



Testing a scalable web and smartphone based intervention to improve depression, anxiety, and resilience: A randomized controlled trial

Acacia C. Parks · Allison L. Williams · Michele M. Tugade · Kara E. Hokes · Ryan D. Honomichl · Ran D. Zilca

Abstract: Psychological interventions targeting wellbeing can reliably increase wellbeing and decrease depressive symptoms. However, only a handful of studies have implemented wellbeing interventions online, and those studies have largely done so in a way that prioritizes experimental control over realism and scalability. We sought to take existing wellbeing interventions with established efficacy and to evaluate their impact when translated into a format that is publicly accessible, scalable, and designed with the goal of engaging users. Participants in this fully online trial were first-time registrants of the Happify platform, a fully automated web and mobile wellbeing intervention grounded in positive psychology, cognitive-behavioral therapy, and mindfulness-based stress reduction, which has offered wellbeing programs to over 3 million registrants to date. Consenting participants were randomly assigned to access the full Happify platform or a psychoeducation comparison condition and further categorized by their usage during the study: recommended usage (a minimum of 2-3 activities per week) or low usage (usage less than the recommended level). Participants were assessed on depressive symptoms, anxiety symptoms, and a composite measure of resilience at baseline and 8 weeks later. Participants who used Happify at the recommended level reported fewer depressive and anxiety symptoms and greater resilience after 8 weeks than participants who used Happify at a low level or participants who used the psychoeducation condition at any level. The Happify group also experienced greater rates of reduction in depression and anxiety symptom severity category, and had a greater net benefit (% users who improved minus % users who deteriorated), compared to the other groups. The results of this study suggest a successful first attempt at implementing and scaling a comprehensive package of lab-tested wellbeing interventions without losing efficacy.

Keywords: online intervention, wellbeing, depressive symptoms, anxiety symptoms, resilience, randomized controlled trial

1. Introduction

Ample research shows that interventions targeting wellbeing can reliably decrease depressive symptoms (Lamers, Bolier, Westerhof, Smit, & Bohlmeijer, 2012; Sin & Lyubomirsky, 2009). A smaller amount of literature has also seen those interventions successfully delivered online as individual exercises (Mongrain & Anselmo-Matthews, 2012; Seligman, Steen, Park, & Peterson, 2005; Sergeant & Mongrain, 2014; Shapira & Mongrain, 2010) and as packaged interventions (Cohn, Pietrucha, Saslow, Hult, & Moskowitz, 2014; Parks, Della Porta, Pierce, Zilca, &



Lyubomirsky, 2012; Schueller & Parks, 2012). Given the many barriers to real-world implementation for in-person mental health interventions, online interventions are a promising avenue for broad dissemination, due to their low per-person cost and ease of accessibility (Muñoz et al., 2016; Wang et al., 2005).

There have been numerous attempts to translate the science of improving wellbeing into a scalable format, usually in the commercial sector. Often, these interventions are science-based, but not science-generating; interventions are not empirically validated using such platforms, but rather, are assumed to work due to their scientific basis. However, we argue, in line with many other thinkers in this field, that it is not enough to be science-based; one must test interventions, in their implemented form, to make sure that they work as expected in a real-world setting (Leykin, Aguilera, Torres, Pérez-Stable, & Muñoz, 2012; Mohr, Cheung, Schueller, Brown, & Duan, 2013; Mohr et al., 2015).

In a previous study, using an online platform called Happify, we sought to implement the science of increasing wellbeing in a real-world setting and verify that these research-based activities remain effective when delivered naturalistically. We found, in an uncontrolled setting, that the use of the Happify platform led to significant improvements in wellbeing among users over 8 weeks (Carpenter et al., 2017). This study gave us initial evidence that the research-based activities we adapted from the literature remained effective in an online, interactive format; however, without a random-assignment, controlled design, it is difficult to infer causality.

In the present study, we seek to take that research a step further by evaluating Happify over an 8-week period using a randomized, controlled design. It is important to control for factors such as natural improvement over time, expectation of improvement, and behavioral activation (simply doing *anything*); the influences of these factors cannot easily be disentangled from “real” observed improvement in an uncontrolled study. Comparison groups are therefore needed to obtain a true measure of an intervention’s effect, and to be sure it had an impact above and beyond natural improvement.

In order to test Happify’s effectiveness using a rigorous experimental design, we crafted an active comparison condition, the *psychoeducation* condition, in which participants receive psychoeducational content—much like an internet user might discover searching for information about wellbeing on their own. Our psychoeducational condition was designed to provide a more robust and more meaningful comparison group than a simple waitlist control. The psychoeducational group was designed to provide a non-clinical equivalent to a “treatment as usual” condition—similar to many existing online platforms that focus on providing users with content and information without any concrete activities. This, therefore, helps us control for the impact of expectation and behavioral activation.

In order to approximate and account for the effects of natural improvement over time, we also examined the moderating impact of platform usage, as measured by the number of activities completed during the 8-week intervention period. We did this by differentiating between participants who followed the recommended usage guidelines by completing a minimum of 2-3 activities per week during the intervention (i.e., a minimum of 16 total activities) and participants who completed fewer activities than this recommended amount (i.e., a maximum of 15 total activities). Although the self-selected nature of this categorization is a limitation, it allows us to separate users who received the recommended dosage from those who did not, and may as a result only be experiencing improvement that would have happened naturally over time, without intervention.

In summary, this study compares change in wellbeing over time among two randomized conditions—Happify and psychoeducation—and also considers usage level (as recommended, or low usage) as a moderator. We specifically looked at three aspects of wellbeing: depression, anxiety, and resilience. Previous research finds that the types of activities that are included on Happify robustly improve depressive symptoms (Bolier et al., 2013; Sin & Lyubomirsky, 2009), and even that Happify itself can reduce depressive symptoms (Carpenter et al., 2017). We seek to replicate our previous finding of Happify’s effect on depression with an RCT design, and explore those effects in more depth. Although depression and anxiety are correlated, and share some common pathways for treatment, only one study that we are aware of has demonstrated positive effects of a wellbeing intervention on anxiety in adults (Schotanus-Dijkstra et al., 2017). This initial result is enough for us to explore Happify’s results on anxiety as well. Lastly, because Happify is designed to not only reduce negative symptomatology, but also to build resistance to future symptoms, we explored resilience. A few studies are emerging that examine the role of mobile technology to build resilience (e.g., Bush, Bosmajian, Fairall, McCann, & Ciulla, 2011; Hoorelbeke & Koster, 2017; Kizakevich et al., 2012; Orr, King, Hawke, & Dalgleish, 2013; Rose et al., 2013; Roy et al., 2017; Roy, Highland, & Costanzo, 2015). To our knowledge, however, few internet-based wellbeing intervention studies to date have directly measured resilience, so we considered this outcome more exploratory.

In addition to examining group differences in outcome change over 8 weeks, we also explored “practical significance” for depression and anxiety, where scale cut-off scores for severity are available. Specifically, we examined rates of participants experiencing improvement, degeneration, or no change in symptom severity levels between recommended usage Happify participants and participants in all other conditions combined. We hypothesized that participants in the Happify condition would improve, compared to participants in the psychoeducation condition, who received some components of Happify (polls and exposure to information about wellbeing) but not others (e.g., interactive activities for improving wellbeing), and that this would be especially true for those who used the platform at the recommended level. We thus expected participants using the Happify condition at the recommended level to show the greatest improvements in symptoms, as well as severity levels, compared to those using Happify at a low rate or those using a psychoeducation comparison platform.

2. Method

2.1 Participant recruitment, onboarding procedures, and screening

Recruitment. This research was conducted under the supervision of the Hiram College Institutional Review Board. Participants were recruited in February and March 2017 from among all new users of the consumer version of Happify.com (see Appendix A). A “freemium” version of Happify is available to the general public for free, with additional premium content available to subscribers. Participants in the current study who were randomized to the Happify condition received the full premium version.

All study experiences took place via smartphone or web, with no human contact outside of assistance with technical problems via email. The only inclusion criterion for intake, besides the level of computer literacy needed in order to find Happify in the first place, was that the user be a new registrant, with no experience on the platform to date. All new users were offered the opportunity to participate in the study at the end of the process of registering for the free consumer site. Participants were informed that the study involved using the Happify wellbeing program two to three times per week, and that their experiences might not necessarily be

reflective of the full Happify product. A total of 7,373 users completed the registration process and consented to participate in this study (see Figure 1 below for a graphical representation of participant recruitment and retention).

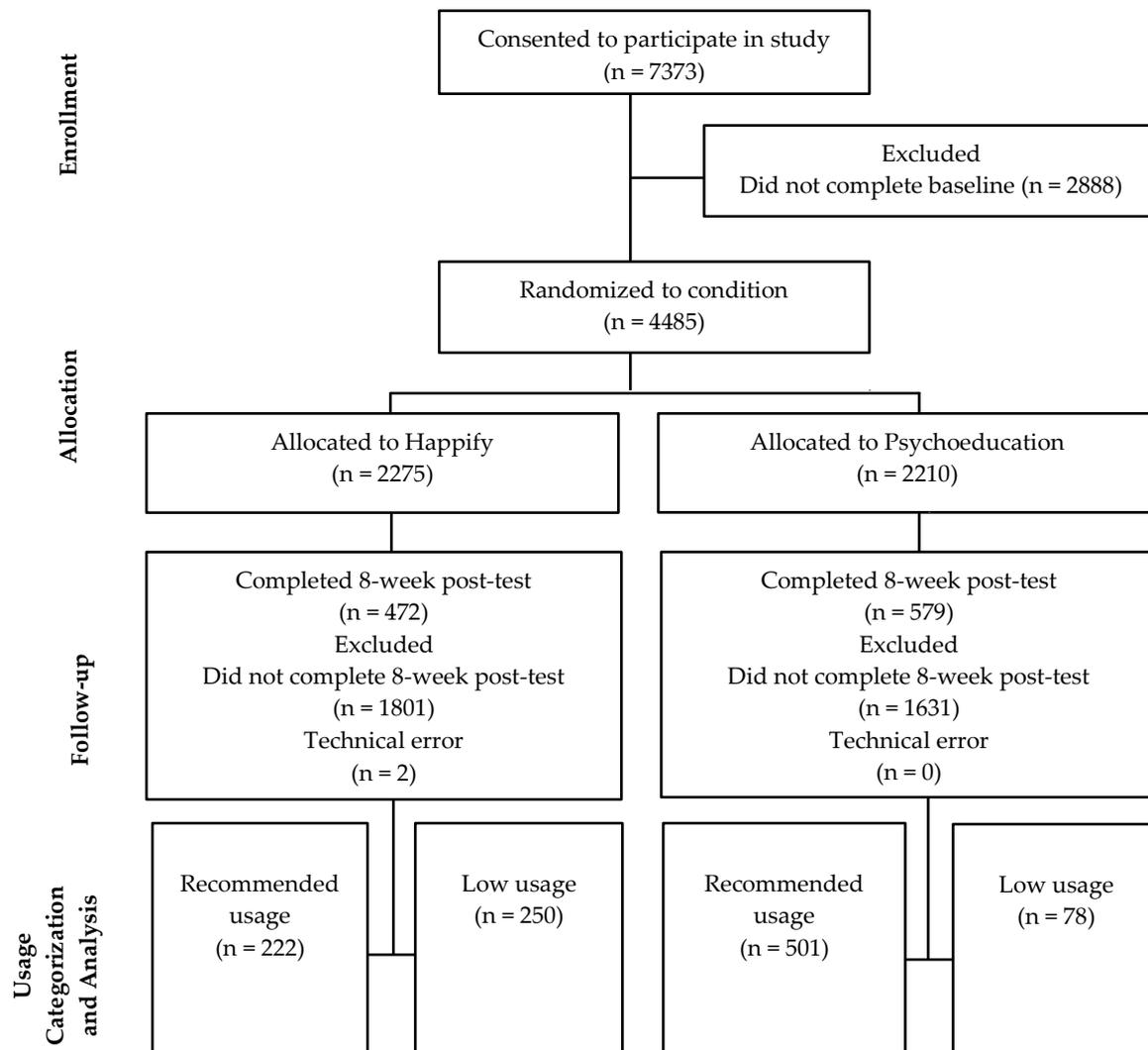
As data security is a concern for any online data collection, we implemented various safety measures to protect user privacy. They are summarized in the consent form (see Appendix B) as follows:

When you register for the study, you will be asked to give us your email address. In our secure database, your email address will be tied to your data by a randomly generated numerical code, which will allow us to connect each new questionnaire you take to the other questionnaires you have taken. Depending on whether or not your email address contains identifying information, this may or may not constitute a risk to privacy were someone to compromise the security of our database. However, because email addresses will be stored separately from questionnaire data, the researchers will never be able to connect your data with any identifying information. However, if you wish to be completely certain that your data will remain private under any circumstances, you may create an anonymous email address for the purposes of this study and have its messages forwarded to your main account.

Allocation. Participants who consented to participate in the study were prompted to complete the baseline assessment via email, as well as push notifications on their mobile device, if they had installed the Happify app. The baseline assessment included measures of depressive symptoms (measured using the Patient Health Questionnaire, PHQ-9; Kroenke et al., 2001), anxiety symptoms (measured using the Generalized Anxiety Disorder scale, GAD-7; Spitzer et al., 2006), stress (measured using the Perceived Stress Scale, PSS; Cohen et al., 1983), positive emotionality (measured using the emotion subscale of the Happify Scale, HS-E; Carpenter et al., 2017), and optimism (measured using the Life Orientation Scale, Revised, LOT-R; Scheier & Carver, 1985), and took about 5 minutes to complete; users were compensated with a \$5 gift card for completing the assessment. Participants who completed the consent process and baseline assessment ($N = 4,485$) were then randomized to condition (Happify $N = 2,275$, psychoeducation $N = 2,210$; 2,888 users did not complete the baseline assessment).

Post-test assessment. The 8-week post-test assessments were administered in April and May 2017; 1,053 participants completed the post-test (Happify $N = 474$, psychoeducation $N = 579$). Due to a technical error, however, two participants from the Happify condition were offered the post-test assessment prematurely, less than 2 weeks after completing the baseline assessment; these participants were excluded from analysis.

Usage categorization and analysis. Because we were interested in testing the impact of research-tested interventions in a new format—not so much on establishing their efficacy, which is already established in previous work, but on seeing if their efficacy persists with a new method of administration—we also examined the moderating impact of platform usage on each of our study outcomes. To do this, we classified participants who engaged with the platform at the recommended minimum rate of two to three activities per week (i.e., a minimum of 16 activities over the course of the 8-week intervention) as *recommended usage* participants, and those who engaged with the platform at less than the recommended rate (i.e., a maximum of 15 activities over the course of the 8-week intervention) as *low usage* participants. Thus, of the 1,051 participants included in analysis, 222 were in the Happify condition and used it at the recommended rate, 250 used Happify at a low rate, 501 used the psychoeducation condition at the recommended rate, and 78 used the psychoeducation condition at a low rate.

Figure 1. Study flow and user retention.

Table 1. Differences in baseline depression, anxiety, and resilience between participants completing at least 2 assessments and those who completed only 1 assessment

	1 Assessment (N = 3434) Mean (SD)	2+ Assessments (N = 1051) Mean (SD)	<i>t</i>	<i>p</i>
Depression (PHQ)	12.9 (6.6)	11.9 (6.6)	4.42	.001
Anxiety (GAD)	11.2 (5.8)	10.3 (5.8)	4.45	.001
Resilience	41.0 (16.8)	42.4 (16.7)	2.27	.022

Table 2. Demographic information for participants completing at least 2 assessments and those who completed only 1 assessment

	1 Assessment N = 3434	2+ Assessments N = 1051	X ²	Cramer's V	df	p
Gender			7.34	.04	2	.025
Male	545 (15.9%)	131 (19.4%)				
Female	2848 (83.1%)	908 (86.4%)				
Complicated	33 (1.0%)	10 (1.0%)				
Age			45.22	.10	5	< .001
18-24	1366 (39.9%)	300 (28.6%)				
25-34	1078 (31.5%)	384 (36.6%)				
35-44	526 (15.4%)	192 (18.3%)				
45-54	288 (8.4%)	113 (10.8%)				
55-64	121 (3.5%)	47 (4.5%)				
65+	48 (1.4%)	14 (1.3%)				
Employment			11.04	.05	5	.051
Employed	1715 (50.0%)	564 (53.7%)				
Homemaker	166 (4.8%)	56 (5.3%)				
Retired	67 (2.0%)	17 (1.6%)				
Self-employed	255 (7.4%)	82 (7.8%)				
Student	918 (26.8%)	230 (21.9%)				
Unemployed	306 (8.9%)	101 (9.6%)				

In order to assess the representativeness of people who completed both pre- and post-questionnaires (those who were included in our study) as compared to the users who only completed baseline (who were excluded for lack of data), we compared these two groups on baseline characteristics and demographic variables. There are some statistically significant differences, and a summary of these differences are presented in Tables 1 and 2. Participants who completed both assessments were significantly less depressed, less anxious, and more resilient than non-completers. On demographic variables (age, employment status, gender),

there was a significant difference on age and gender. Dropouts were more likely to be in the 18-24 age category and less likely to be male.

Table 3. Demographic information for participants included in analyses, by condition and usage

	Happify		Psychoeducation	
	LU (N = 250)	RU (N = 222)	LU (N = 78)	RU (N = 501)
Gender				
Male	39 (15.6%)	15 (6.8%)	13 (16.7%)	64 (12.8%)
Female	208 (83.2%)	205 (92.8%)	63 (80.8%)	432 (86.4%)
Complicated	3 (1.2%)	1 (0.5%)	2 (2.6%)	4 (0.8%)
Age				
18-24	80 (32.0%)	59 (26.6%)	22 (28.2%)	139 (27.8%)
25-34	79 (31.6%)	88 (39.6%)	32 (41.0%)	185 (37.0%)
35-44	44 (17.6%)	42 (18.9%)	11 (14.1%)	95 (19.0%)
45-54	30 (12.0%)	23 (10.4%)	7 (9.0%)	53 (10.6%)
55-64	11 (4.4%)	8 (3.6%)	4 (5.1%)	24 (4.8%)
65+	6 (2.4%)	2 (0.9%)	2 (2.6%)	4 (0.8%)
Employment				
Employed	132 (52.8%)	125 (56.3%)	37 (47.4%)	270 (54.0%)
Homemaker	7 (2.8%)	10 (4.5%)	5 (6.4%)	34 (6.8%)
Retired	7 (2.8%)	3 (1.4%)	1 (1.3%)	6 (1.2%)
Self-employed	15 (6.0%)	17 (7.7%)	7 (9.0%)	43 (8.6%)
Student	60 (24.0%)	50 (22.5%)	23 (29.5%)	97 (19.4%)
Unemployed	29 (11.6%)	17 (7.7%)	5 (6.4%)	50 (10.0%)

LU = Low usage, RU = Recommended usage

The final sample used for analysis contained 1,051 users who completed the post-test assessment without experiencing a technical error. Overall, the sample was relatively

distressed, and experiencing symptoms on the border between subclinical and clinical. At baseline, average PHQ score was 11.9 ($SD = 6.6$), which is in the middle of the moderate depression range (10-14) (Kroenke et al., 2001). Average GAD score was 10.3 ($SD = 5.8$), which indicates moderate (10-14) anxiety (Löwe et al., 2008). Perceived stress scores were in the moderate range ($M = 16.9, SD = 7.1$), which is similar to the average stress levels in the general public (Cohen & Janicki-Deverts, 2012). Optimism scores averaged 11.7 ($SD = 5.3$), which is close to the midpoint of the scale, and positive emotion scores averaged 36.2 ($SD = 18.4$), which, according to internal validation data, is in the bottom 15th percentile of scores for consumers on Happify. In short, the sample was relatively distressed, and experiencing symptoms in the subclinical range.

A breakdown of demographic distributions by condition and usage can be seen in Table 3. Chi square analyses revealed significant differences in gender by condition ($X(6) = 13.42, p < .037$), which appears to be driven by the smaller percentage of male participants in the recommended usage Happify condition, relative to the three other conditions. There were no differences in age range ($p = .77$) or employment status ($p = .28$) as a function of condition. In addition, baseline levels for all continuous variables included in this study, as well as p-values for the condition by usage interaction, are presented in Table 4. We found no baseline differences as a function of condition and/or usage on any of these variables, all main effect and interaction $ps > .15$.

Table 4. Means and standard deviations of baseline continuous variables by condition and usage

Measure	Happify		Psychoeducation	
	LU (N = 250)	RU (N = 222)	LU (N = 78)	RU (N = 501)
Depression (PHQ)	12.3 (6.9)	12.3 (6.3)	12.9 (7.1)	11.4 (6.4)
Anxiety (GAD)	10.3 (6.2)	10.8 (5.6)	10.6 (6.1)	10.1 (5.6)
Resilience	43.5 (17.8)	41.6 (16.0)	42.4 (17.2)	42.2 (16.5)
Optimism (LOT-R)	12.0 (5.2)	11.7 (5.4)	11.6 (5.5)	11.6 (5.4)
Stress (PSS)	17.2 (7.6)	16.5 (6.9)	17.1 (7.2)	16.9 (7.0)
Positive Emotion (HS-E)	37.6 (20.0)	34.7 (17.6)	35.7 (17.7)	36.2 (18.0)

LU = Low usage, RU = Recommended usage, PHQ = Patient Health Questionnaire, GAD = Generalized Anxiety Disorder scale, LOT-R = Life Orientation Test-Revised, PSS = Perceived Stress Scale, HS-E = Emotion subscale of the Happify Scale

2.2 Primary outcomes: Symptoms of depression and anxiety

Two primary outcomes were assessed in this study. First, depressive symptoms were measured using the Patient Health Questionnaire, nine-item scale (PHQ-9; Kroenke et al., 2001). The PHQ-9 is a nine-item checklist that asks participants to rate the extent to which they have experienced a list of depressive symptoms (e.g., “Feeling down, depressed, or hopeless” or “Little interest or pleasure in doing things”) on a scale of 0 (not at all) to 3 (nearly every day).

Second, anxiety symptoms were measured using the Generalized Anxiety Disorder seven-item scale (GAD-7; Spitzer, Kroenke, Williams, & Löwe; 2006). The GAD-7 is a seven-item checklist with the same rating scale as the PHQ-9, with symptoms such as “Feeling nervous, anxious, or on edge” and “Not being able to stop or control worrying.” Consistent with previous research, internal consistency was acceptable for both PHQ-9 ($\alpha = .88$) and GAD-7 ($\alpha = .90$) (Löwe et al., 2008; Spitzer et al., 2006). These measures are both widely used in clinical settings to track symptom severity, and previous research has validated both in online formats (Ahmedani, Belville-Robertson, Hirsch, & Jurayj, 2016; Newby et al., 2017; Titov et al., 2013, 2015). They are known for their sensitivity to change from treatment, making them ideal for this study. Baseline levels of depression and anxiety did not differ significantly by condition or usage (see Table 4).

2.3 Secondary outcome: Resilience

The science of resilience is supported by a host of studies showing that people can maintain mental health and wellbeing despite exposure to psychological or physical adversity. This pattern, conceptualized as “resilience,” has been observed across different populations and various types of adversities (Kalisch et al., 2017). Several lines of research indicate that resilience is a multi-faceted construct that is characterized by three important features: an optimistic outlook, successful adaptation to stress, and the tendency to experience positive emotions in the midst of adversity.

Optimism. One characteristic of resilience is *optimism*, or the tendency to have a positive expectancy for future outcomes (Carver & Scheier, 2014; Carver, Scheier, & Segerstrom, 2010; Klohnen, 1996; Rutter, 2006; Scheier & Carver, 1985; Seligman, 1991; Utsey, Hook, Fischer, & Belvet, 2008). Optimism has a major role to play in adaptation to stressful conditions. When confronting a challenge, optimists show more resilience, even if progress is difficult and slow (Snyder & Lopez, 2002). Optimism facilitates resilience across the lifespan. Recent prospective studies show that one facet of optimism (namely, a positive expectancy for the future) is particularly useful for predicting resilience from depression in adults (Kleiman et al., 2017). Indeed, optimism has been identified as the most influential adolescent cognitive factor to temper the effects of life stressors (Tusaie-Mumford, 2002).

Perceived stress. Resilience is also characterized by *lower perceived distress* and higher wellbeing (Block & Kremen, 1996; Connor & Davidson, 2003; Ford, Eklund, & Gordon, 2000; Willis & Burnett, 2016). Much of the research in this domain examines self-reports of personality characteristics and outcomes from especially challenging life circumstances. For example, among recently bereaved widows, higher levels of resilience and lower levels of perceived stress predicted greater life satisfaction one month following the death of a spouse (Rossi, Bisconti, & Bergeman, 2007). One research study focused on firefighters, who are individuals considered at high risk for developing PTSD symptoms following traumatic stress. Research shows that lower levels of perceived stress characterizes a resilience factor, protecting firefighters from the detrimental impact of highly traumatic stress (Lee, Ahn, Jeong, Chae, & Choi, 2014).

An intervention study examined the outcomes of a resilience-training program (including mindfulness meditation, nutrition, and exercise sessions) in an outpatient clinic. Participants with low levels of baseline perceived stress who underwent resilience training evidenced lower levels of depression and anxiety, compared to those who did not undergo resilience training (Griffin, Johnson, Kitzmann, Kolste, & Dusek, 2015). Another randomized clinical trial study examined the effects of a resilience-training program (focused on attention and interpretation)

among academic physicians. Physicians are an important population to examine because they are especially at high risk for distress because of multiple demands and expectations placed upon them. Findings from this study show that after 8 weeks, participants who participated in the resilience-training program evidenced lower levels of perceived stress (among other variables), compared to wait-list control participants (Sood, Prasad, Schroeder, & Varkey, 2011). Taken together, these findings indicate that higher resilience can predict lower levels of perceived stress.

Positive emotionality. Maintaining and enhancing positive emotions yields important advantages when coping with stress. Several studies have shown that positive emotions can fuel resilience in the face of stress. Theoretically, resilience is defined as the ability to “bounce back” from stressful experiences quickly and effectively. The autonomic correlates of psychological resilience were examined by investigating patterns of cardiovascular recovery from negative emotional arousal (Tugade & Fredrickson, 2004). High-resilient (vs. low-resilient) individuals are more likely to report experiencing positive emotions alongside feelings of lab-induced anxiety and evidence faster cardiovascular recovery from the arousal generated by an anxiety-inducing task. Mediation analyses revealed that the experience of positive emotions contributed, in part, to high-resilient participants’ abilities to achieve efficient emotion regulation, demonstrated by accelerated cardiovascular recovery from negative emotional arousal. These findings indicate that high-resilient individuals are characterized by using positive emotions to cope with stress (Tugade & Fredrickson, 2004; Tugade, Fredrickson, & Barrett, 2004).

Similar patterns have been shown using daily diary techniques. Researchers have found that daily experiences of positive emotion are important for the regulation of negative emotions, a marker of resilience (Zautra, Affleck, Tennen, Reich, & Davis, 2005; Zautra, Johnson, & Davis, 2005). Daily experiences of positive emotion moderate stress reactivity and mediated stress recovery (Tugade et al., 2004; Tugade & Fredrickson, 2004).

Researchers argue that positive emotions are useful in the resilience process for two reasons:

- 1) positive emotions may interrupt the ongoing stress response and
- 2) positive emotions may accelerate one’s ability to adapt to subsequent stressors (Ong, Bergeman, Bisconti, & Wallace, 2006).

Over time, then, the experience of positive emotions functions to assist high-resilient individuals in their ability to recover effectively from daily stress. These findings are in line with other research showing that daily positive emotions accelerate recovery from stress, a pattern of response reflective of resilience (Moskowitz, Folkman, & Acree, 2003).

The resilience index. While there are existing resilience scales, none are considered a “gold standard” and each has its own substantial tradeoffs. In particular, many resilience scales ask about a person’s self-perceptions of their own resilience in stress, which much research has found is unreliable and not strongly related to actual resilient behavior (Windle, Bennett, & Noyes, 2011). An alternative approach would be to assess traits that are known to predict resilient behavior, but that can be reliably self-reported on by participants. In line with this view of resilience assessment, rather than using an existing scale, we developed an index of resilience using the above three important facets of resilience: (low) perceived stress, positive emotionality, and optimism (see Table 5 below). Perceived stress was measured using the Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983), a 10-item questionnaire that asks users to rate on a scale of 0 (never) to 4 (very often) how frequently they experience various symptoms of feeling stress (Cohen et al., 1983; Ezzati et al., 2014). Positive emotionality

was measured using the emotion subscale of the Happify Scale (HS-E; Carpenter et al., 2017), a four-item scale asking participants to rate the extent to which, over the past week, they have experienced positive and negative emotions that are activated or deactivated. Optimism was measured using the Life Orientation Scale, Revised (LOT-R; Scheier & Carver, 1985), a 10-item scale with four filler items; for brevity, we used only non-filler items (Vautier, Raufaste, & Cariou, 2003). Participants rate the extent to which they identify with optimistic or pessimistic beliefs on a scale of 0 (strongly disagree) to 4 (strongly agree). All three of these scales have been validated online in previous published research (Hintz, Frazier, & Meredith, 2015; McHale, Clark, & Tramonte, 2015; Rao & Kemper, 2017; Sergeant & Mongrain, 2014).

Table 5. Sample items from the three scales comprising the resilience index: Life orientation test (LOT-R), perceived stress scale (PSS), and Happify Scale - emotionality (HS-E)

Life Orientation Test	Overall, I expect more good things to happen to me than bad. *If something can go wrong for me, it will.
Perceived Stress Scale	*In the past month, how often have you felt stressed or nervous? In the past month, how often have you felt that you were on top of things?
HS Emotionality Scale	In the past month, how often have you felt joyous, exuberant, inspired, and/or awestruck? *In the past month, how often have you felt sad, guilty, and/or lonely?

Note. That PSS is typically reverse scored in the opposite direction (higher means more stressed), but for a resilience index, we scaled all measures in the same direction so that a larger number corresponded to greater resilience. Reverse-scored items are marked with an asterisk (*).

To compute the resilience composite index, we combined items from each of the three scales, weighted such that each component of resilience contributed equally to the overall composite, regardless of the number of items in a particular scale. To do this, each of the raw survey items from the PSS, HS-E and LOT-R were rescaled from their original scale to run from 0 to 1. Items within each scale were then averaged together, as were each of the three scale averages. For ease of interpretation, this final average was multiplied by 100 to quantify the resilience index, which ranged from 0 to 100. In addition to the theoretical rationale, steps were taken to justify the development of this composite resilience measure empirically. An initial correlational analysis demonstrated that each of these scales was strongly intercorrelated (r s ranged from = .54 to .67), indicating that we had the appropriate correlations to carry out a factor analysis. A principal components analysis with varimax rotation of the 20 statements supported a one-factor solution, with the first factor explaining 41.5% of the variance (eigenvalue = 8.31). Taken together, these results provide rationale for the composite resilience index. At baseline,

resilience and the various resilience components did not differ significantly by condition or usage (see Table 4).

2.4 Experimental conditions

Happify (happify.com) can be accessed via the web, or via an app (Android and iOS), and all study activities took place through the Happify platform. There were two intervention conditions between which participants could be randomly assigned: the full Happify platform ($N = 472$), and a psychoeducational comparison intervention ($N = 579$). As per the consent form (see Appendix B), participants were aware that there were different programs being tested, and that they would be randomly assigned to one program, which they would be unable to change. However, participants were unaware of whether the program to which they were assigned was “experimental” or “comparison.” Since the psychoeducation condition was designed to look like any other Happify program, it is unlikely that users unfamiliar with Happify would be able to determine whether they were in the comparison condition. In the first 30 days, all participants received emails twice per week reminding them to visit the site and complete their assigned program. They did not receive emails after that, except to prompt for follow-up assessment.

Happify condition. The Happify condition received full access to an online platform that offers techniques grounded in positive psychological interventions (PPI; Parks & Schueller, 2014), cognitive-behavioral therapy (CBT; Beck, 2011), and/or mindfulness-based stress reduction (MBSR; Kabat-Zinn, 2013; Carpenter et al., 2017). Each of the core activities on the platform was selected from one or more of these three theoretical traditions, but only if there existed randomized, controlled research demonstrating its efficacy for improving wellbeing. Those core activities are organized into five categories using the acronym STAGE: savor (savoring and mindfulness), thank (gratitude), aspire (goal-setting, reframing/optimism, and meaning), give (prosocial behavior, kindness, and forgiveness), and empathize (perspective-taking and self-compassion).

Some of these activities would be considered “traditional” exemplars of the literature they came from. Gratitude journals, acts of kindness, and using strengths are examples of activities on Happify that are canonical PPIs. Users can do body scan meditations, walking meditations, and sitting meditations very much in line with traditional MBSR. They can also practice thought stopping and thought disputing, as well as goal setting, all of which are staples of traditional CBT. However, because these three approaches overlap in many ways, we do not clearly label each activity as “belonging” to only one area; for example, savoring appears in both positive psychology and mindfulness-based approaches, and many positive psychology-based activities could also be housed under the “behavioral activation” side of CBT.

Happify users have free choice from a growing catalog of over sixty 4-week programs tailored to different goals, such as conquering negative thoughts, feeling more connected to one’s family, or finding meaning in one’s job. Together, these programs contain over 2,500 different activity variants. The design of each program is overseen by an author with expertise in the topic. The author selects a subset of the core activities that are appropriate to the goals of the program, and works with a writer to customize the selected activities for the specific nature of the program. For example, a track on living with chronic pain was overseen by Afton Hassett, a researcher at the University of Michigan who specializes in psychosocial interventions for pain. She sampled interventions from positive psychology (gratitude, kind acts, strengths-building, finding meaning), cognitive-behavioral therapy (goal-setting, thought stopping), and mindfulness-based stress reduction (savoring the present moment, sitting

meditation). Each of these appears in the track in a customized form that mentions pain and helps the user see the activity's relevance to chronic pain.

Psychoeducational comparison. In order to account for regression to the mean and expectation effects, it is important to evaluate active interventions against one or more comparison groups (Mohr et al., 2009). About half of online wellbeing intervention studies use waitlist control conditions, where users do not receive any aspects of the intervention being tested (Bolier et al., 2013a; Bolier et al., 2013b; Schueller & Parks, 2012). Our psychoeducational condition was designed to provide a more robust and more meaningful comparison group than a simple waitlist control. The psychoeducational comparison intervention emulates many existing websites that provide information about wellbeing, where a person might end up if they are searching online.

Psychoeducational participants logged in regularly to an identical looking Happify website or app, included an identical onboarding process, and were offered content that grew and changed over time. However, comparison group users did not experience some of the engagement elements from the main platform—they had no access to social forums, could not post their activity results publicly for others to see and comment on, and could not choose freely between different programs. Like a waitlist group, users do not receive activities to try, but unlike a waitlist group, they do engage with content about wellbeing, and they also have the experience of using the Happify platform. Users in the psychoeducational comparison group completed a series of polls on the topic of wellbeing (Haack, Parks, & Schueller, 2016). In these polls, they were asked survey questions on wellbeing topics, and then were given some social comparison data about where they stand in their opinions compared with other users, as well as information about why the wellbeing topic is important, including references to relevant scientific studies. For example, after responding to a poll question that asked participants to indicate how often they have a deep or meaningful conversation with someone, participants were shown psychoeducational information about the benefits of conversation, including that people who spend less time alone and more time talking to others tend to be happier (Diener & Seligman, 2002), and that the happiest people have much less small talk but many more meaningful conversations as the least happy people (Mehl, Vazire, Holleran, & Clark, 2010).

2.5 Other conditions

In the larger ongoing study from which this paper draws, Happify participants are randomized between three groups who receive different variations of the Happify platform. The purpose of this subdivision is to test theoretical questions outside the scope of this paper, which focuses on main study outcomes. For the purposes of the current paper, therefore, only one Happify-using condition is presented—one in which users have access to the full consumer platform.

2.6 Analytic strategy

We used a 2 (condition: Happify, Psychoeducation) \times 2 (usage: recommended, low) factorial design, controlling for baseline levels of the dependent variable, to examine post-test differences in depressive symptoms, anxiety symptoms, and resilience. Because our analyses included a moderator, our plan was to frame results in terms of the highest order significant effect; in other words, results would be reported in terms of the 2x2 split if the interaction was significant, and if it was not significant, results would be reported in terms of main effects (Aiken & West, 1991).

Due to the large amount of missing data (77.7%), we chose to focus the main analysis only on users for whom we had complete data. Results did not differ, however, when analyses were repeated using multiple imputation by chained equations (MICE) to impute missing data on all three outcomes using SPSS (Biering, Hjollund, & Frydenberg, 2015; Dziura, Post, Zhao, Fu, & Peduzzi, 2013; White, Kalaitzaki, & Thompson, 2011), we generated five imputations with the following predictors: baseline scores on all three outcomes, gender, age, employment status, and condition.

In addition to overall analyses of group differences, we evaluated “practical significance” in two ways. First, we compared rates of improvement in depression and anxiety by one or more severity categories between participants in the recommended usage Happify condition and participants in all other conditions combined. Second, to look at the overall improvement to a population that each intervention offered, we computed a “net improvement” value for each condition; by subtracting the percentage of users who deteriorated from the percentage of users who improved, we can look at the overall improvement to a population that each intervention offered.

3. Results

3.1 General summary

Overall, users in the Happify condition who used the platform at the recommended rate reported fewer depressive and anxiety symptoms, and greater resilience than Happify participants who used the platform at a low rate or participants in the psychoeducation condition, regardless of usage. Baseline and post-test raw means and standard deviations, as well as effect sizes for the baseline to post-test differences, are presented in Table 6 below. Post-test marginal means, standard errors, and 95% confidence intervals, based on the ANCOVA models for each outcome are presented in Table 7 below. Contrast mean difference estimates, standard errors, significance levels, and 95% confidence intervals are shown in Table 8 below.

3.2 Depressive symptoms

On PHQ, the overall ANCOVA revealed a significant interaction between condition and usage, $F(1,1046) = 6.19, p < .013, \eta_p^2 = .02$, a small effect (Cohen, 1988; Miles & Shevlin, 2001). Simple main effects analyses revealed a significant effect of usage in the Happify condition, such that recommended usage participants had lower depression at post-test ($F(1,1046) = 15.23, p < .001$), but no such effect in the psychoeducation condition ($F(1,1046) = 0.03, p = .87$). Moreover, pairwise contrasts on post-test PHQ scores indicated that depressive symptoms for participants in the recommended usage Happify condition were significantly lower than in the low usage Happify condition ($t(1047) = 3.90, p < .001$) or the recommended usage psychoeducation condition ($t(1047) = 3.41, p < .001$). Depressive symptoms in the recommended usage Happify condition were marginally lower than in the low usage psychoeducation condition ($t(1047) = 1.93, p = .053$). All other pairwise contrasts were non-significant, $ps > .27$.

3.3 Anxiety symptoms

On the GAD, the overall ANCOVA indicated a significant interaction between condition and usage, $F(1,1046) = 4.29, p < .039, \eta_p^2 = .004$. Simple main effects analyses indicated that there was a significant effect of usage in the Happify condition ($F(1,1046) = 16.74, p < .001$), but not in the psychoeducation condition ($F(1,1046) = 0.25, p = .62$). Follow-up pairwise contrasts on post-test GAD scores indicated that anxiety symptoms for participants in the recommended usage

Happify condition were significantly lower than in the low usage Happify condition ($t(1047) = 4.09, p < .001$), the low usage psychoeducation condition ($t(1047) = 2.10, p < .036$), and the recommended usage psychoeducation condition ($t(1047) = 2.67, p < .008$). Additional contrasts indicated that anxiety symptoms in the recommended usage psychoeducation condition were significantly lower than in the low usage Happify condition, but that pairwise contrasts comparing anxiety symptoms in the low usage waitlist condition to the low usage Happify condition ($t(1047) = 0.78, p < .44$) and the recommended usage psychoeducation condition ($t(1047) = 0.50, p < .62$) were both non-significant.

Table 6. Baseline and post-test means, standard deviations, percent change from baseline, and effect sizes by condition and usage for depressive symptoms, anxiety symptoms, and resilience

Outcome	Cond	Usage	Baseline		Post-test		% Change from Base	Cohen's <i>d</i>
			Mean	SD	Mean	SD		
Depression	H	LU	12.3	6.9	10.9	6.7	8.6%	0.20
		RU	12.3	6.3	9.2	6.1	25.2%	0.50
	PE	LU	12.9	7.1	10.8	6.9	13.6%	0.30
		RU	11.4	6.4	9.9	6.3	15.0%	0.23
Anxiety	H	LU	10.3	6.2	9.4	6.0	11.2%	0.15
		RU	10.8	5.6	8.1	5.3	24.9%	0.50
	PE	LU	10.6	6.1	9.2	6.2	16.1%	0.24
		RU	10.1	5.6	8.6	5.6	13.0%	0.27
Resilience	H	LU	43.5	17.8	47.1	18.4	8.4%	0.20
		RU	41.6	16.0	50.2	16.3	20.8%	0.54
	PE	LU	42.4	17.2	46.7	17.7	10.2%	0.25
		RU	42.2	16.5	47.5	16.9	12.7%	0.32

Note: H = Happify, PE = psychoeducation, LU = low usage, RU = recommended usage. Cohen's *d* represents the amount of change from baseline to post-test for each of the condition and usage groups.

3.4 Resilience

On the resilience composite, the overall ANCOVA indicated a significant interaction between condition and usage, $F(1,1046) = 3.85, p = .05, \eta_p^2 = .004^1$. Simple main effects analyses revealed a

¹ The resilience composite reported here includes the emotion subscale from the Happify scale, which, along with the life satisfaction subscale of the Happify scale, was used to measure well-being in the Carpenter et al. (2017) article that preceded the current study. Although the full Happify scale was not an outcome of interest in the current study, we find a pattern of results that mirrors those reported above when examining the interactive effects of

significant effect of usage in the Happify condition, such that recommended usage participants had lower depression at post-test ($F(1,1046) = 17.32, p < .001$), but no such effect in the psychoeducation condition ($F(1,1046) = 0.48, p = .49$). Follow-up pairwise contrasts on post-test GAD scores indicated that anxiety symptoms for participants in the recommended usage Happify condition were significantly lower than in the low usage Happify condition ($t(1047) = 4.16, p < .001$), the low usage psychoeducation condition ($t(1047) = 2.67, p < .008$), and the recommended usage psychoeducation condition ($t(1047) = 3.30, p < .001$). All other pairwise contrasts were non-significant, $ps > .13$.

Table 7. Post-test marginal means by condition and usage and usage contrasts for depressive symptoms, anxiety symptoms, and resilience

Outcome	Condition	Usage		LU v. RU contrast
		LU	RU	
Depression	Happify	10.7 (.3)	9.0 (.3)	$F(1,1046) = 15.23, p < .001$
	Psychoeducation	10.2 (.5)	10.3 (.2)	$F(1,1046) = 0.03, p = .87$
Anxiety	Happify	9.4 (.3)	7.8 (.3)	$F(1,1046) = 16.74, p < .001$
	Psychoeducation	9.0 (.5)	8.7 (.5)	$F(1,1046) = 0.25, p = .62$
Resilience	Happify	46.3 (.8)	50.8 (.8)	$F(1,1046) = 17.32, p < .001$
	Psychoeducation	46.7 (1.3)	47.7 (.5)	$F(1,1046) = 0.48, p = .49$

Note: LU = low usage, RU = recommended usage. Standard errors are included in parentheses. Marginal means are based on ANCOVA models that include condition and usage as factors and baseline scores (on the PHQ, GAD, and resilience index) as covariates. The PHQ is scored such that higher scores indicate more depressive symptoms. The GAD is scored such that higher scores indicate more anxiety symptoms. The resilience index is scored such that higher scores indicate greater resilience.

Changes in depressive and anxiety symptoms from baseline to the 8-week post-test for users in the Happify and psychoeducation conditions are displayed graphically by usage level in Figure 2 below. Participants who used Happify at the recommended level demonstrated significantly larger reductions in depressive symptoms (2.7 points) compared to low usage Happify participants (0.9 points, $p < .001$) and recommended usage psychoeducation participants (1.5 points; $p < .001$). Depressive symptom reduction did not differ significantly between low usage psychoeducation participants and recommended usage Happify participants ($p = .14$) or between low usage Happify participants and either low usage psychoeducation participants ($p = .31$) or recommended usage psychoeducation participants ($p = .80$). Similarly, Happify users demonstrated significantly larger reductions in anxiety symptoms (3.1 points) compared to low

condition and usage on the Happify life satisfaction subscale as well as the full Happify scale. In both cases, there is a significant interaction between condition and usage (Life satisfaction: $F(1,1046) = 4.51, p < .033$; Happify scale: $F(1,1046) = 5.34, p < .021$), such that participants showed the greatest increases in life satisfaction (and wellbeing, measured by the full Happify scale) when using Happify at the recommended level, compared to the three other groups.

usage Happify participants (1.4 points, $p < .001$), low usage psychoeducation participants (2.1, $p < .05$) and recommended usage psychoeducation participants (1.5 points; $p < .002$). Anxiety symptom reduction did not differ significantly between low usage Happify participants and low usage psychoeducation participants ($p = .37$) or recommended usage psychoeducation participants ($p = .09$), or between the two psychoeducation conditions ($p = .91$).

Table 8. Contrast estimates for recommended usage Happify versus other conditions

Outcome	Contrast	Mean Difference	SE	<i>t</i>	<i>p</i>	95% CI
Depression	H-RU vs. H-LU	-1.7	.4	3.90	< .001	-2.5, -.8
	H-RU vs. PE-LU	-1.2	.6	1.93	.053	-2.4, .0
	H-RU vs. PE-RU	-1.3	.4	3.41	< .001	-2.0, -.6
Anxiety	H-RU vs. H-LU	-1.6	.4	4.09	< .001	-2.4, -.9
	H-RU vs. PE-LU	-1.2	.6	2.10	.036	-2.3, .1
	H-RU vs. PE-RU	-.9	.4	2.67	.008	-1.6, -.3
Resilience	H-RU vs. H-LU	4.5	1.1	4.16	< .001	2.4, 6.7
	H-RU vs. PE-LU	4.1	1.6	2.67	.008	1.1, 7.2
	H-RU vs. PE-RU	3.1	1.0	3.30	< .001	1.3, 5.0

Note: H-RU = Happify, recommended usage, H-LU = Happify, low usage; PE-LU = psychoeducation, low usage; PE-RU = psychoeducation, recommended usage. PHQ (GAD) is scored such that higher values indicate more severe depressive (anxiety) symptoms; a negative mean difference indicates that the recommended usage Happify condition had a greater reduction in depressive (anxiety) symptoms. Resilience is scored such that higher values indicate more resilience; a positive mean difference indicates that the recommended usage Happify condition had a greater increase in resilience.

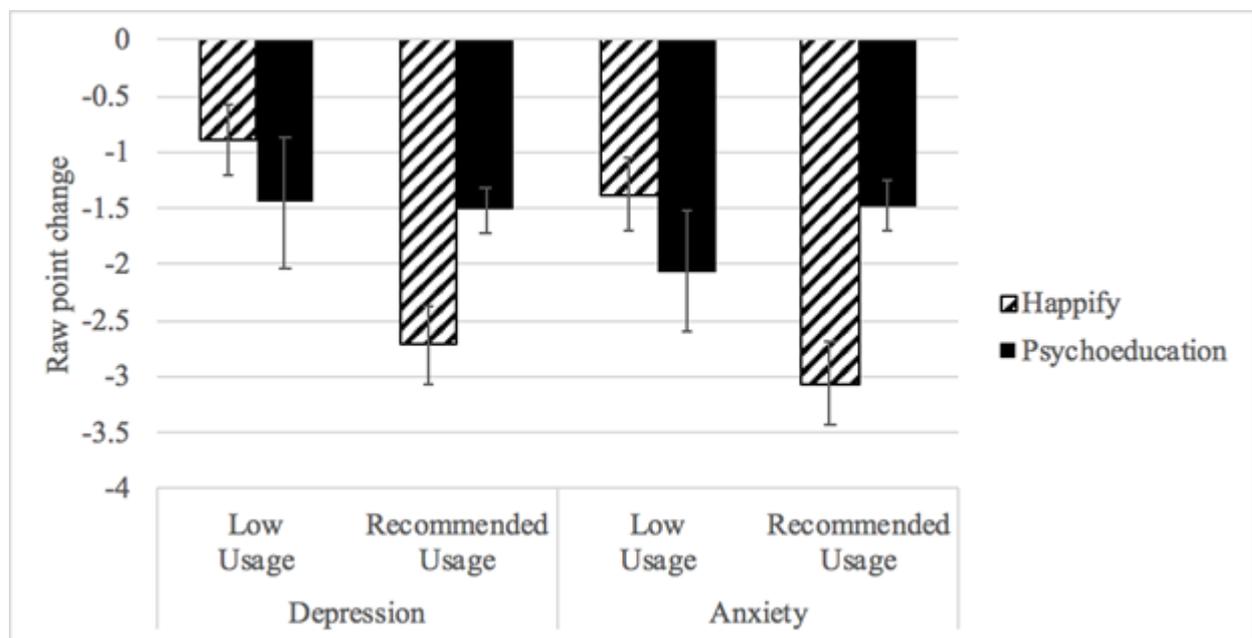
Changes in resilience from baseline to the 8-week post-test for participants, by condition and usage, are shown in Figure 3 below. Recommended usage Happify participants demonstrated significantly larger improvements in resilience (8.7 points) compared to low usage Happify participants (3.7 points; $p < .001$), low usage psychoeducation participants (4.3 points; $p < .009$), or recommended usage psychoeducation participants (5.4 points; $p < .001$). All other conditions did not differ from each other in terms of change from baseline to post-test.

3.5 Practical Significance

The percentages of users in each group who improved, did not change, or deteriorated by at least one clinical symptom category (e.g., changing from moderate to mild) on the PHQ scale and the GAD scale are shown separately for recommended usage Happify participants and participants in all other conditions combined in Figure 4 below. In terms of depressive symptoms, a larger proportion of recommended usage Happify users showed improvement (49.6%) than did participants in all other conditions (37.5%), and a smaller

proportion of recommended usage Happify participants showed deterioration (14.4%) compared to participants in all other conditions (18.3%). The proportions of improvement, no change, and deterioration differed significantly between recommended usage Happify participants and participants in all other conditions, $X^2(2) = 10.59, p < .005$. In terms of anxiety symptoms, a larger proportion of recommended usage Happify users showed improvement (46.9%) than did participants in all other conditions (34.9%), and a smaller proportion of recommended usage Happify participants showed deterioration (10.4%) compared to participants in all other conditions (18.5%). The proportions of improvement, no change, and deterioration differed significantly between recommended usage Happify participants and participants in all other conditions, $X^2(2) = 14.16, p < .001$.

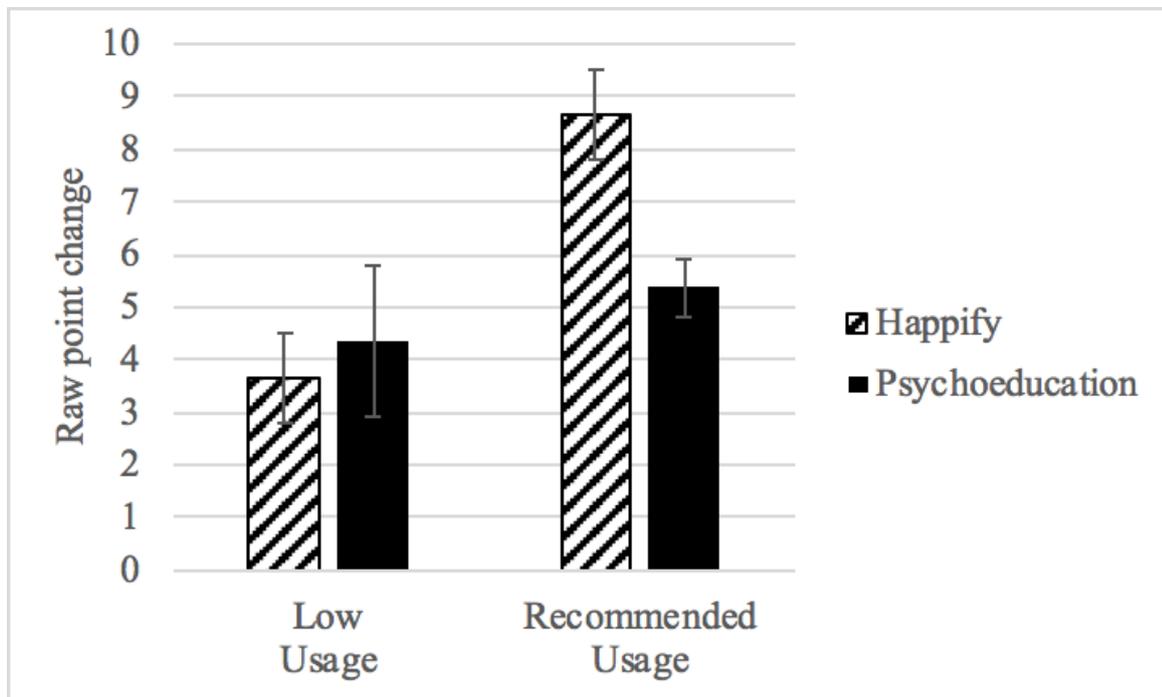
Figure 2. Changes in depressive and anxiety symptoms from baseline to 8 weeks by condition and usage



Note: LU = low usage, RU = recommended usage. The y-axis in this figure represents raw point change from baseline to 8 weeks. Error bars for each condition indicate standard errors of the mean change scores. Depression ranges from 0-27 while anxiety ranges from 0-21.

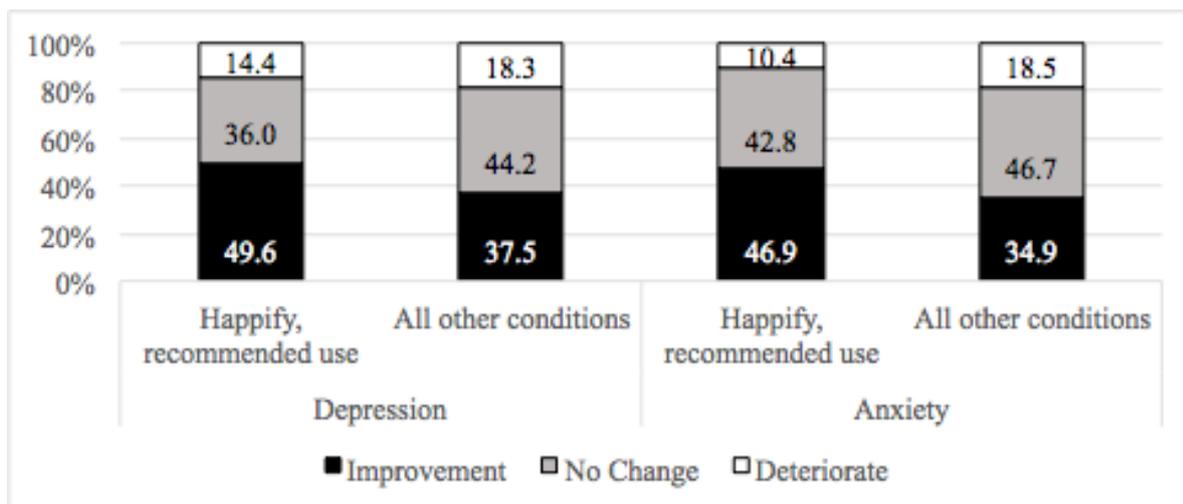
In order to quantify the overall quality of Happify when used at the recommended level, as compared with the other conditions in the study (i.e., low usage Happify, low usage psychoeducation, and recommended usage psychoeducation), we calculated a “net benefit” score – the percentage of users who improved minus the percentage of users who deteriorated. This score allows us to estimate overall intervention quality, taking into account not just average improvement, but the extent to which the number of users experiencing improvement outweighs the number experiencing deterioration. On net benefit, Happify had a net benefit to depression of 35.1% of users, while all other conditions together had a net benefit of 19.2%. Happify had a net benefit to anxiety for 36.5% of users, while all other conditions together had a net benefit of 16.4%.

Figure 3. Change in resilience index from baseline to 8 weeks for Happify, psychoeducational comparison, and no-usage control conditions



Note: The y-axis in this figure represents change from baseline to 8 weeks. Error bars for each condition indicate standard errors of mean change scores. Resilience ranges from 0-100.

Figure 4. Percentages of participants who experienced PHQ and GAD category improvement, no change, or clinical symptom severity category deterioration over 8 weeks by condition



4. Discussion

Building on previous research finding Happify’s effectiveness in an uncontrolled setting, this study shows more convincingly that Happify improves wellbeing (as measured by indicators of depression, anxiety, and resilience), particularly among those who adhere to recommended platform usage guidelines. All hypotheses were supported by the results,

which suggest that the wellbeing techniques made available in the Happify platform survived translation to and remain effective in a less structured context characterized by substantial variation in user engagement and usage. They also suggest that the inclusion of user engagement elements, combined with research-tested activities, lends an advantage above and beyond simple exposure to material on wellbeing (psychoeducation comparison) or the improvement that one would experience naturally over time (low usage sub-groups). Psychoeducation can bring wellbeing concepts to mind, and may even lead individuals towards activities they could do to promote their wellbeing (YouTube videos, for example, may suggest a behavior to try). However, an engaging platform that contains interactive wellbeing activities appears to be, across the board, about twice as powerful as psychoeducation. In short, these findings suggest that although being exposed to wellbeing concepts can be more helpful than nothing, there is an additive benefit to actively teaching and practicing hands-on activities, reinforcing them via conversations with others in forums, and bringing them to life in interactive games. These findings represent an important step towards validating the improvement of wellbeing in real-world settings, on a platform that is scalable and engaging for users.

4.1 Dosage level as a moderator

The analyses in this study distinguished between participants who used their assigned platform as instructed—two to three times per week over the course of the 8-week study—and those who did not. Many clinical trials follow an intent-to-treat approach, including all participants, regardless of participation and adherence, because efficacy is considered to be a combination of effectiveness and palatability (i.e., the most effective intervention cannot be effective if nobody wants to use it). However, since our primary research question revolved around not efficacy but effectiveness, we added usage level as a moderator. We believe that this analytic approach is the best of both worlds, assessing both the overall efficacy of an intervention, while also determining its effectiveness in ideal conditions.

Exploring the impact of Happify on users who used it as recommended allows us to test whether the research-based activities implemented in Happify remained effective in that new format. In the current study, where our emphasis was on simple *effectiveness* in a new environment, our goal was to generate effect sizes and compare them with those of the previous studies from which the content of Happify was adapted. In order to do that, it was important to equalize analytic methodology as much as possible. Many of the studies in the wellbeing literature are published as experiments, not as “clinical trials,” and therefore do not follow clinical trial conventions such as CONSORT reporting or intent-to-treat analysis (see the meta-analysis of Sin & Lyubomirsky (2009) for a broad review; all claims about the wellbeing literature are based on our reading of this meta-analysis). Furthermore, many of these studies were performed in more controlled environments, such as college students completing a study for credit, with credit contingent on participation; the percentage of the sample exhibiting recommended usage patterns would be much higher in an in-person experimental context than in a naturalistic consumer setting. We therefore adapted our analytic strategy to most closely mimic the studies we hoped to make comparisons with—which means the ability to zero in on users who use Happify as recommended. We would argue that comparing overall engagement rates of an online intervention like Happify with an in-person or heavily monitored intervention would be like comparing apples and oranges, as users are much more free to vary in an unstructured clinical trial such as this one. At the same time, our analysis also allows readers to view the impact of low usage by

comparing those who used as recommended with those who used at a lower rate. These data are also valuable in helping organizations understand the importance of encouraging employees, clients, or patients to engage when trying to implement a large-scale intervention.

Second, splitting participants by usage level arguably provides very interesting additional information. By separating out users whose usage was low, we were able to estimate the natural change in wellbeing that can be expected over time. With an estimate of natural change over time, we are able to assess the extent to which the two intervention conditions improved above and beyond the improvement they would have experienced regardless. As many existing (and popular) internet interventions take a psychoeducational approach, providing users with information but not interactive activities, it is useful to have an additional group to see to what extent psychoeducation is helping above and beyond the natural improvement expected over time. While it would have been possible to randomly assign users to a no intervention group, we chose not to—in short, among people very interested in becoming happier, being assigned to do nothing can be problematic, and sometimes deleterious (Mauss, Tamir, Anderson, & Savino, 2011). The moderator split, therefore, allows us to see the wellbeing changes that occur in users who receive a minimal dose of intervention without risking the negative consequences of a no intervention group assignment in this eager population. While the self-selected nature of the low usage participants is a limitation, we felt that limitation was more acceptable than the alternative. We consider this to be a strength of the current study—we have suggested a way to explore the impact of receiving a low intervention dose compared with the as-recommended usage level in a context where doing so might normally be difficult or unethical. That said, we acknowledge that further research is needed to tease apart some of our assumptions about users who were assigned an intervention but did not use it—are they still affected by placebo effect and positive expectation from having enrolled in the study, despite having not actively participated in their assigned intervention (in other words, does this group represent more than natural improvement over time)? How might the results from low usage users differ from or resemble a true waitlist control group, where users do not pursue other treatments? We are eager to continue exploring these questions in future work to assess the validity of this approach.

4.2 Limitations

We believe that the current findings are an interesting first step towards implementing wellbeing interventions in a way that is scalable, without losing the ability to use rigorous research designs. However, we also acknowledge several limitations. This study is just the first in a series of studies that will help to evaluate the ability of an engaging, online platform to improve wellbeing, and future research can help to explore these remaining issues.

4.2.1 Generalizability

This study is specifically focused on users who arrived at Happify naturally—these users are self-selected, and may be different from the general population in some ways that we did not measure. While we do believe our sample was representative of internet wellbeing seekers, it may be harder to generalize to different types of users in other contexts. The current study was conducted in a consumer population, but large-scale implementation efforts often happen in workplaces, among health care payers, or in other contexts where

the users may be quite different. Follow-up research can replicate these findings, as well as exploring other outcomes, in additional samples and contexts. To enhance our ability to study generalizability in the future, we recently launched localized, translated versions of the platform in Spanish, French, Portuguese, German, Chinese, Japanese, and French-Canadian.

4.2.2 Limitation to self-reported, psychological outcomes

The current study reports only on psychological outcomes—depression, anxiety, and resilience. However, there are many other outcomes that a scalable, online wellbeing intervention might impact, and future research can explore this. Applied in populations with chronic health conditions, in primary care settings, or in workplace settings, could the improvements in depressive symptoms that Happify produces, in turn, affect some of these other outcomes that are tied to health care expenses, company profits, or burnout rates?

We further acknowledge a reliance on self-report that is commonplace in the study of wellbeing interventions, but nevertheless suboptimal. With wearable technology and smartphones, it is easier than ever to obtain in-the-moment data about mood, as well as objective data about the exercise, sleep, and even stress levels (using voice analysis) experienced by users. Objective measurement is especially important for a construct like resilience, where self-reported measurement, while the norm in research, can be unreliable at predicting actual resilience during a stressful situation (Razavi, 2001; Schenker & Gentleman, 2001). One objective measure that has particular relevance for resilience is heart rate variability (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012). We are currently in the process of lab validating our smartphone-based relaxation game, *Breather*, which measures heart rate variability. However, there are numerous other laboratory tasks with outcomes tied to resilience that could be implemented as technology-based tests, and we hope to see these developed soon to better the assessment of resilience in future research.

Objective measurement approaches can also provide valuable insight into the mechanisms that underlie wellbeing improvement; cognitive tasks, for example, can allow researchers to assess attentional and/or memory biases towards positive vs. negative stimuli, which may help us to understand what sorts of cognitive shifts take place as users engage with wellbeing promoting activities. Similarly, analysis of text written by users in their social media accounts (e.g., Carpenter et al., 2017) can help us understand the ways that users are thinking about and processing positive and negative emotions and events—and changes in these patterns may help us understand *how* users of wellbeing interventions may change in this regard. Very little is known about the mechanisms by which wellbeing interventions lead to improvements in wellbeing; further research emphasizing objective outcomes could help shed some light on the question.

5. Conclusions

The present study shows, using a randomized, controlled design, that Happify not only promotes wellbeing in users who adhere to recommended usage guidelines, it does so at a level that is substantially higher than those users may have been able to generate on their own. However, as noted above, there are numerous research questions that still remain to be examined. In particular, the rich datasets yielded by internet interventions and the agility afforded by delivering interventions using technology leave room for many research questions above and beyond simple effectiveness.

For example, it would be interesting to zero in on specific components of Happify, such as the games or the forums, and explore their role in improving outcomes, and how they can best be used to enhance user wellbeing. On a related note, internal data analyses suggest that regular use at a rate of two to three times per week yields optimal outcomes compared to other frequency levels. However, Carpenter et al. (2017) found that users experience higher wellbeing during 2-week periods that they use Happify more, and lower wellbeing during 2-week periods that they use Happify less, suggesting that users can boost their wellbeing with even short-term use as needed. More work exploring different types of usage patterns could shed some light on this question. Beyond that, it would be interesting to explore the impact of different person-level variables, such as gender, baseline distress level, or age, on typical usage patterns, and on which usage patterns work best for a given user.

Additionally, it would be valuable to look more specifically at the mechanisms through which Happify is improving wellbeing. Given that Happify targets five different aspects of wellbeing, are different activities working through different mechanisms? For example, are empathize activities improving wellbeing through increasing empathy, while savor activities target mindfulness? Similarly, Happify contains activities that draw from CBT, mindfulness, and positive psychology. Do each of these approaches affect different aspects of wellbeing? Or are all the activities impacting some larger factor, such as attention to positive stimuli? Further work can help as to questions beyond whether Happify works—*how* does it work, and *when* does it work best?

Lastly, as technology evolves over time, new features will be added to technology-based interventions, and internet intervention science will need to evolve as well to accommodate them. For example, with artificial intelligence tools and machine learning knowledge becoming more and more accessible, it is only a matter of time before internet interventions begin to provide highly customized programs for users. What is the best way to study the impact of deep customization when best practices for research rely heavily on the standardization of user experience? With more widespread distribution of smartphones, research in real-world environments is becoming more and more possible. But what is the best way to handle missing data in a consumer marketplace—no doubt, a research environment that is guaranteed to yield high rates of missingness? Beyond that, as reach increases, structure becomes more difficult to uphold—so how can intervention science produce high-quality research in contexts where very little control over users is possible? When is it not feasible to randomly assign some users to a control group? These questions, and many others, remain open for further exploration in coming years.

Authors

Acacia C. Parks
Happify
Hiram College
acacia@happify.com

Allison L. Williams
Happify
University of Minnesota

Michele M. Tugade
Vassar College

Kara E. Hokes
Hiram College

Ryan D. Honomichl
Case Western Reserve University

Ran D. Zilca
Happify

Publishing Timeline

Received dd month year

Accepted dd month year

Published dd month year

References

- Ahmedani, B. K., Belville-Robertson, T., Hirsch, A., & Jurayj, A. (2016). An online mental health and wellness intervention supplementing standard care of depression and anxiety. *Archives of Psychiatric Nursing, 30*(6), 666-670. <https://doi.org/10.1016/j.apnu.2016.03.003>
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Beck, J. S. (2011). *Cognitive behavioral therapy: Basics and beyond* (2nd ed.). New York, NY: Guilford Press.
- Biering, K., Hjollund, N. H., & Frydenberg, M. (2015). Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clinical Epidemiology, 7*, 91-106. <https://doi.org/10.2147/CLEP.S72247>
- Block, J., & Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology, 70*(2), 349-361. <https://doi.org/10.1037/0022-3514.70.2.349>
- Bolier, L., Haverman, M., Kramer, J., Westerhof, G. J., Riper, H., Walburg, J. A., . . . & Bohlmeijer, E. (2013a). An internet-based intervention to promote mental fitness for mildly depressed adults: Randomized controlled trial. *Journal of Medical Internet Research, 15*(9), e200. <https://doi.org/10.2196/jmir.2603>
- Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013b). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health, 13*, 119-139. <https://doi.org/10.1186/1471-2458-13-119>
- Bush, N. E., Bosmajian, C. P., Fairall, J. M., McCann, R. A., & Ciulla, R. P. (2011). Afterdeployment.org: A web-based multimedia wellness resource for the postdeployment military community. *Professional Psychology: Research and Practice, 42*(6), 455-462. <https://doi.org/10.1037/a0025038>
- Carpenter, J., Crutchley, P., Zilca, R. D., Schwartz, H. A., Smith, L. K., Cobb, A. M., & Parks, A. C. (2017). Correction: Seeing the "big" picture: Big data methods for exploring relationships between usage, language, and outcome in internet intervention data. *Journal of Medical Internet Research, 19*(12), e347. <https://doi.org/10.2196/jmir.8099>
- Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Sciences, 18*(6), 293-299. <https://doi.org/10.1016/j.tics.2014.02.003>
- Carver C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review, 30*(7), 879-889. <https://doi.org/10.1016/j.cpr.2010.01.006>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, S., & Janicki-Deverts, D. (2012). Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. *Journal of Applied Social Psychology, 42*(6), 1320-1334. <https://doi.org/10.1111/j.1559-1816.2012.00900.x>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385-396. <https://doi.org/10.2307/2136404>

- Cohn, M. A., Pietrucha, M. E., Saslow, L. R., Hult, J. R., & Moskowitz, J. T. (2014). An online positive affect skills intervention reduces depression in adults with type 2 diabetes. *The Journal of Positive Psychology, 9*(6), 523-534. <https://doi.org/10.1080/17439760.2014.920410>
- Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18*(2), 76-82. <https://doi.org/10.1002/da.10113>
- Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science, 13*(1), 81-84. <https://doi.org/10.1111/1467-9280.00415>
- Dziura, J. D., Post, L. A., Zhao, Q., Fu, Z., & Peduzzi, P. (2013). Strategies for dealing with missing data in clinical trials: From design to analysis. *Yale Journal of Biology and Medicine, 86*(3), 343-358. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3767219/>
- Ezzati, A., Jiang, J., Katz, M. J., Sliwinski, M. J., Zimmerman, M. E., & Lipton, R. B. (2014). Validation of the Perceived Stress Scale in a community sample of older adults. *International Journal of Geriatric Psychiatry, 29*(6), 645-652. <https://doi.org/10.1002/gps.4049>
- Ford, I. W., Eklund, R. C., & Gordon, S. (2000). An examination of psychosocial variables moderating the relationship between life stress and injury time-loss among athletes of a high standard. *Journal of Sports Sciences, 18*(5), 301-312. <https://doi.org/10.1080/026404100402368>
- Griffin, K. H., Johnson, J. R., Kitzmann, J. P., Kolste, A. K., & Dusek, J. A. (2015). Outcomes of a multimodal resilience training program in an outpatient integrative medicine clinic. *The Journal of Alternative and Complementary Medicine, 21*(10), 628-637. <https://doi.org/10.1089/acm.2015.0130>
- Haeck, C., Parks, A. C., & Schueller, S. M. (2016). Impact of psychoeducational content delivered online to a positive psychology aware community. *The Journal of Positive Psychology, 11*(3), 270-275. <https://doi.org/10.1080/17439760.2015.1048819>
- Hintz, S., Frazier, P. A., & Meredith, L. (2015). Evaluating an online stress management intervention for college students. *Journal of Counseling Psychology, 62*(2), 137-147. <https://doi.org/10.1037/cou0000014>
- Hoorelbeke, K., & Koster, E. H. W. (2017). Internet-delivered cognitive control training as a preventive intervention for remitted depressed patients: Evidence from a double-blind randomized controlled trial study. *Journal of Consulting and Clinical Psychology, 85*(2), 135-146. <https://doi.org/10.1037/ccp0000128>
- Kabat-Zinn, J. (2013). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness* (revised ed.). New York, NY: Bantam Books.
- Kalisch, R., Baker, D. G., Basten, U., Boks, M. P., Bonanno, G. A., Brummelmann, E., . . . Kleim, B. (2017). The resilience framework as a strategy to combat stress-related disorders. *Nature Human Behavior, 1*, 784-790. <https://doi.org/10.1038/s41562-017-0200-8>
- Kizakevich, P. N., Hubal, R., Brown, J., Lyden, J., Spira, J., Eckhoff, R., . . . Munoz, G. (2012). PHIT for duty, a mobile approach for psychological health intervention. *Annual Review of Cybertherapy and Telemedicine, 10*, 268-272. <https://doi.org/10.3233/978-1-61499-121-2-268>
- Kleiman, E. M., Chiara, A. M., Liu, R. T., Jager-Hyman, S. G., Choi, J. Y., & Alloy, L. B. (2017). Optimism and well-being: A prospective multi-method and multi-dimensional examination of optimism as a resilience factor following the occurrence of stressful life events. *Cognition and Emotion, 31*(2), 269-283. <https://doi.org/10.1080/02699931.2015.1108284>
- Klohn, E. C. (1996). Conceptual analysis and measurement of the construct of ego-resiliency. *Journal of Personality and Social Psychology, 70*(5), 1067-1079. <https://doi.org/10.1037/0022-3514.70.5.1067>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lamers, S. M. A., Bolier, L., Westerhof, G. J., Smit, F., & Bohlmeijer, E. T. (2012). The impact of emotional well-being on long-term recovery and survival in physical illness: A meta-analysis. *Journal of Behavioral Medicine, 35*(5), 538-547. <https://doi.org/10.1007/s10865-011-9379-8>
- Lee, J-S., Ahn, Y-S., Jeong, K-S., Chae, J-H., & Choi, K-S. (2014). Resilience buffers the impact of traumatic events on the development of PTSD symptoms in firefighters. *Journal of Affective Disorders, 162*, 128-133. <https://doi.org/10.1016/j.jad.2014.02.031>

- Leykin, Y., Aguilera, A., Torres, L. D., Pérez-Stable, E. J., & Muñoz, R. F. (2012). Interpreting the outcomes of automated Internet-based randomized trials: Example of an international smoking cessation study. *Journal of Medical Internet Research*, *14*(1), 102-114. <https://doi.org/10.2196/jmir.1829>
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder screener (GAD-7) in the general population. *Medical Care*, *46*(3), 266-274. <https://doi.org/10.1097/MLR.0b013e318160d093>
- Mauss, I. B., Tamir, M., Anderson, C. L., & Savino, N. S. (2011). Can seeking happiness make people unhappy? Paradoxical effects of valuing happiness. *Emotion*, *11*(4), 807-815. <https://doi.org/10.1037/a0022010>
- McHale, N., Clark, D. A., & Tramonte, L. (2015). Does optimism moderate mood repair? A daily diary study. *Motivation and Emotion*, *39*(3), 409-419. <https://doi.org/10.1007/s11031-014-9464-8>
- Mehl, M. R., Vazire, S., Holleran, S. E., & Clark, C. S. (2010). Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations. *Psychological Science*, *21*(4), 539-541. <https://dx.doi.org/10.1177%2F0956797610362675>
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London, England: Sage.
- Mohr, D. C., Cheung, K., Schueller, S. M., Brown, C. H., & Duan, N. (2013). Continuous evaluation of evolving behavioral intervention technologies. *American Journal of Preventive Medicine*, *45*(4), 517-523. <https://doi.org/10.1016/j.amepre.2013.06.006>
- Mohr, D. C., Schueller, S. M., Riley, W. T., Brown, C. H., Cuijpers, P., Duan, N., . . . & Cheung, K. (2015). Trials of intervention principles: Evaluation methods for evolving behavioral intervention technologies. *Journal of Medical Internet Research*, *17*(7), e166. <https://doi.org/10.2196/jmir.4391>
- Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., . . . & Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, *78*(5), 275-284. <https://doi.org/10.1159/000228248>
- Mongrain, M., & Anselmo-Matthews, T. (2012). Do positive psychology exercises work? A replication of Seligman et al. (2005). *Journal of Clinical Psychology*, *68*(4), 382-389. <https://doi.org/10.1002/jclp.21839>
- Moskowitz, J. T., Folkman, S., & Acree, M. (2003). Do positive psychological states shed light on recovery from bereavement? Findings from a 3-year longitudinal study. *Death Studies*, *27*(6), 471-500. <https://doi.org/10.1080/07481180302885>
- Muñoz, R. F., Bunge, E. L., Chen, K., Schueller, S. M., Bravin, J. I., Shaughnessy, E. A., & Pérez-Stable, E. J. (2016). Massive open online interventions: A novel model for delivering behavioral health services worldwide. *Clinical Psychological Science*, *4*(2), 194-205. <https://doi.org/10.1177/2167702615583840>
- Newby, J., Robins, L., Wilhelm, K., Smith, J., Fletcher, T., Gillis, I., . . . & Andrews, G. (2017). Web-based cognitive behavior therapy for depression in people with diabetes mellitus: A randomized controlled trial. *Journal of Medical Internet Research*, *19*(5), e157. <https://doi.org/10.2196/jmir.7274>
- Ong, A. D., Bergeman, C. S., Bisconti, T. L., & Wallace, K. A. (2006). Psychological resilience, positive emotions, and successful adaptation to stress in later life. *Journal of Personality and Social Psychology*, *91*(4), 730-49. <https://doi.org/10.1037/0022-3514.91.4.730>
- Orr, J. A., King, R. J., Hawke, P., & Dalgleish, J. (2013). Can we text you? A qualitative exploration of young employed job-seekers' attitudes to receiving resilience-building SMS messages. *Advances in Mental Health*, *11*(3), 268-285. <https://doi.org/10.5172/jamh.2013.11.3.268>
- Parks, A. C., Della Porta, M. D., Pierce, R. S., Zilca, R., & Lyubomirsky, S. (2012). Pursuing happiness in everyday life: The characteristics and behaviors of online happiness seekers. *Emotion*, *12*(6), 1222-1234. <https://doi.org/10.1037/a0028587>
- Parks, A., & Schueller, S. (2014). *The Wiley-Blackwell handbook of positive psychological interventions*. West Sussex, UK: Wiley-Blackwell. <https://doi.org/10.1002/9781118315927>
- Rao, N., & Kemper, K. (2017). The feasibility and effectiveness of online guided imagery training for health professionals. *Journal of Evidence-Based Complementary & Alternative Medicine*, *22*(1), 54-58.
- Razavi, T. (2001). *Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research*. Southampton, UK: University of Southampton.

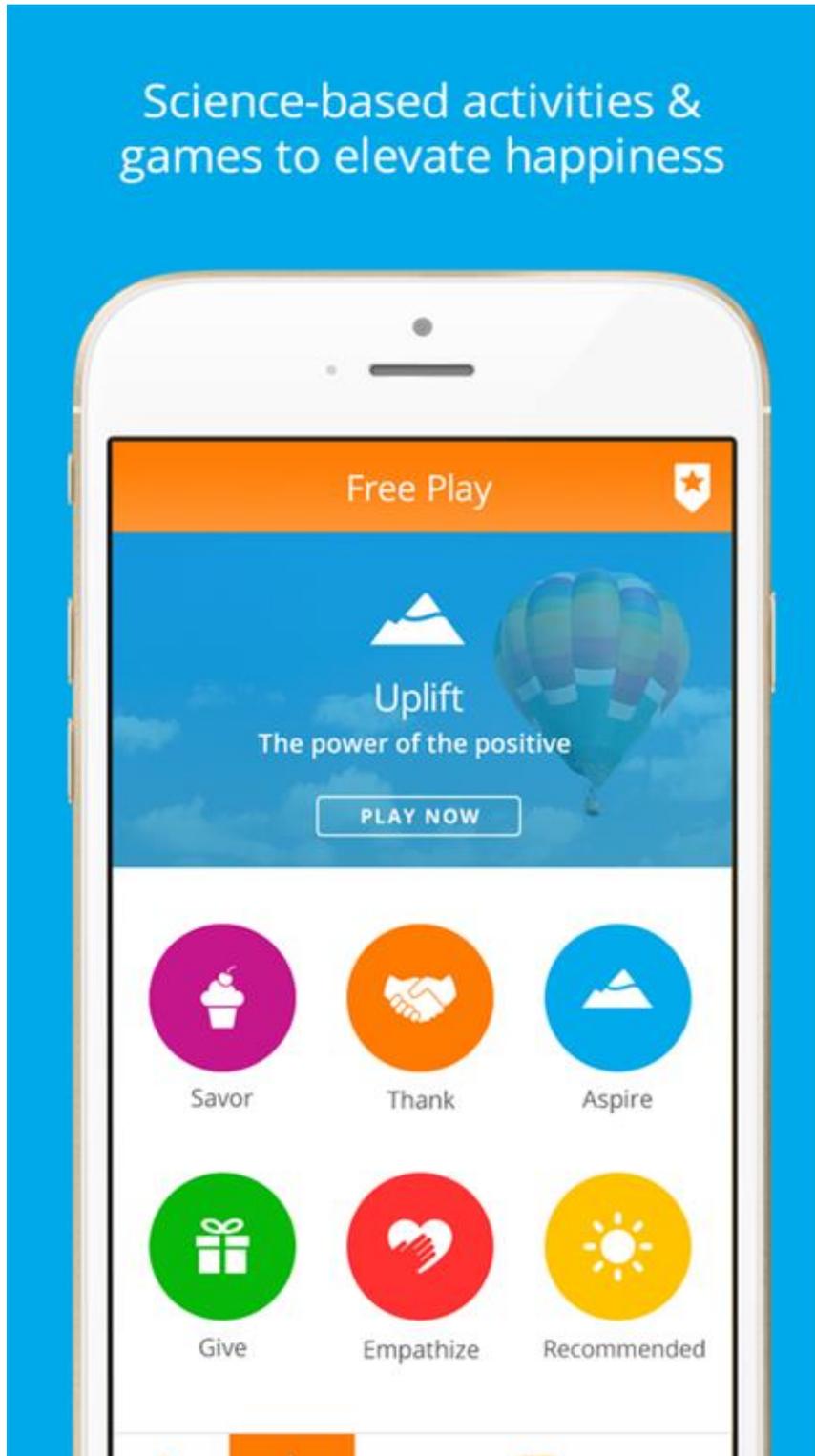
- Rose, R. D., Buckley, J. C., Jr., Zbozinek, T. D., Motivala, S. J., Glenn, D. E., Cartreine, J. A., & Craske, M. G. (2013). A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behaviour Research and Therapy*, *51*(2), 106-112. <https://doi.org/10.1016/j.brat.2012.11.003>
- Rossi, N. E., Bisconti, T. L., & Bergeman, C. S. (2007). The role of dispositional resilience in regaining life satisfaction after the loss of a spouse. *Death Studies*, *31*(10), 863-83. <https://doi.org/10.1080/07481180701603246>
- Roy, M. J., Costanzo, M. E., Highland, K. B., Olsen, C., Clayborne, D., & Law, W. (2017). An app a day keeps the doctor away: Guided education and training via smartphones in subthreshold post traumatic stress disorder. *Cyberpsychology, Behavior, and Social Networking*, *20*(8), 470-478. <https://doi.org/10.1089/cyber.2017.0221>
- Roy, M. J., Highland, K. B., & Costanzo, M. A. (2015). GETSmart: Guided education and training via smart phones to promote resilience. *Annual Review of CyberTherapy and Telemedicine*, *13*, 123-128. <https://doi.org/10.3233/978-1-61499-595-1-123>
- Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences*, *1094*(1), 1-12. <https://doi.org/10.1196/annals.1376.002>
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, *4*(3), 219-247. <https://doi.org/10.1037/0278-6133.4.3.219>
- Schenger, N., & Gentleman, J. F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, *55*(3), 182-186. <https://doi.org/10.1198/000313001317097960>
- Schotanus-Dijkstra, M., Drossaert, C. H. C., Pieterse, M. E., Boon, B., Walburg, J. A., & Bohlmeijer, E. T. (2017). An early intervention to promote well-being and flourishing and reduce anxiety and depression: A randomized controlled trial. *Internet Interventions*, *9*, 15-24. <https://doi.org/10.1016/j.invent.2017.04.002>
- Schueller, S. M., & Parks, A. C. (2012). Disseminating self-help: Positive psychology exercises in an online trial. *Journal of Medical Internet Research*, *14*(3), e63. <https://doi.org/10.2196/jmir.1850>
- Seligman, M. E. (1991). *Learned optimism: How to change your mind and your life*. New York, NY: Pocket Books.
- Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist*, *60*(5), 410-421. <https://doi.org/10.1037/0003-066X.60.5.410>
- Sergeant, S., & Mongrain, M. (2014). An online optimism intervention reduces depression in pessimistic individuals. *Journal of Consulting and Clinical Psychology*, *82*(2), 263-274. <https://doi.org/10.1037/a0035536>
- Shapira, L. B., & Mongrain, M. (2010). The benefits of self-compassion and optimism exercises for individuals vulnerable to depression. *The Journal of Positive Psychology*, *5*(5), 377-389. <https://doi.org/10.1080/17439760.2010.516763>
- Sin, N. L., & Lyubomirsky, S. (2009). Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of Clinical Psychology*, *65*(5), 467-487. <https://doi.org/10.1002/jclp.20593>
- Snyder, C. R., & Lopez, S. J. (2002). *Handbook of positive psychology*. New York, NY: Oxford University Press.
- Sood, A., Prasad, K., Schroeder, D., & Varkey, P. (2011). Stress management and resilience training among department of medicine faculty: A pilot randomized clinical trial. *Journal of General Internal Medicine*, *26*(8), 858-861. <https://doi.org/10.1007/s11606-011-1640-x>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092-1097. <https://doi.org/10.1001/archinte.166.10.1092>

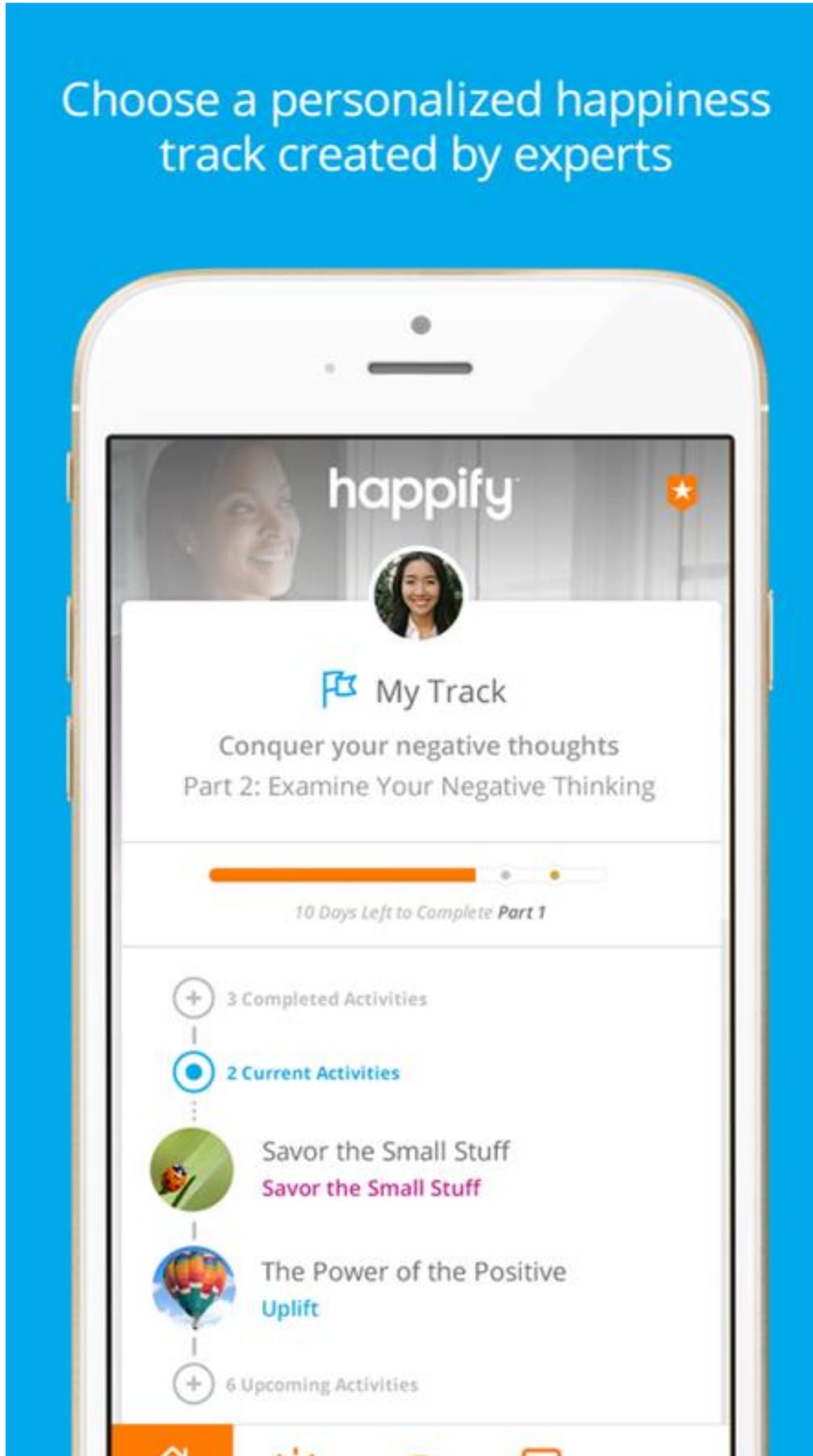
- Thayer, J. F., Åhs, F., Fredrickson, M., Sollers, J. J., III., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36(2), 747-756. <https://doi.org/10.1016/j.neubiorev.2011.11.009>
- Titov, N., Dear, B. F., Johnston, L., Lorian, C., Zou, J., Wootton, B., . . . & Rapee, R. M. (2013). Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: Randomised controlled trial. *PLOS ONE*, 8(7), e62873. <https://doi.org/10.1371/journal.pone.0062873>
- Titov, N., Dear, B. F., Staples, L. G., Bennett-Levy, J., Klein, B., Rapee, R. M., . . . & Nielssen, O. B. (2015). MindSpot clinic: An accessible, efficient, and effective online treatment service for anxiety and depression. *Psychiatric Services*, 66(10), 1043-1050. <https://doi.org/10.1176/appi.ps.201400477>
- Tugade, M. M., & Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *Journal of Personality and Social Psychology*, 86(2), 320-333. <https://doi.org/10.1037/0022-3514.86.2.320>
- Tugade, M. M., Fredrickson, B. L., & Barrett, L. F. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality. Special Issue: Emotions, Personality, and Health*, 72(6), 1161-1190. <https://doi.org/10.1111/j.1467-6494.2004.00294.x>
- Tusaie-Mumford, K. (2002). *Psychosocial resilience in rural adolescents: Optimism, perceived social support and gender differences*. Ann Arbor, MI: ProQuest Information & Learning.
- Utsey, S. O., Hook, J. N., Fischer, N., & Belvet, B. (2008). Cultural orientation, ego resilience, and optimism as predictors of subjective well-being in African Americans. *The Journal of Positive Psychology*, 3(3), 202-210. <https://doi.org/10.1080/17439760801999610>
- Vautier, S., Raufaste, E., & Cariou, M. (2003). Dimensionality of the Revised Life Orientation Test and the status of filler items. *International Journal of Psychology*, 38(6), 390-400. <https://doi.org/10.1080/00207590344000222>
- Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: Results from the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 629-640. <https://doi.org/10.1001/archpsyc.62.6.629>
- White, I. R., Kalaitzaki, E., & Thompson, S. G. (2011). Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an internet-based alcohol trial. *Statistics in Medicine*, 30(27), 3192-3207. <https://doi.org/10.1002/sim.4360>
- Willis, K. D., & Burnett, H. J. (2016). The power of stress: Perceived stress and its relationship with rumination, self-concept clarity, and resilience. *North American Journal of Psychology*, 18, 483-498.
- Windle, G., Bennett, K. M., & Noyes, J. (2011). A methodological review of resilience measurement scales. *Health and Quality of Life Outcomes*, 9, 8-26. <https://doi.org/10.1186/1477-7525-9-8>
- Zautra, A. J., Affleck, G. G., Tennen, H., Reich, J. W., & Davis, M. C. (2005). Dynamic approaches to emotions and stress in everyday life: Bolger and Zuckerman reloaded with positive as well as negative affects. *Journal of Personality*, 73(6), 1511-38. <https://doi.org/10.1111/j.0022-3506.2005.00357.x>
- Zautra, A. J., Johnson, L. M., & Davis, M. C. (2005). Positive affect as a source of resilience for women in chronic pain. *Journal of Consulting and Clinical Psychology*, 73(2), 212-220. <https://doi.org/10.1037/0022-006X.73.2.212>

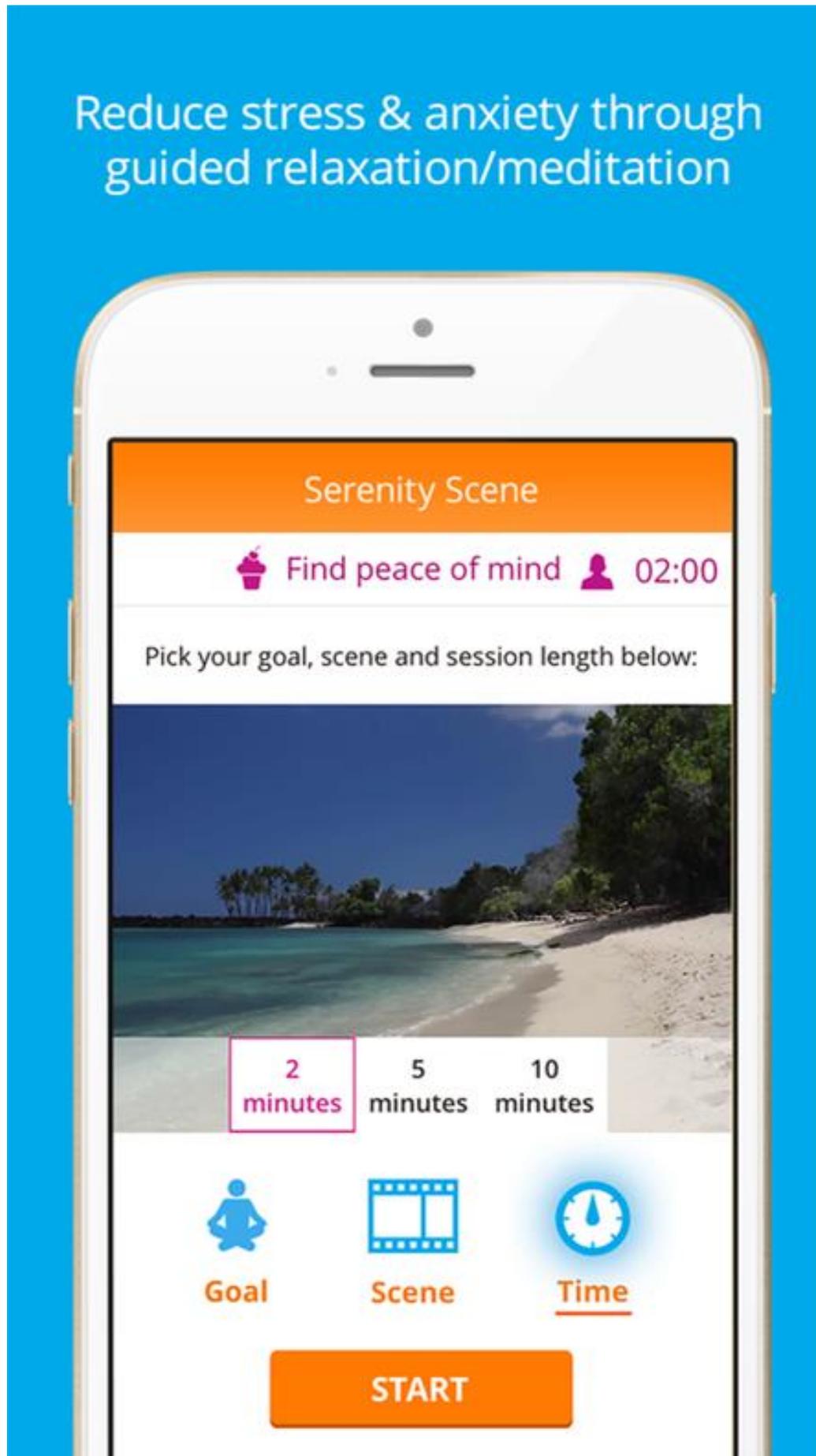
Appendix A. Multimedia screenshots

The following appendix contains screenshots at each stage of participant experience on Happify—from when participants first discover Happify (Part 1; as an example, we show screenshots from the Apple app store, but participants also found us via the Google Play store and the web), to registration (Part 2), to study enrollment and initial assessment (Part 3), to assessment (Part 4), study participation (Part 5), and follow-up prompting (Part 6).

Research process, part 1: App recruitment (advertising images from the Apple app store)





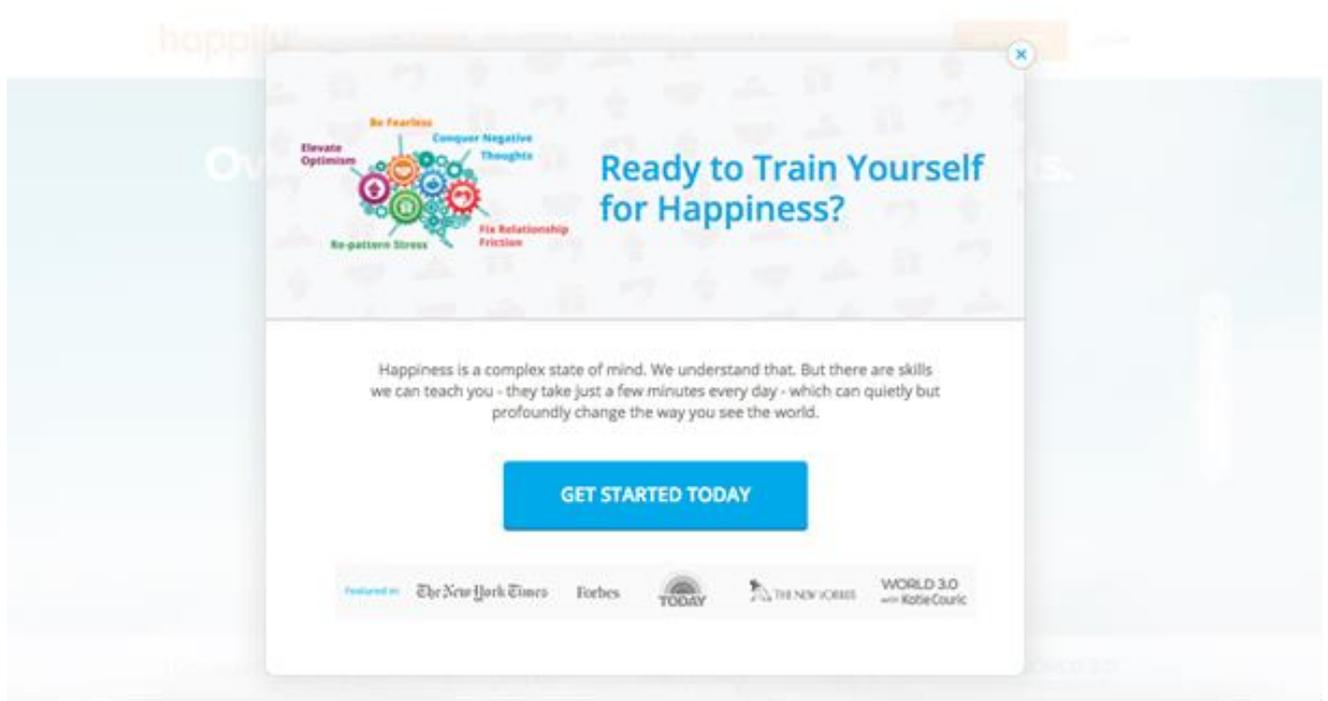




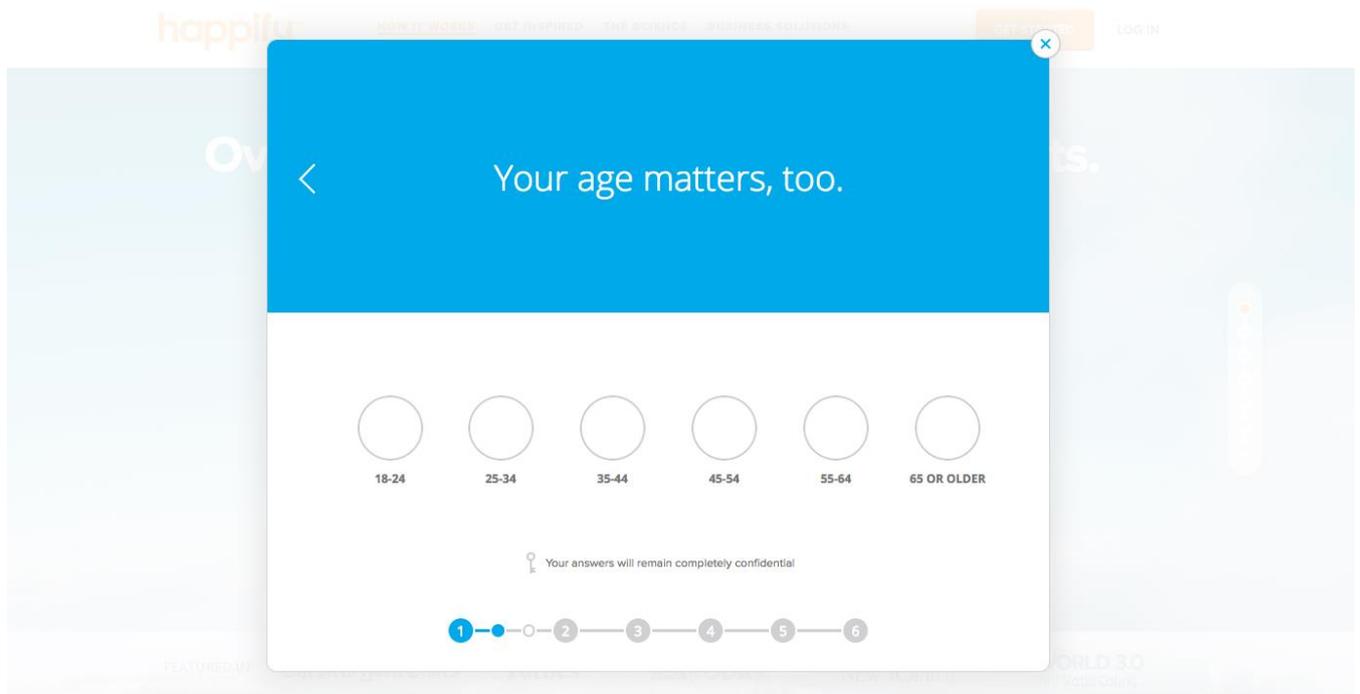


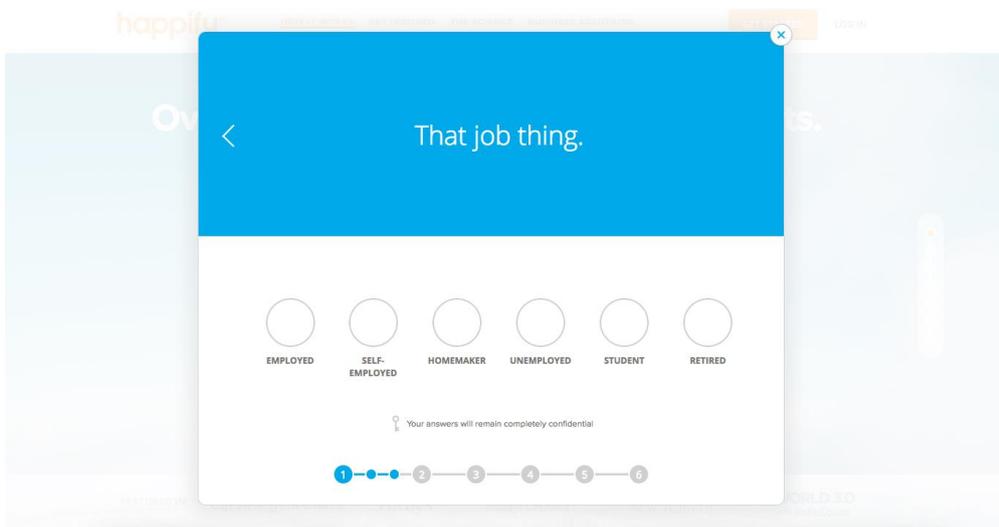
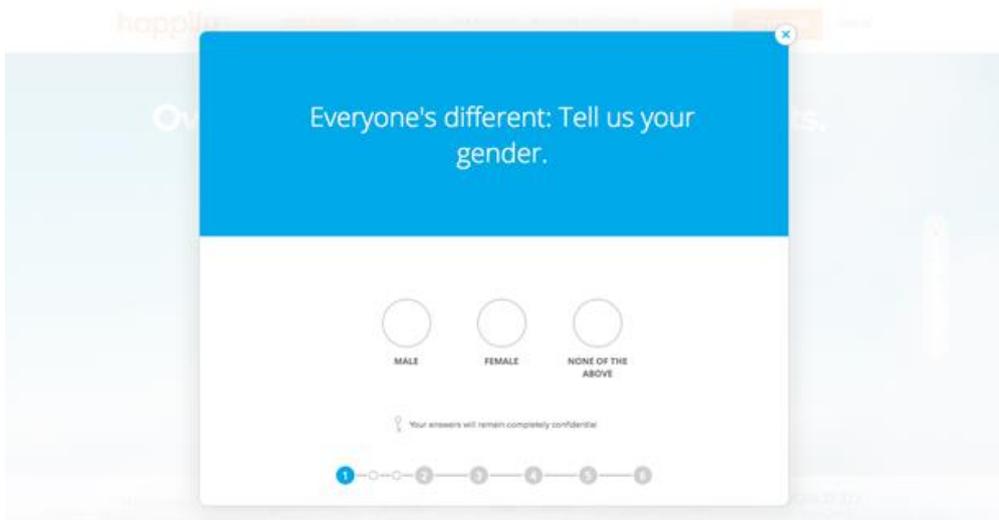
Research process, part 2: Initial user registration (after user downloads app or visits website)

2.1 User is prompted to sign up and create a login to access content.

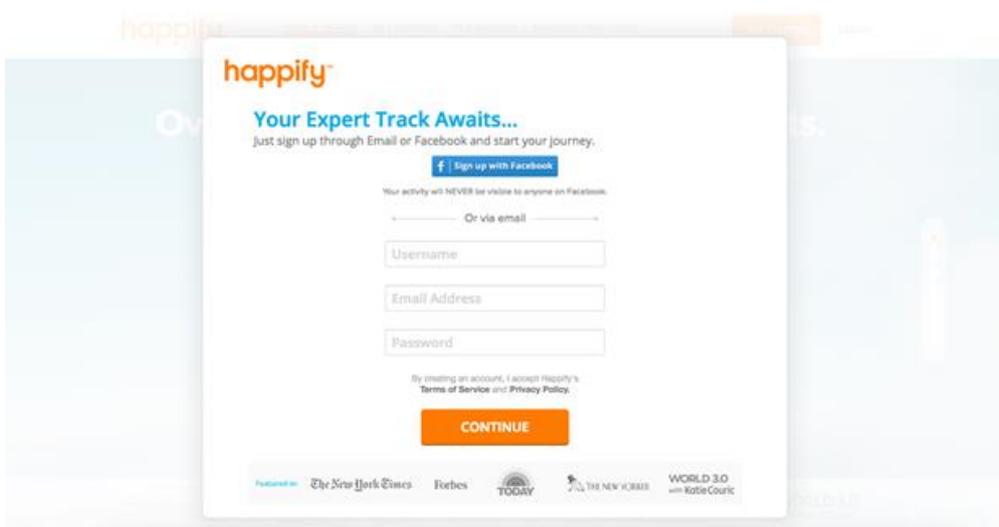


2.2 User completes demographic questions as part of their enrollment in the general site.



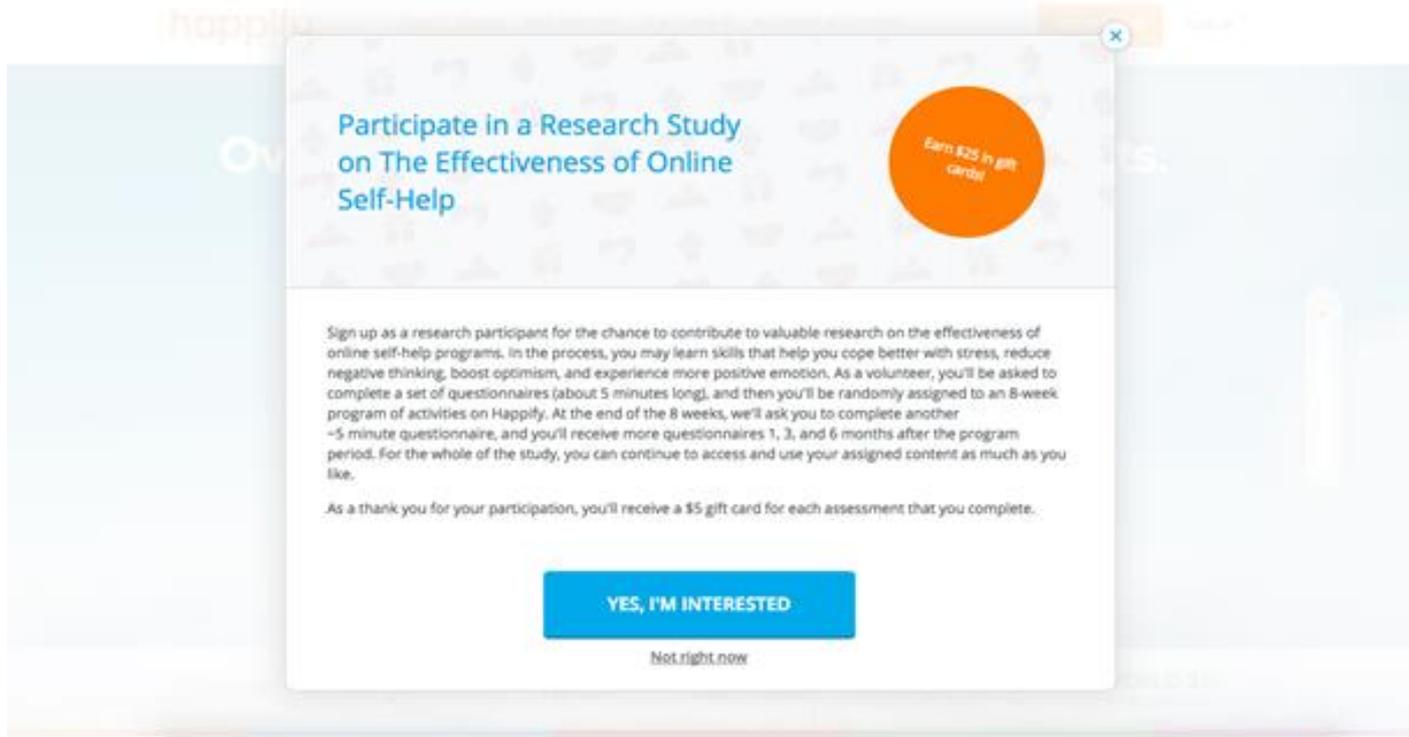


2.3 User is offered the option to create a unique login or to use Facebook to log in.

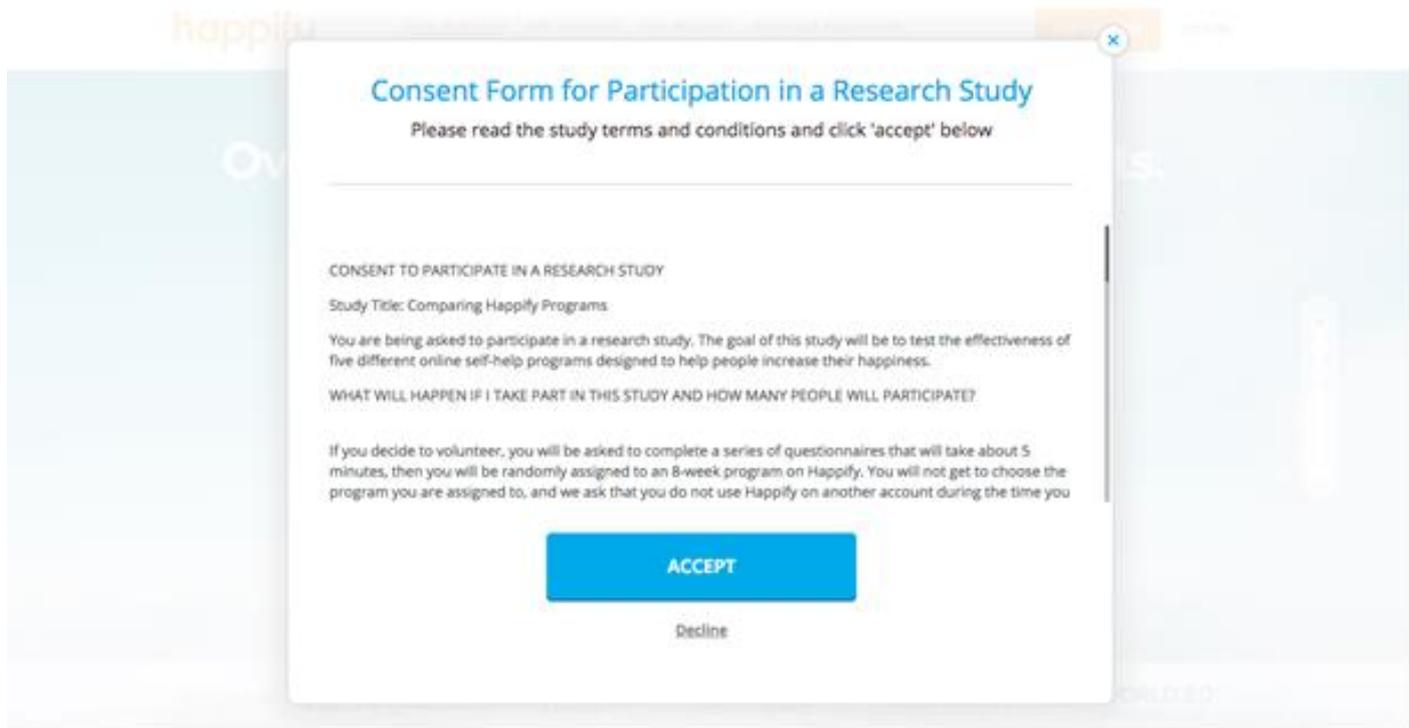


Research process, part 3: Study enrollment (immediately after sign-up)

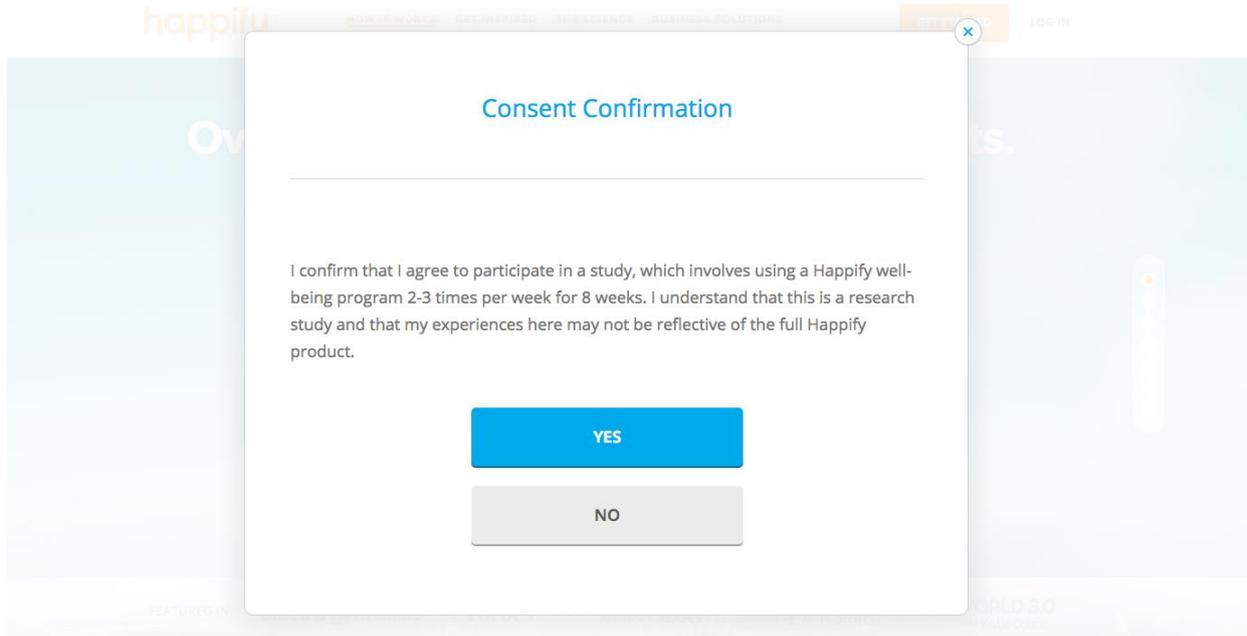
3.1 User reads a brief summary of the study and decides whether they want to learn more.



3.2 User reads a detailed consent form and indicates whether they accept the terms.

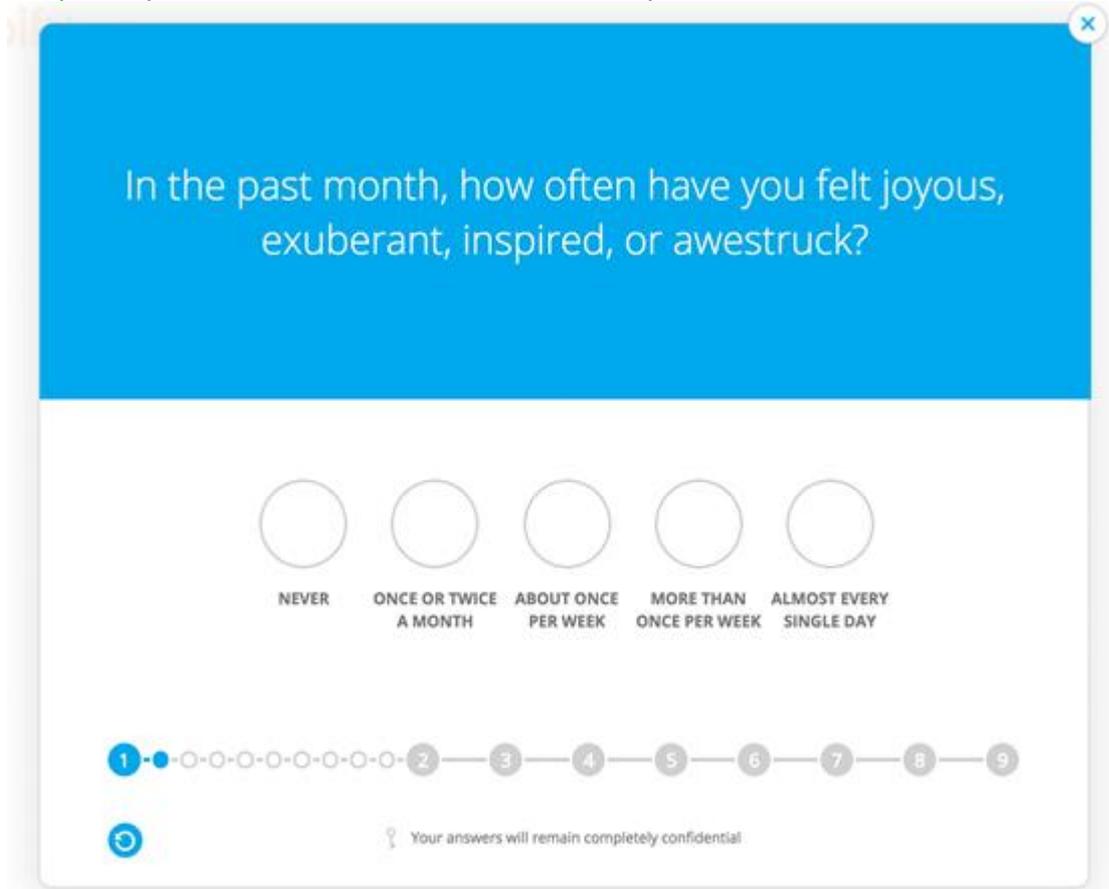


3.3 Because the consent form is lengthy and could be skipped by users, key points of the consent form are reiterated in a pop-up and users again confirm that they wish to participate.

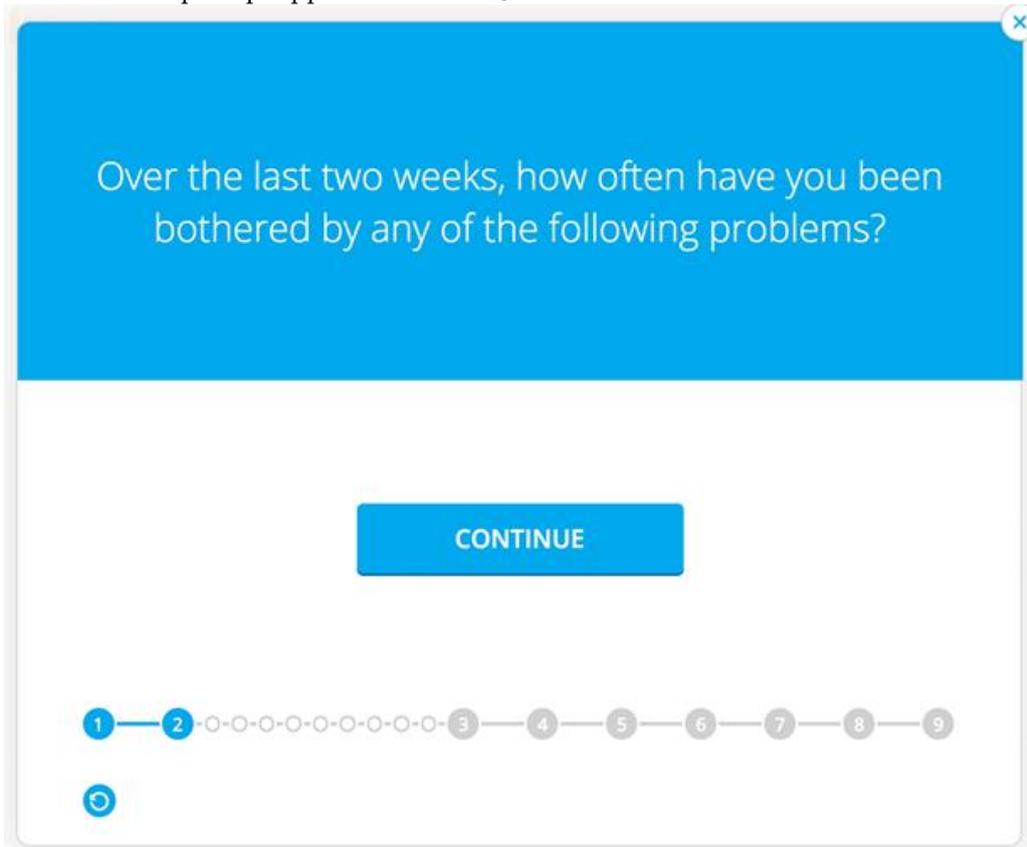


Research process, part 4: Assessment (first one is offered immediately after consent is given)

4.1 The survey used for both baseline and post-test is about 5 minutes long and contains nine different brief questionnaires. A progress bar shows users how many items remain in the current survey, and how many surveys remain in the overall assessment. Surveys are administered one item at a time.



4.2 For surveys with prompts, the prompt was displayed once at the beginning, then individual items followed. This prompt applied to the PHQ-9 and GAD-7 items.



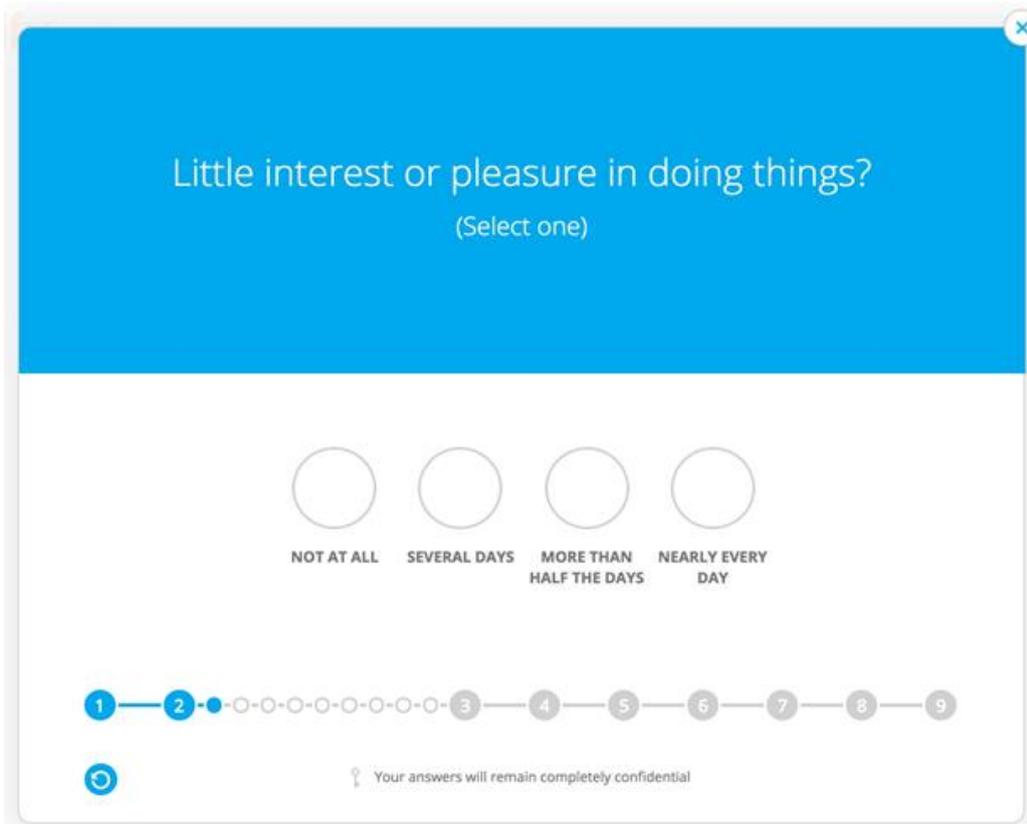
Over the last two weeks, how often have you been bothered by any of the following problems?

CONTINUE

1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9

⌂

This screenshot shows a survey prompt screen. The top half has a blue background with white text asking the user how often they've been bothered by problems over the last two weeks. Below this is a white area with a blue 'CONTINUE' button. At the bottom, there is a progress indicator with nine numbered circles (1-9) and a home icon. Circles 1 and 2 are filled with blue, while circles 3-9 are grey with white outlines.



Little interest or pleasure in doing things?
(Select one)

○ ○ ○ ○

NOT AT ALL SEVERAL DAYS MORE THAN HALF THE DAYS NEARLY EVERY DAY

1 — 2 — 3 — 4 — 5 — 6 — 7 — 8 — 9

⌂

Your answers will remain completely confidential

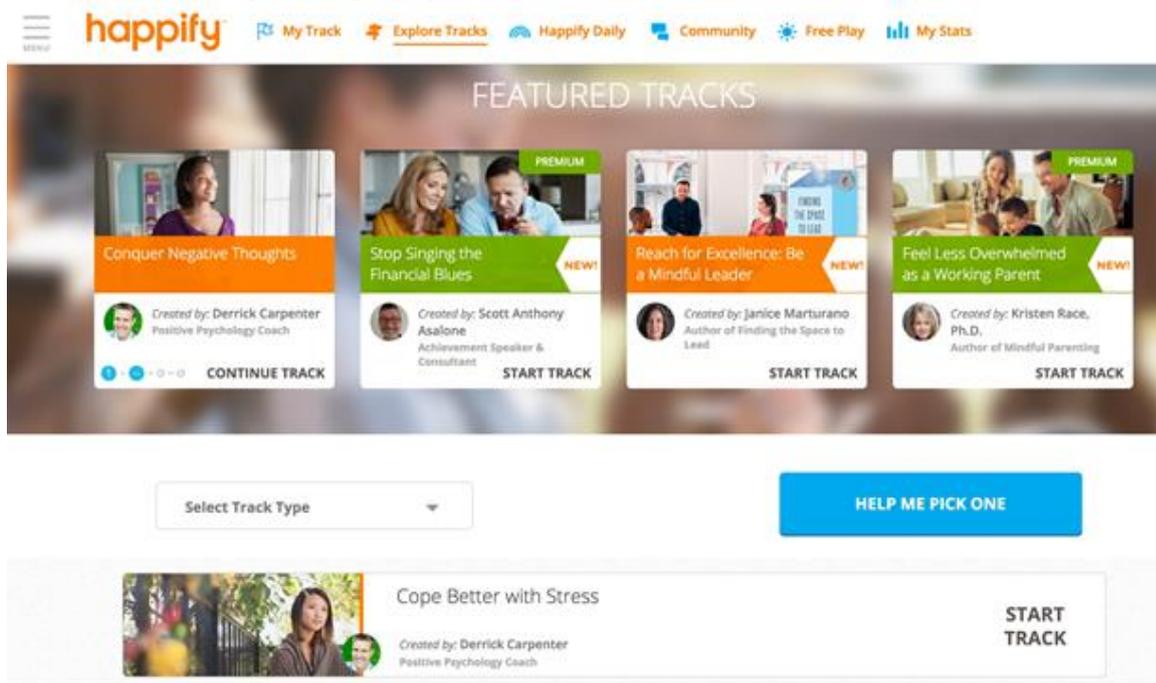
This screenshot shows a survey item screen. The top half has a blue background with white text asking about 'Little interest or pleasure in doing things?' and instructing the user to 'Select one'. Below this are four radio button options: 'NOT AT ALL', 'SEVERAL DAYS', 'MORE THAN HALF THE DAYS', and 'NEARLY EVERY DAY'. At the bottom, there is a progress indicator with nine numbered circles (1-9) and a home icon. Circles 1 and 2 are filled with blue, while circles 3-9 are grey with white outlines. A confidentiality notice is visible at the bottom.

Research process, part 5: Intervention (immediately after baseline assessment is completed)

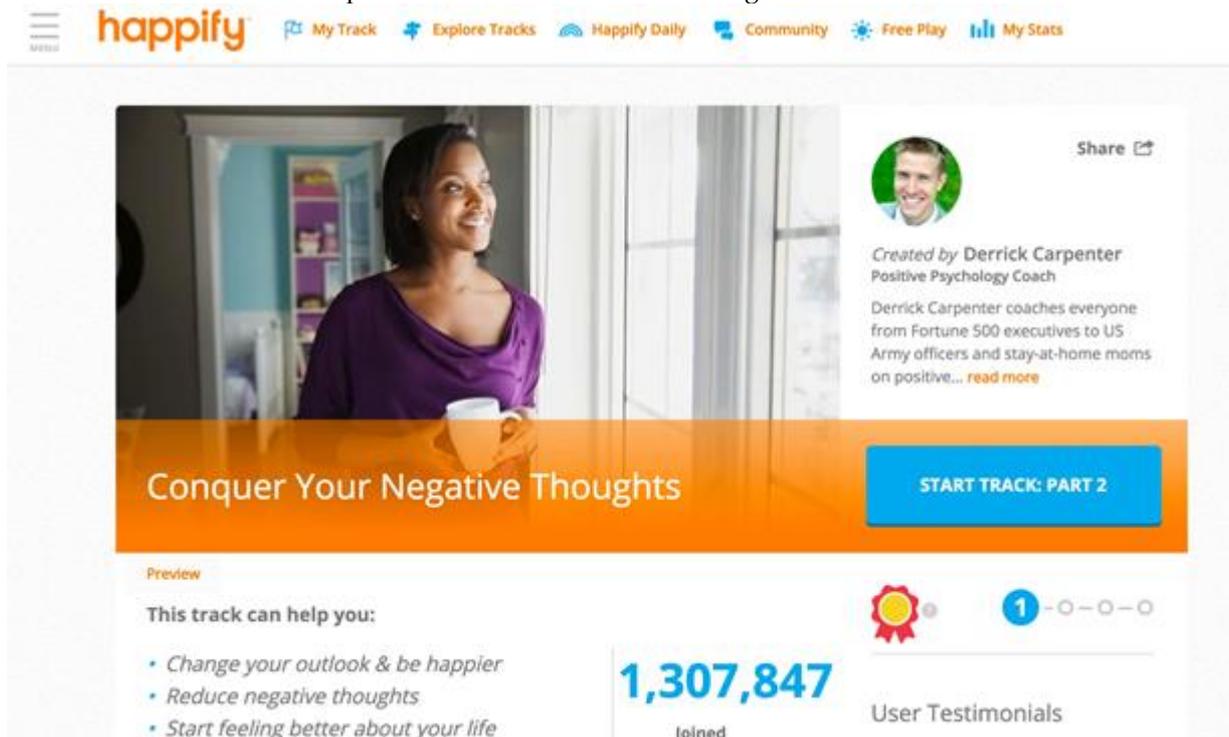
5.1 Happify condition

This section contains a description of the Happify user environment, including key user engagement features.

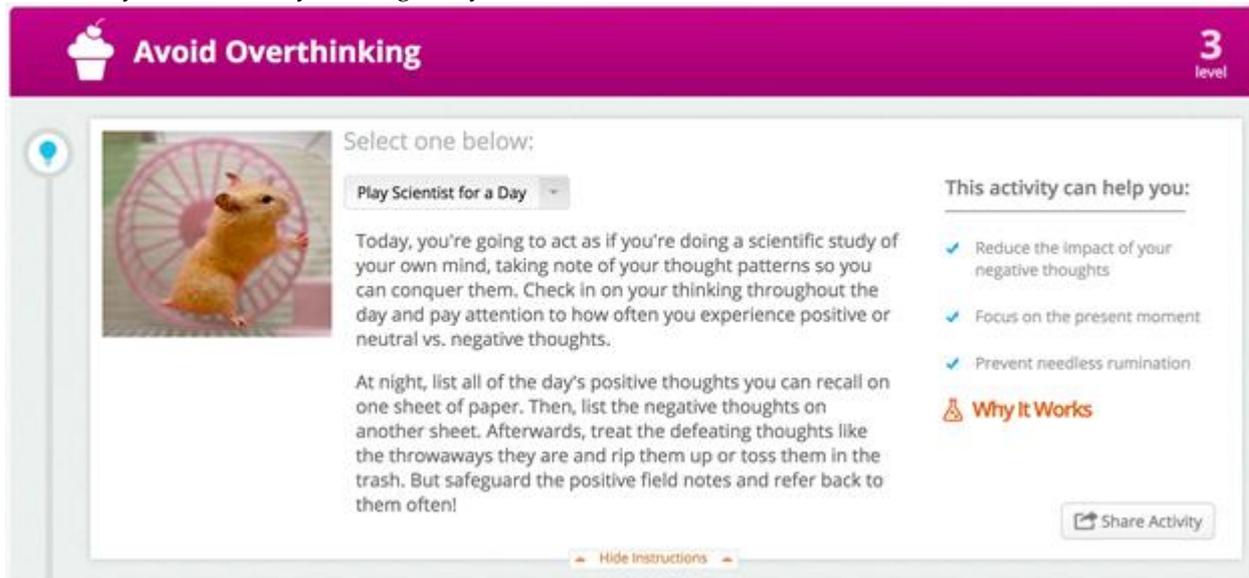
5.1.1. Based on research findings that variety and customization are key to user engagement, Happify users are offered a variety of tracks to choose from. They are also able to take a brief survey to help them match with a track by clicking “Help Me Pick One.”



5.1.2. Users can read a description of each track before choosing to commit to it.

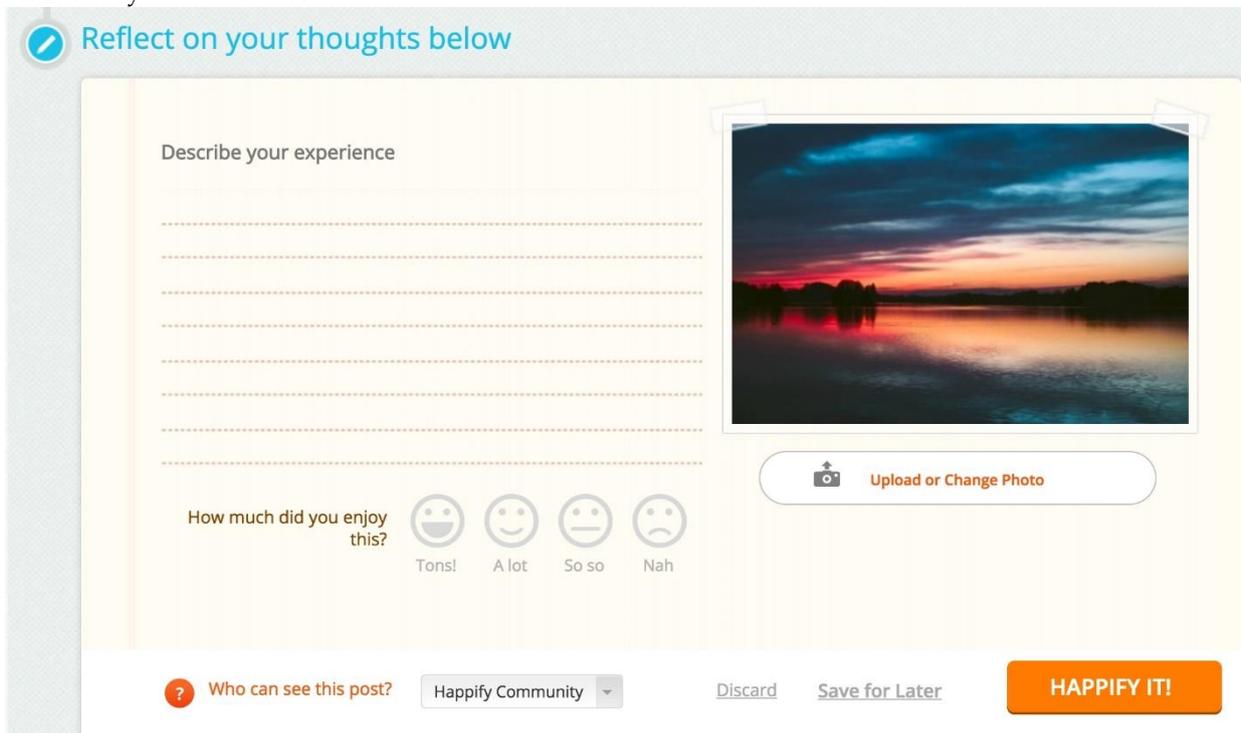


5.1.3. Each activity has 2-3 variations that participants can choose between. In this particular variation of this particular activity, participants practice cognitive reframing. They can learn more about the research the activity is based on by clicking “why it works.”



5.1.4. Once the activity is completed, users write about their experience in provided text boxes and rate their enjoyment of the activity. They can attach a photo to go along with what they did.

Happify users can share their posts with the community of Happify users or post them in a way that is visible only to themselves.



5.1.5 Users can access a community of other Happify users and participate in discussion forums.

Featured Discussion Forums

Overcoming and Thriving

Live & let us live forever! Alleluia! Amen! <3 O:) * (Y)

 **TycoonGoodwin**

Participants 2999 753 discussions

Health and Happiness

I could use a few prayers over the next couple of days. I had a mammogram on...

 **carold0235**

Participants 1891 288 discussions

Forums I'm Following New

Follow Forums on the topics that matter to you

Other Forums	Discussions
Overcoming and Thriving	753 
Why I Happify	357 
Health and Happiness	288 
Inspirations and Aspirations	219
Simple Joys	194
Career and Happiness	156 
Love and Friendship	124

Why I Happify

Hey guys, kinda new to happify. Have been using it now for about 3 weeks starting...

 **liveunderfive**

Participants 936 357 discussions

Career and Happiness

Let us give the Angel his golden measuring rod to measure the holy city, New...

 **TycoonGoodwin**

Participants 779 156 discussions

See the Latest Posts from the Happify Community

Popular
Following
Track Group






5.1.6 Consistent with research findings that generating new content is important to user confidence in an online platform, Happify Daily offers users articles, infographics, and videos related to Happify content.



After Hearing These Anonymous Secrets Read Out Loud, You'll Feel Something Intense

Want More Inspiration?

Subscribe to Happify Daily

Email Address  SIGN UP

Featured Track

 Cope Better with Stress

start now

Recently Shared



Get to Your Calm Place by Listening to This

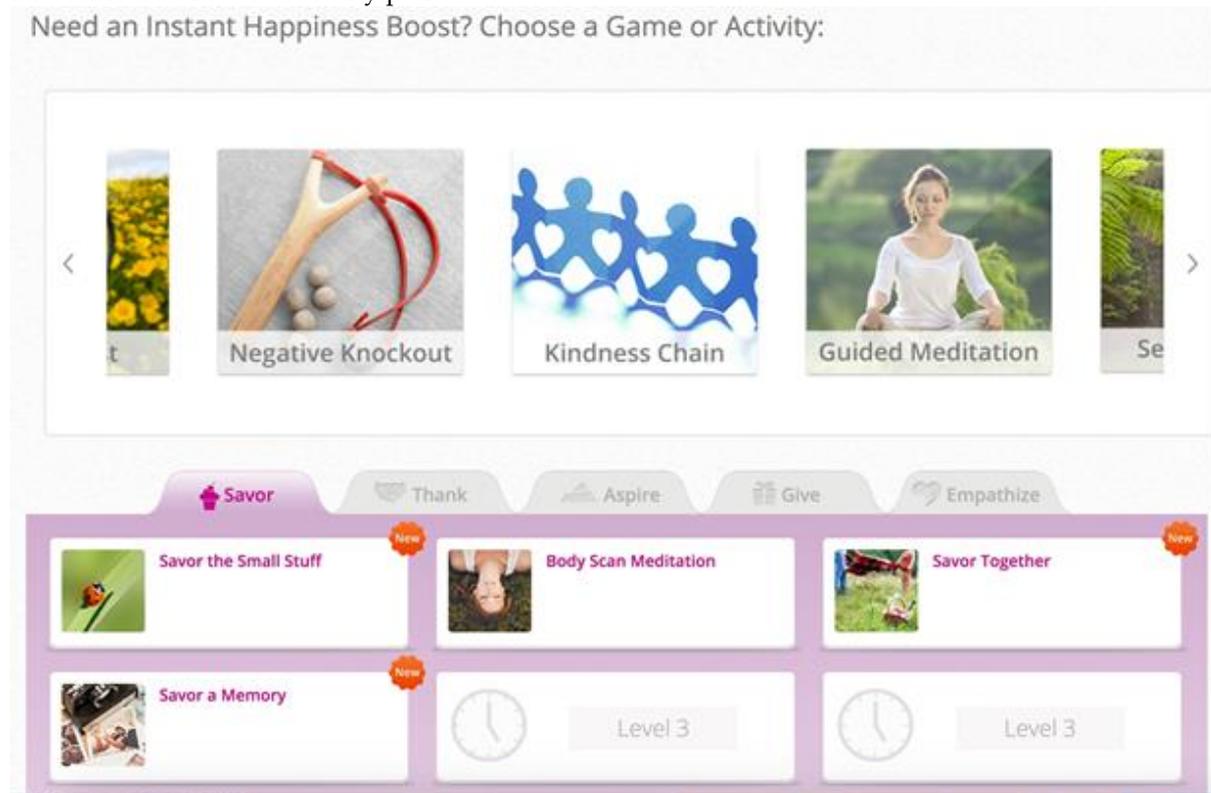


Watch the Profound Effect Yoga Has on These Inmates



7 Surprising Reasons Why Introverts Are Awesome

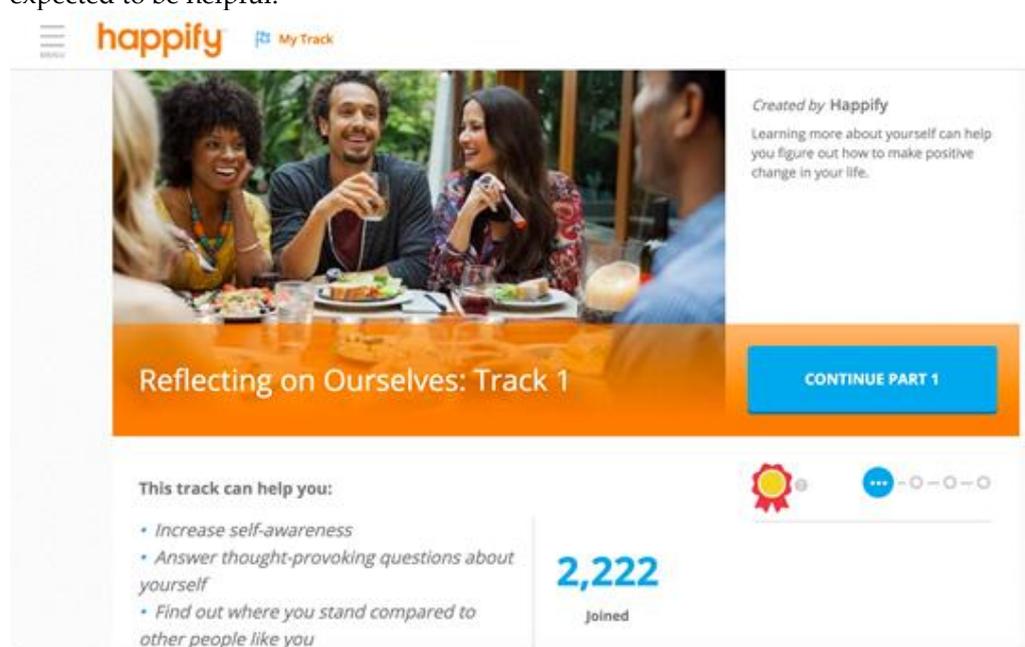
5.1.7 Users can access a “free play” section, which allows them to repeat any game or activity that they have unlocked in a track at any point.



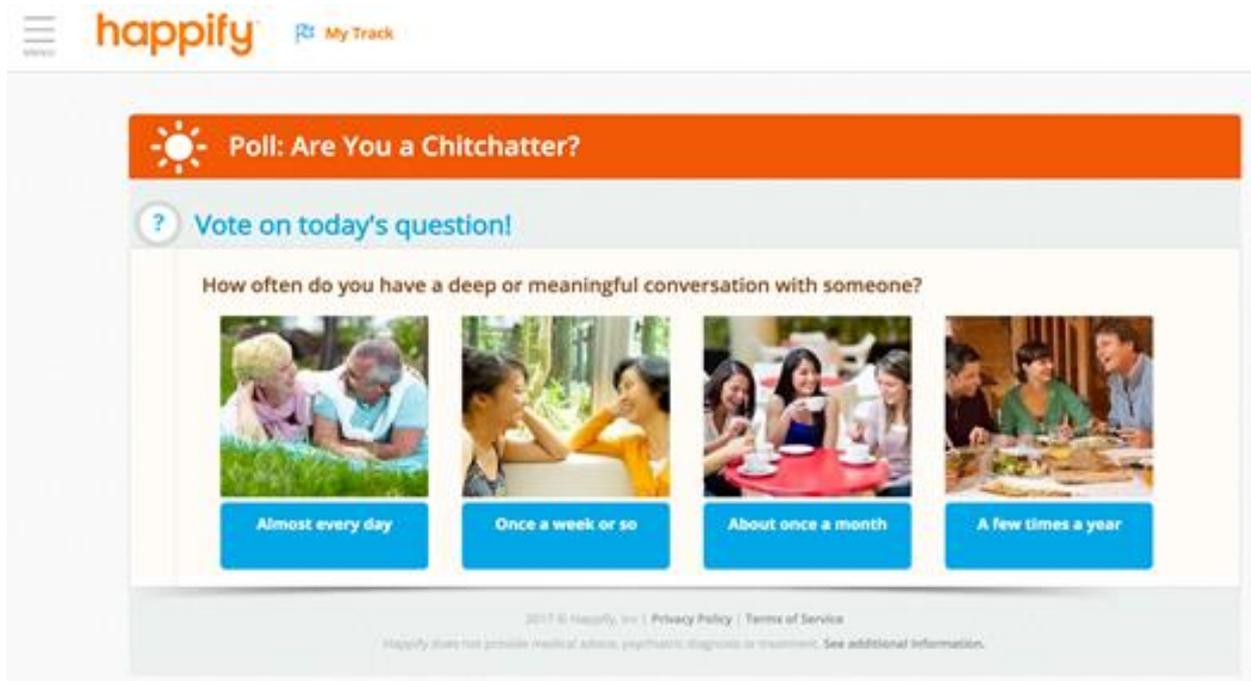
5.2 Psychoeducational comparison group

The control conditions are different from the Happify condition in two key ways: 1) they have no engagement elements (no free play, no community, no Happify Daily) other than generic reminder emails to return and work on their track, and 2) the activities are intended to be hedonically inert.

5.2.1 Users in the psychoeducational comparison group had access to two 4-week tracks that revolve around taking polls. Like the main Happify site, it contains a scientific-sounding rationale for why it is expected to be helpful.



5.2.2 Example poll activity

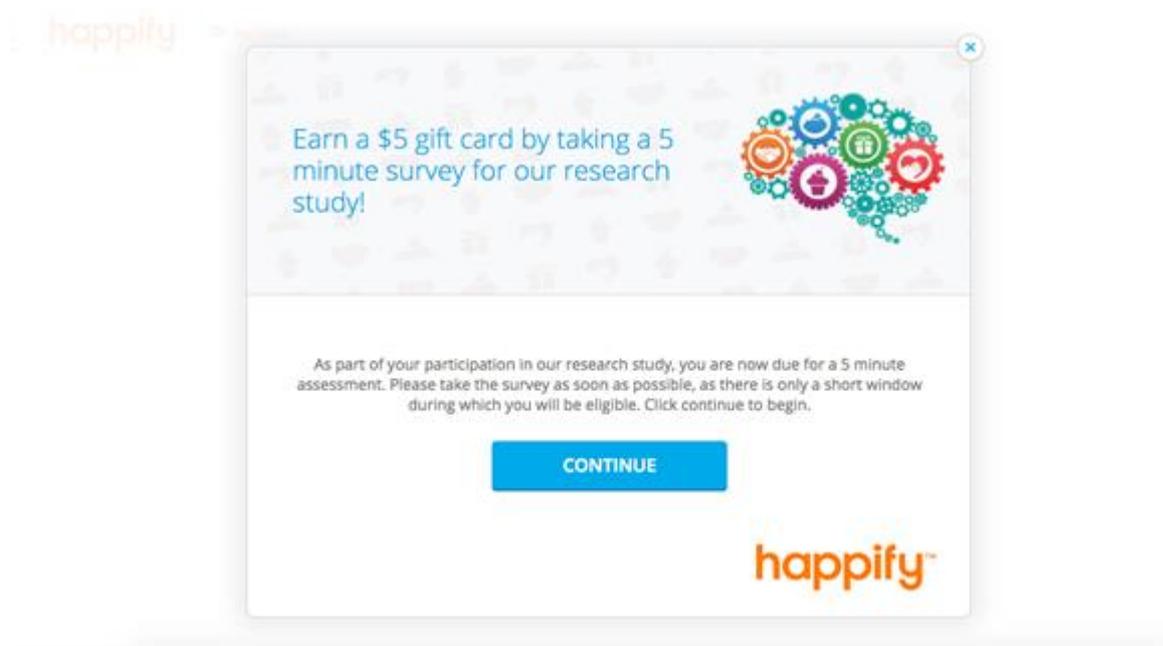


5.2.3 Psychoeducational information that appears after responding to poll



Research process, part 6: Follow-up prompts

When follow-up assessment is due, participants in all conditions are notified via email, and receive the following pop-up as they log in. They will also receive push notifications to their phone if their account is linked to a mobile device.



Appendix B. Consent form

Participate in a Research Study on The Effectiveness of Online Self-Help

Sign up as a research participant for the chance to contribute to valuable research on the effectiveness of online self-help programs. In the process, you may learn skills that help you cope better with stress, reduce negative thinking, boost optimism, and experience more positive emotion. As a volunteer, you'll be asked to complete a set of questionnaires (about 5 minutes long), and then you'll be randomly assigned to an 8-week program of activities on Happify. At the end of the 8 weeks we'll ask you to complete another 5 minute questionnaire, and you'll receive more questionnaires 1, 3, and 6 months after the program period. For the whole of the study, you can continue to access and use your assigned content as much as you like. As a thank you for your participation, you'll receive a \$5 gift card for each assessment that you complete.

CONSENT TO PARTICIPATE IN A RESEARCH STUDY

Study Title: Comparing Happify Programs

You are being asked to participate in a research study. The goal of this study will be to test the effectiveness of five different online self-help programs designed to help people increase their happiness.

WHAT WILL HAPPEN IF I TAKE PART IN THIS STUDY AND HOW MANY PEOPLE WILL PARTICIPATE?

If you decide to volunteer, you will be asked to complete a series of questionnaires that will take about 5 minutes, then you will be randomly assigned to an 8-week program on Happify. You will not get to choose the program you are assigned to, and we ask that you do not use Happify on another account

during the time you are enrolled in this study. At the end of the 8-week period, you will be asked to complete another 5 minute survey. At that point, you will enter the 6-month follow-up period, during which you may continue to use your assigned programs, or not, as you prefer. You will be asked to complete questionnaires 1, 3, and 6 months later. At the completion of the final follow-up (about 8 months from now), your participation will be over. By signing up for this research, you also consent to receive reminder emails throughout the study. A total of 4200 users are expected to participate.

WHAT OTHER CHOICES DO I HAVE IF I DO NOT TAKE PART IN THIS STUDY?

You are not required to participate in this study. If you do not participate, you are free to seek any other self-help, or to sign up for Happify as a regular user.

WHAT RISKS CAN I EXPECT FROM BEING IN THIS STUDY?

There are no anticipated risks to participating in this study.

ARE THERE BENEFITS TO TAKING PART IN THIS STUDY?

By participating in this study, you will gain access to an online program which may benefit you. However, since some of the tracks we are testing are new, the efficacy of which are not yet established, we cannot guarantee that you will benefit directly by participating in this study. However, we are testing these tracks because we believe there is a good chance that they will be helpful to you.

WILL MY INFORMATION BE KEPT PRIVATE?

If you consent to participate in this study, you will be asked to complete a series of web-based questionnaires. When you register for the study, you will be asked to give us your email address. In our secure database, your email address will be tied to your data by a randomly generated numerical code, which will allow us to connect each new questionnaire you take to the other questionnaires you have taken. Depending on whether or not your email address contains identifying information, this may or may not constitute a risk to privacy were someone to compromise the security of our database. However, because email addresses will be stored separately from questionnaire data, the researchers will never be able to connect your data with any identifying information. However, if you wish to be completely certain that your data will remain private under any circumstances, you may create an anonymous email address for the purposes of this study and have its messages forwarded to your main account.

WILL I BE COMPENSATED FOR BEING IN THIS STUDY?

You will receive a \$5 gift card for each of our five minute surveys you complete for up to \$25 total.

WHAT ARE THE COSTS OF TAKING PART IN THIS STUDY?

There is no cost to you beyond the time and effort required to complete the procedure(s) described above.

CAN I STOP BEING IN THIS STUDY?

You may refuse to participate in this study. You may also change your mind about being in the study and quit after the study has started at any time.

WHO CAN ANSWER MY QUESTIONS ABOUT THIS STUDY?

If you have any questions about this research project please contact Chief Scientist at Happify, Acacia Parks, labs@happify.com. By checking the box that says “I accept” and clicking on the “next” button below, I will indicate that I have decided to participate in this study as a research participant. I have read and understand the information above. I understand that I will be emailed a copy of this consent form and the Bill of Rights.

Agreement Pop-Up Text

I confirm that I agree to participate in a study, which involves using a Happify well-being program 2-3 times per week for 8 weeks. I understand that this is a research study and that my experiences here may not be reflective of the full Happify product.

Text When Assigned Programs Are Completed

Congratulations on finishing your assigned tracks! You are now in the “follow-up” period of the study. You'll be asked to complete questionnaires periodically in the coming months, and will receive a \$5 gift card each time you complete a follow-up questionnaire. You're also encouraged to continue using your tracks. For the best results, come back 2-3 times per week!