Application of the Efficiency Model of Human Support for a Depression and Anxiety App Treatment Trial

Emily G Lattie
Department of Preventive Medicine, Northwestern University
750 N Lake Shore Dr., 10th Floor, Chicago, IL, 60611
emily.lattie@northwestern.edu

Kathryn N Tomasino
Department of Preventive Medicine, Northwestern University
750 N Lake Shore Dr., 10th Floor, Chicago, IL, 60611
kathryn.noth@northwestern.edu

David C Mohr
Department of Preventive Medicine, Northwestern University
750 N Lake Shore Dr., 10th Floor, Chicago, IL, 60611
d-mohr@northwestern.edu

ABSTRACT
The inclusion of human support in digital health interventions has repeatedly been shown to be advantageous for usage and clinical outcomes, yet methods of providing this support are varied. In this paper, we describe a coaching protocol used in a single arm field trial of a suite of apps for depression and anxiety called IntelliCare. We discuss the continued evolution of determining an appropriate level of support in order to maximize user experience, system usage and clinical outcomes.

Author Keywords
Mental health; human support; smartphone; computer-mediated communication

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous. J.4 Social and Behavioral Sciences: Psychology.

INTRODUCTION
The provision of support has repeatedly been shown to be an essential feature of successful digital health interventions [1,2]. The Efficiency Model of Support [3] was previously developed to guide individuals (e.g. coaches, clinicians) who support users of technology-enabled health services (e.g. cognitive and behavioral internet interventions, mental health treatment apps). Briefly, this model defines efficiency as the ratio between benefit derived from an intervention relative to the resources devoted to supporting it, and posits that support should be provided when a user encounters a failure point that interferes with obtaining a positive outcome. Failure points include problems with 1) usability (technology problems), 2) engagement (right tool, no use of the tool), 3) fit (wrong tool, use irrelevant), 4) knowledge (right tool, incorrect use), and 5) implementation (right tool, no application). Thus, coaching is intended to improve people's ability to use the digital tools, the quantity of their use, and ultimately their clinical response.

In order to maximize efficiency, and thus the public health potential of technology-enabled health services, it's vital that we carefully assess and iterate upon our methods for providing support. Here, we describe how the Efficiency Model was used in the development of a coaching protocol for an 8-week trial of a suite of mental health apps. We report on participant feedback regarding the coaching protocol, and future directions for providing this type of support.

METHODS AND RESULTS
We conducted a single arm field trial of a suite of apps for depression and anxiety called IntelliCare. The IntelliCare apps present interactive skills for mental health (e.g. behavioral activation, cognitive restructuring, relaxation), and were designed for frequent but short interactions (see [4]). A total of 99 individuals participated in this trial, and all received low-intensity coaching.

Coaches accessed an online dashboard that provided information about the IntelliCare apps installed on each participant's phone, including when they were downloaded, each instance of use, and which apps were selected as the "main" apps to focus on for the week (see Figure 1). Initially, the coaching protocol directed coaches to focus on encouraging IntelliCare system use, and to avoid making recommendations for use of particular apps in the suite. Each participant met with a coach over the phone for approximately 30 minutes at the start of the 8 week trial, and subsequently received approximately 2 text messages per week to provide support and address any suspected failure points.

User Feedback on Initial Coaching Protocol
Approximately one third of the way into the field trial, we reviewed user feedback interviews, which were completed at midpoint (week 4) and end of treatment (week 8). Preliminary thematic analysis revealed 3 common themes –

appreciation of support, *limitations of coaching*, and *coaching confusion*. Many participants noted that they felt supported by their coach (*appreciation of support)* and viewed coaching as an integral part of the program. For example,

"It's nice to have someone to reach out to me to ask about my day. I don't have a lot of outside interactions in the last few months because I'm home so it's nice and encouraging to have someone to answer questions. It makes me feel like someone really cares."

Several participants noted some dissatisfaction with the perceived limitations to what their coach was there to do (*limitations of coaching)*, such as,

"Right now I understand they're a little bit limited in what they can coach you on. They can't see any of your answers. They're not counselors. They're strictly coaching you on the applications. I know that's never gonna happen but…I would see it as a little bit more helpful [if they could do more]*."*

We also found a degree of *coaching confusion* - several participants were unclear about the role of the coach and were not satisfied with the support they received. For example, one participant noted,

"I don't feel any connection with my coach. I'm not clear on the goals of having a coach I'm not sure what's the intended outcome of the relationship,"

while another stated,

"I would hear from her a few times and she would tell me to download the app recommendations but I'm confused about her role in the study and what I should expect from her."

Changes to the Coaching Protocol

This review of user feedback prompted our team to revise the coaching protocol. Compared the original protocol, coaches following the revised protocol took a more active role by making direct recommendations about which app or apps to use when participants did not have a focus in mind, and offered a midpoint phone call (lasting approximately 10-15 minutes) to further assess for failure points, elicit and answer questions, and make recommendations for continued system use. Because coaching is designed to support participants' engagement with the program and application of skills, we expected that improvements to the coaching experience would increase engagement and lead to greater gains.

Feedback on Coaching

At the conclusion of the field trial, 34 participants received coaching that followed the original coach protocol, and 65 participants received coaching that followed the revised coach protocol. In our review of user feedback from those participants who received the revised coach protocol, we found that participants described coaching in clearer and more positive terms. While participants continued to note limitations of the coaching role, these were stated in a more positive light, such as,

"*I thought to myself, wait this whole experience is yes therapeutic but it isn't therapy and my coach is NOT a therapist. And I'm careful to make that distinction because I've been in therapy before so I'm careful to not turn my coach into a therapist or a friend, they're not supposed to be."*

Further, there was no longer any evidence of coaching confusion. For example, one participant noted,

"He makes good suggestions. He's keeping an eye on things. If he sees something that might be helpful he makes a suggestion. I wasn't sure which app to choose, but based on our conversation he gave a recommendation which was good."

and another commented,

"He was very responsive when I had questions and I knew from being told ahead of time that I wouldn't get a response immediately but he always got back to me in a reasonable period of time."

These findings were further supported by responses to the Working Alliance Inventory (WAI) [5] administered at week 4. The WAI is a widely used questionnaire that measures common factors in the relationship with a psychotherapist or counselor. It produces three subscores, including emotional bond, agreement on the goals of treatment, and agreement on tasks. As seen in Table 1, participants who received the revised protocol had significantly higher scores than participants who received the original protocol on the Goal subscale. This difference provides further support that participants had a clearer sense of the goals of treatment following revisions to the coach protocol.

WAI scale	Original M(SD)	Revised M(SD)	Test of difference
Bond	14.8(4.8)	16.3(3.9)	$t(90)=-1.7$, $p=.10$
Goal	13.0(5.5)	16.5(3.4)	$t(89)=-3.7$, $p<.001$
Task	13.1(4.2)	14.3(3.4)	$t(90)=-1.5$, $p=.15$

Table 1. Working Alliance Inventory differences

Usage and Usability

While it appeared that the revised protocol clarified participants' understanding of coaching, this did not result in increased app usage. There were no differences in number of app sessions or in total time spent on IntelliCare apps between participants who received the original protocol compared to those who received the revised protocol (p's > .10) [4].

At weeks 4 and 8, participants completed the USE scale [6], which contains subscales of usefulness, ease of use, ease of learning, and satisfaction. As seen in Table 2, there were no significant differences in usability ratings between participants who received the original protocol and those who received the revised protocol (all p's > .10). This

demonstrates that an improved coaching experience did not translate into differences in overall system usability.

	Week 4		Week 8	
	Orig. M(SD)	Rev. M(SD)	Orig. M(SD)	Rev. M(SD)
Usefulness	4.9 (1.4)	5.1 (1.0)	4.8 (1.5)	5.2 (1.2)
Ease of Use	4.9 (1.3)	5.3 (1.2)	5.3 (1.3)	5.5 (1.2)
Ease of Learning	5.6 (1.3)	5.9 (1.2)	5.8 (1.3)	5.8 (1.2)
Satisfaction	4.7 (1.5)	4.9 (1.2)	4.6 (1.6)	5.1 (1.3)

Table 2. USE scale scores

Clinical Outcomes

As reported in the main outcomes paper for this trial [4], participants demonstrated substantial reductions in both symptoms of depression (measured by the PHQ-9; [7]) and anxiety (measured by the GAD-7; [8]). Because participants appeared to be more engaged and have higher agreement with coaches regarding the goals of the program with the revised protocol, we hypothesized that the second group of participants would attain greater benefit from the IntelliCare program. Yet, there were no significant differences between these groups in changes in symptoms of depression nor were there differences in changes in symptoms of anxiety (*p's* > .10).

It is possible that an improved coaching experience could have translated into better understanding of program goals, leading to greater increases in coping skills. Thus, we examined changes in coping skills, measured both by the Cognitive and Behavioral Response (CB-RSS, [9]; which measures (A) the frequency with which skills were used and (B) their perceived usefulness) and by the Coping Self-Efficacy Scale (CSE, [10]; which measures confidence in ability to perform coping behaviors). However, there were no significant differences between groups on either of these measures (all *p*'s > .10).

	BL M(SD)		Wk 4 M(SD)		Wk 8 M(SD)	
	Orig.	Rev.	Orig.	Rev.	Orig.	Rev.
PHQ-9	12.5 (4.3)	12.6 (4.3)	8.4 (3.9)	8.4 (4.2)	6.5 (3.9)	6.4 (4.5)
GAD-7	10.5 (4.2)	6.9 (4.2)	5.2 (3.7)	11.2 (4.7)	7.2 (3.7)	6.1 (4.2)
CB-RSS A	28.3 (12.3)	24.9 (10.7)	36.6 (10.0)	33.6 (13.2)	36.6 (14.7)	37.9 (12.1)
CB-RSS B	35.0 (15.8)	31.6 (9.9)	48.1 (15.1)	43.7 (15.5)	49.7 (20.2)	49.4 (16.7)
CSE	108.3 (38.9)	100.6 (37.2)	138.1 (43.2)	129.0 (45.3	152.2 (56.8)	148.9 (49.0)

Table 3. Clinical outcome measures

CONCLUSIONS

In this field trial, we found that the original version of our coaching protocol left many participants feeling somewhat confused about the role of coaching. Based on these findings, we revised the protocol to change the quality of the interactions between coaches and participants, and offered additional coaching time to participants (e.g. offering a midpoint phone call). These changes were successful in increasing the clarity surrounding coaching, reflected both in user feedback and WAI reports. However, while coaching is intended to support the patient's ability to use the intervention apps, frequency of use, and treatment response, the increase in clarity resulting from changes in the coaching protocol did not translate to differences in overall system usability ratings, app usage, or psychological outcomes.

The lack of differences in clinical outcomes between these two groups raises the question whether there is any value in improving people's experience with the coaching. From a pure health economic perspective, the addition of coach time and the associated costs cannot be justified in the absence of clinical benefit. However, from a humanistic perspective, having a sense of clarity as to one's role with a care provider might be seen as a common human need that has value and should be respected. From this perspective, user experience in digital mental health experiences could be seen as having value in and of itself, independent of health economic questions.

There are several limitations we should note in considering these results. There may have been differences between participants who joined early in the trial compared to those who joined later (e.g. early participants may have been more eager to participate/be more able to overcome the shortcomings of a less optimal system). Second, participants in this trial committed to complete an 8 week treatment. It is possible that if the IntelliCare program were delivered on a more flexible basis (e.g. without a set time frame), the improved user experience with coaching may have resulted in better engagement and outcomes relative to the initial coaching model. Finally, the coaching protocol revision included both changes to the content (e.g. more direction regarding apps) and increased time (the option of an additional 15 min. call). While we speculate that the improved clarity in patient perceptions of coaching were due primarily to the content, we are unable to definitively disentangle the effects of changes in content and quantity of time.

In conclusion, as digital mental health interventions are developed, designing the coaching services that support them will be critical to their success. In this study, we describe a development approach which led to significantly improved experience, but without a concomitant improvement in clinical outcomes. Because user experience is not always tightly coupled with clinical outcomes, this raises questions as to the role and value of user experience in health and mental health digital interventions. Improving coaching models has costs both to develop and sometimes in

additional coach time. While a strict economic approach would argue that an endeavor that adds cost without improving outcomes is not worth undertaking, we would also argue that patients' comfort in their experience with providers also has an intrinsic value. However, if user experience has value, methods of assessing that value will be required if such an approach is to be accepted in our healthcare system.

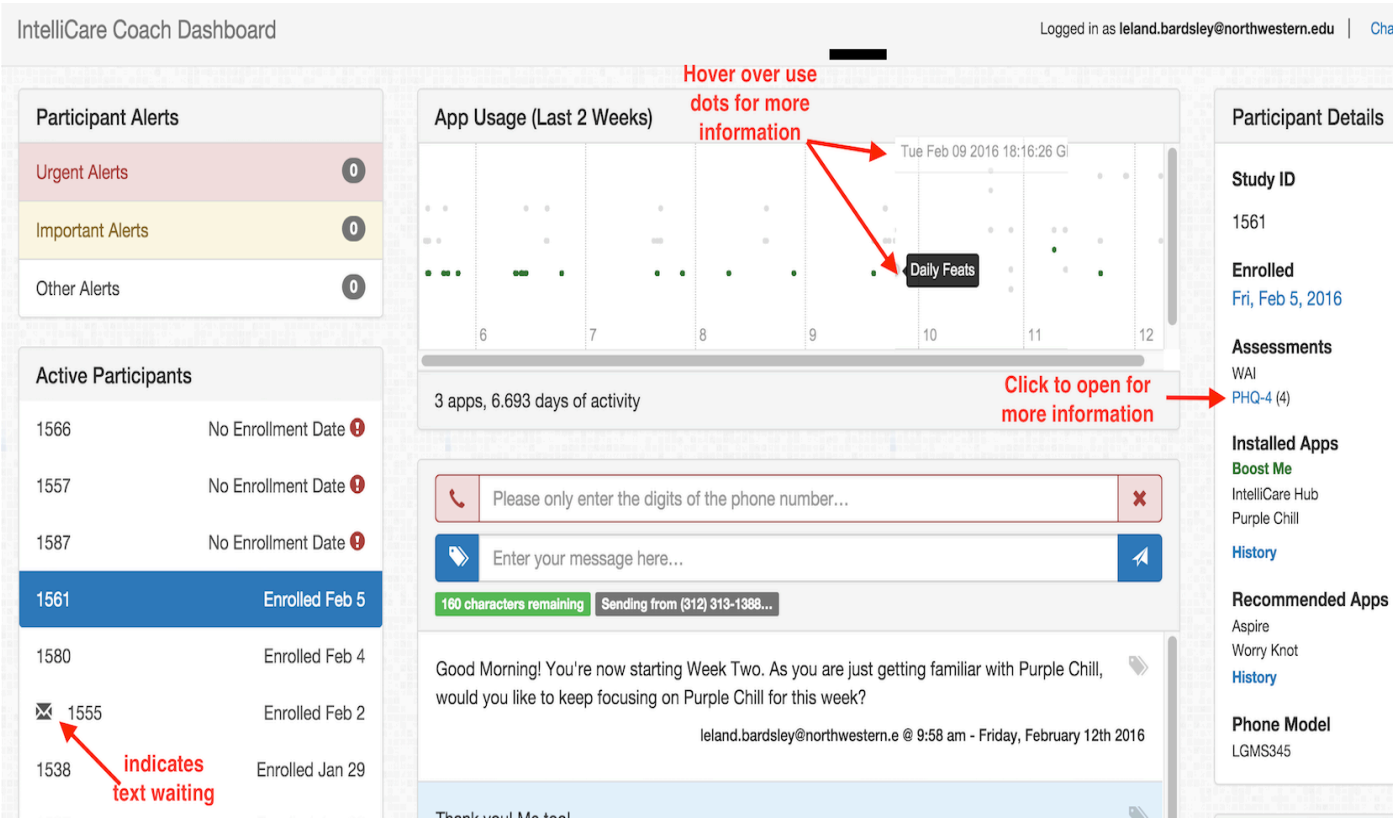

Figure 1: The IntelliCare coach dashboard includes a text messaging interface and displays patient engagement with the app suite, including which apps are being used and when. The dashboard also includes alerts to support efficient triage and clinical decision-making and patients total score on brief symptom questionnaire (PHQ-4).

ACKNOWLEDGMENTS

This research was supported by the United States National Institute of Mental Health grant R01 MH100482.

REFERENCES

1. Andersson, D., and Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther*, 38, (2009), 196-205.
2. Titov, N., Dear, B.F., Staples, L.G., et al. Disorder-specific versus transdiagnostic and clinician-guided versus self-guided treatment for major depressive disorder and comorbid anxiety disorders: A randomized controlled trial. *J Anxiety Disord*, 35, (2015)88-102.
3. Schueller, S.M., Tomasino, K.N. and Mohr, D.C. Integrating human support into behavioral intervention technologies: The Efficiency Model of Support. *Clin Psychol*. 2016.
4. Mohr, D.C., Tomasino, K.N., Lattie, E.G., et al. IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety. *J Med Internet Res*. 2017;19(1):e10.
5. Horvath, A.O., and Greenberg, L.S. Development and validation of the Working Alliance Inventory. *J Counsel Psychol* 1989;36:223-233.
6. Lund, A.M. Measuring usability with the USE Questionnaire. *STC Usability Newsletter*. 2001;8(2).
7. Kroenke K., Spitzer R.L., and Williams J.B. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*, 16, (2001), 606-613.
8. Spitzer R.L., Kroenke K., Williams J.B.W., and Lowe B. A brief measure for assessing generalized anxiety disorder. *Arch Intern Med*, 166, (2006), 1092-1097.
9. Miner, A.S, Schueller, S.M., Lattie, E.G., and Mohr, D.C. Creation and validation of the Cognitive and Behavioral Response to Stress Scale in a depression trial. *Psychiatry Res*. 2015;230(3):819-825.
10. Chesney, M.A., Neilands TB, Chambers DB, Taylor JM, Folkman S. A validity and reliability study of the coping self-efficacy scale. *Brit J Health Psych*. 2006; 11(3):421-37.