

# Randomized Controlled Trial of SuperBetter, a Smartphone-Based/Internet-Based Self-Help Tool to Reduce Depressive Symptoms

Ann Marie Roepke, MA,<sup>1</sup> Sara R. Jaffee, PhD,<sup>1</sup> Olivia M. Riffle, BS,<sup>1</sup> Jane McGonigal, PhD,<sup>2</sup>  
Rose Broome, MA,<sup>2</sup> and Bez Maxwell, BA<sup>2</sup>

## Abstract

**Objective:** Technological advances have sparked the development of computer- and smartphone-based self-help programs for depressed people, but these programs' efficacy is uncertain. This randomized controlled trial evaluated an intervention called SuperBetter (SB), which is accessed via smartphone and/or the SB Web site.

**Materials and Methods:** Online, we recruited 283 adult iPhone® (Apple, Cupertino, CA) users with significant depression symptoms according to the Center for Epidemiological Studies Depression questionnaire (CES-D). They were randomly assigned to one of three conditions: (a) a version of SB using cognitive-behavioral therapy and positive psychotherapy strategies to target depression (CBT-PPT SB); (b) a general SB version focused on self-esteem and acceptance (General SB); or (c) a waiting list control group (WL). The two SB groups were instructed to use SB for 10 minutes daily for 1 month. All participants completed psychological distress and well-being measures online every 2 weeks through follow-up. An intent-to-treat analysis was conducted using hierarchical linear modeling.

**Results:** As hypothesized, SB participants achieved greater reductions in CES-D scores than WL participants by posttest (Cohen's  $d=0.67$ ) and by follow-up ( $d=1.05$ ). Contrary to prediction, CBT-PPT SB did not perform better than General SB; both versions of SB were more effective than the WL control. Differences between SB versions favored General SB but were not statistically significant.

**Conclusions:** These large effect sizes should be interpreted cautiously in light of high attrition rates and the motivated, self-selected sample. Nonetheless, smartphone-based/Internet-based self-help may play an important role in treating depression.

## Introduction

TECHNOLOGICAL ADVANCES HAVE sparked the development of computer- and smartphone-based tools aimed at promoting mental health.<sup>1,2</sup> These tools can augment conventional depression treatment by making therapy homework more convenient and engaging, by serving as a minimally invasive intervention for people with mild symptoms, and by offering treatment where it has been unavailable.<sup>2-4</sup> Of the estimated 350 million people who suffer from depression, fewer than half are treated, and about 30% of those treated do not fully recover.<sup>5,6</sup> High-tech tools present exciting opportunities to address these problems, but do they work?

## Efficacy of computer- and smartphone-based interventions

Online cognitive-behavioral therapy (CBT) programs can alleviate depression, whether they are self-directed, therapist-guided, or video chat-based.<sup>7-12</sup> These programs' effect sizes are small to moderate for self-reported depression, with Cohen's  $d$  ranging from 0.20 to 0.37.<sup>8,9</sup> Their efficacy is perhaps unsurprising, as these programs closely follow traditional therapy, psychoeducation, and/or bibliotherapy models.

Mobile platforms have enabled novel, creative uses of traditional treatment strategies. Mobile applications (apps) have multiple advantages: They are convenient, engaging,

<sup>1</sup>Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania.

<sup>2</sup>SuperBetter Labs, LLC, San Francisco, California.

The views expressed in this publication are those of the authors and do not necessarily reflect the opinions of our funders and supporters.

user-friendly, personalized, and self-paced. Framed as games, apps become potentially powerful tools to promote well-being. "Serious games" are games designed to achieve goals beyond entertainment, such as improved health, cognition, and education.<sup>13</sup> Indeed, playing games is associated with improved mood and decreased physical stress, as well as improved knowledge, attitudes, and behaviors toward health and exercise.<sup>14,15</sup> Additionally, games can build supportive communities through chatrooms, forums, and social networking.<sup>16</sup>

Although well-being apps and serious games have proliferated, their impact is unclear as few have been rigorously evaluated.<sup>17,18</sup> Researchers and app developers can collaborate to identify and evaluate apps with potential to relieve depression and enhance well-being.

### *The present study*

SuperBetter (SB) is an innovative smartphone- and Internet-based tool that uses game mechanics to increase users' drive to accomplish challenging goals and to build social support through online discussion forums and Facebook integration. (More information on SB is available in Supplementary Table S1 [Supplementary Data are available online at [www.liebertonline.com/g4h](http://www.liebertonline.com/g4h)].) The University of Pennsylvania (Philadelphia, PA) and SuperBetter Labs, LLC (San Francisco, CA) collaborated to conduct a randomized controlled trial of SB's ability to relieve depression. A version of SB was developed to specifically target depressive symptoms using principles from successful established therapies, namely, CBT and positive psychotherapy (PPT).<sup>19–21</sup> This version ("CBT-PPT SB") was evaluated alongside an existing version of SB ("General SB"), not specifically designed for depression. Both versions were compared with each other and with a waiting list control group (WL).

We anticipated that SB use would result in decreased depression. We expected CBT-PPT SB to provide the greatest benefit, given its basis in established interventions for depression. According to the cognitive theory of depression, participants should benefit from learning cognitive restructuring techniques that help them identify and correct distorted, negative thoughts about the self, world, and future.<sup>22</sup> In addition, participants should benefit from behavioral activation, a well-established technique that alleviates depression by increasing daily experiences of pleasure and mastery.<sup>23</sup> Finally, PPT has been found to alleviate depression symptoms by increasing positive emotions, meaning, and engagement.<sup>21</sup>

We expected General SB to confer a more modest benefit. First, it should benefit participants by facilitating so-called common factors such as positive expectancy and social support.<sup>24</sup> Second, activities focused on self-esteem should benefit participants by addressing depressive self-devaluation.<sup>22</sup> Finally, third-wave CBT approaches have recently highlighted the value of acceptance-based treatment strategies, which are included in this version of SB.<sup>25</sup>

In addition, we expected that SB use would impact secondary outcomes. We anticipated that SB would diminish participants' anxiety, given the comorbidity of depression and anxiety as well as the commonalities in treatment approaches for depression and anxiety symptoms. We also anticipated that SB would raise participants' overall life

satisfaction, self-efficacy, and perceived social support, as SB was designed to empower participants to make positive changes in their lives and to connect with others. We did not hypothesize that either version of SB would prove superior with regard to these secondary outcomes.

Thus, we tested three hypotheses: (a) Participants using SB will experience greater improvements in depression symptoms compared with WL participants. (b) Participants using CBT-PPT SB will experience greater improvements in depression symptoms compared with those using General SB. (c) Participants using SB will experience greater improvements in secondary outcomes (anxiety, life satisfaction, self-efficacy, and social support) compared with WL participants.

## **Materials and Methods**

### *Participants*

Eligible participants were iPhone® (Apple, Cupertino, CA) owners (as SB was available only on iOS) 18 years of age or older, meeting the criterion score for clinically significant depression (16 or higher) on the Center for Epidemiological Studies Depression questionnaire (CES-D)<sup>26</sup> (Table 1). *A priori* power analyses, completed using G\*Power software, indicated that at least 207 participants would be needed to detect a small effect in a repeated-measures design testing a within-between interaction (e.g., a Time  $\times$  Condition interaction), assuming  $\alpha$  error probability = 0.05 and power ( $1 - \beta$  error probability) = 0.80, with a correlation of 0.40 among repeated measures. Participants ( $n = 283$ ) were recruited online between November 2012 and March 2013 through announcements on the Penn Authentic Happiness Web site and the Craigslist.org community bulletin board. The announcement guided potential participants to [www.Qualtrics.com](http://www.Qualtrics.com) where they completed a CES-D screening and baseline assessment.

Enrollment and random assignment were completed in an automated fashion on the Qualtrics Web site. After completing the baseline assessment, participants were randomly assigned to one of three conditions (using the automated Block Randomizer in Qualtrics): CBT-PPT SB, General SB, or WL. Participants were aware of whether they were assigned to SB or WL, but SB participants were not aware of the version they received or of our specific hypotheses. (Astute participants with knowledge of existing hypotheses may have recognized the CBT and PPT components of the CBT-PPT version.) Information on participant flow is provided in Figure 1.

### *Procedure*

This protocol was approved by the Institutional Review Board (i.e., Human Ethics Committee) at the University of Pennsylvania (protocol number 816882).

**Intervention content.** CBT-PPT SB targeted depression with two sets of activities. These participants first downloaded content adapted from PPT: (a) the 3 Good Things intervention,<sup>27</sup> (b) identification of personal strengths with the Values in Action Inventory,<sup>27</sup> and (c) guidance on using strengths in new ways.<sup>21,27</sup> Upon completion, they were then asked to download content adapted from two classic CBT

TABLE 1. MEASURES

	<i>Variable assessed</i>	<i>Instructions</i>	<i>Item scoring</i>	<i>Score range</i>	<i>Interpretation</i>	<i>Cronbach's alpha</i>
Center for Epidemiological Studies Depression Scale (CES-D) <sup>26</sup>	Depression	Indicate frequency of 20 symptoms in the past 2 weeks	0 to 3	0–60	≥ 16 indicates clinical depression.	0.86
Generalized Anxiety Disorder Scale (GAD-7) <sup>37</sup>	Anxiety	Indicate frequency of seven symptoms in the past 2 weeks	0 to 3	0–21	≥ 10 indicates clinical anxiety.	0.86
Satisfaction with Life Scale (SWLS) <sup>38</sup>	Life satisfaction	Indicate agreement with five statements	7-point Likert scale	5–35	Higher scores indicate higher satisfaction.	0.85
New General Self-Efficacy Scale (NGSE) <sup>39</sup>	Self-efficacy	Indicate agreement with eight statements	5-point Likert scale	8–40	Higher scores indicate higher self-efficacy.	0.89
Multidimensional Scale of Perceived Social Support (MSPSS) <sup>40</sup>	Social support	Indicate agreement with 12 statements	7-point Likert scale	12–84	Higher scores indicate more social support.	0.91
Additional questions	Technology usage	Hours of daily iPhone use; number of applications on iPhone; hours of daily computer use; comfort level with computers; hours per week on Facebook; hours per week spent and enjoyment of playing games on computer, phone, or videogame system				
	Treatment strategies	Past use and present use of (a) therapy/counseling for depression, (b) therapy/counseling for another concern, (c) medication for depression, (d) medication for another mental health concern, and/or (e) life coaching				
	Daily functioning	Number of days this week participant (a) interacted with a friend/partner/family member, (b) exercised, (c) left the house, and (d) worked				

interventions: (a) cognitive restructuring (replacing depressive thoughts with adaptive ones)<sup>22</sup> and (b) behavioral activation (planning and carrying out activities that provide pleasure and mastery).<sup>23</sup>

The General SB program focused on self-esteem and acceptance of the present. For instance, participants were asked to “practice being present” (notice surroundings, breathe deeply, etc.), collect a list of “awesome qualities” others attribute to them, or find a piece of art or music that reminds them to accept life’s ups and downs. Figure 2 shows SB screenshots.

The two versions of SB were otherwise similar. SB users interacted with a game-like platform and were invited to describe a goal (an *epic win*; here, overcoming depression), take recommended steps toward this goal (*quests*), complete recommended mood-boosting activities (*power-ups*), directly address specific obstacles (*battle bad guys*), and enlist social support if desired (invite *allies*). SB users earned points and “leveled up” as they progressed through these activities.

WL participants did not complete any prescribed intervention. They were asked only to complete surveys at 2, 4, and 6 weeks (as SB participants did also). All participants were free to concurrently use other treatment strategies (e.g., psychotherapy, coaching, and/or medication). Data about such activities were collected at each time point and used as time-varying covariates in analyses.

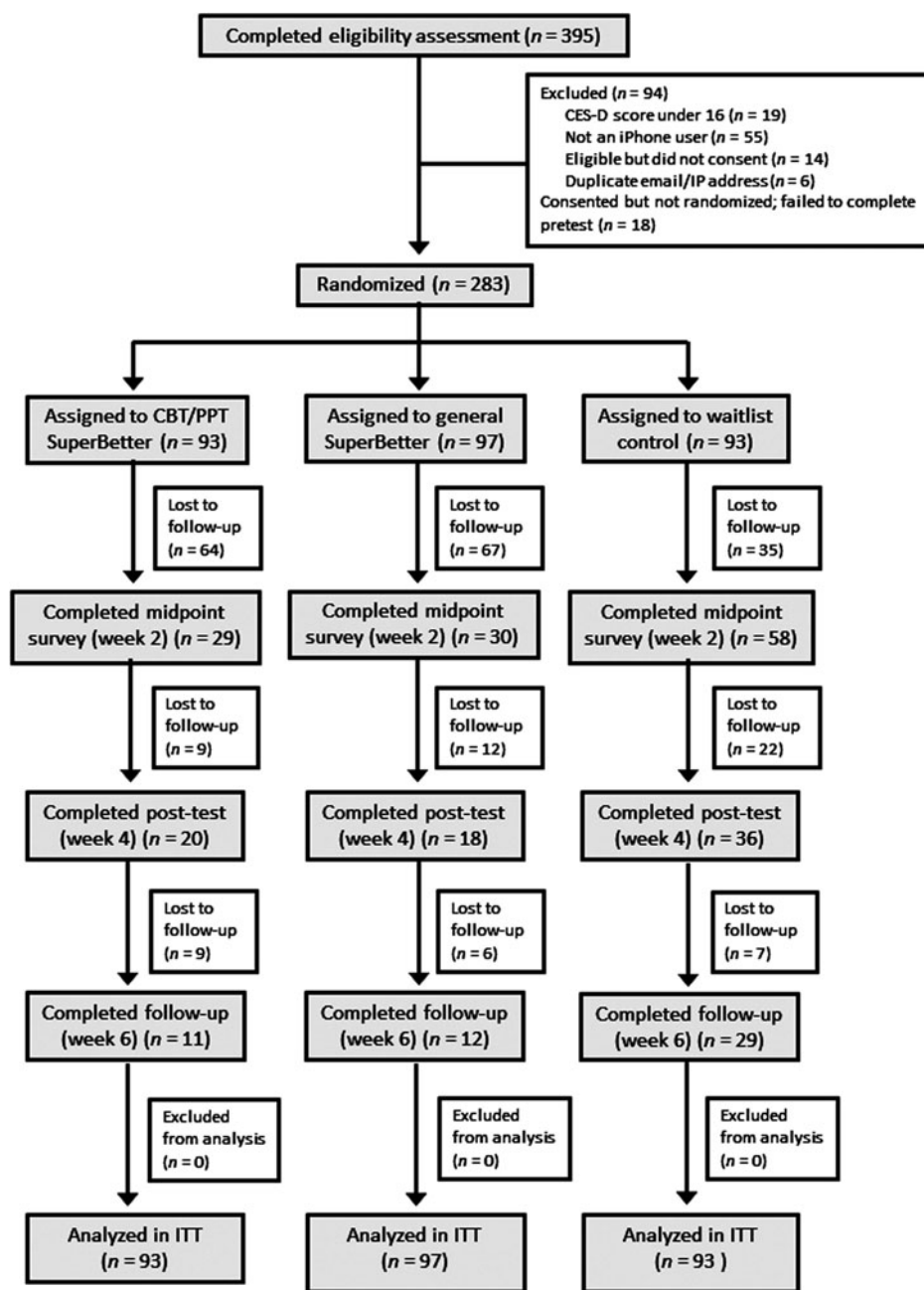
**Intervention procedure.** Both SB groups were instructed to use SB for at least 10 minutes per day for 1 month, as previous literature suggests that interventions lasting ≤ 4 weeks can effectively reduce depression symptoms.<sup>28</sup> The

intervention was targeted to occur on the SB iPhone app, but participants could also use the SB Web site on their personal computers. (The SB app may be somewhat easier to use than the Web version, mainly because participants have their mobile phones accessible when completing SB actions. Thus, they can record their actions in the app immediately. Also, fewer clicks are needed to record each action in the mobile app. On the other hand, some players prefer the accessibility of larger screens generally used with the Web version. However, content and text are identical across versions, and other differences are minor.) They were encouraged to use the forum and recruit Facebook social support (*allies*), which was optional in order to protect privacy and confidentiality.

**Measurement procedure.** Data collection occurred online via Qualtrics surveys and via participants’ iPhones/computers. (SB Labs automatically logged app usage data, such as number of log-ins and what content was downloaded.) Participants engaged in the intervention and surveys in a self-directed manner at their location and time of choice. No incentives were used to increase compliance, but e-mail reminders were sent at 2-, 4-, and 6-week assessment intervals. Data collection lasted through May 2013; the study concluded when the target enrollment was surpassed and the final wave of participants completed their 6-week follow-up assessments.

### Measures

Participants completed a total of four online surveys, each at 2-week intervals (baseline, midpoint, posttest, and follow-up). The survey contained the measures and questions detailed



**FIG. 1.** Study participant flow diagram. CBT/PPT, cognitive-behavioral therapy/positive psychotherapy; CES-D, Center for Epidemiological Studies Depression questionnaire; IP, Internet protocol; ITT, intention-to-treat.

in Table 1. The CES-D was used to assess the primary outcome variable: depression at posttest.<sup>26</sup> The CES-D is a valid measure of depression symptoms for both psychiatric populations and community samples and proved reliable in this study (Cronbach's  $\alpha=0.86$ ).<sup>29</sup> Participants also reported demographic traits: race/ethnicity, gender, age, and education level.

#### Data analysis

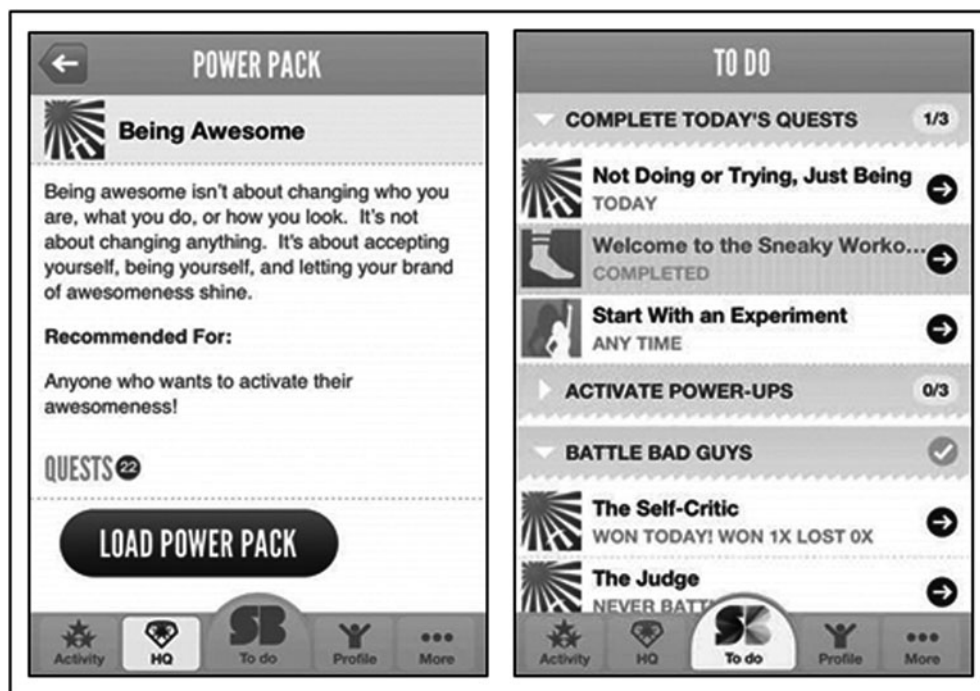
**Preparatory analyses.** We first examined the data distributions and checked that the assumptions of our intended analytic methods were met. We tested for differences in demographic and/or psychosocial variables between conditions

at baseline, using *t* tests and one-way analyses of variance (for continuous variables) and chi-squared tests (for categorical variables). When significant differences were detected, these variables were used as covariates in the main analyses.

We also examined whether intervention usage or fidelity differed across conditions and discovered that it did: 54.41 percent of participants using CBT-PPT SB ( $n=37$ ) downloaded only PPT content and did *not* download or use the CBT content. (Unlike CBT-PPT SB, which comprised these two separate downloads at distinct time points, the General SB condition required only one download to receive all content.)

**Missing data.** We retained 41.34 percent of the original sample at midpoint ( $n=117$ ), 26.15 percent at posttest





**FIG. 2.** Screenshots of the SuperBetter (SB) iPhone interface (including General SB power pack download and a sample General SB to-do list).

( $n=74$ ), and 18.34 percent at follow-up ( $n=52$ ), a high attrition rate typical for Internet-based intervention research.<sup>30</sup> Participants who stayed in the study through posttest were more satisfied with SB (mean=5.26, standard deviation [SD]=1.14) than those who did not (mean=4.24, SD=1.79) [ $t(49)=-2.51$ ,  $P=0.02$ ]. WL participants were less likely to have dropped out by posttest than either of the other groups [ $\chi^2(2, n=283)=11.53$ ,  $P=0.003$ ]. Missing data were accounted for using intent-to-treat analyses estimated with the maximum likelihood method in a hierarchical linear modeling (HLM) framework.

**HLM of outcomes.** A series of models was used to conduct an intent-to-treat analysis, modeling change in participants' depression symptoms (and in secondary outcome variables) over time. We conducted analyses using SAS Enterprise version 9.3 software.<sup>31</sup> The Level-1 model captures within-person change in depression (and other repeatedly measured variables) over four time points. This within-person change is referred to as the slope. The Level-2 model reflects participants' condition (CBT-PPT SB, General SB, or WL) as the between-persons predictor.

For all HLM models (unless otherwise noted in Results), continuous measures in the Level-1 model were centered at pretest (i.e., the intercept), and dichotomous variables were coded 1/0 to allow for meaningful evaluation of parameter estimates. We first tested unconstrained models to confirm that there was significant individual variation about the slope and intercept before accounting for random assignment to condition. Treatment effects were evaluated by examining the Time  $\times$  Condition interaction, which reflects group differences in improvement over time and is represented by the  $\beta$  coefficient associated with treatment condition in the Level-2

model. The  $\beta$  coefficient represents how much the slope of the dependent variable (e.g., depression symptoms) changes with every 1-unit change in the independent variable of interest (here, condition), controlling for any other variables in the Level-2 model. We calculated effect sizes (Cohen's  $d$ ) for between-group changes using the procedure recommended by Feingold.<sup>32</sup> In all models, we included key covariates: participants' use of other treatment (medication, therapy, and coaching) during the trial, age at baseline, and gender.

Post hoc analysis examining the impact of CBT content. As noted above, 54.41 percent of CBT-PPT SB users downloaded only half of the intended content. We suspect that this was due not to meaningful systematic differences across individuals, but rather to confusion about how to download content: Unlike General SB participants, CBT-PPT SB participants were asked to download content on two separate occasions. To better understand the impact of PPT versus CBT content, we conducted another HLM analysis in which we separately examined the impact of General content, PPT content, and CBT content compared with WL (a treatment-on-treated analysis).

## Results

### Participant flow

There were 283 participants randomly assigned to the CBT-PPT SB ( $n=93$ ), General SB ( $n=97$ ), or WL ( $n=93$ ) groups. (Three individuals were excluded from analyses because they enrolled in the study twice and were assigned to two different conditions.) Of these, 117 completed the midpoint assessment, 74 completed the posttest, and 52 completed the follow-up (Fig. 1).

### Demographics and clinical characteristics

The sample was primarily white ( $n=238$ , 84.1 percent), female ( $n=197$ , 69.60 percent), and educated at the Associate's Degree (i.e., 2-year college degree) level or higher ( $n=261$ , 92.2 percent). Mean age was 40.15 ( $SD=12.40$ ) years. Comparing the three conditions, age differences approached conventional levels of statistical significance [ $F(2, 280)=2.89$ ,  $P=0.06$ ], as did gender differences [ $\chi^2(2, n=280)=5.57$ ,  $P=0.06$ ]. Thus, age and gender were used as covariates in subsequent analyses. Other demographic characteristics, recruitment source, and key clinical characteristics did not significantly differ across groups. Table 2 provides detailed demographic and clinical information.

At baseline, participants' mean CES-D score of 33.39 ( $SD=9.41$ ) reflected clinically significant levels of depression symptoms. At baseline, 35.7 percent ( $n=101$ ) of participants were using therapy as a strategy to treat depression and/or another condition, 43.8 percent ( $n=124$ ) were using medication, and 7.4 percent ( $n=21$ ) were using life coaching. Overall, 61.1 percent of participants ( $n=173$ ) were using one or more of these strategies. Depression symptoms and other psychosocial variables did not significantly differ across groups (Table 2).

### Treatment fidelity

This study prioritized external validity and made SB usage as naturalistic as possible. Treatment adherence was lower than might be expected in traditional clinical randomized controlled trials, likely because of the absence of incentives. Of 190 participants assigned to use SB, 75.80 percent ( $n=144$ ) logged in at least once. Number of log-ins ranged from 1 to 274 total, with a mean of 21.53 ( $SD=34.27$ ) and median of 9.50. Relatively few participants used the optional forum ( $n=21$ ) or invited allies via Facebook ( $n=6$ ), and this did not significantly differ across the two SB conditions. Similarly, the two SB conditions did not significantly differ in the number of times they signed in or used various SB features (power-ups, quests, battles, and extra powerpacks).

Some CBT-PPT SB participants did not receive all intended SB content, perhaps owing to confusion about technical aspects of the app. Of the 93 participants assigned to CBT-PPT SB, 72 (77.4 percent) logged in to SB. Of the 68 (73.12 percent) who downloaded content, 37 (54.41 percent) downloaded PPT content only, whereas 31 (45.59 percent) downloaded both PPT and CBT as intended. (In contrast, of the 97 participants assigned to General SB, 72 [74.23 percent] logged in, and 64 of these [88.89 percent] downloaded all the intended content for this group.) We further examine this issue below.

### Primary outcome: Changes in depression by posttest

SB's effects were evaluated by examining the significance of the difference between the rates of change (slopes) in CES-D scores for the CBT-PPT SB, General SB, and WL conditions. All models controlled for participants' age, gender, and use of other treatment (medication, therapy, and coaching). (Note that medication, therapy/coaching, and gender did not predict change in depression. Older participants showed slightly greater decreases in depression [Age coefficient =  $-0.13$ ,  $t(276)=-2.83$ ,  $P<0.001$ ].) Table 3 gives means and SDs, and Tables 4 and 5 give HLM parameter estimates and significance tests. Generally, participants became less depressed over time (i.e., the main effect of time was significant). SB users achieved greater relief from depression symptoms than WL participants (Table 4), with an effect size (Cohen's  $d$ ) of 0.67 by posttest. The difference between SB and WL groups was statistically significant at posttest [Condition coefficient =  $-6.13$ ,  $t(276)=-3.90$ ,  $P<0.001$ ] (the Condition coefficient was estimated in the model using scores centered at posttest).

Participants using CBT-PPT SB did not achieve greater relief from symptoms than participants using General SB, contrary to prediction (Fig. 3). Both groups showed significantly faster rates of improvement than WL (Table 5). CBT-PPT SB participants were less depressed than WL participants at posttest [Condition coefficient =  $-3.92$ ,  $t(275)=-2.06$ ,  $P=0.04$ ], as were General SB participants

TABLE 2. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS

Characteristic	CBT-PPT SB (n=93)	General SB (n=97)	Waiting list control (n=93)
<b>Demographic</b>			
Age (years) [mean (SD)]	42.28 (12.56)	37.99 (11.31)	40.27 (13.06)
Sex (percentage female)	57 (61.29)	72 (74.23)	71 (76.34)
Ethnicity [number (percent)]			
Arab	0	1 (1.03)	0
Asian/Pacific Islander	2 (2.15)	6 (6.19)	2 (2.15)
Black/African American	0	3 (3.09)	3 (3.23)
Hispanic or Latino	8 (8.60)	5 (5.16)	5 (5.38)
White/Caucasian	77 (82.80)	75 (77.32)	76 (81.72)
Other	4 (4.30)	3 (3.09)	0
Unknown	0	1 (1.03)	1 (1.08)
Multiracial	2 (2.15)	3 (3.09)	6 (6.45)
<b>Clinical (at baseline) [number (percent)]</b>			
Medication	35 (37.63)	48 (49.49)	41 (44.09)
Therapy	34 (36.56)	31 (31.96)	36 (38.71)
Coaching	7 (7.53)	7 (7.23)	7 (7.53)

CBT-PPT, cognitive-behavioral therapy/positive psychotherapy; SB, SuperBetter; SD, standard deviation.

TABLE 3. MAIN STUDY OUTCOMES ACROSS BASELINE AND 2-, 4-, AND 6-WEEK ASSESSMENTS

<i>Outcome</i>	<i>Baseline</i>	<i>Midpoint (2 weeks)</i>	<i>Posttest (4 weeks)</i>	<i>Follow-up (6 weeks)</i>
Depression <sup>a</sup>				
CBT-PPT SB	34.48 (9.24)	25.66 (12.93)	23.55 (13.73)	18.73 (13.19)
General SB	33.07 (8.81)	23.77 (10.81)	19.06 (10.30)	16.83 (9.63)
Waiting list control	32.62 (10.15)	28.34 (10.60)	27.36 (10.63)	25.14 (15.14)
Anxiety <sup>b</sup>				
CBT-PPT SB	12.49 (4.75)	9.10 (5.63)	8.20 (6.01)	8.18 (5.02)
General SB	10.99 (4.98)	7.80 (4.80)	6.94 (4.09)	4.33 (3.39)
Waiting list control	11.55 (5.15)	10.84 (5.01)	9.86 (5.39)	9.28 (5.99)
Life satisfaction <sup>c</sup>				
CBT-PPT SB	13.20 (6.04)	15.41 (7.43)	16.60 (7.78)	19.64 (7.80)
General SB	14.12 (6.38)	17.37 (7.17)	18.56 (7.16)	18.42 (7.60)
Waiting list control	14.17 (6.07)	14.60 (5.67)	15.08 (6.86)	14.45 (6.63)
Self-efficacy <sup>d</sup>				
CBT-PPT SB	24.32 (5.90)	27.10 (6.57)	28.05 (4.81)	29.73 (5.78)
General SB	24.92 (6.54)	28.40 (5.17)	28.94 (6.39)	28.08 (6.87)
Waiting list control	25.35 (5.69)	26.09 (5.42)	26.44 (5.12)	25.68 (6.78)
Social support <sup>e</sup>				
CBT-PPT SB	50.04 (15.53)	49.31 (15.18)	52.75 (14.65)	60.46 (10.92)
General SB	52.70 (14.51)	57.86 (16.02)	63.06 (13.14)	65.50 (11.21)
Waiting list control	52.61 (16.05)	54.04 (15.69)	53.14 (15.36)	52.57 (15.57)
Sample size				
CBT-PPT SB	93	29	20	11
General SB	97	30	18	12
Waiting list control	93	58	36	29

Data are mean (standard deviation) values. They are derived from raw data and represent descriptive statistics of each subsample *n* rather than estimated means from the hierarchical linear modeling model.

<sup>a</sup>Measured by the Center for Epidemiological Studies Depression Scale.

<sup>b</sup>Measured by the Generalized Anxiety Disorder Scale.

<sup>c</sup>Measured by the Satisfaction with Life Scale.

<sup>d</sup>Measured by the New General Self-Efficacy Scale.

<sup>e</sup>Measured by the Multidimensional Scale of Perceived Social Support.

CBT-PPT, cognitive-behavioral therapy/positive psychotherapy; SB, SuperBetter;

[Condition coefficient =  $-8.37$ ,  $t(275) = -4.37$ ,  $P < 0.001$ ] (with the Condition coefficient estimated using scores centered at posttest). Neither version of SB proved superior; the estimated difference between the two SB groups' rates of change (1.06), favoring General SB, was not significant [ $t(237) = 0.82$ ,  $P = 0.41$ ]. CBT-PPT SB yielded an effect size of  $d = 0.43$ , and General SB yielded an effect size of  $d = 0.92$

by posttest, in comparison with WL. These calculations should be interpreted cautiously in light of high attrition rates and the motivated, self-selected sample (discussed below).

#### Secondary outcomes

**Depression by follow-up.** By the follow-up assessment, SB users again reported significantly greater changes in

TABLE 4. MAIN EFFECTS AND TIME  $\times$  CONDITION INTERACTION EFFECTS ANALYSIS FOR STUDY OUTCOMES USING HIERARCHICAL LINEAR MODELS: SUPERBETTER COMPARED WITH WAITING LIST CONTROL

<i>Outcome</i>	<i>Time coefficient</i>	<i>t ratio</i>	<i>df</i>	<i>P</i>	<i>Time <math>\times</math> Condition coefficient</i>	<i>t ratio</i>	<i>df</i>	<i>P</i>
Depression <sup>a</sup>	-2.10	-3.22	238	0.002	-3.62	-3.94	238	<0.001
Anxiety <sup>b</sup>	-0.41	-1.63	237	0.11	-1.42	-3.93	237	<0.001
Life satisfaction <sup>c</sup>	0.14	0.48	237	0.63	1.56	3.81	237	<0.001
Self-efficacy <sup>d</sup>	0.22	0.80	235	0.43	1.43	3.73	235	<0.001
Social support <sup>e</sup>	0.06	0.13	232	0.90	1.62	2.39	232	0.02

Negative coefficients indicate that SuperBetter users had greater decreases over time compared with waiting list control participants. Positive coefficients indicate that SuperBetter users had greater gains over time compared with waiting list control participants. Coefficients represent the effect of condition controlling for age at baseline, gender, and medication and therapy/coaching usage (measured at each time point).

<sup>a</sup>Measured by the Center for Epidemiological Studies Depression Scale.

<sup>b</sup>Measured by the Generalized Anxiety Disorder Scale.

<sup>c</sup>Measured by the Satisfaction with Life Scale.

<sup>d</sup>Measured by the New General Self-Efficacy Scale.

<sup>e</sup>Measured by the Multidimensional Scale of Perceived Social Support.

TABLE 5. MAIN EFFECTS AND TIME  $\times$  CONDITION INTERACTION EFFECTS ANALYSIS FOR STUDY OUTCOMES USING HIERARCHICAL LINEAR MODELS: COGNITIVE-BEHAVIORAL THERAPY/POSITIVE PSYCHOTHERAPY SUPERBETTER AND GENERAL SUPERBETTER COMPARED WITH WAITING LIST CONTROL

Outcome	Time coefficient	t ratio	df	P	Time $\times$ Condition coefficient	t ratio	df	P
Depression <sup>a</sup>	−2.09	−3.24	237	0.001				
CBT-PPT SB					−3.12	−2.80	237	0.01
General SB					−4.17	−3.73	237	<0.001
Anxiety <sup>b</sup>	−0.41	−1.65	236	0.10				
CBT-PPT SB					−1.08	−2.48	236	0.01
General SB					−1.78	−4.10	236	<0.001
Life satisfaction <sup>c</sup>	0.14	0.47	236	0.64				
CBT-PPT SB					1.77	3.55	236	<0.001
General SB					1.36	2.71	236	0.007
Self-efficacy <sup>d</sup>	0.21	0.79	234	0.43				
CBT-PPT SB					1.68	3.59	234	<0.001
General SB					1.19	2.55	234	0.01
Social support <sup>e</sup>	0.06	0.12	231	0.90				
CBT-PPT SB					1.77	2.12	231	0.04
General SB					1.51	1.80	231	0.07

Negative coefficients indicate that SuperBetter (SB) users had greater decreases over time compared with waiting list control participants. Positive coefficients indicate that SB users had greater gains over time compared with waiting list control participants. Scores were centered at pretest for these analyses, gender and age at baseline were treated as covariates, and medication and therapy/coaching usage (measured at each time point) were treated as time-varying covariates.

<sup>a</sup>Measured by the Center for Epidemiological Studies Depression Scale.

<sup>b</sup>Measured by the Generalized Anxiety Disorder Scale.

<sup>c</sup>Measured by the Satisfaction with Life Scale.

<sup>d</sup>Measured by the New General Self-Efficacy Scale.

<sup>e</sup>Measured by the Multidimensional Scale of Perceived Social Support.

CPT-PPT, cognitive-behavioral therapy/positive psychotherapy.

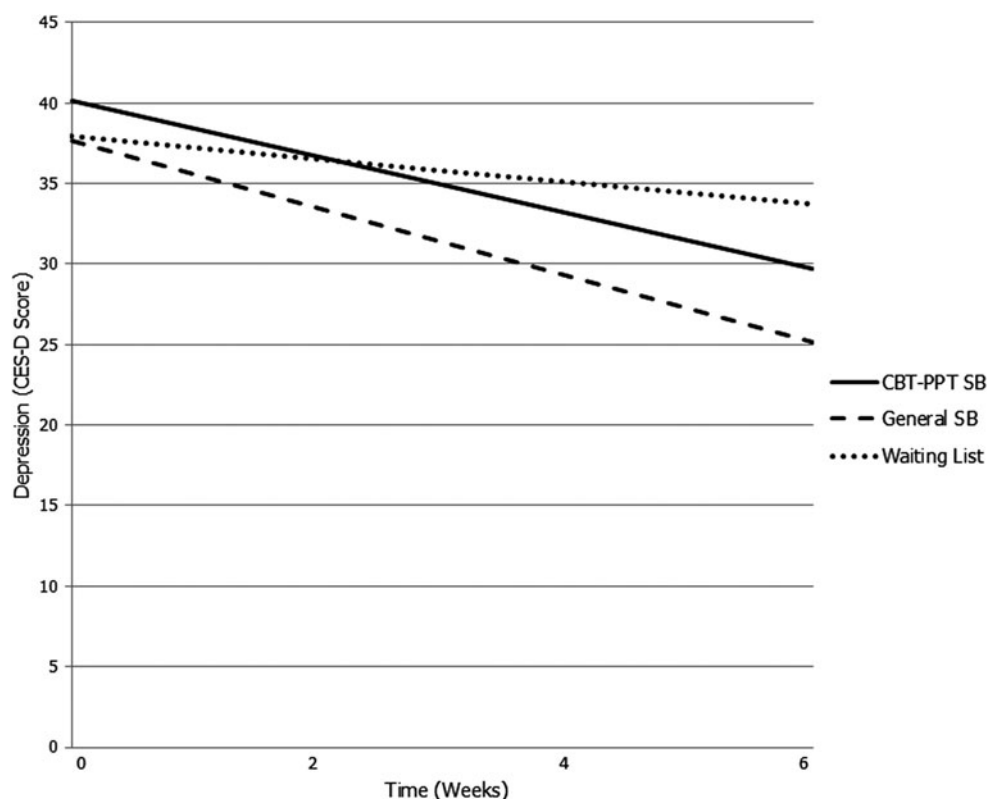


FIG. 3. Change in depression symptoms over a 6-week period across conditions. CBT-PPT, cognitive-behavioral therapy/positive psychotherapy; CES-D, Center for Epidemiological Studies Depression questionnaire; SB, SuperBetter.



depression than WL participants ( $d=1.05$ ) (Table 4). Again, both CBT-PPT SB ( $d=0.76$ ) and General SB ( $d=1.36$ ) participants improved more rapidly than WL participants (Table 5). Both CBT-PPT SB [Condition coefficient =  $-6.99$ ,  $t(275) = -2.66$ ,  $P=0.008$ ] and General SB [Condition coefficient =  $-12.42$ ,  $t(275) = -4.70$ ,  $P<0.001$ ] participants were less depressed than WL participants at follow-up (with this Condition coefficient estimated using scores centered at posttest). Again, these models controlled for age, gender, and use of other treatment (medication, therapy, and coaching).

Anxiety, life satisfaction, self-efficacy, and social support. Detailed information about SB's impact on secondary outcome variables is provided in Tables 4 and 5. Of note is that SB users experienced greater decreases in anxiety than WL and greater improvements in life satisfaction, self-efficacy, and social support than WL. Again, both versions of SB were generally effective, and neither version proved superior to the other. Here too, all models controlled for age, gender, and use of other treatment.

#### Exploratory analysis of impact of CBT content

We conducted a treatment-on-treated analysis to determine the impact of PPT content versus combined CBT-PPT content. We used a hierarchical linear model similar to those above to investigate the impact of having actually downloaded PPT content only ( $n=37$ ), CBT and PPT content ( $n=31$ ), or General SB content ( $n=64$ ). In comparison with WL participants who did not use SB at all ( $n=79$ ), participants who actually downloaded General SB or the complete CBT-PPT content achieved significantly greater decreases in depression (Table 6). In contrast, those who downloaded PPT only did not fare better than participants who did not use SB at all. (In 14 cases it was uncertain whether a WL participant had downloaded SB content before or after the WL period had elapsed. These individuals were excluded here in order to make this a clean treatment-on-treated analysis.)

#### Discussion

Participants who used the SB tool achieved decreases in depression symptoms. Our first hypothesis was supported: SB users demonstrated significantly fewer depressive symptoms by posttest than control participants. These findings are consistent with past research on the positive impact of online programs for reducing depressive symptoms.<sup>7-10</sup>

Furthermore, these findings suggest that mobile platforms and game mechanics provide opportunities for creative and effective uses of clinical knowledge.

Our second hypothesis was not supported. Even though the CBT-PPT SB version was based on empirically supported depression interventions, CBT-PPT SB users did not experience greater decreases in depression than General SB users. Rather, both groups of SB users fared better than WL participants, and neither SB version proved superior; General SB's apparent advantage was not statistically significant.

Why was CBT-PPT SB no more effective than General SB? First, treatment fidelity issues help to explain the pattern of results. About half of CBT-PPT SB users failed to download the CBT content, and so they did not receive the full intervention as intended. We would not expect that completing just two PPT exercises would have a large impact on depression symptoms. Indeed, the treatment-on-treated analysis indicated that participants who downloaded the combined CBT-PPT content fared significantly better than the WL, whereas those who downloaded only PPT content did not. There may also be substantive issues with the CBT-PPT SB content that can explain why it was not even more effective. For instance, CBT-PPT SB users may have found this newly developed content less user-friendly and engaging compared with the General SB participants who used the more refined, established, popular SB content. In addition, the more complex skills involved in CBT may take longer to master or may require more direct and intensive guidance. Moreover, depressed individuals can have motivational and cognitive deficits that cause them to be frustrated and discouraged by challenging, reflective activities; in contrast, easy, pleasant activities offer a simpler path to improved mood.<sup>33</sup>

On the other hand, why was General SB more effective than anticipated? First, this version of SB was chosen because it was a well-established favorite of previous SB users, and so it may have been much more engaging. Second, negative self-evaluations are central to depression, so General SB's self-esteem content may have conferred important benefits.<sup>22</sup> Third, there is increasing evidence that acceptance is a useful approach for dealing with depression.<sup>34</sup> General SB's acceptance-based content may have proved accessible and helpful—and it may have been easier to adapt to the mobile format in comparison with complex cognitive restructuring skills. Fourth, it is possible that both versions of SB exert their effects through common mechanisms such as increasing positive expectancy or helping users feel empowered.

TABLE 6. TREATMENT-ON-TREATED ANALYSIS: IMPACT OF SUPERBETTER CONTENT ON DEPRESSION SYMPTOMS USING HIERARCHICAL LINEAR MODELING

Content downloaded	Time coefficient	t ratio	df	P	Time × Condition coefficient	t ratio	df	P
	-2.91	-3.73	177	<0.001				
General SB					-3.67	-2.97	177	0.003
PPT SB only					-0.39	-0.25	177	0.80
PPT and CBT SB					-2.97	-2.03	177	0.04

Negative coefficients indicate that SuperBetter (SB) users had greater decreases over time compared with waiting list control participants. Scores were centered at pretest for these analyses, age at baseline and gender were treated as covariates, and medication and therapy/coaching usage (measured at each time point) were treated as time-varying covariates.

CBT, cognitive-behavioral therapy; PPT, positive psychotherapy.

Effect sizes (Cohen's  $d$ ) ranged from 0.43 to 1.36. The binomial effect size display can help make sense of Cohen's  $d$ .<sup>35</sup> The binomial effect size display presents hypothetical differences in outcome rates across the intervention and control groups. An effect size of  $d = 0.67$  (SB versus WL at posttest) is equivalent to a scenario in which 66 percent of intervention participants experience meaningful change and only 34 percent of WL participants do. An effect size of  $d = 1.36$  (General versus WL at follow-up, the greatest effect found here) is equivalent to 78 percent of intervention participants experiencing meaningful change compared with 22 percent of WL participants.

### Limitations

These effect sizes should be interpreted cautiously in light of several important limitations of the study, concerning treatment fidelity and dosage, attrition, measurement, and generalizability.

**Treatment fidelity and dosage.** About half (54.41 percent) of participants using CBT-PPT SB ( $n = 37$ ) did not receive the complete intervention, as explained above. This limits conclusions that can be drawn about differences between the SB versions. To more accurately determine whether one version of SB is superior, it would be necessary to ensure that (in a large sample with low attrition) all participants received all intended content. It would also be beneficial to measure any moderators that may make a particular version of SB more effective for a particular set of individuals.

Also, although all SB participants received the same instructions about how often to use SB, treatment dosage (i.e., number of log-ins) varied from 1 to 274 (mean = 21.53). The majority of participants did not in fact log in daily for 1 month. It is notable that several participants commented that they would have appreciated having more content to download. SB encourages users to create their own content (*quests, battles, bad guys, power-ups*, etc.) if/when they complete all the preprogrammed content. However, it may be that some users disengage when they run out of novel predesigned content. Our log-in rates hint that to maximize SB's effectiveness, it may be necessary to spur users to return to the site daily by creating a higher volume of novel content or developing other motivational strategies.

**Attrition.** This study had low retention rates, with only 26.15 percent of the sample intact at posttest and 18.34 percent at follow-up. These large attrition rates are not uncommon in Internet-based research.<sup>30</sup> The attrition rate is due in part to our prioritization of ecological validity. For naturalistic SB usage, incentives for compliance were not used. The large attrition rate limits the conclusions we can draw, in spite of our efforts to alleviate some of these concerns by using intention-to-treat with HLM. The relatively small number of individuals who stayed in the study through follow-up strongly influenced the parameter estimates. These retained participants might have been especially motivated or high-functioning.

**Measurement.** Depression was measured using a validated self-report instrument, not a clinical assessment. Thus, participants may not have met criteria for a diagnosis of

Major Depressive Disorder, and these findings may not generalize to such individuals. (Nonetheless, the sample's mean CES-D score was well above the CES-D's standard clinical cutoff, and a majority of participants were receiving treatment.)

**Generalizability.** This sample of iPhone owners may not represent the general population, as iPhone users tend to be more educated, liberal, and of higher income than the average individual.<sup>36</sup> In addition, SB's effects may vary for users with greater or lesser comfort/expertise with mobile and online technology. Because of the short follow-up, we also cannot generalize about the long-term effects of the intervention.

It is notable that our sample was largely recruited from a self-help Web site, Authentic Happiness. Individuals who are actively seeking self-help tools may be an especially motivated, hopeful, or proactive subset of the depressed population. Furthermore, our participants were aware of whether they were in WL or intervention group. Intervention participants may have enjoyed positive expectancies, hope, and/or increased self-efficacy owing to the knowledge that they had accessed a new self-help resource. These factors could have increased the effect sizes found here.

Thus, our findings may not generalize to all depressed people, particularly those in traditional clinical settings. Our findings are more applicable to depressed individuals seeking self-help resources. Participants did not receive any incentives to take part or to comply with the intervention guidelines, and this increases the study's external validity.

### Implications for research

There are exciting future directions for this research. Positive changes were effectively elicited through an innovative technological format unlike traditional therapy and psychoeducation. This suggests that psychological interventions may benefit from embracing the creative opportunities provided by mobile platforms and game-like formats. In particular, it is important to understand how evidence-based treatment strategies can translate into novel formats without sacrificing fidelity or effectiveness. It will be valuable for researchers to replicate this study but with an emphasis on internal validity, namely, by minimizing attrition rates and maximizing treatment fidelity. We also need to understand the mechanisms of change in these new interventions and the populations for whom these interventions are most appropriate.

### Implications for clinical work

These findings suggest that smartphone-based apps provide promising opportunities for mental health interventions. The tools could be integrated with traditional psychotherapy approaches. For example, therapists could use mobile apps to assign, monitor, and collaborate with clients on homework between sessions. Although these tools cannot achieve or replace what individual therapists do, they can expand access to treatment. People unable to receive treatment owing to stigma or other barriers (cost, location, long waiting lists) could be helped by technology-assisted interventions. Additionally, smartphone-based tools could serve as appropriate and cost-effective interventions for people with minimal symptoms not yet requiring medication or therapy. It is our

hope that a new wave of innovative, evidence-based online and mobile interventions will help to alleviate depression symptoms and raise well-being.

### Acknowledgments

The authors gratefully acknowledge Dr. Eli Tsukayama and Dr. Margaret (Peggy) Kern, who provided valuable guidance on data analysis. We also deeply appreciate the support of Ms. Elizabeth F. Silverman, whose private donation to the University of Pennsylvania helped make this research possible.

### Author Disclosure Statement

A.M.R. has served as a Science Advisor to SuperBetter Labs, LLC, on a strictly *pro bono* basis. J.M. founded the company SuperBetter, LLC, a for-profit organization. SuperBetter was developed under the auspices of SuperBetter, LLC. R.B. and B.M. were formerly employed by the company SuperBetter, LLC. A.M.R., S.R.J., and O.M.R. declare no competing financial interests exist.

### References

- Barak A, Grohol JM. Current and future trends in internet-supported mental health interventions. *J Technol Hum Serv* 2011; 29:155–196.
- Jorm AF, Morgan AJ, Malhi GS. The future of e-mental health. *Aust N Z J Psychiatry* 2013; 47:104–106.
- Meglić M, Ivanovski M, Marušić A. Information technology tools to improve treatment of patients with depression: Focus on guidelines implementation. *Psychiatr Danub* 2008; 20:239–244.
- Espie CA. "Stepped care": A health technology solution for delivering cognitive behavioral therapy as a first line insomnia treatment. *Sleep* 2009; 32:1549–1558.
- World Health Organization. Depression Fact Sheet. October 2012. [www.who.int/mediacentre/factsheets/fs369/en/index.html](http://www.who.int/mediacentre/factsheets/fs369/en/index.html) (accessed January 29, 2014).
- Rupp A, Gause E, Regier D. Research policy implications of cost-of-illness studies for mental disorders. *Br J Psychiatry Suppl* 1998; 36:19–25.
- Spek V, Nyklíček I, Smits N, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: A random controlled clinical trial. *Psychol Med* 2007; 37:1797–1806.
- Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: A meta-analysis. *Cogn Behav Ther* 2009; 38:196–205.
- Richards D, Richardson T. Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clin Psychol Rev* 2012; 32:329–342.
- Powell J, Hamborg T, Stallard N, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: Randomized controlled trial. *J Med Internet Res* 2013; 15:e2.
- Kenter R, Warmerdam L, Brouwer-Dudokde Wit C, et al. Guided online treatment in routine mental health care: An observational study on uptake, drop-out and effects. *BMC Psychiatry* 2013; 13:43.
- Santhiveeran J, Grant B. Use of communication tools and fee-setting in E-therapy: A web site survey. *Soc Work Mental Health* 2005; 4:31–45.
- Michael DR, Chen SL. *Serious Games: Games That Educate, Train, and Inform*. Mason, OH; Muska & Lipman/Premier-Trade; 2005.
- Russoniello CV, O'Brien K, Parks JM. The effectiveness of casual video games in improving mood and decreasing stress. *J Cyberther Rehabil* 2009; 2:53–66.
- Papastergiou M. Exploring the potential of computer and video games for health and physical education: A literature review. *Comput Educ* 2009; 53:603–622.
- Schott G, Hodgetts D. Health and digital gaming: The benefits of a community of practice. *J Health Psychol* 2006; 11:309–316.
- Burns JM, Webb M, Durkin LA, Hickie IB. Reach Out Central: A serious game designed to engage young men to improve mental health and wellbeing. *Med J Aust* 2010; 192(11 Suppl):S27–S30.
- Merry SN, Stasiak K, Shepherd M, et al. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ* 2012; 344:e2598.
- Beck AT. The current state of cognitive therapy: A 40-year retrospective. *Arch Gen Psychiatry* 2005; 62:953–959.
- Hollon SD, Ponniah K. A review of empirically supported psychological therapies for mood disorders in adults. *Depress Anxiety* 2010; 27:891–932.
- Seligman MEP, Rashid T, Parks AC. Positive psychotherapy. *Am Psychol* 2006; 61:774–788.
- Beck AT. *Cognitive Therapy and the Emotional Disorders*. New York: Penguin Inc.; 1979.
- Cuijpers P, Van Straten A, Warmerdam L. Behavioral activation treatments of depression: A meta-analysis. *Clin Psychol Rev* 2007; 27:318–326.
- Asay TP, Lambert MJ. The empirical case for the common factors in therapy: Quantitative findings. In: Hubble MA, Duncan BL, Miller SD, eds. *The Heart and Soul of Change: What Works in Therapy*. Washington, DC; American Psychological Association; 1999: 23–55.
- Hayes SC, Luoma JB, Bond FW, et al. Acceptance and commitment therapy: Model, processes and outcomes. *Behav Res Ther* 2006; 44:1–25.
- Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; 1:385–401.
- Park N, Peterson C, Seligman MEP. Strengths of character and well-being. *J Soc Clin Psychol* 2004; 23:603–619.
- Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *J Clin Psychol* 2009; 65:467–487.
- Weissman MM, Sholomskas D, Pottenger M, et al. Assessing depressive symptoms in five psychiatric populations: A validation study. *Am J Epidemiol* 1977; 106: 203–214.
- Eysenbach G. The law of attrition. *J Med Internet Res* 2005; 7:e11.
- SAS Institute Inc. Administering SAS® Enterprise Guide 4.3 (Version 9.3) [computer software]. 2010. [www.sas.com/en\\_us/software/enterprise-guide.html](http://www.sas.com/en_us/software/enterprise-guide.html) (accessed January 29, 2014).
- Feingold A. Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychol Methods* 2009; 14:43–53.
- Sin NL, Della Porta MD, Lyubomirsky S. Tailoring positive psychology interventions to treat depressed individuals. In:

- Donaldson SI, Csikszentmihalyi M, Nakamura J, eds. *Applied Positive Psychology: Improving Everyday Life, Health, Schools, Work, and Society*. New York; Routledge; 1999: 79–96.
34. Forman EM, Herbert JD, Moitra E, et al. A randomized controlled effectiveness trial of acceptance and commitment therapy and cognitive therapy for anxiety and depression. *Behav Modif* 2007; 31:772–799.
  35. Rosenthal R, Rubin DB. A simple, general purpose display of magnitude of experimental effect. *J Educ Psychol* 1982; 74:166–169.
  36. Hunch. Android vs. iPhone: Battle of the Mobile Operating Systems [blog post]. August 15, 2011. <http://blog.hunch.com/?p=51781> (accessed January 29, 2014).
  37. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med* 2006; 166:1092–1097.
  38. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess* 1985; 49(1):71–75.
  39. Chen G, Gully SM, Eden D. Validation of a new general self-efficacy scale. *Organ Res Methods* 2001; 4:62–83.
  40. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. *J Pers Assess* 1988; 52:30–41.

Address correspondence to:

Ann Marie Roepke, MA

University of Pennsylvania Positive Psychology Center

3701 Market Street, Suite 200

Philadelphia, PA 19104

E-mail: aroepke@sas.upenn.edu