, and showed interest in practicing breathing retraining in a gaming context.

The results from this feasibility and efficacy randomized controlled pilot trial suggest that an mHealth application operationalizing breathing retraining exercises for behaviorally targeted gaming can be an engaging and acceptable intervention for anxiety disorders. However, we were unable to demonstrate clinical efficacy—our trial results indicate that ''Flowy'' did not produce significant decreases in anxiety, panic, and self-report hyperventilation scores. We hypothesize various explanations for this, the first being that realistically, a short, lightweight, self-directed game-based app is unlikely to change established clinical psychopathology within a 4-week time frame. Second, in aiming for an early evaluation of ''Flowy'' to inform future design and development decisions, we may have trialed an immature version of ''Flowy'' that was not robust enough as a game to increase breathing retraining adherence and consequently have clinical impact. Third, our decision to mitigate against high attrition rates in the control group by sending out weekly psychoeducational newsletters may have turned our waitlist control into an active control; this may explain the lack of an interaction effect for our three clinical outcome measures. Fourth, our final sample size of 63 was not sufficiently powered to detect small effect sizes and prevented an accurate elucidation of how ''Flowy'' affected clinical symptomology. Although we obtained all study measures remotely through both app log data and an in-house online assessment platform, thereby relieving participants of having to attend traditional face-to-face measurement visits, we did not take advantage of this methodology to increase scalability and recruit a large sample size capable of detecting small effect sizes for our clinical measures.

Finally, our decision to use an RCT design as an evaluation framework for ''Flowy'' was potentially inappropriate to establish true clinical impact. We encountered many of the methodological barriers highlighted by researchers[89] that hindered our ability to accurately measure the clinical impact of ''Flowy'' on the studied sample. Our intervention group experienced high attrition, which we suspected was due to hardware compatibility issues and software errors affecting gameplay; this would have made it difficult for participants to initially engage with ''Flowy'' and may have discouraged them from playing further. However, we were unable to verify this hypothesis because of a methodological requirement to freeze development during the trial in order to maintain internal validity. We identified issues with suboptimal frequency and duration of play, which may have been alleviated through adjustments to the intensity of gameplay, breathing indicator, or gamification techniques, but there was no scope in our study design for a more critical analysis of

the process to determine where problems existed and where improvements could be made to make the game efficacious. This methodological intolerance for responding in real time to the usage behaviors exhibited in-trial by participants is problematic given the continuously evolving nature of both digital interventions and their users, and also regrettable given the ease and flexibility of change inherent to digital interventions that allows them to be rapidly iterated. Our future research will build on the methodological limitations identified from this study and explore alternative framework for evaluating mHealth interventions.

Evaluating quality of life and usability was integral to the study design due to a strong focus on having ''Flowy'' be applicable as a management application and not just a clinical intervention. Broad intervention applicability and availability were at the forefront of game design prioritization, which meant quality of life improvement would ascertain this design decision. We observed a statistically significant improvement in quality of life scores for participants randomly assigned to the intervention group, alongside positive usability feedback; this is noteworthy given the 4-week duration of the trial and shows that ''Flowy'' can effect positive change in a short period of time. ''Flowy'' may have promoted increased awareness of panic attack triggers, which potentially contributed to greater self-monitoring and reframing of the panic attack experience and consequently improved overall quality of life.

The self-determination theory of Deci and Ryan[90] predicts increased player engagement with games fosters a sense of autonomy, competence, and social relatedness.[91] The positive improvements in quality of life from our study are consistent with this: when given a game that facilitated the process of managing their condition, our participants felt motivated to take better care of themselves and consequently felt happier and healthier. Our findings align with previous studies that highlight self-efficacy as being particularly sensitive to game-based intervention[92–95]; the observed high levels of non-enforced usage demonstrate promising validation for this method of intervention delivery.

Gamified breathing retraining may have special value to individuals with other chronic conditions involving impaired perceptions of their own breathing. Similar effects and value could apply for treating asthma, chronic obstructive pulmonary disease, cystic fibrosis, and other chronic lung diseases. Additional research is required to evaluate the scope of behavioral processes that are amenable to change through this intervention platform.

Limitations to this pilot study must be acknowledged. First, using self-report screening and not formal diagnostic instruments to confirm existing anxiety symptomology means that without a formal diagnosis, no claims about prevention of anxiety or mental disorders can be made. Non-uniform severity of symptoms at trial entry may have reduced statistical power but likely improved the generalizability of study findings. Several of the outcome questionnaires used in this study have not yet been validated for online use, which may change their psychometric properties. The open-recruitment strategy used in this study was chosen to optimize recruitment, but the cost of such a strategy is that the attracted population may not be representative of the actual clinical population. Analysis was done on an ''intention-to-treat'' basis, which is more common in RCTs. However, for future studies analysis should be undertaken on a ''per protocol'' basis; study attrition and loss to follow-up were not correlated with lack of improvement but rather with practical and logistic issues associated with hardware and software failure. Suboptimal adherence to the intervention resulted in high attrition rates in the intervention group; to mitigate against this, future research will not enforce usage but will provide participants with the recommended frequency and usage parameters. Finally, generalizability of the findings may be limited because of the requirement of a smartphone to join the study. This may have economically biased the study findings and may not reflect patient access to similar technology on a global level.

In summary, the feasibility and efficacy demonstrated in this pilot trial provided preliminary empirical support for an mHealth game called ''Flowy'' in managing symptomology for a chronically ill population suffering from anxiety, panic, and hyperventilation. Although more research is required to demonstrate the clinical efficacy of ''Flowy'' in promoting significant symptom amelioration, this study represents an important first step toward highlighting the potential use of ''Flowy'' in hospitals and homes for anxiety management. It is hoped that ''Flowy'' may one day constitute one component of a broader integrative approach to health care that holistically combines physiological and behavioral interventions to aid patients in the prevention, detection, treatment, and recovery from CMHDs.

Acknowledgments

We would like to thank Erica Simon for reviewing this manuscript. We wish to acknowledge the contribution of the team that developed ''Flowy'': Adam Hodgson, Anders Furevik, Khi-Jon Chua, Jean-Baptise Krohn, and Meg Jayanth. We extend our gratitude to the 63 people who participated in the trial and took a leap of faith in allowing ''Flowy'' into their lives.

Author Disclosure Statement

This research was a collaboration between industry partner Playlab London and academic partner Wolfson Institute for Preventive Medicine. Q.P. is the current Lead Researcher for Playlab London and at the time of the trial an MSc candidate at the Wolfson Institute. She conceived, designed, and implemented this study alongside support from her academic department and Playlab London. S.F. and T.G. co-founded Playlab London and conceived, designed, and developed ''Flowy.''

References

1. Mathers C, Fat DM, Boerma JT; World Health Organization. *The Global Burden of Disease: 2004 Update*. Geneva: World Health Organization; 2008.
2. Pilling S, Whittington C, Taylor C, Kendrick T. Identification and care pathways for common mental health disorders: Summary of NICE guidance. BMJ 2011; 342:d2868.
3. World Health Organization. *Integrating the Response to Mental Health Disorders and Other Chronic Diseases in Health Care Systems*. Geneva: World Health Organization; 2014.
4. Skapinakis P, Lewis G, Davies S, et al. Panic disorder and subthreshold panic in the UK general population: Epidemiology,

comorbidity and functional limitation. Eur Psychiatry 2011; 26:354–362.
5. Institute for Health Metrics and Evaluation. *Financing Global Health 2010: Development Assistance and Country Spending in Economic Uncertainty*. Seattle, WA: Institute for Health Metrics and Evaluation; 2010.
6. Department of Health. *No Health Without Mental Health: A Cross-Government Mental Health Outcomes Strategy for People of All Ages. Supporting Document—The Economic Case for Improving Efficiency and Quality in Mental Health.* London: Department of Health; 2011.
7. Beamish PM, Granello PF, Granello DH, et al. Outcome studies in the treatment of panic disorder: A review. J Counsel Dev 1996; 74:460–467.
8. Clark DM. A cognitive approach to panic. Behav Res Ther 1986; 24:461–470.
9. Ley R. The efficacy of breathing retraining and the centrality of hyperventilation in panic disorder: A reinterpretation of experimental findings. Behav Res Ther 1991; 29:301–304.
10. Ley R. Pulmonary function and dyspnea/suffocation theory of panic. J Behav Ther Exp Psychiatry 1998; 29:1–11.
11. Lum LC. Hyperventilation and anxiety state. J R Soc Med 1981; 74:1–4.
12. Meuret AE, Wilhelm FH, Roth WT. Respiratory feedback for treating panic disorder. J Clin Psychol 2004; 60:197–207.
13. Mueser KT, Lu W, Rosenberg SD, Wolfe R. The trauma of psychosis: Posttraumatic stress disorder and recent onset psychosis. Schizophr Res 2010; 116:217–227.
14. Rapee RM, Ancis JR, Barlow DH. Emotional reactions to physiological sensations: Panic disorder patients and nonclinical Ss. Behav Res Ther 1988; 26:265–269.
15. Rapee RM, Brown TA, Antony MM, Barlow DH. Response to hyperventilation and inhalation of 5.5% carbon dioxide-enriched air across the DSM-III-R anxiety disorders. J Abnorm Psychol 1992; 101:538–552.
16. Kim S, Wollburg E, Roth WT. Opposing breathing therapies for panic disorder: A randomized controlled trial of lowering vs raising end-tidal P_{CO2}. J Clin Psychiatry 2012; 73:931–939.
17. Sanchez-Meca J, Rosa-Alcazar AI, Marin-Martinez F, Gomez-Conesa A. Psychological treatment of panic disorder with or without agoraphobia: A meta-analysis. Clin Psychol Rev 2010; 30:37–50.
18. Young C, Hunte A, Newell J, Valian P. Coping with panic. In: Cambridgeshire and Peterborough NHS Foundation Trust, ed. *Improving Access to Psychological Therapies 2011, Vol. 4*. Cambridge, United Kingdom: Cambridgeshire and Peterborough NHS Foundation Trust; 2011: 27–28.
19. Saxena S, Setoya Y. World Health Organization's comprehensive mental health action plan 2013–2020. Psychiatry Clin Neurosci 2014; 68:585–586.
20. Saha S, Beach MC, Cooper LA. Patient centeredness, cultural competence and healthcare quality. J Natl Med Assoc 2008; 100:1275–1285.
21. Fox S, Duggan M. Mobile Health 2012. Pew Internet & American Life Project. 2012. www.pewinternet.org/2012/11/08/mobile-health-2012/ (accessed July 29, 2015).
22. U.S. Food and Drug Administration. Mobile Medical Applications. 2014. www.fda.gov/downloads/MedicalDevices/…/UCM263366.pdf (accessed July 29, 2015).
23. Cotton R, Irwin J, Wilkins A, Young C. *The Future's Digital: Mental Health and Technology*. London: NHS Confederation Mental Health Network; 2014.
24. World Health Organization. *mHealth: New Horizons for Health Through Mobile Technologies*. Geneva: World Health Organization; 2011.
25. McCurdie T, Taneva S, Casselman M, et al. mHealth consumer apps: The case for user-centered design. Biomed Instrum Technol 2012;(Suppl):49–56.
26. Cafazzo JA, Casselman M, Hamming N, et al. Design of an mHealth app for the self-management of adolescent type 1 diabetes: A pilot study. J Med Internet Res 2012; 14:e70.
27. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining ''gamification.'' In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. Tampere, Finland: ACM Press; 2011.
28. Zichermann G, Cunningham C. *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. Sebastopol, CA: O'Reilly Media, Inc.; 2011.
29. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010; 15:1–39.
30. Brown SJ, Lieberman DA, Germeny BA, et al. Educational video game for juvenile diabetes: Results of a controlled trial. Med Inform (Lond) 1997; 22:77–89.
31. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. Pediatrics 2008; 122:e305–e317.
32. Kato PM, Beale IL. Factors affecting acceptability to young cancer patients of a psychoeducational video game about cancer. J Pediatr Oncol Nurs 2006; 23:269–275.
33. Stinson JN, Jibb LA, Nguyen C, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013; 15:e51.
34. Bingham PM, Bates JH, Thompson-Figueroa J, Lahiri T. A breath biofeedback computer game for children with cystic fibrosis. Clin Pediatr (Phila) 2010; 49:337–342.
35. Bingham PM, Lahiri T, Ashikaga T. Pilot trial of spirometer games for airway clearance practice in cystic fibrosis. Respir Care 2012; 57:1278–1284.
36. Davis MA, Quittner AL, Stack CM, Yang MC. Controlled evaluation of the STARBRIGHT CD-ROM program for children and adolescents with cystic fibrosis. J Pediatr Psychol 2004; 29:259–267.
37. Asthma Advisory Committee Meeting. Asthma Self-Management Application Project. In: University of Western Ontario Centre for Global e-Health Innovation, Queen's University. The Lung Association. 2014. www.on.lung.ca/document.doc?id=2109 (accessed July 29, 2015).
38. Krishna S, Francisco BD, Balas EA, et al. Internet-enabled interactive multimedia asthma education program: A randomized trial. Pediatrics 2003; 111:503–510.
39. *Developing Breathe: A Self-Management Web App for Asthma*. Ottawa: Centre for Global eHealth Innovation, University Health Network; 2013.
40. Garde A, Umedaly A, Abulnaga SM, et al. Assessment of a mobile game (''MobileKids Monster Manor'') to promote physical activity among children. Games Health J 2015; 4:149–158.
41. Lin J-H. ''Just Dance'': The effects of exergame feedback and controller use on physical activity and psychological outcomes. Games Health J 2015; 4:183–189.

42. Moller CA, Majewski S, Standish M, et al. Active fantasy sports: Rationale and feasibility of leveraging online fantasy sports to promote physical activity. JMIR Serious Games 2014; 2:e13.
43. Sato K, Kuroki K, Saiki S, Nagatomi R. Improving walking, muscle strength, and balance in the elderly with an exergame using Kinect: A randomized controlled trial. Games Health J 2015; 4:161–167.
44. Fish MT, Russoniello CV, O'Brien K. The efficacy of prescribed casual videogame play in reducing symptoms of anxiety: A randomized controlled study. Games Health J 2014; 3:291–295.
45. Leutwyler H, Hubbard EM, Dowling GA. Adherence to a videogame-based physical activity program for older adults with schizophrenia. Games Health J 2014; 3:227–233.
46. Roepke AM, Jaffee SR, Riffle OM, et al. Randomized controlled trial of SuperBetter, a smartphone-based/Internet-based self-help tool to reduce depressive symptoms. Games Health J 2015; 4:235–246.
47. Shepherd M, Fleming T, Lucassen M, et al. The design and relevance of a computerized gamified depression therapy program for indigenous Māori adolescents. JMIR Serious Games 2015; 3:e1.
48. Gutierrez B, Kaatz A, Chu S, et al. ''Fair Play'': A videogame designed to address implicit race bias through active perspective taking. Games Health J 2014; 3:371–378.
49. Chu SKW, Kwan ACM, Reynolds R, et al. Promoting sex education among teenagers through an interactive game: Reasons for success and implications. Games Health J 2015; 4:168–174.
50. D'Cruz J, Santa Maria D, Dube S, et al. Promoting parent–child sexual health dialogue with an intergenerational game: Parent and youth perspectives. Games Health J 2015; 4:113–122.
51. Driesmans K, Vandenbosch L, Eggermont S. Playing a videogame with a sexualized female character increases adolescents' rape myth acceptance and tolerance toward sexual harassment. Games Health J 2014; 4:91–94.
52. Jozkowski KN, Ekbia HR. ''Campus Craft'': A game for sexual assault prevention in universities. Games Health J 2014; 4:95–106.
53. Gorini A, Marzorati C, Casiraghi M, et al. A neurofeedback-based intervention to reduce post-operative pain in lung cancer patients: Study protocol for a randomized controlled trial. JMIR Res Protoc 2015; 4:e52.
54. Kottink AIR, Prange GB, Krabben T, et al. Gaming and conventional exercises for improvement of arm function after stroke: A randomized controlled pilot study. Games Health J 2014; 3:184–191.
55. Lozano-Quilis J-A, Gil-Gómez H, Gil-Gómez J-A, et al. Virtual rehabilitation for multiple sclerosis using a Kinect-based system: Randomized controlled trial. JMIR Serious Games 2014; 2:e12.
56. Proffitt R, Sevick M, Chang C-Y, Lange B. User-centered design of a controller-free game for hand rehabilitation. Games Health J 2015; 4:259–264.
57. Szturm T, Reimer KM, Hochman J. Home-based computer gaming in vestibular rehabilitation of gaze and balance impairment. Games Health J 2015; 4:211–220.
58. Granic I, Lobel A, Engels RCME. *The Benefits of Playing Video Games*. Nijmegen, The Netherlands: Radboud University Nijmegen; 2013.
59. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: Lack of evidence undermines quality and safety. Evid Based Med 2013; 18:90–92.
60. Haffey F, Brady RR, Maxwell S. A comparison of the reliability of smartphone apps for opioid conversion. Drug Saf 2013; 36:111–117.
61. Misra S, Lewis TL, Aungst TD. Medical application use and the need for further research and assessment for clinical practice: Creation and integration of standards for best practice to alleviate poor application design. JAMA Dermatol 2013; 149:661–662.
62. Wolf JA, Moreau JF, Akilov O, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol 2013; 149:422–426.
63. McCartney M. How do we know whether medical apps work? BMJ 2013; 346:f1811.
64. Rosser BA, Eccleston C. Smartphone applications for pain management. J Telemed Telecare 2011; 17:308–312.
65. Ferrero NA, Morrell DS, Burkhart CN. Skin scan: A demonstration of the need for FDA regulation of medical apps on iPhone. J Am Acad Dermatol 2013; 68:515–516.
66. Lewis LT, Wyatt CJ. mHealth and mobile medical apps: A framework to assess risk and promote safer use. J Med Internet Res 2014; 16:e210.
67. Donker T, Petrie K, Proudfoot J, et al. Smartphones for smarter delivery of mental health programs: A systematic review. J Med Internet Res 2013; 15:e247.
68. Piantadosi S. *Clinical Trials: A Methodologic Perspective*. New York: Wiley; 2005.
69. Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998; 66:7–18.
70. Rounsaville BJ, Carroll KM, Onken LS. A stage model of behavioral therapies research: getting started and moving on from stage I. Clin Psychol Sci Pract 2001; 8:133–142.
71. Balas EA, Boren SA. Managing clinical knowledge for healthcare improvements. In: van Bemmel JH, McCray AT, eds. *Yearbook of Medical Informatics 2000.* Stuttgart: Schattauer; 2000: 65–70.
72. Brown CH, Kellam SG, Kaupert S, et al. Partnerships for the design, conduct, and analysis of effectiveness, and implementation research: Experiences of the Prevention Science and Methodology Group. Adm Policy Ment Health 2012; 39:301–316.
73. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am J Public Health 2003; 93:1261–1267.
74. Collins LM, Murphy SA, Strecher V. The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART): New methods for more potent eHealth interventions. Am J Prev Med 2007; 32(5 Suppl):S112–S118.
75. Mohr CD, Schueller MS, Riley TW, et al. Trials of intervention principles: Evaluation methods for evolving behavioral intervention technologies. J Med Internet Res 2015; 17:e166.
76. Mohr DC, Cheung K, Schueller SM, et al. Continuous evaluation of evolving behavioral intervention technologies. Am J Prev Med 2013; 45:517–523.
77. Wyatt J. Assessing and improving evidence based health informatics research. Stud Health Technol Inform 2010; 151:435–445.
78. Eysenbach G; CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and standardizing evaluation

reports of Web-based and mobile health interventions. J Med Internet Res 2011; 13:e126.
79. Eysenbach G. CONSORT-EHEALTH: Implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013; 192:657–661.
80. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 2006; 166:1092–1097.
81. Norman SB, Cissell SH, Means-Christensen AJ, Stein MB. Development and validation of an Overall Anxiety Severity and Impairment Scale (OASIS). Depress Anxiety 2006; 23:245–249.
82. Taylor S, Zvolensky MJ, Cox BJ, et al. Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. Psychol Assess 2007; 19:176–188.
83. Shear MK, Rucci P, Williams J, et al. Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. J Psychiatr Res 2001; 35:293–296.
84. van Dixhoorn J, Duivenvoorden HJ. Efficacy of Nijmegen Questionnaire in recognition of the hyperventilation syndrome. J Psychosom Res 1985; 29:199–206.
85. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: A new measure. Psychopharmacol Bull 1993; 29:321–326.
86. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006; 8:e27.
87. Everly GS, Lating JM. *A Clinical Guide to the Treatment of the Human Stress Response*, 3rd ed. New York: Springer; 2013.
88. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
89. Kumar S, Nilsen WJ, Abernethy A, et al. Mobile health technology evaluation: The mHealth Evidence Workshop. Am J Prev Med 2013; 45:228–236.
90. Deci EL, Ryan RM. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum; 1985.
91. Przybylski AK, Deci EL, Rigby CS, Ryan RM. Competence-impeding electronic games and players' aggressive feelings, thoughts, and behaviors. J Pers Soc Psychol 2014; 106:441–457.
92. Chao YY, Scherer YK, Wu YW, et al. The feasibility of an intervention combining self-efficacy theory and Wii Fit exergames in assisted living residents: A pilot study. Geriatr Nurs 2013; 34:377–382.
93. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. Pediatrics 2008; 122:e305–317.
94. Lyons EJ, Hatkevich C. Prevalence of behavior changing strategies in fitness video games: Theory-based content analysis. J Med Internet Res 2013; 15:e81.
95. Schneider KL, Ferrara J, Lance B, et al. Acceptability of an online health videogame to improve diet and physical activity in elementary school students: ''Fitter Critters.'' Games Health J 2012; 1:262–268.

Address correspondence to:
Quynh Pham, MSc
Playlab London
71B St John Street
London, United Kingdom, EC1M 4NJ

E-mail: quynh@playlablondon.co.uk