

Comparing in-the-moment skill coaching effects from tailored versus non-tailored Acceptance
and Commitment Therapy mobile apps in a non-clinical sample

Michael E. Levin ^{a*}, Cynthia Navarro ^a, Rick A. Cruz ^a & Jack Haeger ^a

^a Utah State University, Department of Psychology, 2810 Old Main Hill, Logan, UT 84322.

* Corresponding author. Utah State University, 2810 Old Main Hill, Logan, UT 84322, United States. Phone: +001 (541) 531-3892; Fax: +001 (435) 797-1448, E-mail address:

Michael.Levin@usu.edu.

Abstract

Mobile apps are promising for teaching how to practice psychological skills in high-risk and in vivo momentary situations, but there has been minimal research on the immediate effects of app-based skill coaching on mental health in-the-moment. This study analyzed the mobile app data in a non-clinical sample of 39 adults participating in a larger randomized controlled trial, with participants randomized to an acceptance and commitment therapy (ACT) mobile app that tailors skill coaching based on in-the-moment variables ($n = 17$) or an app that provides randomly selected skill coaching ($n = 22$). Data was collected before and after each ACT skill coaching session on proximal outcome (depression, anxiety, engagement in meaningful activity) and ACT process variables. Multilevel models indicated significant immediate improvements on average following ACT skill coaching sessions on all proximal outcome and ACT process variables, although with relatively small effects ranging between .17 and .27 *SD* units change. Larger immediate pre-to- post effects from ACT coaching sessions were found for anxiety, experiential avoidance, and cognitive fusion in the tailored app versus random app condition. Overall, results suggest an ACT app can have immediate, in-the-moment effects on psychological functioning, which may be enhanced by tailoring skills to current context.

Keywords: mHealth; psychological inflexibility; mindfulness; micro-interventions; just-in-time adaptive interventions.

Comparing in-the-moment skill coaching effects from tailored versus non-tailored Acceptance and Commitment Therapy mobile apps in a non-clinical sample

Mobile apps offer a promising method for delivering in-the-moment psychological skills to improve mental health. The vast majority (77%) of adults in the United States own a smartphone (Pew Research Center, 2017), providing ready access to mental health resources throughout people's day in a convenient, and in some cases, anonymous format. Although research on mobile apps for mental health is still in its early stages, preliminary findings support their potential efficacy for addressing problems including depression, anxiety, and smoking cessation (Firth et al., 2017; Torous, Levin, Ahern & Oser, 2017).

Beyond the potential for mobile apps to increase access to alternate, affordable mental health supports, this modality offers unique opportunities for intervention. One of the most notable features is the capacity for mobile apps to tailor interventions in-the-moment to provide the right strategy, at the right time, based on personal data (often referred to as ecological momentary interventions or just-in-time adaptive interventions; Nahum-Shani et al., 2016). This affords unique opportunities to provide high frequency, low intensity interventions in an efficient way such that people receive the most effective skills at the most effective times to improve mental health. However, this is a new area and there has been minimal research to-date on how to implement such tailored interventions and their effects on mental health.

Acceptance and commitment therapy (ACT; Hayes, Strosahl & Wilson, 2012) is one treatment that has been adapted to mobile app formats that deliver tailored skills in the moment (*removed for blind review*). The ACT model specifies a set of pathological processes that can be targeted through ACT skills to improve mental health including acceptance skills to reduce experiential avoidance (i.e., excessive efforts to avoid or control unwanted inner experiences),

cognitive defusion skills to reduce cognitive fusion (i.e., excessive impact of thoughts on actions), present moment skills to reduce inflexible or limited attention to the present, and values skills to improve clarity regarding personal values (i.e., meaningful qualities of action). There is a large body of research indicating that these ACT components can impact mental health (Levin et al., 2012) and that reducing these pathological processes with ACT mediates treatment effects for problems including depression and anxiety disorders (Twohig & Levin 2017).

Based on this model, our research team developed an ACT mobile app that tailors ACT skills in-the-moment in relation to immediate pathological processes relevant to the app user. Two previous pilot open trials were conducted with a prototype of ACT Daily with clients on a waitlist for therapy (*removed for blind review*) and clients currently receiving ACT from a therapist (*removed for blind review*). These studies found that clients' significantly improved on depression and anxiety during a two-week period using the ACT Daily app (although note this was a single condition, open trial without a comparison condition). ACT Daily was designed to also assess whether in-the-moment variables collected before the skill coaching improved immediately following skill coaching to more precisely examine the immediate effects of skill coaching. Consistent with predictions, in-the-moment depression, anxiety, cognitive fusion, experiential avoidance, values disconnect, and lack of present moment awareness all improved following ACT skill coaching sessions (*removed for blind review*). Furthermore, the pilot trial with clients on a waitlist found that the immediate pre-to-post effects of each skill coaching session improved over days using the app (i.e., there were larger in the moment effects from each coaching session as participants used the app for more days) (*removed for blind review*). This suggests the immediate effects of ACT Daily on mental health may improve over time using the app. However, these results are very preliminary due to the sample size and open trial design.

A randomized controlled trial (RCT) was recently conducted in a non-clinical sample of adults seeking to improve their mental health to further evaluate ACT Daily (*removed for blind review*). Participants were randomized to standard ACT Daily (i.e., tailored skill condition), a version of ACT Daily where skill coaching sessions were not associated with check-in assessments (i.e., random skill condition), or a condition where users completed daily check-ins without any skill coaching (i.e., ecological momentary assessment [EMA] condition). This study found that the tailored skill version of ACT Daily led to significant improvements in mental health over four weeks relative to both the random skill and EMA conditions. Surprisingly, the random skill version of ACT Daily did not improve outcomes relative to EMA only. These results suggest that the tailoring of ACT skills is particularly relevant to the impact of ACT Daily on mental health, and surprisingly, delivering ACT skills that are not tailored to current pathological processes may lead to inert effects.

The intensive longitudinal data afforded by ACT Daily offers an opportunity to further examine the immediate, in-the-moment effects of the mobile app. This provides a more fine-grained method to further confirm whether the tailored app skill coaching does improve mental health and pathological processes in-the-moment. Even more importantly, this data allows for examination of why the random skill coaching version of ACT Daily failed to outperform an EMA-only version of the app (i.e., Were there any immediate improvements in-the-moment from completing non-tailored skill sessions? Were these effects weaker relative to skill coaching provided by the tailored app?). Finally, this data could be used to examine the effects of the app over time to inform future app design (e.g., how to maintain or enhance effects over time) and implementation (e.g., how long to recommend using an app).

The current study tested three predictions from the intensive longitudinal data collected in the larger ACT Daily RCT (*removed for blind review*). First, that completing any ACT skill coaching sessions would lead to immediate improvements in psychological flexibility and mental health (i.e., depression, anxiety, engagement in meaningful activity) in-the-moment. Second, that completing ACT skill coaching sessions through a tailored app would lead to greater in-the-moment improvements relative to an app that delivered random (non-tailored) ACT skills. A third exploratory prediction was that the within and between condition in-the-moment effects of the apps would remain consistent over four weeks (rather than weakening or strengthening over time), as indicated by number of days in the intervention failing to moderate effects.

Method

Participants

This study used data from a larger RCT conducted with 69 adults interested in using a mobile app to improve mental health. Inclusion criteria were being 18 years of age or older, owning an iPhone or Android, and reported interest in using a mobile app to improve well-being, with participants excluded for participating in previous clinical trials from our research lab. Recruitment primarily occurred at a university in the Mountain West region of the United States through flyers and online advertisements. Recruitment materials used the headline “test out a new mobile app for promoting well-being and learning psychological skills.” Due to the focus on in-the-moment effects from ACT Daily skill coaching, analyses were limited to the 39 participants who were assigned to the Tailored ACT Daily or Random ACT Daily app conditions. This excluded 30 participants who were either in the EMA-only condition (without skill coaching; $n = 24$) or never completed a skill session ($n = 6$).

Within the sub-sample of 39 participants included for the secondary analysis, 60% were female with a mean age of 21.85 ($SD = 5.18$). The sample was primarily White, non-Hispanic (86%), with the remaining participants identifying as Hispanic or Latino (8%), Black/African American (3%) and American Indian/Alaska Native (3%). Recruitment efforts largely occurred on the local university campus, which led to all but one participant being a college student.

This study included a broad sample of individuals who might choose to use a mobile app to address mental health, meaning that there were no inclusion criteria related to a specific form/degree of psychological disorder. In addition to increasing generalizability of the results to any adult who might use the app, this approach matches the transdiagnostic use of ACT, which is designed to reduce suffering and improve functioning across a wide spectrum of types and levels of psychological problems (Hayes, Pistorello & Levin, 2012). Although participants were not recruited or selected for based on level or form of distress, Depression, Anxiety and Stress Scale cutoff scores (DASS; Lovibond & Lovibond, 1995) can be used to provide a general estimate of the prevalence of notable psychological distress. Based on empirically established cutoff scores with the DASS (Lovibond & Lovibond, 1995), the majority of the sample (70%) endorsed moderate or greater symptoms of anxiety (≥ 10), depression (≥ 14), and/or stress symptoms (≥ 19).

Procedures

Overview. After completing an online baseline assessment, participants were randomized to one of three versions of the mobile app. Participants were oriented to using the mobile app over the following 4 weeks. A final online follow up assessment was completed 4 weeks after baseline (for more details regarding the study procedures see *removed for blind review*).

ACT Daily Design. The mobile app was delivered as a non-native app through Qualtrics and was only accessible online through a mobile website browser. Our previous research has found that this online survey platform (Qualtrics) provides sufficient capabilities to deliver engaging and efficacious online self-help interventions (e.g., *removed for blind review*). Furthermore, this platform delivers surveys within a responsive-design format, creating a visually appealing, easy-to-complete coaching session on any smartphone regardless of screen size or operating system.

Check-In Assessments. In both study conditions, notifications to complete an app check-in were sent via text messages twice a day over the four-week app period. Participants were also instructed on how to save a bookmarked icon on their phone's home screen to access the non-native app at any time. All app sessions (for both study conditions) began with completion of a check-in assessment. The check-in included 7 items assessing psychological functioning and relevant pathological processes for practicing ACT skills (see Measure section for further details).

Skill Coaching. For both the random and tailored app conditions, participants had three options after completing a check-in – 1) end the session, 2) quick skill, or 3) depth skill. Selecting “quick skills” directed participants to a very brief, text instructed skill to practice in the moment, with an option to see another quick skill or end the session. Selecting “depth skills” directed participants to a list of potential longer ACT skill training sessions to choose from (e.g., interactive exercises that span multiple pages, audio-guided exercises). Selecting “end session” would close the session without any skill training for users who did not want to practice skills at the time.

The ACT Daily skill coaching sessions were each organized by ACT component (acceptance, cognitive defusion, present moment awareness, and values). When a user selected depth or quick skills, they would be directed to a set of skills for only one ACT component for that app session (e.g., a user might have a cognitive defusion session where they can practice one or more defusion quick skills or select from a defusion depth skill to practice). Each ACT component included a library of 28 quick skills and 6 depth skills. For example, values quick skills included brief questions one could consider related to personal values, metaphors to help consider how values might apply in the moment, and suggestions for how to move towards one values with actions in the current moment. Values depth skills included an audio guided exercise to reflect on a valued moment (“sweet spot” exercise) and a values card sort interactive exercise to identify one’s top values. A 15-minute online training was provided for participants in the tailored and random skill coaching conditions immediately after assignment to treatment condition in order to ensure a basic understanding of the check-in assessment and ACT component skills delivered in the app.

Random Versus Tailored ACT Skill Assignment. The two skill coaching conditions differed in terms of how participants were directed to a specific ACT component skill set to practice. In the random app condition, participants were assigned to one of four ACT component skill sets randomly each app session. In other words, if a participant selected quick or depth skill after completing a check-in, they were randomly assigned to practice either acceptance, cognitive defusion, values, or present moment skills, rather than based on the data provided in their app check-in assessment.

In the tailored app condition, participants were assigned to one of four ACT component skill sets to practice based on their responses provided in the app check-in assessment that was

just completed. Tailoring was decided based on whatever relevant pathological process was rated the highest in the current app check-in assessment. If “fighting your feelings” was rated the highest, then they were directed to acceptance skills. Participants were directed to cognitive defusion if “stuck in thoughts” was rated highest, present moment if “running on autopilot” was highest, and values if “disconnected from values” was highest. Thus, the ACT component skill users were directed to was based on the relevant pathological process they reported struggling with the most in that moment. In all other ways, the random and tailored ACT Daily apps were identical.

Post Assessment. The last page of every ACT Daily skill coaching session, was a post skill coaching assessment. This post assessment was identical to the pre check-in assessment, except users also rated how helpful the skill coaching was. The post skill coaching assessment was designed to examine if there were any changes in psychological functioning or pathological processes immediately following practice of an ACT skill. Of note, app skill coaching sessions guided participants through practicing a relevant ACT skill in the moment (e.g., following an audio guided mindfulness exercise, completing an interactive worksheet, practicing a quick skill), and thus the post skill assessment was delivered immediately following actual practice of an ACT skill. The post skill assessment was only provided when a user selected to practice a depth or quick skill.

Measures

The pre and post check-in assessments were delivered through the ACT Daily mobile app. Items were rated on a visual analogue scale from 0 (none) to 100 (extremely) in relation to how much they were experiencing each of 7 variables. The first 3 items assessed proximal psychological outcome variables including depression (“feeling sad/depressed”), anxiety

(“feeling anxious/afraid”), and valued action (“unable to do what matters”). The last 4 items assessed relevant ACT pathological processes that were also used for tailoring in the tailored app condition including experiential avoidance (“fighting your feelings”), cognitive fusion (“stuck in thoughts”) inattention to the present (“running on autopilot”) and lack of values connection (“disconnected from values”). A brief description of what each item aimed to assess was provided during the online orientation to the app. Items were rated in relation to the prompt “how much are you...” pre skill session and “now, how much are you...” post skill session. The post check-in assessment included the exact same items, with an additional item assessing the helpfulness of the skills practiced in that session.

All of the check-in items had previously been tested in two initial pilot trials with the exception of the “unable to do what matters” item, which was added for the current study. The previous pilot trials found that check-in items all improved in response to ACT daily app coaching sessions (*removed for blind review*). To further support their validity, one study found targeted effects between what ACT skill component was practiced in the moment, and what pathological processes improved (e.g., fighting with feelings improved more after practicing acceptance skills than other ACT skills) (*blinded for review*).

Analytic Plan

Prior to examining treatment effects, preliminary analyses examined the rate and predictors of missing data. Preliminary analyses also tested for any between condition differences on the average pre check-in process or outcome items using a series of *t*-tests.

The central research questions examined average changes in dimensions of psychological functioning and pathological processes across time in the context of two mobile app experimental interventions. Scores were nested within individuals over time, thus, we used

multilevel modeling (MLM) to estimate average change and individual variability in the outcomes over time. Multiple pre-post intervention periods per day were possible. The exact time of intervention completion was used to center the time variable for each individual (i.e., group-mean centering), using *days* as the time metric.

To test the MLMs, we used R version 3.4.2 (R Development Core Team, 2015) within the R Studio version 1.1.383 (R Studio Team, 2017) environment. Restricted Maximum Likelihood (REML) estimation methods were used for all models, tested using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015). The analysis process included creating models for each outcome, which consisted of first fitting null-models to calculate the Intraclass Correlation Coefficient (ICC) value and determine how much variance of the outcomes were explained by individual differences in level (at the centered day timepoint) and in slope (change over time). Since individual differences explained a significant amount of variance for all outcomes measured (see Table 2), we included a random effect for participants and for time.

The final models included fixed effects for pre-post assessment (pre- versus post- score), study condition (random versus tailored), and time (number of days since baseline) as well as all two- and three-way interactions between these variables. The pre-post variable was a binary code (0 = pre, 1 = post) representing the degree of change occurring within each intervention period. This pre-post fixed effect tested the first prediction that relevant psychological outcomes (depression, anxiety, valued action) and pathological processes (experiential avoidance, cognitive fusion, inattention, disconnect with values) would improve on average from immediately before to immediately after completing skill coaching (indicating in-the-moment improvements following use of an app skill). The two-way pre-post by condition interaction tested the second prediction that the tailored app skill coaching condition would produce greater

pre-to-post improvements in-the-moment on average relative to the random app condition. To test the third exploratory predictions, interaction effects were examined between time and pre-post (do pre-to-post improvements get larger or smaller over time with the app) and a three-way interaction between time, pre-post, and condition (do differences between conditions in pre-to-post improvements change over time). Additional fixed effects in the model tested whether conditions differed on average check-in scores (condition effect), whether average check-in scores changed over time (time effect) and whether average check-in scores changed over time by condition (time * condition effect). Bootstrapped 95% confidence intervals were generated for all model parameters.

Effect sizes were calculated in two ways. First, we used the *MuMin* package (Bartón, 2017) to calculate marginal R^2 values, which represent the total variance explained by the model across fixed and random effects, for each model. Second, since the pre- to post-score effect was a central value of interest, we calculated change in standard deviation units by dividing the unstandardized estimate for the pre-post effect by the standard deviation for the outcome variable. This provided a common metric by which to compare the magnitude of change for each outcome variable.

Results

Preliminary Analyses

Overall, participants started 871 skills coaching sessions out of 1,375 app check-ins, selecting the “end session” option for the remaining 504 check-in assessments. Given the focus on pre to post skill coaching effects, analyses were only conducted with the check-in assessments that led to a skill coaching session (referred to as pre-intervention observations). Only 17 (2%) post-assessment scores were missing out of 871 pre-intervention assessments. These missing

post-assessments were due primarily to exiting the app in the middle of a skill coaching session. Given there was minimal missing post-intervention assessment data, we deleted pre-assessment observations that did not have a post-assessment score. When examining this by treatment group, the tailored group had 312 pre-intervention observations and 310 post-intervention observations (99% of total data), while the random treatment group had 559 pre-intervention observations and 544 post-intervention observations (97% of total data). Pre-intervention assessment observations that had missing post-outcome scores had higher mean levels on “feeling anxious/afraid” ($t [869] = 2.46, p = .01$) and “running on autopilot” ($t [869] = 3.17, p = .002$), but lower scores on “stuck in thoughts” ($t = [869] = -2.09, p = .04$), relative to complete pre-post observations. After removing observations with missing post-outcome data, we had a total of 1708 pre-post observations between 39 participants (Tailored condition = 17 participants, 620 observations, Random condition = 22 participants, 1,088 observations).

According to a series of t-tests, average pre-intervention scores did not significantly differ between the tailored and random treatment groups ($p > .10$). This was supplemented by the final estimated models in which treatment condition fixed effects were not significantly related to the outcomes (see Table 2), suggesting that conditions were equal at pre-intervention at their centered midpoint.

MLM Analyses

A series of models tested study predictions with each check-in variable including proximal outcomes (depression, anxiety, valued action) and pathological processes (experiential avoidance, cognitive fusion, inattention, values disconnect). Descriptive statistics for each check-in variable at pre and post intervention are reported in table 1 with model results in table 2. The three-way interaction was trimmed from all models given that it resulted in poorer model fit,

indicated by increases in Bayesian Information Criteria and Akaike Information Criteria relative fit indices.

For all proximal outcome and pathological process variables, scores significantly decreased from pre- to post-skill training as indicated by significant pre-post fixed effects (see Table 2). On average, there was a 0.27 *SD* units decrease in “feeling sad/depressed”, 0.17 *SD* decrease in “feeling anxious/afraid”, 0.25 *SD* decrease in “fighting your feelings” and “stuck in thought”, 0.24 *SD* decrease in “running on autopilot”, and 0.19 *SD* units decrease in “disconnected from values.” These pre-post effects were consistent with the prediction that completing ACT Daily skill coaching sessions would lead to immediate, in-the-moment improvements in psychological functioning and pathological processes targeted in ACT.

Significant condition by pre-post interactions were found for one proximal outcome (anxiety) and two pathological processes (experiential avoidance and cognitive fusion). Consistent with predictions, the random app condition reported smaller decreases in anxiety, "fighting feelings", and “stuck in thoughts” after completing a coaching session, compared to a larger decrease in the tailored app condition (see Table 2 and Figures 1 - 3). For the other four variables, the non-significant pre-post * condition interaction suggested there were equivalent reductions in outcomes across groups. These findings provided mixed support for the prediction that a tailored version of ACT Daily would produce stronger in-the-moment effects on psychological functioning and pathological processes.

For this primary study question, we conducted post-hoc power analyses with 500 simulated models to evaluate achieved power for condition * pre-post effects using the *simr* package (Green & McLeod, 2016). This tested whether increasing sample size would be adequate to detect the observed estimates as significant, assuming that the groups were truly

different in their average pre-post scores. Corresponding with the significant observed effects, we found that power was adequate (.80) for anxiety and “fighting feelings” at the current sample size, although markedly lower for “stuck in thoughts” (.50). For the other four outcome variables, even relatively large increases in the sample size (75 per condition) would not achieve adequate power to detect the observed effects as significant (see Figure 4 for example with the “running on autopilot” outcome).

To test the third exploratory prediction, time by pre-post interactions tested whether the size of pre to post skill training effects varied over days with the app irrespective of condition (i.e., whether app effects remained consistent over time in the 4 week intervention). Of the seven check-in variables, only one pathological process (cognitive fusion) was significant. Irrespective of app condition, participants reported a smaller decrease in cognitive fusion (“stuck in thoughts”) after completing a skill coaching session as the 4-week app period progressed (see Table 2 and Figure 3). Pre-post skill effects on the remaining variables did not change over the app usage period, suggesting most effects neither strengthened or weakened over time.

There were no significant condition main effects indicating that average check-in scores were equivalent between conditions at the sample grand-mean centered pre-test time point, and there were no significant main effects for time, suggesting that pre-scores did not increase or decrease on average across the study period. In addition, there were no differences in average check-in scores over time by condition, indicated by non-significant time * condition effect in all models.

Discussion

This study examined the in-the-moment effects of providing tailored (or non-tailored) ACT skill coaching through a mobile app with a nonclinical sample of adults interested in

improving their wellbeing. Results supported the prediction that proximal psychological outcomes and pathological processes would improve on average immediately after practicing ACT skills, albeit effect sizes were relatively small. Results were more mixed for the prediction that the tailored version of ACT Daily would produce greater pre-post effects relative to a random (non-tailored) version, which was only the case for anxiety, cognitive fusion, and experiential avoidance. Effects were generally stable over days using ACT Daily, with the only exception being that effects on cognitive fusion diminished over time across study conditions. Overall, these findings from mobile app data add to the results of the RCT (*removed for blind review*), indicating that the ACT app skill coaching sessions produce small in-the-moment improvements in psychological functioning and related processes, particularly when skill coaching is tailored to current challenges.

The research literature on mobile apps for mental health is still relatively new, with only a few RCTs published on cognitive behavioral apps, or ACT more specifically (Torous et al., 2017). Although well-controlled RCTs are essential for determining whether such apps are efficacious, there are a wide range of additional methodologies afforded by mobile apps that could shed further light on their causal impact, mechanisms of action, and areas for future refinement. The current study is an example of the additional information and opportunities to test research questions provided by intensive longitudinal data collected in an ACT mobile app.

The ACT Daily mobile app provided immediate data before and after each skill session, allowing for a precise test of the proximal effects of the app. This can further clarify how ACT Daily works. For example, in the RCT the random app did not outperform the assessment-only condition (*citation removed for blind review*). Interestingly, it appears that the random app still produced in-the-moment effects on proximal variables, suggesting that non-tailored skill

coaching was not inert. This raises additional questions to further understand the impact of tailoring such as whether non-tailored skill training may produce immediate, but not long lasting effects, and why that may occur. For example, it may be that teaching *any* ACT skill is useful in the moment, but, as a result, individuals do not learn when to use what skills in the future so they have less effect over time. Alternatively, these pre-post effects across conditions are likely to also include placebo and related method effects, raising the question of whether non-tailored ACT skill coaching is no more effective than any other seemingly helpful app intervention.

The stronger effects from the tailored app condition on anxiety, experiential avoidance, and cognitive fusion suggests that an app can produce greater effects when skills match current context and challenges. These findings lend further support for the active impact of the tailored ACT Daily app, given effects were stronger than a plausible alternative condition without tailoring. It is worth noting that in the RCT the tailored and random app had similar high satisfaction rates and participants were equally likely to believe the random app actually tailored skill coaching based on their assessment responses despite these sessions being random (*removed for blind review*). In addition, post hoc power analyses indicated that the observed pre-post * condition effect sizes on depression, valued action, inattention, and being disconnected from values were small enough that they would continue to be non-significant even with a large sample size of 75 participants per condition. Thus, results seem to indicate that tailoring ACT skills improve in-the-moment mobile app effects for a targeted subset of variables including anxiety, experiential avoidance, and cognitive fusion, and that these effects are not readily ruled out by potential confounds such as app usage or credibility.

One question this study raises is when and how ACT skills need to be tailored in-the-moment. These results suggest it is particularly important to provide acceptance skills when high

in experiential avoidance and cognitive defusion skills when high in cognitive fusion, while alternatively any ACT skill is useful when inattention or disconnect from values are the current issue in-the-moment. Momentary experiential avoidance and cognitive fusion are arguably the two most central, driving aspects of psychological inflexibility, defining when behavior is rigidly under the control of cognitions or attempts to avoid/escape aversive internal states (Bardeen & Fergus, 2016). In contrast, inattention to the present and a lack of connection with values are more indicative of deficits in aspects of psychological flexibility, the primary therapeutic change process in ACT. Thus, strengthening in-the-moment deficits in psychological flexibility (e.g., values clarification, present moment attention) may not require tailoring ACT skills, or at least not tailoring based on corresponding ACT components. In contrast, reducing in-the-moment excesses of psychological inflexibility defined by experiential avoidance and cognitive fusion may require tailoring so that the most directly relevant ACT skill is used to reduce the most salient aspect of inflexibility in-the-moment. This is somewhat consistent with a previous pilot trial of ACT Daily in which acceptance skill coaching was found to have a larger in-the-moment effect on experiential avoidance than other ACT components, suggesting this component is most effective for reducing momentary avoidance (*removed for blind review*). Future research is needed to further clarify when and how best to tailor ACT coaching skills in-the-moment.

The stronger in-the-moment effects for a tailored ACT app on anxiety further raises the question of whether tailoring is more effective for a subset of mental health problems related to anxiety. It is possible based on the current findings that non-tailored ACT apps could be equally effective for depression, at least on in-the-moment effects. Future RCTs with clinical samples defined by more specific presenting problems are needed to examine which types of clients benefit from tailored skill coaching.

It is important to note that the effect sizes for skill coaching in-the-moment were small, indicating that ACT coaching sessions had only a modest immediate impact on proximal outcome and process variable. This may have been due in part to the use of a non-clinical sample, which might have created floor effects on measures as indicated by the relatively low average pre check-in scores. Alternatively, the in-the-moment effects from this app (and similar apps) may actually be fairly small. Consistent with these findings, a dialectical behavior therapy (DBT) coach app that similarly assessed effects immediately after each coaching session found only modest changes in emotional intensity (pre coaching $M = 6.83$, $SD = 2.09$ vs. post coaching $M = 5.69$, $SD = 2.31$) in a sample of outpatient DBT clients with comorbid substance use and borderline personality disorder (Rizvi et al., 2011). This is in some ways to be expected given the high frequency, low intensity intervention approach used in these apps, in which the strategy is to engage users in frequent, small changes that produce large effects on mental health over time. With both ACT Daily and the DBT coaching app, continued completion of these coaching skills led to large improvements in mental health outcomes over time (*removed for blind review*; Rizvi et al., 2011). Thus, previous research indicates that although the immediate, in-the-moment effects of such mobile apps are relatively small, they have an ongoing impact on mental health that leads to robust effects over time.

This study also examined whether the effects of the ACT Daily app changed over time, based on previous pilot research suggesting effects may strengthen as participants use the app over more days (*removed for blind review*). The current study found that the pre-post in-the-moment effects within and between mobile app conditions did not change over days using the app, suggesting effects were relatively stable. The only exception was a significant pre-post * time effect for cognitive fusion, which suggested there were smaller momentary effects on fusion

over time across the two app conditions. Such time-related analyses are important for cognitive behavioral mobile app research to begin answering practical questions such as how long it will take to see benefits from an app, when might using an app start to have diminishing returns, and how long to recommend using an app for. A related method used in ACT mobile app research has been to test relations between program usage patterns and outcomes, which in addition to finding specific exercises that predict stronger outcomes (Heffner et al., 2015), have also found that usage of apps over more weeks predicts stronger outcomes (Mattila et al., 2016). Future studies might test ACT Daily for longer to further examine such time effects including if/when the app starts to produce diminishing effects and whether tailoring helps address this.

This study had notable limitations, particularly due to sample characteristics and measurement issues. The study included a non-clinical, fairly homogenous sample of mostly White college students. Furthermore, there were no inclusion criteria related to level or form of psychological distress, although the majority (70% reported moderate or greater distress). As a result, it is unclear the degree to which in-the-moment coaching effects and tailoring might generalize or vary with more diverse or specific clinical samples. A tailored ACT app might function quite differently with clients actively seeking help for a significant mental health concern, including coaching either having less of an effect (due to not being impactful enough for more significant current challenges) or more of an effect (due to ACT skills being more applicable when significant challenges occur). Effects may also vary by client presenting problem as indicated by differences between conditions on anxiety, but not depression in the current study. As further indication of limitations with the non-clinical sample, low average pre check-in scores were found, which might have reduced effect sizes and increased the potential for false negatives. Thus, the results from this study should be interpreted cautiously and it is

important future research replicate and extend these analyses of in-the-moment effects from ACT mobile apps in more distressed and specific clinical presentations.

The study also used a set of unvalidated items designed to inform the app coaching sessions. However, these items have been used in previous trials and found to relate to differential effects of specific ACT components (*removed for blind review*), suggesting the check-in items provide a valid assessment of distinct pathological processes. Future research using rigorous psychometric methods is needed to determine the degree to which the intended constructs are being reliably and validly assessed with the items used in this study.

Finally, the timing of the post coaching check-in assessment immediately following a coaching skill may have reduced sensitivity to detecting effects. Previous laboratory-based component studies have found improvements in variables immediately following a brief ACT intervention (e.g., Healy et al., 2008; Levin et al., 2012). Yet, completing an assessment directly after practicing an ACT skill may have been too brief of a time period to fully detect effects on proximal outcome and process variables. Future research might include a follow up EMA assessment some time after the coaching session (e.g., 15 minutes later), which may be more sensitive to detecting the effects of skill coaching at their maximal impact and could help clarify the duration of coaching effects over time.

In conclusion, this study adds to a growing body of literature indicating the potential efficacy of mobile apps based on ACT and other forms of cognitive behavior therapy. Overall, the results suggest that ACT mobile apps can have an immediate effect in-the-moment on psychological flexibility and mental health, albeit with relatively small effect sizes. Furthermore, it appears that tailoring ACT skills produces greater effects on some key aspects of psychological flexibility and mental health. Although the study is limited by the use of a non-clinical sample,

results suggest ACT mobile apps can have immediate effects on psychological functioning, particularly when tailored based on current psychological challenges in-the-moment.

References

- Bardeen, J.R., & Fergus, T.A. (2016). The interactive effect of cognitive fusion and experiential avoidance on anxiety, depression, stress and posttraumatic stress symptoms. *Journal of Contextual Behavioral Science*, 5, 1–6.
- Bartoń, K. (2017). MuMIn: Multi-Model Inference. R package version 1.40.0. <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1-48.
- Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S. & Sarris, J. (2017) Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 218, 15-22.
- Green, P., & MacLeod, C. J. (2016). simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493-498.
- Hayes, S.C., Pistorello, J. & Levin, M.E. (2012). Acceptance and Commitment Therapy as a unified model of behavior change. *The Counseling Psychologist*, 40, 976-1002.
- Hayes, S.C., Strosahl, K.D., & Wilson, K.G. (2011). *Acceptance and Commitment Therapy: The process and practice of mindful change*. New York: The Guilford Press.
- Healy, H., Barnes-Holmes, Y., Barnes-Holmes, D., Keogh, C., Luciano, C. & Wilson, K. (2008). An experimental test of a cognitive defusion exercise: Coping with negative and positive self-statements. *The Psychological Record*, 58, 623-640.
- Heffner, J.L., Vilardaga, R., Mercer, L.D., Kientz, J.A. & Bricker, J.B. (2015). Feature-level Analysis of a Novel Smartphone Application for Smoking Cessation. *The American Journal of Drug and Alcohol Abuse*, 41, 68-73

- Levin, M.E., Hildebrandt, M., Lillis, J. & Hayes, S.C. (2012). The impact of treatment components suggested by the psychological flexibility model: A meta-analysis of laboratory-based component studies. *Behavior Therapy, 43*, 741-756.
- Lovibond, S.H. & Lovibond, P.F. (1995). *Manual for the Depression Anxiety Stress Scales, (2nd ed.)*. Sydney, AU: Psychology Foundation of Australia.
- Mattila, E., Lappalainen, R., Välikynen, P., Sairanen, E., Lappalainen, P., ... Ermes, M. (2016). Usage and dose response of a mobile acceptance and commitment therapy app: Secondary analysis of the intervention arm of a randomized controlled trial. *JMIR Mhealth Uhealth, 4*, e90.
- Nakagawa, S, Schielzeth, H. (2013). A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. *Methods in Ecology and Evolution, 4*, 133–142.
- Nahum-Shani, I., Smith, S.N., Spring, B.J., Collins, L.M., Witkiewitz, K., Tewari, A. & Murphy, S.A. (2016). Just-in-time adaptive interventions (JITAI) in mobile health: Key components and design principles for ongoing health support. *Annals of Behavioral Medicine, 1-17*.
- Pew Research Center (2017). Mobile Fact Sheet. Retrieved June 26, 2017, from <http://www.pewinternet.org/fact-sheet/mobile/>
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rizvi, S.L., Dimeff, L.A., Skutch, J., Carroll, D. & Linehan, M.M. (2011). A pilot study of the DBT coach: An interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. *Behavior Therapy, 42*, 589-600.

RStudio Team (2017). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA

URL <http://www.rstudio.com/>.

Torous, J.B., Levin, M.E., Ahern, D. & Oser, M. (2017). Cognitive behavioral mobile applications:

Research literature, marketplace data, and evaluation guidelines. *Cognitive and Behavioral Practice, 24*, 215-225.

Twohig, M.P. & Levin M.E. (2017). Acceptance and commitment therapy as a treatment for

anxiety and depression: A review. *Psychiatric Clinics, 40*, 751-770.

Table 1. Descriptive statistics for average pre- and post-intervention check-in scores by study condition.

	<u>Tailored App</u>		<u>Random App</u>	
	Pre <i>M</i> (<i>SD</i>)	Post <i>M</i> (<i>SD</i>)	Pre <i>M</i> (<i>SD</i>)	Post <i>M</i> (<i>SD</i>)
“Feeling sad/depressed”	17.80 (21.62)	13.41 (16.74)	16.09 (21.48)	13.31 (18.69)
“Feeling anxious/afraid”	21.95 (22.10)	15.01 (16.93)	19.03 (21.10)	15.59 (20.10)
“Unable to do what matters”	18.25 (18.98)	13.33 (15.57)	17.18 (19.38)	13.60 (16.91)
“Fighting your feelings”	18.53 (21.85)	12.15 (17.38)	15.40 (17.40)	12.12 (15.45)
“Stuck in thoughts”	23.87 (23.86)	17.23 (20.82)	17.33 (19.72)	13.79 (17.69)
“Running on autopilot”	18.35 (22.64)	11.55 (15.43)	17.62 (21.36)	12.83 (16.59)
“Disconnected from values”	11.84 (15.18)	8.17 (11.12)	9.61 (16.70)	6.51 (12.22)

Table 2. MLM results for each proximal outcome and pathological process.

Parameter	Depression <i>b</i> 95% <i>CI</i>	Anxiety <i>b</i> 95% <i>CI</i>	Valued Action <i>b</i> 95% <i>CI</i>	Experiential Avoidance <i>b</i> 95% <i>CI</i>	Cognitive Fusion <i>b</i> 95% <i>CI</i>	Inattention <i>b</i> 95% <i>CI</i>	Values Disconnect <i>b</i> 95% <i>CI</i>
Intercept	19.21*** [10.76, 26.72]	23.03*** [14.08, 32.54]	19.99*** [13.00, 26.88]	20.86*** [13.95, 29.51]	25.45*** [18.02, 33.52]	20.19*** [14.75, 24.65]	12.59*** [5.04, 19.51]
Condition	-1.38 [-12.16, 8.75]	-0.96 [-13.22, 11.44]	3.02 [-5.96, 13.59]	-2.28 [-12.72, 8.25]	-4.91 [-15.42, 4.50]	-1.48 [-7.57, 5.18]	2.46 [-5.97, 12.22]
Time	0.06 [-.28, .39]	-0.26 [-.58, .06]	-0.24 [-.59, .11]	0.00 [-.35, .37]	-0.31 [-.67, .06]	-0.10 [-.56, .34]	-0.21 [-.46, .01]
Pre-Post	-4.32*** [-6.69, -2.08]	-6.80*** [-8.94, -4.52]	-4.78*** [-7.20, -2.50]	-6.38*** [-8.33, -4.51]	-6.37*** [-8.56, -3.80]	-6.78*** [-9.16, -4.39]	-3.59*** [-5.39, -1.85]
Condition * Pre-Post	1.57 [-1.17, 4.52]	3.43** [.61, 5.98]	1.27 [-1.52, 4.09]	3.10* [.47, 5.81]	2.97* [.14, 5.75]	1.99 [-1.33, 5.27]	0.53 [-1.82, 2.83]
Time * Pre-Post	0.07 [-.09, .23]	0.11 [-.03, .25]	0.13 [-.02, .30]	0.00 [-.13, .13]	0.23** [.05, .40]	0.02 [-.15, .21]	0.07 [-.06, .19]
Time * Condition	-0.12 [-.50, .32]	0.35 [-.09, .77]	0.23 [-.20, .60]	-0.10 [-.63, .40]	-0.04 [-.51, .44]	-0.05 [-.67, .57]	0.03 [-.28, .35]
Null Model ICC	0.52	0.62	0.48	0.53	0.47	0.25	0.56
Marginal R^2	0.59	0.69	0.55	0.64	0.56	0.41	0.62

Note. *** $p < .001$, ** $p < .01$, * $p < .05$. Significant effects are bolded for clarity. *b* = unstandardized model estimate; *CI* = confidence interval; ICC = Intraclass Correlation Coefficient. Marginal R^2 = Total variance explained by model fixed and random effects.

Figure Caption

Figure 1. Plot of Estimated Anxiety Scores Over Time (Days) by Study Condition with 95% Confidence Interval Around Trajectories.

Figure 2. Plot of Estimated “Fighting Your Feelings” Scores Over Time (Days) by Study Condition with 95% Confidence Interval Around Trajectories.

Figure 3. Plot of Estimated “Stuck in your Thoughts” Scores Over Time (Days) by Study Condition with 95% Confidence Interval Around Trajectories.

*Figure 4. Simulated Power Curve for Observed Estimate of Condition * Pre-Post Interaction Effect for Inattention (“Running on Autopilot”) Outcome.*







