

Evaluating an mHealth Intervention for First Responders, Military Personnel and Veterans: Randomized Controlled Trial

Morgan K Dunphy, Heather J. Nuske

Submitted to: Journal of Medical Internet Research
on: January 10, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 50

Figures..... 51

 Figure 1 54

 Figure 2 55

Multimedia Appendixes..... 60

 Multimedia Appendix 0 61

 Multimedia Appendix 0 61

 Multimedia Appendix 1 61

 Multimedia Appendix 2 61

Evaluating an mHealth Intervention for First Responders, Military Personnel and Veterans: Randomized Controlled Trial

Morgan K Dunphy^{1*}; Heather J. Nuske^{1*}

¹Department of Psychiatry Perelman School of Medicine University of Pennsylvania Philadelphia US

*these authors contributed equally

Corresponding Author:

Heather J. Nuske
Department of Psychiatry
Perelman School of Medicine
University of Pennsylvania
3535 Market St
3rd Floor
Philadelphia
US

Abstract

Background: First responders, military personnel and veterans face a disproportionate risk for mental health and wellness issues due to the stress of their occupations. Stigma and confidentiality concerns are common barriers to use of traditional mental health services. Digital interventions offer a promising alternative, as they can be anonymous, convenient, and cost-effective.

Objective: This study aimed to test GUIDE, a digital wellness app designed for first responders, military personnel, and veterans. We explored the impact of GUIDE on various aspects of wellness, emotional wellbeing, mental health, social connectedness, and personal growth.

Methods: This randomized waitlist-controlled trial enrolled 115 participants allocated into 3 groups: GUIDE with financial incentives (GUIDE+Incentives, n=37), GUIDE-only (n=39), or waitlisted GUIDE control (waitlist, n=39). Surveys assessed baseline and post-trial wellness (PERMA total score, WHO-5 score, Personal Wellbeing Score, PERMA Health sub-scale score), emotional wellbeing (PERMA Positive, Negative, and Happiness sub-scale scores; Difficulties in Emotion Regulation Scale), mental health/health (PHQ-8 for depression and GAD-7 for anxiety), social connectedness (PERMA Relationships and Loneliness sub-scale scores), and personal growth (PERMA Accomplishment, Meaning, and Engagement sub-scale scores). App engagement and technical merit were also evaluated.

Results: Overall, 93.04% of enrolled participants (107/115) completed the post-trial assessment and were included in our intent-to-treat analyses. In repeated measures ANOVAs, there were no significant group $\times$ time interactions. In post-hoc pairwise comparisons of pre-post deltas per group, the GUIDE-only group improved significantly over the Waitlist in wellness, emotional wellbeing, and mental health. No group differences were significant for social connectedness or personal growth measures. Of participants allocated to GUIDE intervention groups, 67.10% (51/76) completed at least three activities per week for all four weeks. GUIDE+Incentives completed significantly more activities, posts and replies compared to the GUIDE-only group – but not lessons, likes, or mood surveys. App engagement was correlated with improvements on wellness, emotional wellbeing, mental health, and personal growth measures, driven mostly by lessons, posts and replies – the educational and peer support aspects of GUIDE. Participants who engaged more felt more accomplished, regardless of feature used. Participants gave GUIDE good appropriateness and feasibility scores, and satisfactory acceptability and usability scores. Exploratory subgroup analyses suggest the app may be most beneficial to military-affiliated individuals and males.

Conclusions: This trial indicates that GUIDE-only participants improved significantly compared to the waitlist and that GUIDE is a feasible and appropriate intervention with the potential to improve mental health and wellbeing of first responders, veterans and active military. Financial incentives increased engagement with peer support aspects of GUIDE, but did not lead to significant improvements over the waitlist. Future research should test the effect of longer-term GUIDE use and whether improvements sustain. Clinical Trial: clinicaltrials.gov (NCT06336967)

(JMIR Preprints 10/01/2025:71155)

DOI: <https://doi.org/10.2196/preprints.71155>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://www.jmir.org/>

Original Manuscript

Evaluating an mHealth Intervention for First Responders, Military Personnel and Veterans: Randomized Controlled Trial

Morgan Dunphy [1] & Heather J. Nuske [1]

[1] Penn Center for Mental Health, University of Pennsylvania, Philadelphia, PA, USA

Preprint
JMIR Publications

Abstract

Background: First responders, military personnel and veterans face disproportionate risk for mental health and wellness issues due to occupational stress. Stigma and confidentiality are common barriers to traditional mental health services. mHealth interventions offer an anonymous, convenient, and cost-effective alternative.

Objective: This study tested GUIDE, an mHealth intervention, by assessing its impact on wellness, emotional wellbeing, mental health, social connectedness and personal growth in first responders, military personnel and veterans.

Methods: In this unblinded randomized waitlist-controlled trial, 115 participants recruited online and offline were virtually enrolled and allocated into 3 groups by the research coordinator: GUIDE with financial incentives (GUIDE+Incentives, $n=37$), GUIDE-only ($n=39$), or waitlisted GUIDE control (waitlist, $n=39$). Web-based surveys assessed baseline and post-trial wellness (PERMA Overall score, WHO-5 Well-being, Personal Wellbeing Score, PERMA Health score), emotional wellbeing (PERMA Positive emotion, Negative emotion, and Happiness scores; Difficulties in Emotion Regulation Scale), mental health (PHQ-8 for depression and GAD-7 for anxiety), social connectedness (PERMA Relationships and Loneliness scores), and personal growth (PERMA Accomplishment, Meaning, and Engagement scores). App engagement and technical merit were also evaluated.

Results: Overall, 93.04% of enrolled participants (107/115) completed the post-trial assessment and were included in our main analysis. In repeated measures ANOVAs, there were no significant group $\times$ time interactions for wellness ($\eta p^2=.02-.04$), emotional wellbeing ($\eta p^2=.01-.12$), mental health ($\eta p^2=.01-.02$), social connectedness ($\eta p^2=.03-.04$) or personal growth ($\eta p^2=.01-.02$). In post-hoc pairwise comparisons of pre-post deltas per group, the GUIDE-only group improved significantly over the Waitlist in wellness ($d=.05-.31$), emotional wellbeing ($d=.06-.31$), and depression ($d=-.30$) and anxiety ($d=-.29$) symptoms. No group differences were significant for social connectedness ($d=-.20-.20$) or personal growth ($d=-.04-.06$). Of participants allocated to GUIDE interventions, 67.10% (51/76) completed at least three activities per week for all four weeks. GUIDE+Incentives completed significantly more activities ($t(72)=2.01$, $P=.05$), posts ($t(72)=2.15$, $P=.04$) and replies ($t(72)=3.40$, $P=.001$) compared to the GUIDE-only group – but not lessons ($t(72)=.73$, $P=.47$), likes ($t(72)=1.22$, $P=.23$), or mood surveys ($t(72)=.48$, $P=.63$). App engagement correlated with improvements on wellness, emotional wellbeing, mental health, and personal growth measures, driven mostly by lessons, posts and replies – the educational and peer support aspects of GUIDE. Participants who engaged more felt more accomplished, regardless of feature used ($r=.23-.36$). GUIDE received good appropriateness ($M=4.01$) and feasibility ($M=4.01$), and satisfactory acceptability ($M=3.81$) and usability ($M=70.62$), scores. Exploratory subgroup analyses suggest the app may be most beneficial to military-affiliated individuals and males.

Conclusions: GUIDE-only participants improved significantly compared to the waitlist. GUIDE is a feasible and appropriate intervention with the potential to improve mental health and wellbeing of first responders, veterans and active military. Financial incentives increased engagement with peer support features, but did not lead to significant improvements over the waitlist. Future research should test longer-term GUIDE use and whether improvements sustain.

Trial Registration: Clinicaltrials.gov (NCT06336967)

Keywords: *digital health, mHealth, first responders, military, veterans, wellness, resilience, depression, anxiety, randomized trial*

Introduction

There is an urgent need to address mental suffering, isolation, and burnout among first responders, active-duty military and veterans. Due to the nature of their occupations, they are disproportionately exposed to trauma, stress, and sleep disturbances, which increase their risk of mental health conditions including depression[1–4], anxiety[5–9], post-traumatic stress disorder (PTSD)[10–12], and substance abuse[13–20]. In the wake of the COVID-19 pandemic, professional stress has only compounded for first responders, who report that heightened workloads, isolation, increased intensity, and negative public sentiment have come at emotional, physical and mental costs[21,22]. These conditions in turn put both military and first responders at an elevated risk for suicide as compared to the general civilian population[23–25], to the extent that police officers and firefighters are more likely to die by suicide than in the line of duty[26]. This crisis comes with urgent costs to the public, since mental health and suicidality can result in absenteeism, problems with retention, higher error rates, and decreased job performance – which can lead to security risks at the national and local levels[27,28].

Despite the severity and prevalence of mental health issues, as well as growing awareness of the importance of mental health from military personnel and first responders[29], services to address this crisis continue to go under-utilized, including by those who need them most[30]. First responders and military personnel share concerns about stigma, and fear being ostracized by peers, family and leadership for seeking treatment or asking for help[31,32]. They pride themselves on strength and do not want to be perceived as weak or soft, which leads them to neglect symptoms until they are severe[22,30]. Additionally, the time and cost associated with traditional mental health care can be prohibitive for shift workers and those with busy schedules[33].

Research indicates that education and training about mental health and wellness can play an important role in reducing stigma, increasing awareness, and cultivating a more resilient mindset[22,30,31,34,35]. The challenge comes from making mental health and self-care training accessible and appealing to first responders and military personnel specifically, and importantly, emphasizing its confidentiality to address any privacy concerns. Peer support, another key strategy for building resilience and promoting wellbeing among warriors, faces similar challenges, since stigma and concerns about professional ramifications inhibit first responders and military from reaching out to one another about their concerns, especially face-to-face[34,36–38].

Overcoming Barriers with a Digital Solution

Existing solutions for mental health and wellness are not adequate in addressing the needs of first responders, veterans and active-duty military. Warriors are in need of interventions that take their unique needs and barriers into consideration. mHealth interventions offer a promising and scalable alternative to existing solutions, with several advantages over traditional clinical models. Namely, they can be anonymous, confidential, convenient, cost-effective, and asynchronous– which in combination can counteract stigma, time and cost barriers[39,40]. While digital interventions are relatively new, studies indicate that they can be efficacious in reducing depressive symptoms, anxiety levels and isolation[41–44], and that they can play a proactive role in promoting wellbeing and preventing potential mental health issues from becoming severe[45–47].

One digital mental health intervention built explicitly for the warrior community is The GUIDE App[48]. GUIDE is a smartphone application that uses educational videos, prompts, small discussion groups, and surveys to build resilience and promote wellness among first responders and military personnel. The GUIDE team has conducted extensive research with stakeholders from across first responder and military communities to build a compelling platform that hinges upon two evidence-based approaches to wellness: education and peer support. Crucially, it does so with complete anonymity. The platform does not store personal identifiable information, so users can engage with the app without confidentiality concerns.

GUIDE provides wellness education through its learning management system of courses that focus on topics relevant to first responders and military (e.g., Financial Wellness, Communication, Stress

Reduction, Physical Wellness, and Meditation). Each course is made up of one to four micro-lessons. GUIDE users can discuss the content of their lessons, as well as anything that is on their mind, in an assigned small group chat – the peer support element of The GUIDE App. The small group chat is comprised of 15-30 peers from across the country and is moderated by GUIDE employees, so members can get the support they need in a positive environment, from people who understand.

In a 4-week quasi-experimental pilot study with 16 participants from a U.S. police department, GUIDE with financial incentives showed promising results in promoting wellbeing, reducing anxiety and encouraging personal growth[49]. Paired-samples pre-post *t*-tests on PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Overall and Personal Wellbeing Score (PWS) scores showed that users experienced statistically significant improvements in overall wellbeing ($P=.007$) and accomplishment ($P<.001$), and reductions in anxiety ($P=.002$).

Current Study

In this study, we aimed to test the impact of GUIDE – with and without financial incentives – on the wellness, emotional wellbeing, mental health, social connectedness, and personal growth of first responders, military personnel, and veterans through a fully powered randomized waitlist-controlled trial. For our Theory of Change model, see Figure 1 in Appendix 1. Participants were assigned to one of three groups (GUIDE access with financial incentives [‘GUIDE+Incentives’], GUIDE access without financial incentives [‘GUIDE-only’], waitlisted GUIDE access [‘Waitlist’]) for the 4-week trial. Additional objectives included exploring app engagement, assessing the influence of financial incentives on app usage, and reporting on implementation outcomes and technical merit (e.g., feasibility and user-friendliness). Subgroup analyses were conducted to explore whether there were variations in outcomes for first responders versus those with military affiliation, as well as by sex.

Methods

Study Design

This study is a fully powered, unblinded randomized waitlist-controlled trial designed to evaluate how The GUIDE App – a digital health intervention for first responders, military personnel, and veterans – impacts wellness, emotional wellbeing, mental health, social connectedness, and personal growth. Participants were allocated to one of three groups (GUIDE access with financial incentives [‘GUIDE+Incentives’], GUIDE access without financial incentives [‘GUIDE-only’], waitlisted GUIDE access [‘Waitlist’]) for the 4-week trial.

Recruitment

Recruitment for this study took place online and in-person through social media posts, email lists, existing GUIDE partnerships, and local events for first responders/military. Recruitment materials invited potential participants to take an eligibility screener survey on REDCap, a HIPAA-Compliant web application for building and managing online surveys and databases. To be eligible, respondents had to be first responders, veterans, or military personnel in the U.S. with full-time employment, willing to download the GUIDE app on their personal device, not receive regular compensation from Nlyten Corp/GUIDE, and not have previously used GUIDE. Respondents who previously used GUIDE, received regular compensation from GUIDE/Nlyten Corp, expressed disinterest in downloading GUIDE on their phone, were employed less than full time, or were not first responders, veterans or military personnel were considered ineligible.

Because this was an app-based study with web mitigated interactions, technical literacy was also a “de facto” eligibility criterion. To avoid selection bias against non-digital or older first responders, we used in-person recruitment meetings where GUIDE personnel provided helped first responders, veterans and military sign up for the study via iPad.

mHealth intervention: The GUIDE App

GUIDE is a digital wellness app built with warriors in mind. It uses a small group chat, a learning management system, a drivers of behavior change model, and an anonymous user experience to

promote social connectedness, personal growth, resilience and wellbeing in a way that is accessible to the target population. See Figure 1 below for a summary of the GUIDE model. GUIDE has been implemented in several use cases and care settings, including police departments and fraternal orders, veterans' non-profits, veteran service organizations, government (e.g. military), and other organizations in support of the delivery of mental health programming. As of December 2024, GUIDE has over 18,000 users, which they refer to as members.



Figure 1. The GUIDE App Model.

During app onboarding, members personalize GUIDE by answering questions about their wellness interests and experience, as well as their background. To ease concerns about confidentiality, GUIDE does not collect Personal Identifying Information (PII), and instead safeguards these data with Okta, an industry leader in identity and login privacy. The anonymity of GUIDE is a key strength of the app, since it gives first responders, military personnel and veterans predisposed to stigma concerns a way access to mental health and health interventions, without risking exposing their vulnerabilities to family, peers, or leaders.

After set-up, members reach the app homepage, where they fill out a brief questionnaire to calculate their baseline wellness score which modulates as tasks are completed. Members may retake the initial questionnaire at any point to see how their wellness score compares. Underneath, the homepage contains a personalized Daily Plan, designed to take 15 minutes or less, which aggregates recommended content from across the app (e.g., lessons, surveys, peer posts) into a single list to help members get the most out of GUIDE. Beyond the homepage, GUIDE is organized into four tabs—Group, Courses, Practices, and Profile—as described in Table 2 of Appendix 1. Micro-lessons were developed by a team of experienced professionals specialized in social work, trauma-informed care, and product design. Please see Appendix 2 for a full list of micro-lessons, organized by category.

GUIDE recommends that members complete 3 lessons and 3 mood tracker surveys per week, as well as 3 posts, 3 replies and 3 likes in their small group chat, to get the most benefits from the app. Members may turn on notifications to promote sustained engagement.

Primary Outcome Measures

Wellness

PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Overall score (PERMA Overall), our primary outcome variable, is a composite score from the PERMA-Profiler designed to measure self-reported flourishing in adults across five factors/sub-scales: positive emotion, engagement, relationships, meaning, and accomplishment[50]. It contains 15 items rated on an 11-point Likert scale, from 0 to 10. Scores are calculated as the average of the items comprising each factor/sub-scale, with higher scores indicate higher flourishing. A score ≤ 6.4 is considered sub-optimal, while a score ≥ 8.0 is considered high[51]. It has demonstrated high internal consistency

across settings ($\alpha=.92-.95$)[52].

Personal Wellbeing Score (PWS) is a self-reported measure based on the Office of National Statistics four subjective well-being questions (ONS4) and thresholds. It contains 4 items, rated on a scale of 0 to 4, where 0="Disagree," 1="Neutral," 2="Agree," and 3="Strongly Agree." The summary score is calculated as an aggregate of the four items, transformed to a 0-100 scale, with higher scores indicating higher wellbeing. It shows good internal reliability ($\alpha=0.90$)[53].

World Health Organization Well-being Index (WHO-5) is a self-reported measure of current mental wellbeing[54]. It consists of 5 statements rated on a 5-point scale in relation to the previous two weeks, where 5="All of the time," 4="Most of the time," 3="More than half of the time," 2="Less than half of the time," 1="Some of the time," and 0="At no time." Scores are calculated by adding up each item and multiplying the total by 4, with 0 representing the worst imaginable well-being and 100 representing the best imaginable well-being. A clinical cutoff score ≤ 50 indicates low well-being, while ≤ 28 indicates clinical depression[54]. It has high internal consistency and reliability ($\alpha=.81-.90$), is valid in screening for depression, has strong construct validity among adolescents and adults, and is sensitive to change in interventions[54,55].

PERMA Health sub-scale (Health) is a three-item filler measure from the larger PERMA-Profiler, used to assess perceived physical health and vitality. Each item is rated from 0 to 10. Scores are the items averaged, with higher scores indicating better perceived health. It is not a part of the main PERMA model and is instead used to provide additional context. It has high reliability across settings and contexts ($\alpha=.86-.93$)[52].

Emotional Wellbeing

PERMA Positive emotion sub-scale (Positive emotion) is a three-item subscale from the PERMA-Profiler, designed to measure positive emotions, including joy, hope, pleasure, happiness and contentment[50]. Each item is ranked from 0 to 10. Scores are the average of items, with higher scores indicating greater positivity. A score ≤ 6.4 is considered sub-optimal, while ≥ 8.0 is considered high[51]. It has demonstrated good internal consistency across settings ($\alpha=.71-.89$)[52].

PERMA Happiness sub-scale (Happiness) is a single filler item from the PERMA-Profiler used alongside the primary PERMA model to provide additional context: "Taking all things together, how happy would you say that you are?" It is rated from 0 to 10, with a higher score indicating more happiness[52].

PERMA Negative sub-scale (Negative emotion) is a three-item filler measure from the PERMA-Profiler used to assess tendencies toward feeling sad, anxious or angry. Each item is rated from 0 to 10. Scores are calculated as the average of items. Lower scores are better (i.e., less negativity). It is not a part of the main PERMA model and is instead used to provide additional context. It has acceptable reliability and has been shown to be negatively correlated with wellbeing measures ($\alpha=.63-.75$)[52].

Difficulties in Emotion Regulation Scale–Short Form (DERS-SF; Emotion dysregulation) is an 18-item self-reported measure that assesses for emotion dysregulation. It includes six subscales with three items each, including Strategies, Non-acceptance, Impulse, Goals, Awareness, and Clarity. Each item is rated on a 5-point scale of frequency ranging from 1 ("Almost Never 0-10%") to 5 ("Almost Always 91-100%"). Items are summed, with the Awareness items reverse coded, to create a total score of 18 to 90, with higher scores indicating greater difficulties with emotion regulation[56]. It is a validated measure for assessing emotion regulation problems and possesses high internal consistency ($\alpha=.89-.91$)[57].

Mental Health

Patient Health Questionnaire (PHQ-8) is an 8-item diagnostic and severity measure for assessing depression disorders in large clinical settings[58]. Statements about the previous two weeks are rated on 4-point scale ranging from 0 ("not at all") to 3 ("nearly every day"). Scores are reported on a scale of 0 to 24, with higher scores indicating a greater severity in depressive symptoms. Scores above 10 indicate clinically significant depressive symptoms. It has high internal consistency reliability ($\alpha =$

0.82) and high validity; 96.5% of those who score above clinical cutoff have been found to have a depression diagnosis[58,59].

Generalized Anxiety Disorder-7 (GAD-7) consists of 7 items rated on a 4-point Likert scale ranging from 0 ("not at all") to 3 ("nearly every day") with the same options as in the PHQ-8[60]. Items are summed to create a total severity score ranging from 0 to 21, with higher scores indicating greater symptom severity. Scores above 10 indicate clinically significant anxiety symptoms. It has high internal consistency, reliability ($\alpha=0.89$) and validity, with moderate correlations with related constructs such as depression and self-esteem[61].

Social Connectedness

PERMA Relationships sub-scale (Relationships) consists of 3 items from the larger PERMA-Profiler that measure whether a person feels cared about, authentically connected, and securely connected to others[50]. Each item is rated on a scale from 0 to 10. Scores are the average of items, with higher scores indicating higher social connectedness. A score ≤ 6.4 is considered sub-optimal, while ≥ 8.0 is considered high[51]. It has demonstrated high internal consistency across settings ($\alpha=.75-.85$)[52].

PERMA Loneliness sub-scale (Loneliness) is a single filler item from the PERMA-Profiler used alongside the primary PERMA model to provide additional context: "How lonely do you feel in your daily life?" It is rated from 0 to 10, with a lower score indicating less loneliness. It is negatively correlated with the PERMA wellbeing measures[52].

Personal Growth

PERMA Accomplishment sub-scale (Accomplishment) contains 3 items from the PERMA-Profiler that measure drive toward personal goals[50]. Each item is rated from 0 to 10. Scores the average items, with higher scores indicate higher accomplishment. A score ≤ 6.4 is considered sub-optimal, while ≥ 8.0 is considered high.[51] It has demonstrated good internal consistency across settings ($\alpha=.70-.86$)[52].

PERMA Meaning sub-scale (Meaning) contains 3 items from the PERMA-Profiler that measure sense of purpose.[50] Each item is rated from 0 to 10. Scores are the items averaged, with higher scores indicate greater sense of meaning. A score ≤ 6.4 is considered sub-optimal, while ≥ 8.0 is considered high.[51] It has demonstrated high internal consistency across settings ($\alpha=.85-.92$)[52].

PERMA Engagement sub-scale (Engagement) contains 3 items from the PERMA-Profiler that measure absorption, interest and involvement in the world.[50] Each item is rated from 0 to 10. Scores the items averaged, with higher scores indicate greater engagement. A score ≤ 6.4 is considered sub-optimal, while ≥ 8.0 is considered high.[51] It has acceptable internal consistency ($\alpha=.60-.81$)[52].

Secondary Outcome Measures

App Engagement

App engagement is assessed by the total and weekly sum of activities completed in the app during the trial period, calculated as the sum of lessons, posts, replies, likes, and mood surveys.

Weekly sums are assessed relative to recommended benchmarks from GUIDE (≥ 3 lessons, posts, replies, likes, and mood surveys), as well as minimum usage standards (≥ 3 activities).

Implementation Outcomes

Acceptability of Intervention Measure (AIM) assesses how agreeable or likeable an intervention is to stakeholders. It contains 4 items rated on a 5-point Likert Scale from 1 ("Completely disagree") to 5 ("Completely agree"). Scores are the average of items, with higher scores indicating greater acceptability. It has been shown to have good internal consistency and test-retest reliability ($\alpha=.83-.85$)[62].

Intervention Appropriateness Measure (IAM) assesses the perceived fit of an intervention to a specific setting. It contains 4 items rated on a 5-point Likert Scale from 1 ("Completely disagree") to 5 ("Completely agree"). Scores are the average of items, with higher scores indicating greater appropriateness. It has been shown to have high internal consistency and test-retest reliability ($\alpha=.87-.91$)[62].

Feasibility of Intervention Measure (FIM) assesses how easily an intervention can be carried out within a given setting. It contains 4 items rated on a 5-point Likert Scale from 1 ("Completely disagree") to 5 ("Completely agree"). Scores are the average of items, with higher scores indicating greater feasibility. It has been shown to have good internal consistency and test-retest reliability ($\alpha=.88-.89$)[62].

Technical Merit

System Usability Scale (SUS) assesses perceptions of usability of a given system (in this case, GUIDE)[63]. It contains 10 items rated on a 5-point Likert Scale from 1 ("Strongly disagree") to 5 ("Strongly agree"). Responses are transformed to get a total range of possible values from 0-100, with higher scores indicating better usability. It has high internal consistency and reliability ($\alpha=0.91$) [64]. Scores above 68 are regarded as a satisfactory or above average, while scores above 80 are considered good[65].

Procedure

Potential participants from across the US were instructed to take an eligibility survey on REDCap. To be eligible, respondents had to be first responders, veterans, or military personnel in the U.S. with full-time employment, willing to download the GUIDE app on their personal device, not receive regular compensation from Nlyten Corp/GUIDE, and not have previously used GUIDE. Respondents deemed eligible were immediately directed to a copy of an IRB-approved informed consent and HIPAA authorization for virtual signature. According to institutional guidelines, the informed consent stressed that this study would not influence their employment status, that they could opt out at any time without consequence. A copy of this informed consent has been uploaded to Clinicaltrials.gov.

After providing consent, REDCap automatically emailed baseline assessment survey link, which included demographic questions and our self-report outcome measures (the PERMA-Profiler, WHO-5, PWS, GAD-7, PHQ-8, and DERS-SF). Consented respondents received a separate welcome email sent manually by the research team, instructing them to schedule a virtual intake meeting. Participants were asked to complete the baseline assessment prior to their scheduled intake meeting. Prior to intake, participants were randomly assigned to one of three study arms (GUIDE+Incentives, GUIDE-only, or Waitlist control). The clinical research coordinator conducted a manual block randomization procedure in Microsoft Excel, using the RAND() function to generate predetermined blocks at the beginning of the study. Participants were assigned to blocks and treatment groups sequentially based on the order in which they signed up for an intake meeting. If a participant missed their intake meeting, their assigned spot was given to the next enrolled participant to maintain allocation balance. Since app onboarding and financial incentives were explained to participants during the intake meeting, this study was unblinded to both participants and study personnel.

Intake meetings were conducted by the clinical research coordinator over Zoom. The coordinator ensured that each participant understood what the study entailed and answered outstanding questions before informing them of their assigned treatment group. Regardless of assigned group, all participants used this time to schedule a post-trial follow-up meeting for the end of the trial period, approximately 4 weeks out from the intake meeting. Afterward, participants from the Waitlist condition were informed that they would receive access to GUIDE during the post-trial meeting and dismissed.

Participants assigned to GUIDE+Incentives and GUIDE-only groups were instructed to download and set-up the app on their personal device. They were provided with an invitation code associated with the research study, since GUIDE is an invite-only platform. After setting up the app, participants were given a brief overview of GUIDE features, including the home, group, courses and practices tabs. Participants were told that if a higher level of care was required while using the app, that they should navigate to the in-app SOS button and use the available crisis hotlines and resources.

Next, participants were informed about weekly usage benchmarks recommended by the GUIDE team (at least 3 lessons, small group posts, replies to peers, likes to peer posts, and mood surveys),

while those assigned to the GUIDE+Incentives group were additionally told that they would receive compensation for reaching these recommended benchmarks. GUIDE+Incentives participants earned Amazon gift cards of increasing value for each week they met GUIDE's criteria: \$10 for week 1 of the trial, \$20 for week 2, \$30 for week 3, and \$40 for week 4. Gift cards were emailed to participants by GUIDE on the Tuesday after they were earned.

After finishing this intake meeting, participants were considered enrolled in the study. All enrolled participants received \$50 from the research team for completing their intake meeting, regardless of treatment group. The day of the intake meeting was considered day 1 of the 28-day trial period. During the trial, participants in the GUIDE+Incentives and GUIDE-only treatment groups were asked to use GUIDE as recommended to support their wellness. Participants in these conditions received text message reminders from the GUIDE team if they did not meet the minimum weekly usage by day 5 of each trial week.

On the last day of the trial, participants automatically received a link from REDCap to their post-trial assessment. The post-trial assessment was comprised of the same primary outcome measures from the baseline assessment (PERMA-Profilier, WHO-5, PWS, GAD-7, PHQ-8, and DERS-SF). Additionally, assessments for participants assigned to GUIDE+Incentives and GUIDE-only conditions also included SUS, AIM, IAM, and FIM measures. After completing the post-trial assessment, participants received a \$50 payment from the research team for their time.

During the virtual post-trial follow-up meeting with the clinical research coordinator, participants in the waitlist control received access to The GUIDE App and were provided the same overview that the GUIDE+Incentives and GUIDE-only conditions received during intake; while participants in the GUIDE+Incentives and GUIDE-only conditions provided feedback about their experience using the app, which the research team aggregated and provided to the GUIDE team.

There were no changes to our methods after trial commencement and deviations to the registered protocol did not occur; however, periodic bugs impacting participant access and use of GUIDE were mentioned in open-ended survey responses and during post-trial meetings.

Ethical Considerations

The study was conducted in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines for digital health interventions[66]. Ethical approval was obtained from the University of Pennsylvania Institutional Review Board (IRB) and is registered on Clinicaltrials.gov (NCT06336967). We obtained informed consent and HIPAA authorization via virtual signature for all participants. The informed consent stressed that this study would not influence their employment status, that they could opt out at any time without consequence. At the conclusion of the trial, data was anonymized and de-identified for future analysis.

Participants were compensated \$100 for their time and effort – \$50 for their completed baseline assessment and intake meeting, and \$50 for completing their post-trial assessment.

Data analysis

To test how GUIDE access impacted our primary outcomes of wellness, emotional wellbeing, mental health, social connectedness and personal growth, we ran a three-way repeated measures analysis of variance (ANOVA) (3 groups: GUIDE+Incentives, GUIDE-only, waitlist) for each measure. Qualification for this analysis was based on participant completion of both the baseline and post-trial assessments (and therefore did not include participants lost-to follow-up).

We applied Bonferroni corrections to control for Type I errors (false positives) arising from multiple comparisons, given the large number of outcome measures in this study. Bonferroni was chosen for its strict alpha control requirements, ensuring that the risk of any false positive findings remained below the conventional significance threshold of 0.05 when performing multiple statistical tests as was required for this analysis[68]. However, it is acknowledged that the Bonferroni correction is a conservative method, which increases the risk of Type II errors (false negatives), particularly when many tests are conducted.

To test how GUIDE app engagement related to wellness, emotional wellbeing, mental health, social

connectedness and personal growth, we conducted an exploratory analysis of the relationships between app engagement (total, weekly, and by feature) for each of our outcome measures (change between pre-post) using Pearson correlation coefficients. To assess the extent to which participants engaged with GUIDE, we ran descriptive statistics for each of our app engagement outcome variables. To test whether financial incentives influenced engagement, we ran an independent samples *t*-test to assess how financial incentives impacted engagement with GUIDE features. To assess GUIDE for implementation outcomes and technical merit, we reported on descriptive statistics.

In addition to our main analysis, we conducted post-hoc exploratory subgroup analyses to assess whether first responder status, military affiliation or sex impacted outcomes by re-running three-way repeated measures analyses of variance (ANOVAs) (3 groups: GUIDE+Incentives, GUIDE-only, Waitlist) for select primary outcome measures.

We included all randomized participants who completed their post-trial assessment in our analysis, regardless of app usage during the trial period. App engagement, implementation outcomes and technical merit analyses included fewer participants, since only the GUIDE+Incentives and GUIDE-only groups completed these measures. Participants who did not complete their post-trial assessments were included in app engagement analyses, but excluded from analyses of primary outcomes, implementation outcomes and subgroup analyses. All analyses were conducted in SPSS 30, while data visualizations were created in Excel.

To achieve 80% power for three groups based on lowest effect size ($d=0.40$) and correlation ($r=.47$) of the pre-post *t*-tests from the pilot study[49], our trial required at least 68 participants (~23 per group) to 108 participants (~36 per group). Based on these parameters, we were powered for our main analysis, as well as our first responder and male subgroup analyses – but not for our military or female subgroup analyses. Results from underpowered subgroups should be interpreted with caution.

Results

Demographics

Data was analyzed for skewness, kurtosis, and outliers by calculating z-scores. Based on the established 3.29 cutoff [67], PHQ-8, GAD-7, and DERS-SF scores were positively skewed, therefore we log transformed these variables for our main analysis. App engagement metrics (total and by feature) were also positively skewed and therefore log transformed for later analysis. After transformation, our outcome measures were normally distributed and therefore parametric tests were used.

Next, data were assessed for differences between treatment groups. We found that there were differences between groups at baseline for some of our outcome measures, see Table 1 below. In our analyses, we controlled for baseline where differences between group were significant ($p<.05$).

Table 1. Outcome measures at baseline by treatment group

Baseline Outcome Measure	GUIDE+Incentives N=39	GUIDE only N=37	Waitlist N=39	One-Way ANOVA
	Mean (SD)	Mean (SD)	Mean (SD)	P
<i>Wellness</i>				
PERMA Overall	6.57 (1.43)	7.16 (1.65)	6.45 (1.43)	.09~
PWS	67.52 (22.52)	77.25 (21.57)	60.9 (24.04)	.01*
WHO-5	53.03 (21.32)	57.3 (21.87)	52.21 (20.66)	.54
Health	6.01 (1.69)	6.80 (1.83)	5.54 (1.77)	.01**
<i>Emotional wellbeing</i>				
Positive emotion	6.07 (1.81)	6.61 (2.09)	5.89 (1.64)	.21
Happiness	6.64 (1.91)	6.89 (1.90)	6.31 (1.81)	.40
Negative emotion	4.45 (2.18)	3.81 (1.93)	4.22 (1.74)	.36
DERS-SF	36.54 (12.59)	31.19 (10.17)	39 (12.46)	.02*
<i>Mental health</i>				

GAD-7	13.36 (5.73)	11 (4.34)	13.54 (4.32)	.04
PHQ-8	7.36 (5.15)	5.14 (5.05)	7.51 (5.19)	.08~
<i>Social connectedness</i>				
Relationships	6.56 (2.07)	7.43 (2.1)	6.14 (1.94)	.02*
Loneliness	3.69 (2.99)	3.57 (2.77)	5.13 (3.10)	.04*
<i>Personal growth</i>				
Accomplishment	6.83 (1.6)	7.6 (1.44)	6.74 (1.6)	.03*
Meaning	6.91 (2.03)	7.77 (1.95)	6.74 (1.88)	.06*
Engagement	6.46 (1.56)	6.39 (1.55)	6.72 (1.39)	.60

Note. ** indicates $P < .01$. * indicates $P < .05$. ~ indicates $P = .05-.10$. PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

All 115 enrolled participants were first responders, active-duty military and/or veterans and were block randomized into one of three treatment groups during enrollment: GUIDE+Incentives, GUIDE-only, or waitlist. Treatment groups were matched on all participant demographic variables (age, race, sex, ethnicity, education, years with employer, responder type, and receipt of behavioral healthcare) indicating that the randomization to the treatment versus control groups was successful, see Table 2 below.

Table 2. Participant characteristics by treatment group

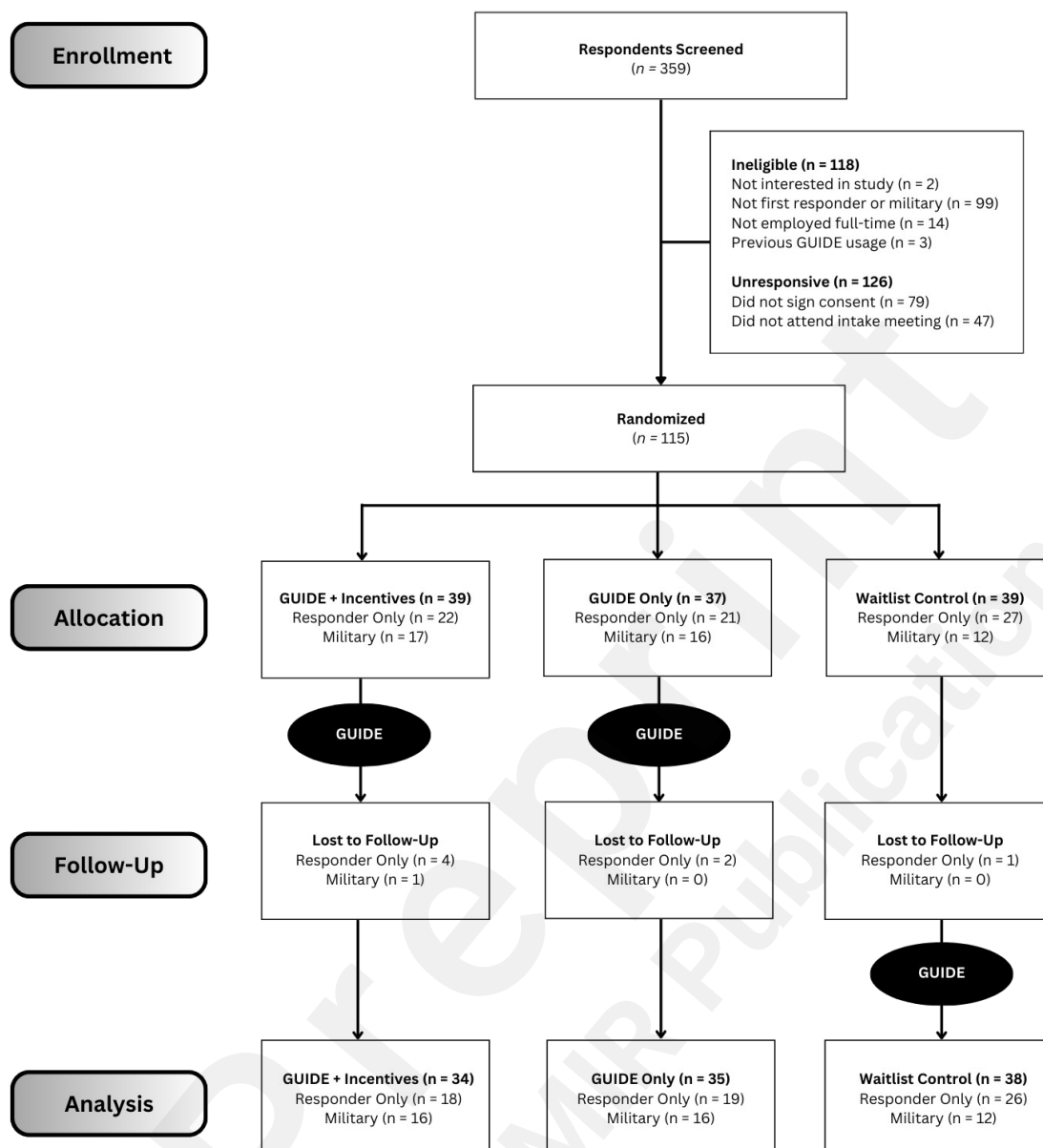
	GUIDE+Incentives N=39	GUIDE Only N=37	Waitlist Control N=39	Comparison Statistics
	Mean (SD) / % (n)	Mean (SD) / % (n)	Mean (SD) / % (n)	P / χ^2
Age	42.77 (10.28)	42.22 (11.43)	43.69 (10.82)	.84
Race				
White	87.18% (34)	86.49% (32)	92.31% (36)	.68
Black or African American	10.26% (4)	8.11% (3)	2.56% (1)	.39
American Indian/Alaska				.37
Native	2.56% (1)	0.00% (0)	0.00% (0)	
Hispanic/Latino	5.13% (2)	10.81% (4)	0.00% (0)	.11
Asian	0.00% (0)	0.00% (0)	5.13% (2)	.14
Refused to Answer	0.00% (0)	2.70% (1)	0.00% (0)	.35
Sex				.78
Male	69.23% (27)	75.68% (28)	69.23% (27)	
Female	30.77% (12)	24.32% (9)	30.77% (12)	
Ethnicity				.11
Hispanic or Latino	5.13% (2)	10.81% (4)	0.00% (0)	
Education				.47
High School Diploma	8.11% (3)	7.69% (3)	7.69% (3)	
Some College	16.22% (6)	23.08% (9)	30.77% (12)	
College Graduate	45.95% (17)	53.85% (21)	33.33% (13)	
Graduate/Professional				
Training	29.73% (11)	15.38% (6)	28.21% (11)	
Years w/ Employer				.49
< 1 year	16.22% (6)	17.95% (7)	15.38% (6)	
1 - 2 years	2.70% (1)	5.13% (2)	12.82% (5)	
2 - 5 years	13.51% (5)	15.38% (6)	15.38% (6)	
5 - 10 years	21.62% (8)	25.64% (10)	25.64% (10)	
10 - 15 years	8.11% (3)	17.95% (7)	15.38% (6)	
15 - 20 years	10.81% (4)	5.13% (2)	0.00% (0)	
20 + years	27.03% (10)	12.82% (5)	15.38% (6)	
Responder Type				.29
Responder Only	56.76% (21)	56.41% (22)	69.23% (27)	
Military Only	24.32% (9)	25.64% (10)	7.69% (3)	
Military AND Responder	18.92% (7)	17.95% (7)	23.08% (9)	
Household Income				.59
\$25,000-\$34,999	2.56% (1)	2.70% (1)	2.56% (1)	
\$35,000-\$49,999	2.56% (1)	5.41% (2)	2.56% (1)	
\$50,000-\$74,999	15.38% (6)	8.11% (3)	10.26% (4)	
\$75,000- \$99,999	10.26% (4)	18.92% (7)	10.26% (4)	
\$100,000-\$149,999	33.33% (13)	24.32% (9)	41.03% (16)	
\$150,000-\$199,999	17.95% (7)	8.11% (3)	15.38% (6)	
\$200,000 or more	7.69% (3)	27.03% (10)	10.26% (4)	
Prefer not to answer	10.26% (4)	5.41% (2)	7.69% (3)	
Behavioral Healthcare				.29
Yes	46.15% (18)	45.95% (17)	61.54% (24)	
No	51.28% (20)	56.76% (21)	38.46% (15)	

Note. From one-way ANOVA / chi-squared test.

Enrollment & Dropout Analysis

Of the 115 participants enrolled, 93.04% ($n=107$) completed a post-trial assessment. Please see Figure 2 below for a CONSORT flow diagram of enrollment and dropout rates by assigned treatment group.

Changes to self-reported primary outcome measures could not be assessed for participants that did not complete a post-trial assessment ($n=8$), and therefore these cases were excluded from the analysis of our primary outcomes. While imputing data or using pairwise deletions may have been a stronger approach[69,70], we used listwise deletions in this instance since our investigation was fully powered without the participants lost to follow-up. To determine whether participants who dropped out were significantly different from those who completed their post-trial assessments, Welch's t -tests and chi-squared tests were run to compare demographic features, baseline outcome measures, and overall app activity level. Only two significant differences were found. First, that participants who dropped out were less likely to have received Behavioral Healthcare in the past, $\chi^2(1)=5.18$, $P=.02$; and second, that among participants assigned to GUIDE+Incentives and GUIDE-only treatment groups, those who dropped out used the app significantly less over the 4-week trial period than those who completed their post-trial assessment, $t(72)=-8.18$, $P<.001$. These differences suggest that dropout may have been influenced by baseline characteristics and engagement levels, which could introduce bias by overestimating intervention effects, since participants less likely to seek help for mental health and less likely to use the intervention were not included in our analysis of primary outcomes.



Primary Outcomes

To assess how GUIDE influenced wellness, emotion wellbeing, mental health, social connectedness and personal growth, repeated measures ANOVAs were conducted to examine the effect of time and treatment group on each outcome measure. Post-hoc pairwise comparisons were included to explore between-group and within-group differences. See Table 3 below for a complete summary of statistics from all three tests.

Table 3. Effect sizes for repeated measures ANOVAs and pairwise comparisons for all participants.

ANOVA Effects	Pairwise Comparisons (Group*Time)			Pairwise Group Comparisons of Pre-Post Deltas		
Time	T i m e x	Incentives	GUIDE only Waitlist	Incentives vs. WL	GUIDE only vs. WL	Incentives vs. GUIDE only

Outcome measure	$\eta p2^a$	Group $\eta p2^a$	M		$\eta p2^a$		MD^c		d^b		MD^c		d^b		MD^c	
			$\eta p2^a$	MD^c	$\eta p2^a$	MD^c	d^b	MD^c	d^b	MD^c	d^b	MD^c	d^b	MD^c		
Wellness																
PERMA Overall	.07**	.0	.07**	.0	.02	.00	0.09	.12	0.35	.31*	0.91*	-.19	-.056			
PWS †	.16**	.0	.04~	.0	.00	.00	-0.10	.17	2.61	.05	0.83	.11	0.09			
WHO-5	.08**	.0	.03~	.0	.07**	.00	1.47	.08	3.08	.27~	10.25	-.19	-.015			
Health †	.16**	.0	.05*	.0	.07**	.00	0.00	.20	0.20	.22	0.23	-.03	-.003			
Emotional wellbeing																
Positive emotion	.10**	.0	.06**	.0	.05*	.01	0.14	.15	0.51	.31*	1.09*	-.17	-.013			
Happiness	.04*	.1	.01	.0	.06*	.00	-0.05	.16	0.55	.30*	1.06*	-.14	-.051			
Negative emotion	.15**	.0	.10**	.0	.03~	.04*	0.46*	.04	0.15	-.15	-0.51	.19	0.66			
In_DERS-SF †	.11**	.0	.08**	.0	.00	.01	-0.03	-.17	-0.04	.06	0.01	-.23	-.005			
Mental health																
In_PHQ-8	.14**	.0	.08**	.0	.04*	.04~	0.17	-.05	-0.07	-.30*	-0.44*	.25	0.37			
In_GAD-7 †	.12**	.0	.02	.0	.04*	.00	-0.00	-.23	-0.06	-.29*	-0.07*	.06	0.01			
Personal growth																
Accomplishment †	.16**	.0	.03	.0	.02	.01	0.13	.10	0.11	.05	0.06	.05	0.04			
Meaning †	.08**	.0	.04*	.0	.00	.00	-0.03	.16	0.20	.06	0.06	.13	0.10			
Engagement	.01	.0	.05	.0	.13	.00	0.00	.00	0.01	-.04	-0.11	-.04	-.011			
Social Connectedness																
Relationships †	.13**	.0	.04*	.0	.04*	.00	-0.08	.22	0.25	.20	0.24	.01	0.01			
Loneliness †	.07**	.0	.02	.0	.07**	.00	-0.10	-.09	-0.19	-.20	-0.43	.11	0.24			

Significance level with Bonferroni adjustment for multiple comparisons: *** $P < .001$, ** $P < .01$, * $P < .05$, ~ $P = .05$ -.1. PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model [50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

† controlled for baseline.

a Interpretation of $\eta p2^a$: .01=small effect, .06=moderate effect, .11=large effect. [71,72]

b Interpretation of Cohen's d : .2=small effect, .5=moderate effect, .8=large effect. [72,73]

c Based on adjusted means.

Overall, the main effect of time was highly significant with moderate to large effect sizes for all measures except Engagement. Time×group interactions were not significant for any outcome measures; however, time×group was borderline significant for DERS-SF. This result was driven by the GUIDE+Incentives group, which showed a significant reduction in DERS-SF (i.e., emotion dysregulation) over time with a large effect size, while GUIDE-only and Waitlist groups did not.

In post-hoc comparisons by group, GUIDE+Incentives showed greater improvement than the waitlist on 93.33% of measures (14/15), though none of these differences were significant. GUIDE-only

showed greater improvement than the waitlist on 86.67% of measures (13/15), which were significant for wellness (PERMA Overall), emotional wellbeing (Positive emotion, Happiness), and mental health (PHQ-8 for depression, GAD-7 for anxiety). While differences between GUIDE+Incentives and GUIDE-only groups were not significant in any between-group pairwise comparisons, overall, the GUIDE-only group outperformed GUIDE+Incentives on 66.68% of measures (10/15) (particularly for emotional wellbeing and mental health).



Wellness

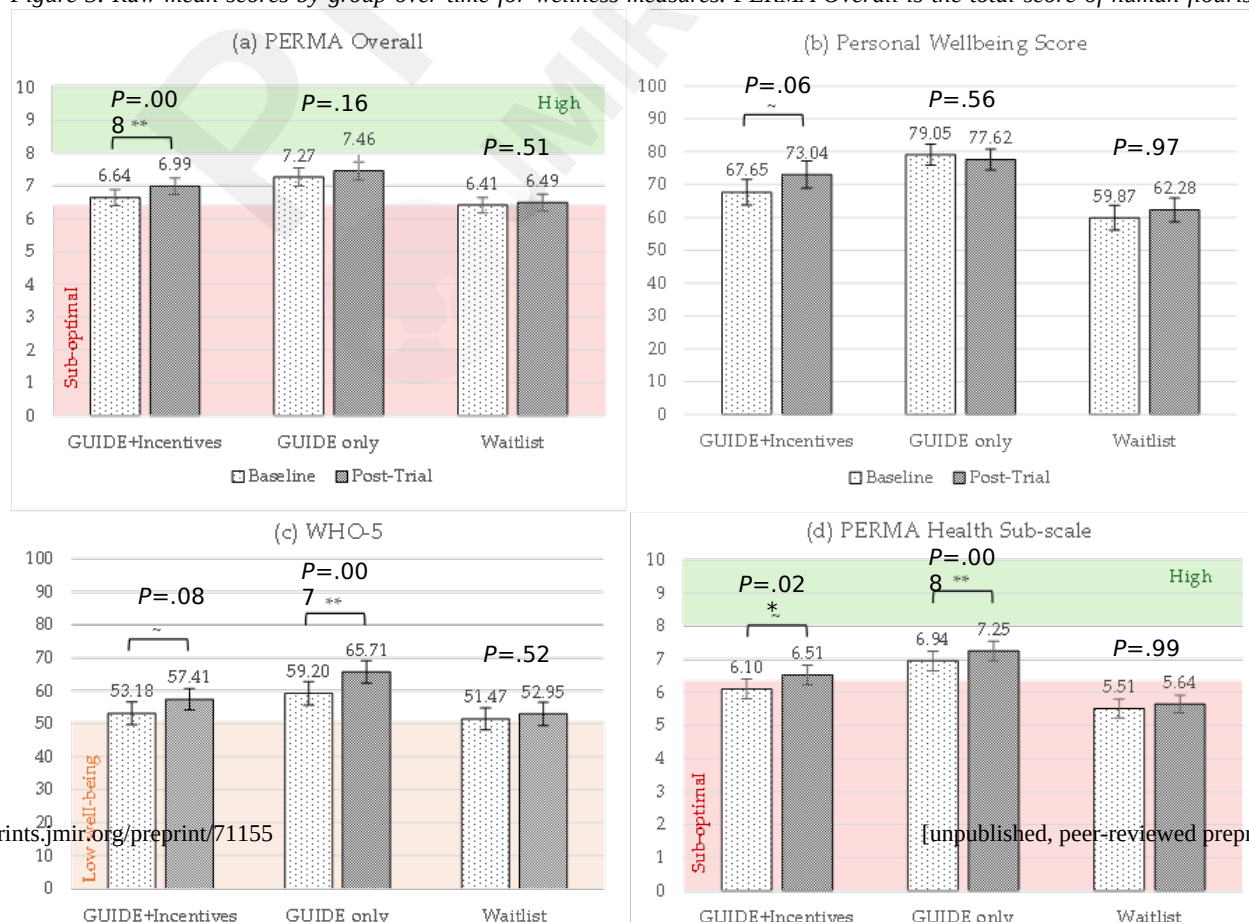
For our primary outcome variable, PERMA Overall, the main effect of time was significant with a moderate effect size $F(1,104)=7.78$, $P<.01$, $\eta_p^2=.07$. The time $\times$ group interaction was not significant $F(2,104)=1.15$, $P=.32$, $\eta_p^2=.02$. Post-hoc pairwise comparisons revealed that the GUIDE-only group experienced a significantly greater improvement on PERMA Overall when compared to the waitlist control (*raw mean* .11, $P=.03$, $d=0.31$), a difference which translates to a 1% percentile point increase in raw PERMA Overall score, or a 9.1% percentile-point increase when controlling for time. Raw scores for all three groups fell in the moderate wellbeing range for all time points, see Figure 3(a) below.

For the PWS, the main effect of time was significant with a large effect size $F(1,103)=19.22$, $P<.001$, $\eta_p^2=.16$ when controlling for baseline score. The time $\times$ baseline interaction effect was also significant with a large effect size $F(2,103)=17.19$, $P<.001$, $\eta_p^2=.14$. The time $\times$ group interaction was not significant. See Figure 3(b) below for raw scores by treatment group over time.

For the WHO-5, the main effect of time was significant with a moderate effect size $F(1,104)=9.09$, $P=.003$, $\eta_p^2=.08$. The time $\times$ group interaction was not significant. In post-hoc pairwise comparisons, there was a borderline significant mean difference between the GUIDE-only and waitlist, with the GUIDE-only group on average improving more (*raw MD*=5.03, $P=.07$, $D=.27$), a difference which translates to a 5.04% percentile point relative increase in raw score, or a 10.25% improvement controlling for time. Raw mean scores were in the moderate well-being range for treatment groups across time, see Figure 3(c) below.

For (perceived) Health, the main effect of time was significant with a large effect size $F(1,103)=19.48$, $P<.001$, $\eta_p^2=.16$. The time $\times$ baseline interaction was significant with a large effect size $F(2,103)=14.03$, $P<.001$, $\eta_p^2=.12$. The time $\times$ group interaction was not significant. In post-hoc pairwise comparisons, there were no significant differences. See Figure 3(d) for raw mean scores by group over time.

Figure 3. Raw mean scores by group over time for wellness measures. PERMA Overall is the total score of human flourishing



from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. Brackets indicate significant differences over time for each group, where * $P < .05$, ** $P < .01$, *** $P < .001$, ~ $P = .05-.10$.

Emotional wellbeing

For Positive emotion, the main effect of time was significant with a moderate effect size $F(1,104)=11.88$, $P < .001$, $\eta_p^2=.10$. The time $\times$ group interaction was not significant. In post-hoc pairwise comparisons, the GUIDE-only group showed significant improvement compared to the waitlist control (*raw MD* .29, $P=.03$, $D=.31$), translating to a 10.9% increase when controlling for time. See Figure 4(a) below for raw mean scores by group over time.

For Happiness, the main effect of time was significant with a small effect size, $F(1,104)=4.14$, $P=.04$, $\eta_p^2=.04$. The time $\times$ group interaction was not significant. In post-hoc pairwise comparisons, the GUIDE-only group showed significant improvement on this item relative to the waitlist group (*raw MD*=1.24, $P=.04$, $D=.30$), a 12.4 percentile point relative improvement which translates to 10.6% when controlling for time. See Figure 4(b) for raw scores by group over time.

For Negative emotion, the main effect of time was significant with a large effect size $F(1,104)=18.02$, $P < .001$, $\eta_p^2=.16$. The time $\times$ group interaction was not significant. All three treatment groups decreased in negative emotion over time, see Figure 4(c).

For Emotion dysregulation (DERS-SF), the main effect of time was significant with a moderate effect size $F(1,103)=13.14$, $P<.001$, $\eta_p^2=.11$ when controlling for baseline. The time $\times$ baseline interaction was also significant with a large effect size $F(2,103)=20.31$, $P<.001$, $\eta_p^2=.17$. The time $\times$ group interaction was borderline significant with a moderate effect size $F(2,103)=2.88$, $P=.06$, $\eta_p^2=.05$. In group $\times$ time post-hoc pairwise comparisons, only the GUIDE+Incentives group showed a significant reduction in emotion dysregulation over time with a moderate effect size (*raw MD*=-4.56, $P<.001$, $\eta_p^2=.09$), indicating that the 6.33% percentile point reduction in raw emotion dysregulation for this group drove the borderline significant time $\times$ group interaction effect. Post-hoc pairwise comparisons were not significant. See Figure 4(d) for raw scores by treatment group over time.

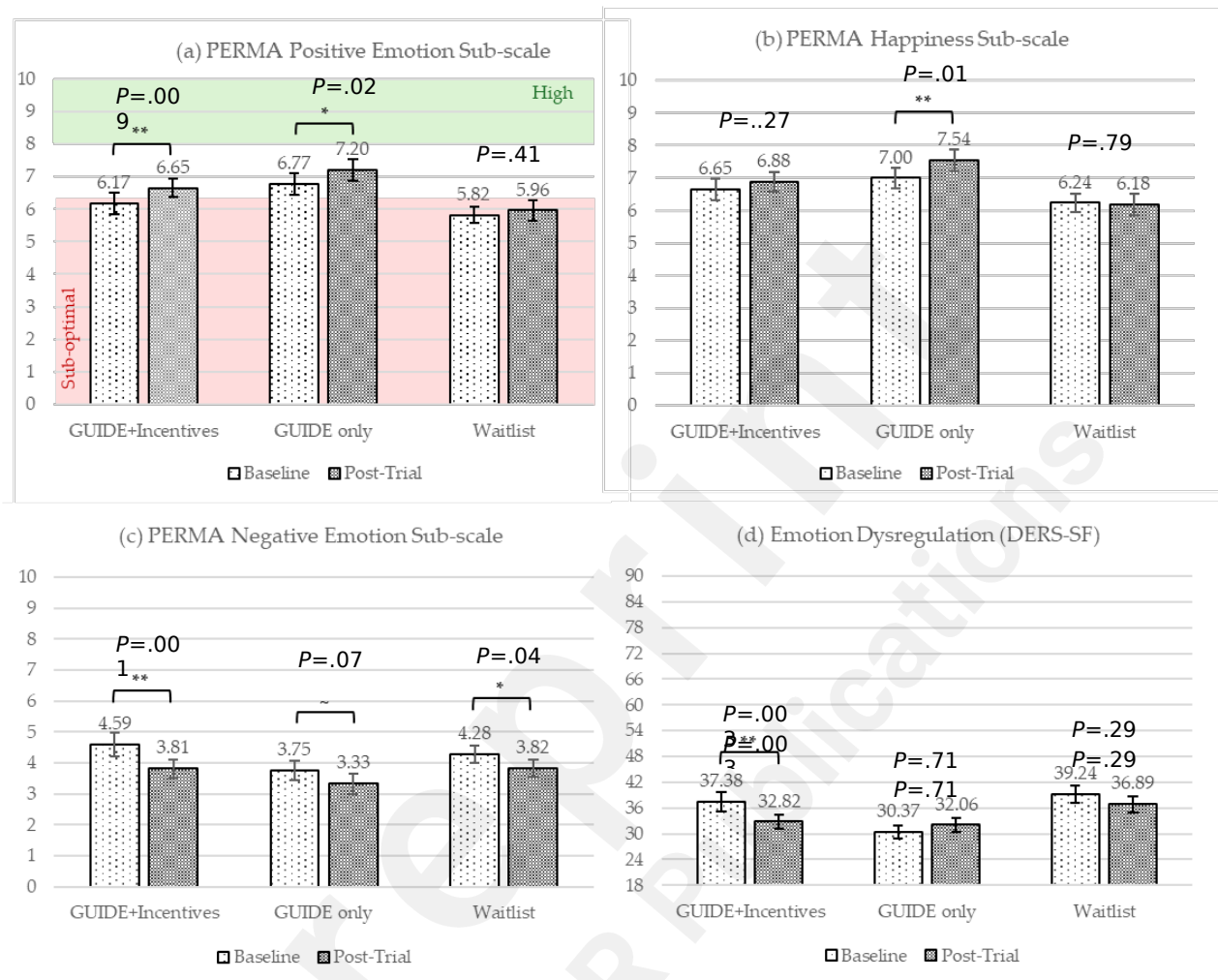


Figure 4. Raw mean scores over time by group for emotional wellbeing measures. Brackets indicate significant differences over time for each group, where * $P<.05$, ** $P<.01$, *** $P<.001$, ~ $P=.05-.10$.

Mental health

For depression symptoms (PHQ-8), the main effect of time was significant with a large effect size $F(1,104)=16.99$, $P<.001$, $\eta_p^2=.14$. The time $\times$ group interaction was not significant. In post-hoc pairwise comparisons, the GUIDE-only group showed a significantly greater reduction in depression symptoms compared to the waitlist (raw mean -1.34, $P=.04$, $d=-.30$), which translates to an additional 5.58% percentile point reduction in raw depression score for the GUIDE-only group. Per Figure 5(a), the average depression score for all groups fell below clinical levels for both time periods; however, a substantial portion of our participants did meet clinical criteria: 26.96% at baseline (31/107) and 16.82% at post-trial (18/107). A majority of participants had symptoms of sub-clinical or clinical depression (≥ 5) at baseline (59.13%, 68/115), while nearly half did at post-trial (49.53%, 53/107). See Figure 2 of Appendix A for depression severity by participant across time.

For anxiety symptoms (GAD-7), the main effect of time was significant with a large effect size when controlling for baseline $F(1,103)=14.33$, $P<.001$, $\eta_p^2=.12$. The time $\times$ baseline interaction was also significant with a larger effect size $F(2,103)=23.92$, $P<.001$, $\eta_p^2=.19$. The time $\times$ group

interaction was not significant. In post-hoc pairwise comparisons, the GUIDE-only group showed a significant relative reduction in anxiety compared to the waitlist (*raw MD* = -.5, $P = .05$, $d = -.29$). Mean raw anxiety scores exceeded the clinical cutoff for all three groups over time, per Figure 5(b) below, indicating a high level of anxiety present in our sample. A majority of participants met the criteria for clinical anxiety at both baseline (71.30%, 82/115) and post-trial (65.42%, 70/107), and 100% of participants met the criteria for mild anxiety at both time periods. See Figure 3 of Appendix A for anxiety level by participant across time.

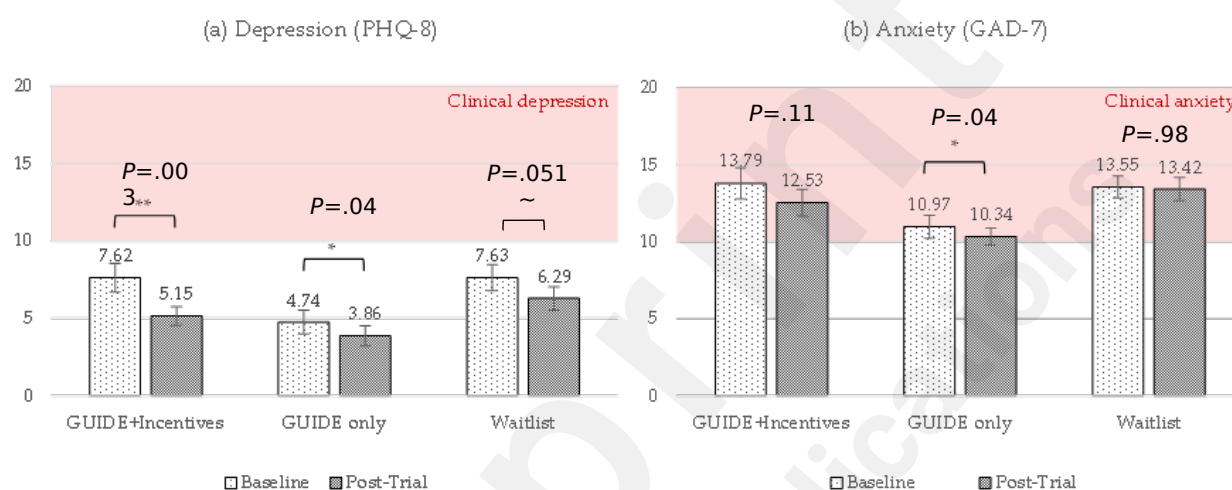


Figure 5. Raw mean scores over time by group for mental health measures. Brackets indicate significant differences over time for each group, where * $P < .05$, ** $P < .01$, *** $P < .001$, ~ $P = .05-.10$.

Social connectedness

For Relationships, the main effect of time was significant with a large effect size when controlling for baseline $F(1,103) = 14.86$, $P < .001$, $\eta_p^2 = .13$. The time $\times$ baseline interaction was also significant with a moderate effect size $F(1,103) = 11.40$, $P = .001$, $\eta_p^2 = .10$. The time $\times$ group was not significant. Post-hoc pairwise comparisons were not significant. See Figure 6(a) for raw mean scores by treatment group over time.

For Loneliness, the main effect of time was significant with a moderate effect size $F(1,103) = 7.55$, $P = .007$, $\eta_p^2 = .07$. The time $\times$ baseline interaction was significant with a large effect size $F(1,103) = 26.78$, $P < .001$, $\eta_p^2 = .21$. The time $\times$ group interaction was not significant. See Figure 6(b) for raw mean scores by treatment group over time.

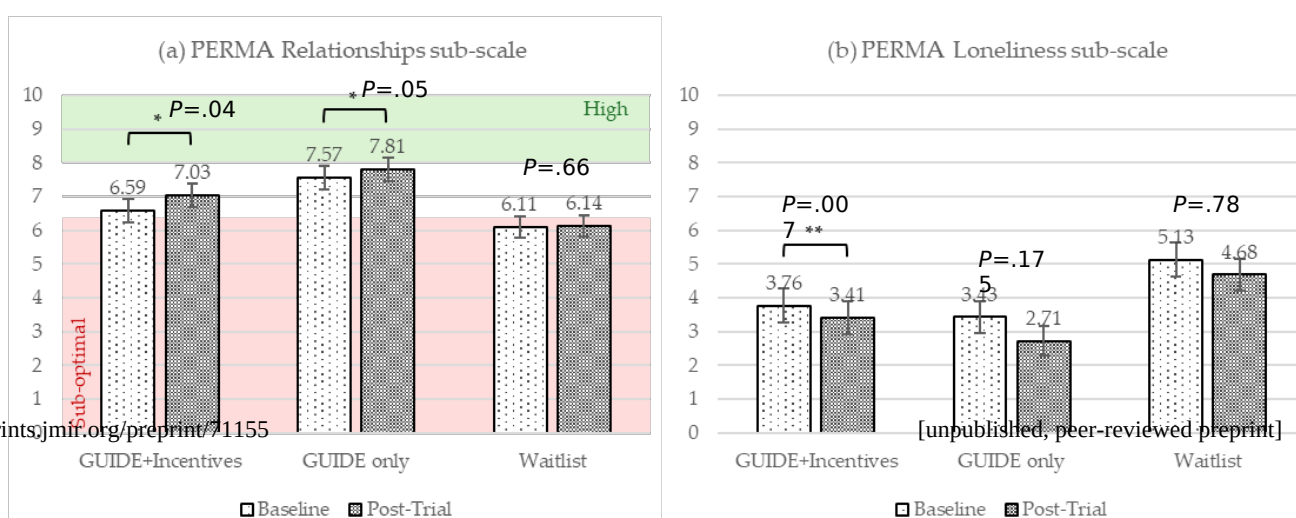


Figure 6. Raw mean scores over time by group for social connectedness measures. Brackets indicate significant differences over time for each group, where * $P < .05$, ** $P < .01$, *** $P < .001$, ~ $P = .05-.10$.

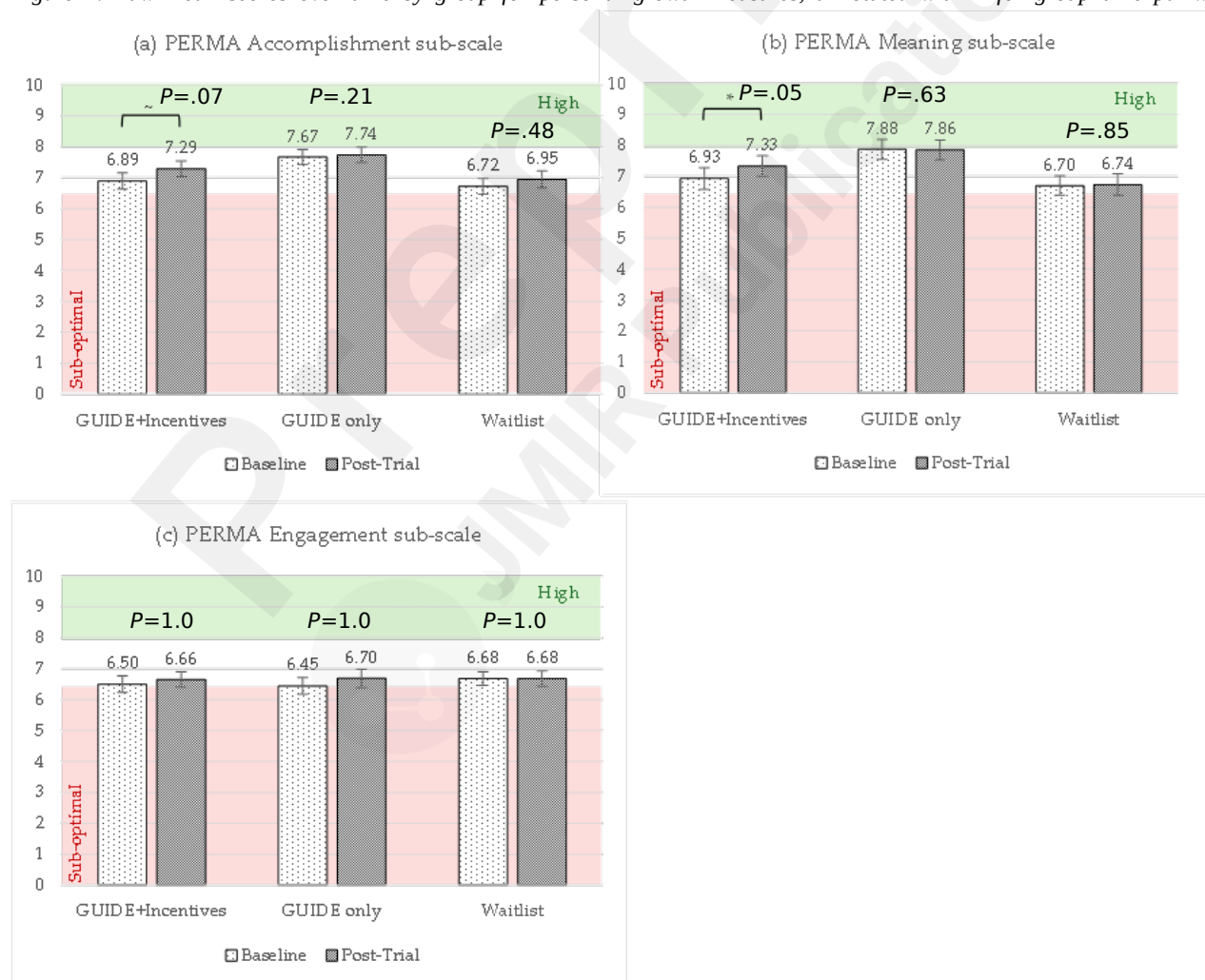
Personal accomplishment

For Accomplishment, the main effect of time was significant with a large effect size when controlling for baseline $F(1,103)=19.16$, $P < .001$, $\eta_p^2 = .16$. The time $\times$ baseline interaction was also significant with a large effect size $F(2,103)=16.00$, $P < .001$, $\eta_p^2 = .13$. The time $\times$ group interaction was not significant. Post-hoc pairwise comparisons were not significant. See Figure 7(a) for raw mean scores by treatment group over time.

For Meaning, the main effect of time was significant with a moderate effect size when controlling for baseline score $F(1,103)=9.33$, $P = .003$, $\eta_p^2 = .08$. The time $\times$ baseline interaction was significant with a moderate effect size $F(2,103)=7.86$, $P = .006$, $\eta_p^2 = .07$. The time $\times$ group interaction was not significant. Post-hoc pairwise comparisons were not significant. See Figure 7(b) for raw mean scores by treatment group over time.

For Engagement, neither time nor time $\times$ group were significant. Mean scores for all three groups across both time periods fell near the sub-optimal cutoff, per Figure 7(c).

Figure 7. Raw mean scores over time by group for personal growth measures, annotated with P for group*time pairwise



comparison. Brackets indicate significant differences over time for a group, where * $P < .05$, ** $P < .01$, *** $P < .001$, ~ $P = .05-.10$.

Subgroup Analyses

First Responders

Next, we explored outcomes for active first responders ($n=93$). See Table 2 in Appendix A for a breakdown by responder type. We reran our repeated measures ANOVAs for select outcome measures using data from active first responders only. We excluded ‘filler’ items from the PERMA-Profiler (i.e., Negative emotion, Happiness, Health, Loneliness), as well as measures without significant differences in our main analysis (i.e., Engagement). See Table 4 below for a summary of our findings.

Table 4. Effect sizes for repeated measures ANOVAs and pairwise comparisons for first responders.

Outcome measure	ANOVA Effects		Pairwise Comparisons (Group*Time)						Pairwise Group Comparisons of Pre-Post Deltas					
	Time	Time x Group	Incentives		GUIDE only		Waitlist		Incentives vs. WL		GUIDE vs. WL		Incentives vs. GUIDE only	
	ηp^2 ^a	ηp^2 ^a	ηp^2 ^a	MD ^c	ηp^2 ^a	MD	ηp^2 ^a	MD ^c	d ^b	MD	d ^b	MD ^c	d ^b	MD ^c
Wellness														
PERMA Overall	.04~	.01	.04	0.28	.00	0.06	.01	0.12	.17	0.48	.33*	0.94*	-.16	-0.46
PWS †	.14**	.04	.05*	6.36*	.00	0.20	.00	-0.41	.22	3.39	.02	0.30	.21	3.08
WHO-5	.06*	.03	.03	4.32	.05*	5.54*	.00	0.34	.13	4.71	.25	9.34	-.13	-4.63
Emotional well-being														
Positive emotion	.11**	.02	.08*	0.53*	.02	0.26	.02	0.18	.18	0.61	.29~	0.97~	-.11	-0.36
DERS-SF †	.08**	.05~	.06*	-3.33*	.01	1.18	.00	-0.47	-.19	-1.43	.10	0.82	-.30	-2.26
Mental health														
PHQ-8	.12**	.06~	.12*	-2.16**	.00	-0.08	.06*	-1.23*	-.13	-1.05	-.33	-2.65	.20	1.60
GAD-7 †	.10**	.02	.03	-0.87	.01	-0.14	.00	0.00	-.16	-0.44	-.02	-0.07	-.13	-0.36
Personal growth														
Accomplishment †	.11**	.01	.02	0.28	.00	0.10	.00	0.10	.08	0.09	.00	0.00	.08	0.09
Meaning †	.05*	.02	.02	0.26	.00	-0.04	.00	-0.02	.14	0.14	-.01	-0.01	.15	0.15
Social Connectedness														
Relationships †	.08**	.01	.01	0.23	.03	0.31	.03	0.03	.09	0.10	.12	0.14	-.04	-0.04

Significance level with Bonferroni adjustment for multiple comparisons: *** $P < .001$, ** $P < .01$, * $P < .05$, ~ $P = .05-.1$. PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

† controlled for baseline.

a Interpretation of ηp^2 : .01=small effect, .06=moderate effect, .11=large effect.[71,72]

b Interpretation of Cohen's d : .2=small effect, .5=moderate effect, .8=large effect.[72,73]

c Based on adjusted means.

Time was significant with moderate to large effect sizes for all outcome measures except the PERMA Overall, which was borderline significant with a small effect size. There were no significant time×group interactions; however, two interactions were borderline significant with moderate effect sizes – emotion dysregulation (DERS-SF), $F(2,82)=2.35$, $P=.10$, $\eta_p^2=.05$, and depression symptoms (PHQ-8), $F(2,82)=2.68$, $P=.07$, $\eta_p^2=.06$. In group×time pairwise comparisons for these outcome measures, the GUIDE+Incentives group showed the most significant improvement, with a 4.63% reduction in emotion dysregulation and 9.00% reduction in depression symptoms based on adjusted means, with moderate and large effect sizes

respectively. Depression symptoms decreased by 5.13% for the waitlist condition, but the effect sizes and mean differences were half as large as those of the GUIDE+Incentives group.

In pairwise comparisons by group, the GUIDE+Incentives group performed better than the waitlist for all outcomes, but the mean differences between these groups was not significant. The GUIDE-only group improved more than the waitlist across wellness, emotional wellbeing, mental health and social connectedness measures; however, the difference between groups was only significant for wellness (PERMA Overall) and borderline significant for emotional wellbeing (Positive emotion). The GUIDE-only group did not improve significantly over the waitlist on either personal growth measure. When comparing the GUIDE+Incentives and GUIDE-only groups, results were mixed, indicating that half of our outcomes improved more for participants who received financial incentives, while the other half improved more without – though none of the differences between our two interventions were statistically significant. GUIDE+Incentives showed greater improvement on Personal Growth measures, while GUIDE-only showed greater improvement on Wellness and Emotional Wellbeing.

During our intake meetings, many participants disclosed that they were leaders in their organization or involved in wellness for their community. We did not include an item to capture this in our demographic survey, so we selected participants with less education (college or less, $n=63$) and fewer years of employment (<15 years, $n=66$) as proxies for less leadership/wellness experience in a sensitivity analysis. See Tables 3 and 4 in Appendix A for our findings.

In group×time pairwise comparisons for responders with less education/experience, the GUIDE+Incentives group had more significant improvements and larger effect sizes than in the analysis for all first responders. While these sensitivity analyses were not powered, they suggest that financial incentives may lead to greater improvements for less experienced first responders than GUIDE alone, particularly in reducing emotion dysregulation, depression and anxiety; but that the GUIDE-only group still shows more significant improvement compared to the waitlist for wellness and emotional wellbeing.

Military Affiliation

To explore outcomes for participants with military affiliation, we conducted repeated measures ANOVAs for veterans and active-duty military ($n=45$). We excluded ‘filler’ items from the PERMA-Profilier, as well as Engagement, like in the responder analysis. See Table 5 below for our findings.

Table 5. Effect sizes for repeated measures ANOVAs and pairwise comparisons for participants with military affiliation.

ANOVA Effects	Pairwise Comparisons (Group*Time)				Pairwise Group Comparisons of Pre-Post Deltas			
Time	Time	Incentives	GUIDE only	Waitlist	Incentives vs. WL	GUIDE only vs. WL	Incentives vs. GUIDE only	

Outcome measure	ηp^2 a	e x G r o u p ηp^2 a	M ηp^2 a D c		M ηp^2 a D c		M ηp^2 a D c		d b MD		d b MD c		d b MD c	
			ηp^2	a	ηp^2	a	ηp^2	a	d	MD	d	MD	d	MD
Wellness														
PERMA Overall	.06~	.1	.13*	.05	.02	-0.19	.08	0.40	.27~	1.32~	.20	-0.93		
PWS †	.16**	.0	.02	.03	.00	-.33	.07	1.68	.10	2.68	-.04	-1.01		
WHO-5	.19**	.0	.07~	.14*	.03	4.33	.04	2.42	.29*	18.29*	-.28~	-15.88		
Emotional wellbeing														
Positive emotion	.11*	.0	.10*	.12*	.00	-.06	.13	.74	.33*	1.92*	.22	-1.18		
DERS-SF †	.29** *	.0	.24**	.06	.01	-1.33	-.20	-1.92	-.05	-0.50	-.16	-1.42		
Mental health														
PHQ-8	.30** *	.0	.24**	.10*	.06	-1.58	-.02	-.27	.29*	-4.37*	.30*	4.09*		
GAD-7 †	.26** *	.1	.02	.25**	.00	.32	-.09	-.46	.32*	-1.60*	.23	1.14		
Personal growth														
Accomplishment †	.19**	.0	.03	.03	.02	-.25	.16	.28	.16	0.29	-.01	-0.02		
Meaning †	.15*	.0	.04	.01	.03	-.40	.20	.39	.13	0.27	.06	0.12		
Social Connectedness														
Relationship †	.31** *	.1	.13*	.09*	.04	-.46	.31*	.60*	.27*	0.54*	.03	0.05		

Significance level with Bonferroni adjustment for multiple comparisons: *** $P < .001$, ** $P < .01$, * $P < .05$, ~ $P = .05-.1$. PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

† controlled for baseline.

a Interpretation of ηp^2 : .01=small effect, .06=moderate effect, .11=large effect.[71,72]

b Interpretation of Cohen's d : .2=small effect, .5=moderate effect, .8=large effect.[72,73]

c Based on adjusted means.

The main effect of time was significant for all wellness, emotional wellbeing, mental health, social connectedness, and personal growth measures – except for PERMA Overall, which was near significance.

Unlike our main analysis, two measures had significant time×group interactions with large effect sizes: the GAD-7 for anxiety, $F(2,40)=3.85$, $P=.03$, $\eta_p^2=.16$, and Relationships, $F(2,40)=3.85$, $P=.03$, $\eta_p^2=.16$.

In post-hoc group×time pairwise comparisons of depression symptoms, the GUIDE+Incentives group improved significantly over time with a large effect size (raw $MD=-1.44$, $P=.001$, $\eta_p^2=.25$), translating to a 6.86% percentile point reduction in raw anxiety score, or a 13.76% reduction when controlling for baseline. In post-hoc between-group pairwise comparisons, the GUIDE-only group showed a significantly greater reduction in anxiety symptoms relative to the waitlist (raw $MD=-1.69$, $P=.03$, $D=-.31$), a difference which translates to an 8.05% greater percentile point reduction in raw anxiety, or 7.62% when controlling for baseline.

For Relationships, post-hoc group×time pairwise comparisons show that both GUIDE+Incentives (*raw MD*=.81; *SD*=.76) and GUIDE-only (*raw MD*=.33, $P=.05$, $\eta_p^2=.13$) improved significantly over time, with large and moderate effect sizes respectively, while the waitlist condition (*raw MD*=-.22, $P=.02$, $\eta_p^2=.09$) did not. In post-hoc between-group pairwise comparisons, both intervention groups showed a significantly greater improvement on this measure when compared to the waitlist. This indicates that social connectedness improved for both groups with GUIDE access, but not the waitlist control, and that these results were significant with moderate to large effect sizes.

Beyond anxiety and social connectedness, the GUIDE-only group also improved significantly over the waitlist in wellness (WHO-5), emotional wellbeing (Positive emotion), and mental health (PHQ-8 for depression), though the effect sizes were small. Wellness as measured by PERMA Overall, our primary outcome variable, was also near significant for GUIDE-only compared with the waitlist. The GUIDE-only group also had significantly reduced depression symptoms compared to the GUIDE+Incentives group, and near significant relative improvement on wellbeing (WHO-5). GUIDE+Incentives only showed a significant improvement over the waitlist for social connectedness. Together, these results suggest that GUIDE alone may perform better without financial incentives for military personnel and veterans.

In our analysis of participants without military affiliation (i.e., first responders only), none of the time×group interactions were significant. See Table 5 in Appendix A.

Sex differences

To explore how sex influenced participant outcomes, we conducted exploratory analyses of select outcome measures by rerunning our repeated measures ANOVAs for male ($n=76$) and female ($n=31$) participants, since previous research has shown significant variation between sexes on self-report measures related to health and mental health.[74] Results for male participants were powered and are reported in Table 6 below.

Table 6. Effect sizes for repeated measures ANOVAs and pairwise comparisons for male participants.

Outcome measure	ANOVA Effects		Pairwise Comparisons (Group*Time)						Pairwise Group Comparisons of Pre-Post Deltas					
	Time	Time×Group	Incentives		GUIDE only		Waitlist		Incentives vs. WL		GUIDE only vs. WL		Incentives vs. GUIDE only	
	$\eta p2^a$	$\eta p2^a$	$\eta p2^a$	MD^c	$\eta p2^a$	MD	$\eta p2^a$	MD^c	d^b	MD	d^b	MD^c	d^b	MD^c
Wellness														
PERMA Overall	.08*	.05	.10*	.51*	.02	.19	.00	.04	.11	.37	.18	.63	-.07	-.26
PWS †	.20**	.03	.32	4.82	.07	1.26	.07	-1.41	.17	3.1	.07	1.34	.10	1.78
WHO-5	.12**	.03	.08*	6.78*	.07*	5.93*	.01	1.85	.06	2.76	.16	7.04	-.09	-4.28
Emotional wellbeing														
Positive emotion	.07*	.05	.08*	.52*	.04~	.33~	.00	-.03	0.16	0.70	.04	0.16	-.04	-0.16
DEERS-SF †	.15**	.06~	.10**	-4.45**	.00	.23	.01	-1.14	-.18	-1.65	.08	.69	-.26	-2.34
Mental health														
PHQ-8	.18**	.04	.14**	-2.70**	.03	-.96	.05	-1.35	-.25	-3.27	-.34*	-5.02*	.12	1.76
GAD-7 †	.19**	.03	.02	-.82	.05*	-1.18*	.00	.02	-.11	-.42	-.17	-.60	.05	.18
Personal growth														
Accomplishment †	.17**	.02	.04	.41	.02	.23	.00	.08	.12	.17	.06	.08	.07	.09

Meaning †	.09**	.06	.07*	.58*	.00	.11	.00	-.12	.25	.34	.16	.12	.09	.22
<i>Social Connectedness</i>														
Relationships †	.18**	.09*	.10**	.70**	.02	.28	.00	-.17	.30*	.44*	.23	.23	.15	.21

Significance level with Bonferroni adjustment for multiple comparisons: *** $P < .001$, ** $P < .01$, * $P < .05$, ~ $P = .05-.1$. PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

† controlled for baseline.

a Interpretation of η^2 : .01=small effect, .06=moderate effect, .11=large effect.[71,72]

b Interpretation of Cohen's d : .2=small effect, .5=moderate effect, .8=large effect.[72,73]

c Based on adjusted means.

The main effect of time was significant across all wellness, emotional wellbeing, mental health, personal growth and social connectedness measures included in the analysis. For Relationships, the time×group interaction was significant with a moderate effect size when controlling for baseline. In post-hoc group×time pairwise comparisons, only the GUIDE+Incentives condition showed significant improvement. Post-hoc pairwise comparisons by group corroborate these findings and indicate that the GUIDE+Incentives group performed significantly better than the waitlist on this measure. Altogether, these findings indicate that male participants may feel more socially connected when using GUIDE with financial incentives, rather than without.

For emotion dysregulation (DERS-SF), the time×group interaction was near significance. In post-hoc group×time pairwise comparisons, only the GUIDE+Incentives condition showed significant reduction in emotion dysregulation over time when controlling for baseline – though in post-hoc pairwise comparisons by group, there were no significant differences. GUIDE with financial incentives may improve emotion regulation for male users, but these findings are not as strong as for Relationships.

For females, results are reported in Appendix A, Table 6. The main effect of time was only significant for the included emotional wellbeing and mental health measures. None of the time×group interactions were significant, but unlike the male sample, these analyses were not powered and so should be interpreted cautiously. In post-hoc pairwise comparisons by group, we found that the GUIDE-only group showed a significant reduction in depression symptoms (PHQ-8) relative to the waitlist condition, indicating that GUIDE usage without financial incentives may have helped reduce these symptoms relative to the control. There was also a near significant difference between the GUIDE-only and waitlist conditions for Positive emotion in post-hoc pairwise comparisons.

In short, the effect of GUIDE and financial incentives on wellness, emotional wellbeing, and mental health may differ by sex. Male participants benefitted more from the GUIDE+Incentives condition, while females may have benefitted from the GUIDE-only condition, signaling that males may be more motivated by financial incentives than females.

Secondary Outcomes

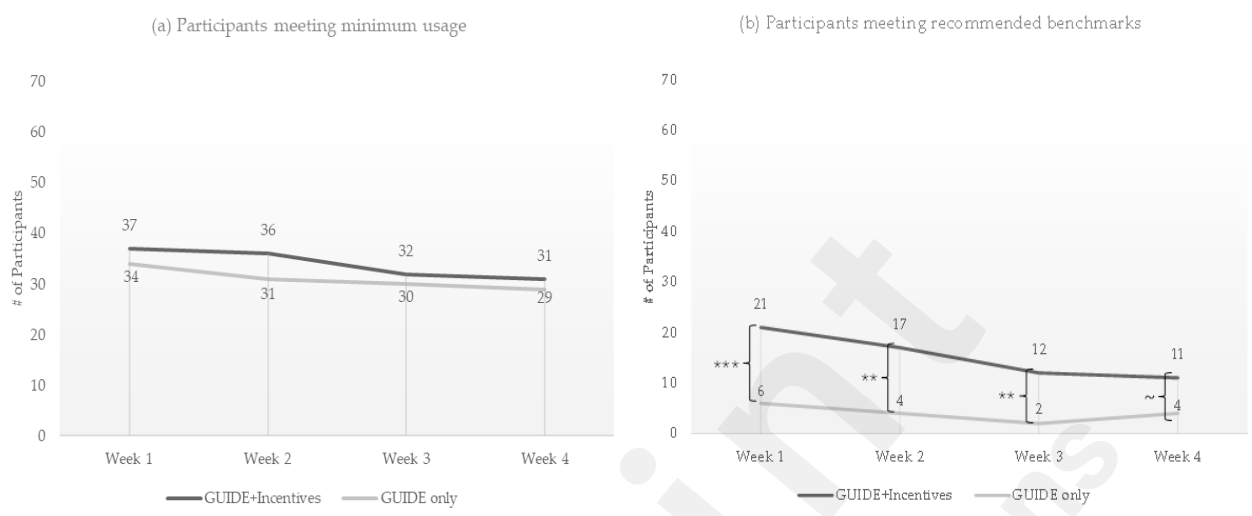
App Engagement

(a) App engagement by minimum usage and recommended benchmarks

Of participants allocated to GUIDE+Incentives and GUIDE-only intervention conditions, 97.37% (74/76) were included in the app engagement analysis. Excluded participants experienced technical difficulties and were unable to use the app (and thus were considered lost to follow-up). By week 4, 15.79% of participants from GUIDE+Incentives (6/38) and 11.11% of participants from GUIDE-only (4/36) had stopped using GUIDE, for a combined attrition rate of 13.51%.

To assess for adherence criterion as recommended by the literature[75], we assessed app engagement based on minimum usage standards (≥ 3 weekly activities) vs. recommended usage benchmarks (≥ 3 weekly lessons/posts/replies/likes/mood surveys) across all weeks. Overall, 67.10% of participants (51/76) met the minimum usage standards for all four weeks, while only 7.89% of participants (6/76) met the recommended benchmarks for all four weeks. To compare how financial incentives influenced whether a given participant met the minimum vs. recommended benchmarks, we conducted an independent samples *t*-test. There was not a significant difference in average weeks meeting minimum usage standards for GUIDE+Incentives (mean 3.58, *SD* .72) compared to GUIDE-only (mean 3.44, *SD* 1.03), indicating that financial incentives did not impact engagement at this minimum; however, when it came to number of weeks meeting the more rigorous recommended usage benchmarks, GUIDE+Incentives (mean 1.61, *SD* 1.46) completed the recommended usage significantly more often than GUIDE-only (mean .44, *SD*=.81), $t(72)=4.19$, $P<.001$.

This pattern held in analysis of weekly app engagement, see Figure 8(a-b) below. There were no significant differences between intervention groups in minimum usage for any week, while differences between GUIDE+Incentives and GUIDE-only were highly significant for recommended benchmark across the first three weeks of the trial, but only borderline significant for Week 4 – indicating that financial incentives impacted whether participants met recommended benchmarks, but that the influence of these incentives decreased over time.



(b) Impact of treatment group on app engagement

To assess whether financial incentives influenced app engagement, *t*-tests were conducted to compare app usage (total and by feature) between GUIDE+Incentives and GUIDE-only conditions. The descriptive statistics and *t*-test results are reported in Table 7 below. Participants in GUIDE+Incentives had significantly higher total app activity compared to GUIDE-only, driven by more small group posts and replies, indicating that financial incentives influenced engagement with the peer support features of GUIDE. There was not a significant difference between lessons completed, likes, or mood surveys.

Table 7. App engagement by treatment group

App metric	GUIDE+Incentives		GUIDE only		t-test
	M	SD	M	SD	
Lessons	27.53	27.81	22.61	30.47	0.73
Posts	83.66	71.58	52.92	48.50	2.15*
Replies	19.05	18.04	7.75	8.71	3.40**
Likes	27.05	30.93	18.83	27.01	1.22
Mood surveys	13.47	7.86	12.50	9.69	0.48
Total activity	170.76	133.12	114.61	104.75	2.01*

Indicates ** $P < .01$, * $P < .05$

(c) Correlations between app engagement and primary outcome measures

To explore how app engagement influenced pre-post scores for our wellness, emotional wellbeing, mental health, social connectedness and personal growth, Pearson correlation coefficients were calculated. See Table 8 below for our findings.

Table 8. Pearson correlations of app engagement and pre-post outcome measures

Pre-post outcome measure	M	SD	Lessons	Posts	Replies	Likes	Mood surveys	Total activity
Wellness								
PERMA Overall	0.21	0.81	.19	.14	.17	.16	-.04	.16
WHO-5	4.00	13.99	.26*	.13	.16	.23~	.03	.25*
PWS	2.10	16.59	.36**	.32**	.28*	.31*	.19	.25*
Health	0.28	1.05	.07	.06	.06	.09	-.05	.07

Emotional wellbeing

Positive emotion	0.34	1.05	.26*	.13	.16	.23	.03	.21~
Happiness	0.23	1.24	.13	.04	-.14	.21	-.12	.09
Negative emotion	-0.55	1.34	-.23	-.33**	-.23	-.21	-.17	-.35**
DERS-SF	-1.73	8.60	-.26*	-.39**	-.30*	.22~	-.19	-.14
<i>Mental health</i>								
PHQ-8	-1.55	3.54	-.19	-.22~	-.27*	-.35**	.09	-.12
GAD-7	-0.65	3.45	-.12	-.18	-.23~	-.22~	-.01	-.05
<i>Social connectedness</i>								
Relationships	0.23	1.18	.11	.09	.14	.15	-.04	.11
Loneliness	-0.50	2.29	.05	.01	-.09	-.17	-.04	-.03
<i>Personal growth</i>								
Meaning	0.13	1.12	.18	.09	.19	.15	-.10	.12
Accomplishment	0.23	1.16	.34**	.33**	.23~	.36**	.30*	.14
Engagement	0.13	1.22	-.03	.00	-.08	-.10	-.04	-.01

PERMA Overall is the total score of human flourishing from PERMA (Positive emotion, Engagement, Relationships, Meaning, Accomplishment) Model[50]. PWS and WHO-5 assess for wellbeing. Health is a measure of perceived health. DERS-SF assesses for emotion dysregulation. GAD-7 assesses for anxiety symptoms. PHQ-8 assesses for depression symptoms.

Interpretation of correlation coefficient: <.1=negligible correlation, .10=weak correlation, .40=moderate correlation [76]
 ** indicates $P < .01$ *, $P < .05$ ~ $P = .05$ -.10

Total activity was weakly correlated with improvements on two wellness measures (WHO-5, PWS), but PERMA Overall. In terms of emotional wellbeing, there was a borderline significant positive correlation between total activity and Positive emotion. There were no significant correlations between total app activity and any mental health, personal growth, or social connectedness measures.

Lessons completed was positively correlated with two wellness measures (WHO-5, PWS). It was also weakly correlated with an improvement in emotional wellbeing measures (Positive emotion, DERS-SF). For personal growth, more lessons completed was moderately correlated with an improvement on Accomplishment. Mental health and social connectedness were not correlated with lessons completed.

Total posts was moderately and positively correlated with wellness (PWS) and personal growth (Accomplishment). Additionally, posts were moderately and negatively correlated with emotion dysregulation (DERS-SF).

Replies to peer posts was weakly correlated with an improvement in wellness (PWS). There was a moderate and negative correlation between replies and emotion dysregulation (DERS-SF). In terms of mental health, an increase in replies was weakly correlated with a reduction in depression symptoms (PHQ-8). There was also a weak and borderline significant correlation with reduction in anxiety symptoms (GAD-7). There was a borderline significant and weak correlation between replies and personal growth (Accomplishment). Replies were not significantly correlated with changes in social connectedness.

Likes were significantly and moderately correlated with an improvement in wellness (PWS), mental health (PHQ-8 for depression), and personal growth (Accomplishment). There were also borderline significant and weak correlations between total likes and improvement on WHO-5 well-being, emotion dysregulation (DERS-SF), and anxiety symptoms (GAD-7). Likes to peer posts was not correlated with change in social connectedness.

Mood surveys completed was only correlated with an improvement in personal growth (Accomplishment).

Implementation Outcomes & Technical Merit

To assess for implementation outcomes and technical merit, we relied on descriptive statistics for acceptability (mean 3.81, SD 0.89, min 1, max 5), appropriateness (mean 4.01, SD 0.75, min 2,

max 5), feasibility (mean 4.01, SD 0.77, min 1, max 5), and usability (mean 70.62, SD 18.82, min 7.5, max 100).

While there are no agreed-upon cutoffs for acceptability, appropriateness, and feasibility, for the purposes of this investigation, we classify a mean score ≥ 4 to be good, since it indicates that the participant on average agreed with each item. With this classification in mind, the appropriateness and feasibility measures are above the cutoff, indicating that GUIDE is well suited for the population and feasible to implement. The mean acceptability score was 3.81, below our threshold, indicating that participants on average rated GUIDE somewhere between the neutral and agree options on a 5-point Likert scale. This likely means that the app is adequately acceptable but that there is room for improvement.

Mean usability was satisfactory, or slightly above the average cutoff score of 68, indicating that the app is generally usable; however, like with acceptability, there is room for improvement, since 39% of participants ($n=27$) rated the app below this cutoff.

Overall, GUIDE received scores above industry standards for appropriateness, feasibility, and usability, suggesting strong performance on these indicators; however, the acceptability score was slightly below the desired threshold, indicating there may be room for improvement in terms of likability.

Discussion

Population

The prevalence of clinical depression and anxiety symptoms in our sample indicates that we managed to recruit a population in need of an app like GUIDE. In the 2022 National Health Interview Survey, depression (PHQ-8) and anxiety (GAD-7) symptoms met clinical criteria for 7.5% and 6.7% of the general US population respectively.[77] The rates of clinical depression (16.80–26.96%) and clinical anxiety (65.42–71.30%) in our sample were substantially higher, and even exceeded rates reported in a recent systematic review of first responder mental health.[78] The severity of reported depression and anxiety becomes more apparent when including those with mild or sub-clinical levels. These figures indicate cause for concern and corroborate the notion that there is a mental health crisis among first responders, veterans and active-duty military.

Primary Outcomes

Overall, the main effect of time was highly significant with moderate to large effect sizes for wellness, emotional wellbeing, mental health, social connectedness and personal growth primary outcome measures. While this cross-measure improvement over time aligns with the reported pre-post t -tests from the quasi-experimental GUIDE pilot study[29], our findings indicate that, since time $\times$ group interactions were not significant, pre-post t -test results may have been greatly influenced by time.

When controlling for time, there was one borderline significant time $\times$ group interaction for emotion dysregulation (DERS-SF). This interaction effect was driven by the GUIDE+Incentives group, which showed a significant reduction in emotion dysregulation over time with a large effect size, while the GUIDE-only and waitlist groups did not, indicating that GUIDE plus financial incentives may be most effective as an intervention for improving emotion regulation. This result is encouraging, since emotion regulation plays an important role in ethical decision-making and acts as a protective factor in stressful contexts, like those encountered by first responders and military.[79,80]

Across most wellness, emotional wellbeing, mental health, social connectedness and personal growth measures, GUIDE+Incentives and GUIDE-only conditions showed larger and more

consistent improvements over time compared to the waitlist control. In terms of post-hoc pairwise group×time comparisons, active intervention groups showed significant improvement on 53.33% measures (8/15) (GUIDE+Incentives: PERMA Overall, Positive emotion, DERS-SF, Negative emotion, Depression, Health, Meaning, Relationships) (GUIDE-only: WHO-5, Positive emotion, Happiness, PHQ-8 for depression, GAD-7 for anxiety, Health, Relationships, Loneliness), while the waitlist only showed significant improvement on a single measure (6.67%) (Negative emotion). This ratio suggests that participants in active intervention groups, on average, improved significantly over time, while the waitlist improved at a rate just above random chance (i.e., $P < .05$).

While both GUIDE+Incentives and GUIDE-only groups showed improvement over time on a majority of wellbeing measures when compared to the waitlist, in between-group post-hoc pairwise comparisons, only GUIDE-only showed statistically significant improvements over the waitlist for wellness (PERMA Overall), emotional wellbeing (Positive emotion, Happiness), and mental health (PHQ-8 depression and GAD-7 anxiety symptoms). The mean difference for WHO-5 well-being was also near significance. These findings suggest that GUIDE has a positive influence on wellness, emotional wellbeing and mental health among warriors – without financial incentives. This result is encouraging, since it suggests that financial incentives are not required to see improvements among GUIDE members. This aligns with previous research indicating that financial incentives did not impact mental health/wellness measures over time for a sample of college students using a digital mental health intervention [81].

Given the significant improvements the GUIDE-only condition showed over the waitlist, GUIDE is just as, if not more effective, than interventions included in a recent systematic review of wellbeing and resilience programs for first responders (including in-person interventions), half of which reported no significant differences within or between treatment groups on outcome measures.[82] While the percentile point improvements for our sample were modest with small effect sizes, they are still encouraging since these mental health and wellness changes were observable after only 4 weeks and were aligned with other RCTs testing digital mental health apps considered effective.[83]

Observable trends and directionality of between-group post-hoc pairwise comparisons were also encouraging, even if they were not all statistically significant. GUIDE+Incentives improved more than the waitlist on 93.33% of measures (14/15) though none were significant; while GUIDE-only improved more than the waitlist on 86.67% of measures (13/15), 33.33% of which were significant (5/15). The directionality of these results suggests that GUIDE (with and without financial incentives) can positively influence wellness and mental health for first responders, veterans, and military over 4 weeks – a result which, given research indicating how sustained engagement leads to larger improvements for digital health users, may become more apparent over a longer time horizon.[83]

Subgroup Analyses

Active First Responders

Like in our main analysis, the main effect of time for active first responders was significant with moderate to large effect sizes for all of reported outcome measures (excepting being PERMA Overall, which was borderline significant with a small effect size). There were no significant time×group interactions; however, the interaction effects for emotion dysregulation and depression symptoms were borderline significant. GUIDE+Incentives showed the most significant improvement on these measures, with a 4.63% percentile point reduction in emotion dysregulation and 9.00% percentile point reduction in anxiety symptoms across the two time

periods based on adjusted means, with moderate and large effect sizes respectively. Depression symptoms also decreased by 5.13% for the waitlist condition, but the effect sizes and mean differences were half as large as those of the GUIDE+Incentives group.

In between-group comparisons, the GUIDE+Incentives group performed better than the waitlist group for all outcomes, but the mean differences between these groups were not significant. The GUIDE-only group improved more than the waitlist across wellness, emotional wellbeing, mental health and social connectedness measures; however, the difference between groups was only significant for wellness (PERMA Overall) and borderline significant for emotional wellbeing (Positive emotion). Notably, the GUIDE-only group did not improve significantly over the waitlist on personal growth measures. When comparing GUIDE+Incentives and GUIDE-only groups, results were mixed, with half of outcomes improving more for responders who received financial incentives, while the other half improving more without – though differences were not statistically significant – corroborating again that financial incentives are not integral to efficacy. In sensitivity analyses selecting for first responders with less education and experience, group×time pairwise comparisons showed that GUIDE+Incentives had more significant improvements and larger effect sizes compared to the analysis of all responders. While these subgroup sensitivity analyses were not powered, they imply that financial incentives may act as a more powerful motivator for first responders with less education and experience, particularly for reducing emotion dysregulation, depression and anxiety. In terms of wellness and emotional wellbeing, however, GUIDE-only still performed significantly better, indicating these outcomes may be less influenced by financial incentives, regardless of experience or education.

Military Affiliation

Veterans and active-duty military who used GUIDE showed a greater improvement on outcome measures over time when compared to participants without military affiliation. These improvements were more significant with larger effect sizes than for participants overall, indicating that GUIDE may be more effective in improving wellness, emotion wellbeing, mental health and social connectedness for veterans and active-duty military. In between-group post-hoc pairwise comparisons, GUIDE-only improved significantly over time compared to the waitlist for wellness (WHO-5), emotional wellbeing (Positive emotion), mental health (GAD-7 anxiety and PHQ-8 depression symptoms), and social connectedness (Relationships) – while GUIDE+Incentives only showed a significant improvement over the waitlist for social connectedness (Relationships). Of note, GUIDE-only showed a significant and near significant improvement in PHQ-8 depression symptoms and WHO-5 well-being respectively when compared to GUIDE+Incentives, strengthening the case that GUIDE alone leads to better wellness outcomes than GUIDE with financial incentives for military-affiliated participants. These findings are extremely encouraging and suggest that GUIDE may just as if not more effective than existing mental health-related apps created by the Veterans Affairs (VA) or Department of Defense (DoD)[84]. While these sub-group results are promising, they are underpowered and must be interpreted with caution. Future research should include a larger number of veterans and military personnel to validate our findings.

Sex Differences

Our analyses by sex reveal that GUIDE may be an especially good fit for males, which is promising since males make up about 75% of first responders[85], 82.5% of active-duty military[86], and 89% of veterans[87]. Additionally, males are more likely to commit suicide[88], are less likely to seek help or avoid treatment due to stigma[89].

Male participants allocated to GUIDE+Incentives improved significantly over the waitlist on

social connectedness (Relationships), and had the most significant improvements over time for our selected outcome measures (PERMA Overall, WHO-5, Positive emotion, Emotion dysregulation, Depression, Anxiety, Meaning, Relationships), indicating that they may have responded more to financial incentives than females, a finding which aligns with previous research[90]. This suggests that incorporating financial incentives may yield more positive outcomes and app engagement for male GUIDE users.

Secondary Outcomes

App engagement

Overall 67.10% of participants (51/76) met our minimum usage standards for all weeks, while only 7.89% of participants (6/76) met the recommended benchmarks for all weeks. This indicates that the recommended benchmarks may be high for this population, since individuals willing to volunteer for a clinical trial are likely more motivated to persist than average.

The app attrition rate of 13.31% ($n=10$) was higher than the trial attrition rate of 6.09% ($n=7$)—both of which were substantially lower than the weighted meta-analytic attrition rate of 24.7% from a recent study of mindfulness apps, a promising sign about the uptake of GUIDE among first responders and military.[91] Attrition was higher in GUIDE conditions than the waitlist (which is typical for digital mental health interventions), and was highest for the GUIDE+Incentives condition, indicating that app usage and financial incentives may have been considered burdensome by some participants.

Financial incentives significantly increased the average number of posts and replies in the small group chat of GUIDE. This suggests that financial incentives may drive engagement with the peer support element of GUIDE, or that the peer support feature has room to improve on usability or acceptability to keep users engaged. Financial incentives did not lead to a significantly more lessons completed, indicating that the education feature keeps users engaged and that they are motivated to complete these lessons with or without opportunity for compensation.

When assessing how app engagement (overall and by feature) was correlated with outcome measures, lessons, posts and replies were associated with improvements on wellness and emotional wellbeing measures. Additionally, all features were associated with an improvement on Accomplishment, indicating participants who engaged with the app more felt more accomplished, regardless of feature used. Replies and likes, indicative of peer-to-peer interaction, were significantly correlated with a reduction in depression symptoms, indicating that these interactions may be beneficial for mental health. This finding is important to keep in mind, since the number of replies per participant was highly dependent on whether they received financial incentives. Figuring out a way to increase engagement and interaction between users may lead to an improvement across these outcomes, especially for mental health measures. Financial incentives may be an effective way to encourage use for some participants.

While financial incentives effectively increased app engagement, particularly in small group chats, they did not lead to corresponding improvements in primary outcomes. This suggests that intrinsic motivation may be essential for first responders, veterans, and military personnel to benefit from an app like GUIDE. According to self-determination theory, individuals respond differently to financial incentives based on their intrinsic values and beliefs. Because financial materialism is a less reliable motivator than financial altruism—especially in service-oriented professions—gamification strategies that foster intrinsic motivation may be more effective in driving sustained app engagement.[92] This explanation is supported by our dropout analysis, which found that participants with no prior behavioral healthcare experience and lower app

usage during the trial (i.e., those less intrinsically motivated to engage with mental health interventions) were significantly less likely to complete the post-trial assessment, regardless of financial incentives.

Implementation Outcomes & Technical Merit

Appropriateness and feasibility, the suitability of an intervention for a given population and the practicality of implementation in real-world conditions, were good according to our participants. This indicates that GUIDE is a good fit for first responders and military, and that it would be reasonable to implement given available resources.

Acceptability, or the degree to which an intervention is viewed positively, was satisfactory. High acceptability is important for digital mental health interventions, since it has been identified as a factor that increases both uptake and adherence among individuals with depression and anxiety. [93] Exploring opportunities to increase acceptability may improve outcomes for GUIDE members.

Usability was above average for GUIDE, but scores indicate there is room for improvement on this measure as well. Usability is critical when designing an effective digital mental health intervention, since users are more likely to stop using an app due to usability issues (e.g., bugs and poor user interaction design, data loss, battery and memory usage, and lack of explanation or guidance).[94] Usability scores for this trial may have been impacted by app outages experienced during the second half of the trial.

Overall, GUIDE performs above average across implementation outcomes and technical merit. Focusing on acceptability and usability can help take this app to the next level, which in turn could help drive adherence, engagement, and ultimately, further improvement on wellness and mental health outcomes.

Limitations

Our study has several limitations. First, we relied exclusively on subjective self-report data in the trial. Although we considered using objective measures related to work performance to complement our results, we avoided this for the sake of confidentiality so that participants could join the trial without informing their employers. In the future, a study like this could limit recruitment to a specified number of agencies and gather workplace metrics for all employees to avoid the breach of confidentiality. Second, our sample includes only those motivated enough to voluntarily participate in a clinical trial and also was limited to those interested in testing a mental health and wellness app, which may not be representative of the larger first responder, veteran and military population. Third, several aspects unique to our sample may limit generalizability and external validity of our findings. Participants were higher than average in anxiety and depression symptoms as compared to other studies of first responders and military, indicating that generalizability to a larger 'warrior' population may be difficult. Additionally, participation required a certain amount of digital literacy and accessibility, meaning that these results may not be generalizable to non-digital populations, especially older veterans and first responders. Demographic features, including a mostly male sample, limit generalizability to underrepresented groups like female first responders. Fourth, we did not include an item about leadership status/rank in our demographic survey, a factor that may have influenced outcomes. Future research should be sure to include an item that assesses for whether a participant is in a leadership or managing position. Fourth, 4-weeks may not have been enough time to see substantial change in our measures. Research on other digital wellness apps suggests that efficacy improves over time[83]. Without assessing long-term adherence, the study does not establish whether GUIDE fosters sustainable behavior change. Future studies should assess

outcome measures across a longer time horizon, and should incorporate a follow-up assessment 3 or 6 months after the intervention. Fifth, we relied on listwise deletions in our primary outcome analysis for participants who did not complete their post-trial assessment, which is not the most robust method of estimation. Research indicates that listwise deletion may lead to an overestimate of treatment effectiveness and bias results. Future research should consider pairwise deletions or imputation methods to assess.

Sixth, this study employed a waitlist-control design rather than using a direct comparator. Since this was the first randomized controlled trial of GUIDE, we chose a non-active comparison to assess the effects of GUIDE use relative to no intervention. As a result, outcomes may have been influenced by expectancy effects of participants anticipating treatment. Additionally, due to this design choice, the effectiveness of GUIDE cannot be directly compared to similar interventions. Future research should include a comparator to evaluate these effects against a similar mHealth product. Seventh, neither the participant nor the research team were blinded to the assignment. Future iterations of this study should include a comparator so that at least the participant is not aware of their assigned treatment group. Finally, this study relied on a multiplicity of analyses, which may have increased the likelihood of Type I Error. We addressed this risk using Bonferroni corrections, though this conservative method may have increased the risk of Type II Errors. Future research should use the Benjamini-Hochberg procedure in a sensitivity analysis to assess for the effects of Type II Error.

Conclusions

This trial indicates that The GUIDE App is a feasible and appropriate intervention with the potential to improve mental health and wellbeing of first responders, veterans and military. GUIDE-only participants improved significantly compared to the waitlist on wellness (PERMA Overall), emotional wellbeing (Positive emotion), and mental health (PHQ-8 for depression and GAD-7 for anxiety symptoms). Financial incentives increased engagement with peer support aspects of GUIDE, but did not lead to significant improvements over the waitlist in between group comparisons. More lessons, posts and replies completed were correlated with improvements in wellness and emotional wellbeing. Increased use of any feature was correlated with increased personal growth. Exploratory subgroup analyses suggest GUIDE as a mental health and wellness intervention may be effective for active first responders, those with military affiliation, and males. Future research should test the effect of longer-term GUIDE use and whether improvements sustain.

Acknowledgements

We would like to acknowledge all the incredible staff at The GUIDE App that supported the implementation of this trial, including but not limited to (in alphabetical order) Shawn Carey, Andy Roke, and Patrick Sandone.

Footnotes

Contributorship: Conceptualization: HN. Data curation: MD. Formal analysis: MD and HN. Funding acquisition: HN. Investigation: MD. Methodology: HN. Project administration: MD. Supervision: HN. Validation: HN. Visualization: MD. Writing – original draft: MD. Writing – review & editing: HN and MD. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Ethical approval: The ethics committee of the University of Pennsylvania approved this study (IRB number: 855282).

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The GUIDE App/Nlyten Corp.

Conflict of Interest: This study was funded by Nlyten Corp, the developer of The GUIDE app. While Nlyten Corp provided financial support for the research, it had no role in the study design, data collection, analysis, interpretation of data, or manuscript preparation. The authors declare no other conflicts of interest related to this study.

Guarantor: HJN.

ORCID

Morgan

Dunphy

iD:

<https://orcid.org/0009-0007-3701-2979>

Heather J Nuske <https://orcid.org/0000-0001-6618-3768>

Supplemental material: Supplemental material for this article is available online and published with this article. The full protocol can be accessed through Clinicaltrials.gov (NCT06336967).

Data availability: Deidentified data supporting the findings of this study are available as a supplementary file accompanying this manuscript. Additional details can be provided upon request.

Abbreviations:

ANOVA	–	Analysis	of	Variance
CONSORT	–	Consolidated	Standards	of Reporting Trials
DERS-SF	–	Difficulties in	Emotion Regulation	scale, Short form
GAD-7	–	Generalized Anxiety Disorder	7-item scale	for anxiety symptoms
PERMA	–	Positive emotion, Engagement, Relationships, Meaning, Accomplishment		
PHQ-8	–	Patient Health Questionnaire	8-item scale	for depression symptoms
PWS	–	Personal	Wellbeing	Score
WHO	–	World	Health	Organization
WHO-5 – World Health Organization 5-item well-being index				

References

1. Fullerton CS, Ursano RJ, Wang L. Acute Stress Disorder, Posttraumatic Stress Disorder, and Depression in Disaster or Rescue Workers. *Am J Psychiatry* 2004 Aug 1;161(8):1370–1376. doi: 10.1176/appi.ajp.161.8.1370
2. Moradi Y, Dowran B, Sepandi M. The global prevalence of depression, suicide ideation, and attempts in the military forces: a systematic review and Meta-analysis of cross sectional

- studies. BMC Psychiatry 2021 Dec;21(1):510. doi: 10.1186/s12888-021-03526-2
3. Cardozo BL, Holtz TH, Kaiser R, Gotway CA, Ghitis F, Toomey E, Salama P. The Mental Health of Expatriate and Kosovar Albanian Humanitarian Aid Workers: The Mental Health of Expatriate and Kosovar Albanian Humanitarian Aid Workers. Disasters 2005 Jun;29(2):152–170. doi: 10.1111/j.0361-3666.2005.00278.x
 4. Thériault FL, Gardner W, Momoli F, Garber BG, Kingsbury M, Clayborne Z, Cousineau-Short DY, Sampasa-Kanyinga H, Landry H, Colman I. Mental Health Service Use in Depressed Military Personnel: A Systematic Review. Mil Med 2020 Aug 14;185(7–8):e1255–e1262. doi: 10.1093/milmed/usaa015
 5. Courtney JA, Francis AJP, Paxton SJ. Caring for the Country: Fatigue, Sleep and Mental Health in Australian Rural Paramedic Shiftworkers. J Community Health 2013 Feb;38(1):178–186. doi: 10.1007/s10900-012-9599-z
 6. Macdonald-Gagnon G, Stefanovics EA, Potenza MN, Pietrzak RH. Generalized anxiety and mild anxiety symptoms in U.S. military veterans: Prevalence, characteristics, and functioning. J Psychiatr Res 2024 Mar;171:263–270. doi: 10.1016/j.jpsychires.2024.02.013
 7. Black DW, Carney CP, Peloso PM, Woolson RF, Schwartz DA, Voelker MD, Barrett DH, Doebbeling BN. Gulf War Veterans With Anxiety: Prevalence, Comorbidity, and Risk Factors. Epidemiology 2004 Mar;15(2):135–142. doi: 10.1097/01.EDE.0000103188.18059.21
 8. Bennett P, Williams Y, Page N, Hood K, Woollard M. Levels of mental health problems among UK emergency ambulance workers. Emerg Med J British Association for Accident and Emergency Medicine; 2004 Mar 1;21(2):235–236. PMID:14988360
 9. Aldrich RS, Cerel J. Occupational Suicide Exposure and Impact on Mental Health: Examining Differences Across Helping Professions. OMEGA - J Death Dying SAGE Publications Inc; 2022 May 1;85(1):23–37. doi: 10.1177/0030222820933019
 10. Walker A, McKune A, Ferguson S, Pyne DB, Rattray B. Chronic occupational exposures can influence the rate of PTSD and depressive disorders in first responders and military personnel. Extreme Physiol Med 2016 Dec;5(1):8, s13728-016-0049–x. doi: 10.1186/s13728-016-0049-x
 11. EMT AS MSN, APN, CRNP, FNP-BC. First Responders and PTSD: A Literature Review. JEMS EMS Emerg Med Serv - Train Paramed EMT News. 2020. Available from: <https://www.jems.com/mental-health-wellness/first-responders-and-ptsd-a-literature-review/> [accessed Dec 1, 2024]
 12. Whitworth J, Galusha J, Carbajal J, Ponder WN, Schuman DL. Affective Depression Mediates PTSD to Suicide in a Sample of Treatment-Seeking First Responders. J Occup Environ Med 2023 Mar;65(3):249–254. doi: 10.1097/JOM.0000000000002724
 13. Smith LJ, Zegel M, Bartlett BA, Lebeaut A, Vujanovic AA. Posttraumatic Stress and

- Alcohol Use Among First Responders: In: Bowers CA, Beidel DC, Marks MR, editors. *Adv Psychol Ment Health Behav Stud IGI Global*; 2019. p. 32–48. doi: 10.4018/978-1-5225-9803-9.ch003 ISBN:978-1-5225-9803-9
14. Arble E, Daugherty AM, Arnetz BB. Models of First Responder Coping: Police Officers as a Unique Population. *Stress Health J Int Soc Investig Stress* 2018 Dec;34(5):612–621. PMID:29882624
 15. Beauchamp AM, Weerakoon SM, Ponder WN, Jetelina KK. Possible substance use disorders among first responders during the COVID-19 era: a quasi-experimental study of personal and residential vulnerability. *Am J Drug Alcohol Abuse* Taylor & Francis Ltd; 2022 Nov;48(6):724–733. doi: 10.1080/00952990.2022.2088376
 16. Bonumwezi JL, Tramutola D, Lawrence J, Kobezak HM, Lowe SR. Posttraumatic stress disorder symptoms, work-related trauma exposure, and substance use in first responders. *Drug Alcohol Depend* 2022 Aug 1;237:109439. doi: 10.1016/j.drugalcdep.2022.109439
 17. Wagner TH, Harris KM, Federman B, Dai L, Luna Y, Humphreys K. Prevalence of substance use disorders among veterans and comparable nonveterans from the National Survey on Drug Use and Health. *Psychol Serv* 2007 Aug;4(3):149–157. doi: 10.1037/1541-1559.4.3.149
 18. Sleep Problems, Depression, Substance Use, Social Bonding, and Quality of Life in Professional Firefighters. Available from: <https://oce-ovid-com.proxy.library.upenn.edu/article/00043764-201108000-00015/HTML> [accessed Nov 20, 2024]
 19. Teeters J, Lancaster C, Brown D, Back S. Substance use disorders in military veterans: prevalence and treatment challenges. *Subst Abuse Rehabil* 2017 Aug;Volume 8:69–77. doi: 10.2147/SAR.S116720
 20. Lan C-W, Fiellin DA, Barry DT, Bryant KJ, Gordon AJ, Edelman EJ, Gaither JR, Maisto SA, Marshall BDL. The Epidemiology of Substance Use Disorders in US Veterans: A Systematic Review and Analysis of Assessment Methods. *Am J Addict* 2016 Jan;25(1):7–24. PMID:26693830
 21. Zolnikov TR, Furio F. Stigma on first responders during COVID-19. *Stigma Health* 2020 Nov;5(4):375–379. doi: 10.1037/sah0000270
 22. Fisher MP, Lavender CD. Ensuring Optimal Mental Health Programs and Policies for First Responders: Opportunities and Challenges in One U.S. State. *Community Ment Health J* 2023;59(7):1341–1351. PMID:36947386
 23. Stanley IH, Hom MA, Joiner TE. A systematic review of suicidal thoughts and behaviors among police officers, firefighters, EMTs, and paramedics. *Clin Psychol Rev* 2016 Mar;44:25–44. doi: 10.1016/j.cpr.2015.12.002
 24. Morral AR, Schell TL, Smart R. Comparison of Suicide Rates Among US Veteran and

- Nonveteran Populations. *JAMA Netw Open* 2023 Jul 18;6(7):e2324191. doi: 10.1001/jamanetworkopen.2023.24191
25. Lawrence DS, Dockstader J. Law Enforcement Deaths by Suicide.
 26. The Ruderman White Paper Update on Mental Health and Suicide of First Responders. Ruderman Fam Found. Available from: https://rudermanfoundation.org/white_papers/the-ruderman-white-paper-update-on-mental-health-and-suicide-of-first-responders/ [accessed Nov 17, 2024]
 27. Kleim B, Westphal M. Mental health in first responders: A review and recommendation for prevention and intervention strategies. *Traumatology* 2011 Dec;17(4):17–24. doi: 10.1177/1534765611429079
 28. Pflanz SE, Ogle AD. Job Stress, Depression, Work Performance, and Perceptions of Supervisors in Military Personnel. *Mil Med* 2006 Sep 1;171(9):861–865. doi: 10.7205/MILMED.171.9.861
 29. University of Phoenix Survey Finds 93 Percent of First Responders Say Mental Health is as Important as Physical Health - University of Phoenix. Available from: <https://www.phoenix.edu/press-release/university-of-phoenix-survey-finds-93-percent-of-first-responders-say-mental-health-is-as-important-as-physical-health.html> [accessed Dec 1, 2024]
 30. Ben-Zeev D, Corrigan PW, Britt TW, Langford L. Stigma of mental illness and service use in the military. *J Ment Health* Routledge; 2012 Jun;21(3):264–273. doi: 10.3109/09638237.2011.621468
 31. Crowe A, Glass JS, Lancaster MF, Raines JM, Waggy MR. A Content Analysis of Psychological Resilience Among First Responders and the General Population. Sage Open SAGE Publications; 2017 Jan 1;7(1):2158244017698530. doi: 10.1177/2158244017698530
 32. Sharp M-L, Fear NT, Rona RJ, Wessely S, Greenberg N, Jones N, Goodwin L. Stigma as a Barrier to Seeking Health Care Among Military Personnel With Mental Health Problems. *Epidemiol Rev* 2015 Jan 1;37(1):144–162. doi: 10.1093/epirev/mxu012
 33. Bowers CA, Beidel DC, Marks MR. Obstacles to mental health treatment: Similarities and differences among first responder groups. *J Community Saf Well-Being* 2022 Jun 16;7(2):42–46. doi: 10.35502/jcswb.248
 34. Jones S, Agud K, McSweeney J. Barriers and Facilitators to Seeking Mental Health Care Among First Responders: “Removing the Darkness.” *J Am Psychiatr Nurses Assoc* 2020 Jan;26(1):43–54. doi: 10.1177/1078390319871997
 35. Rafferty LA, Wessely S, Stevelink SAM, Greenberg N. The journey to professional mental health support: a qualitative exploration of the barriers and facilitators impacting military veterans’ engagement with mental health treatment. *Eur J Psychotraumatology* 2019 Dec 31;10(1):1700613. doi: 10.1080/20008198.2019.1700613

36. Horan KA, Marks M, Ruiz J, Bowers C, Cunningham A. Here for My Peer: The Future of First Responder Mental Health. *Int J Environ Res Public Health Multidisciplinary Digital Publishing Institute*; 2021 Jan;18(21):11097. doi: 10.3390/ijerph182111097
37. Mercier J-M, Hosseiny F, Rodrigues S, Friio A, Brémault-Phillips S, Shields D, Dupuis G. Peer Support Activities for Veterans, Serving Members, and Their Families: Results of a Scoping Review. *Int J Environ Res Public Health* 2023 Feb 18;20(4):3628. doi: 10.3390/ijerph20043628
38. Royle L, Keenan P, Farrell D. Issues of stigma for first responders accessing support for post traumatic stress. *Int J Emerg Ment Health US: Chevron Publishing*; 2009;11(2):79–86.
39. Hilty DM, Chan S, Hwang T, Wong A, Bauer AM. Advances in Mobile Mental Health: Opportunities and Implications for the Spectrum of E-Mental Health Services. *Focus* 2018 Jul;16(3):314–327. doi: 10.1176/appi.focus.16301
40. Moghimi E, Knyahnytska Y, Omrani M, Nikjoo N, Stephenson C, Layzell G, Frederic Simpson AI, Alavi N. Benefits of Digital Mental Health Care Interventions for Correctional Workers and Other Public Safety Personnel: A Narrative Review. *Front Psychiatry* 2022 Jul 8;13:921527. doi: 10.3389/fpsyt.2022.921527
41. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord* 2017 Aug 15;218:15–22. doi: 10.1016/j.jad.2017.04.046
42. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, Pulkki-Råback L, Baumeister H, Sander LB. Digital interventions for the treatment of depression: A meta-analytic review. *Psychol Bull* 2021 Aug;147(8):749–786. doi: 10.1037/bul0000334
43. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, Sarris J. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017;16(3):287–298. doi: 10.1002/wps.20472
44. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 2019;18(3):325–336. doi: 10.1002/wps.20673
45. Spijkerman MPJ, Pots WTM, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. *Clin Psychol Rev* 2016 Apr 1;45:102–114. doi: 10.1016/j.cpr.2016.03.009
46. Seppälä EM, Bradley C, Moeller J, Harouni L, Nandamudi D, Brackett MA. Promoting Mental Health and Psychological Thriving in University Students: A Randomized Controlled Trial of Three Well-Being Interventions. *Front Psychiatry Frontiers*; 2020 Jul 15;11. doi: 10.3389/fpsyt.2020.00590

47. Groot J, MacLellan A, Butler M, Todor E, Zulfiqar M, Thackrah T, Clarke C, Brosnan M, Ainsworth B. The Effectiveness of Fully Automated Digital Interventions in Promoting Mental Well-Being in the General Population: Systematic Review and Meta-Analysis. *JMIR Ment Health* 2023 Oct 19;10(1):e44658. doi: 10.2196/44658
48. GUIDE App | Build a Life You Love in 15 Minutes a Day. GUIDE. Available from: <https://theguideapp.com/> [accessed Mar 9, 2025]
49. Stewart V, Merron K. The GUIDE Pilot Study White Paper. GUIDE * 201 Lackawanna Avenue, Scranton, PA 18503; 2023.
50. Seligman MEP. Flourish: A visionary new understanding of happiness and well-being. New York, NY, US: Free Press; 2011. p. xii, 349. ISBN:978-1-4391-9075-3
51. Kern P. Questionnaires. Available from: <https://www.peggykern.org/questionnaires.html> [accessed Dec 1, 2024]
52. Butler J, Kern ML. The PERMA-Profiler: A brief multidimensional measure of flourishing. *Int J Wellbeing* 2016 Oct 13;6(3). doi: 10.5502/ijw.v6i3.526
53. Benson T, Sladen J, Liles A, Potts HWW. Personal Wellbeing Score (PWS)—a short version of ONS4: development and validation in social prescribing. *BMJ Open Qual* 2019 Apr;8(2):e000394. doi: 10.1136/bmjopen-2018-000394
54. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychother Psychosom* 2015 Mar 28;84(3):167–176. doi: 10.1159/000376585
55. Sischka PE, Costa AP, Steffgen G, Schmidt AF. The WHO-5 well-being index – validation based on item response theory and the analysis of measurement invariance across 35 countries. *J Affect Disord Rep* 2020 Dec 1;1:100020. doi: 10.1016/j.jadr.2020.100020
56. Kaufman EA, Xia M, Fosco G, Yaptangco M, Skidmore CR, Crowell SE. The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): Validation and Replication in Adolescent and Adult Samples. *J Psychopathol Behav Assess* 2016 Sep;38(3):443–455. doi: 10.1007/s10862-015-9529-3
57. Burton AL, Brown R, Abbott MJ. Overcoming difficulties in measuring emotional regulation: Assessing and comparing the psychometric properties of the DERS long and short forms. *Cogent Psychol* 2022 Dec 31;9(1):2060629. doi: 10.1080/23311908.2022.2060629
58. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009 Apr;114(1–3):163–173. doi: 10.1016/j.jad.2008.06.026
59. Pressler SJ, Subramanian U, Perkins SM, Gradus-Pizlo I, Kareken D, Kim J, Ding Y, Sauve MJ, Sloan R. Measuring Depressive Symptoms in Heart Failure: Validity and Reliability of

- the Patient Health Questionnaire-8. *Am J Crit Care* 2011 Mar 1;20(2):146–152. doi: 10.4037/ajcc2010931
60. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092–1097. doi: 10.1001/archinte.166.10.1092
61. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, Herzberg PY. Validation and Standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the General Population. *Med Care* 2008 Mar;46(3):266–274. doi: 10.1097/MLR.0b013e318160d093
62. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, Boynton MH, Halko H. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci IS* 2017 Aug 29;12(1):108. PMID:28851459
63. Brooke J. SUS - A quick and dirty usability scale.
64. Bangor A, Kortum PhilipT, Miller JamesT. An Empirical Evaluation of the System Usability Scale. *Int J Hum-Comput Interact* Taylor & Francis Ltd; 2008 Aug;24(6):574–594. doi: 10.1080/10447310802205776
65. Lewis JR, Sauro J. Item Benchmarks for the System Usability Scale - JUX. JUX - J User Exp. 2018. Available from: <https://uxpajournal.org/item-benchmarks-system-usability-scale-sus/> [accessed Nov 25, 2024]
66. Eysenbach G, Group C-E. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. *J Med Internet Res* 2011 Dec 31;13(4):e1923. doi: 10.2196/jmir.1923
67. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod* 2013 Feb;38(1):52–54. PMID:23495371
68. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom* 2014 Sep;34(5):502–508. PMID:24697967
69. Nich C, Carroll KM. ‘Intention-to-treat’ meets ‘missing data’: implications of alternate strategies for analyzing clinical trials data. *Drug Alcohol Depend* 2002 Oct 1;68(2):121–130. PMID:12234641
70. (PDF) Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: A review. *ResearchGate* 2024 Oct 22; doi: 10.1179/174328809X405928
71. Richardson JTE. Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev* 2011 Jan 1;6(2):135–147. doi: 10.1016/j.edurev.2010.12.001
72. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 0 ed. Routledge; 2013. doi: 10.4324/9780203771587ISBN:978-1-134-74270-7

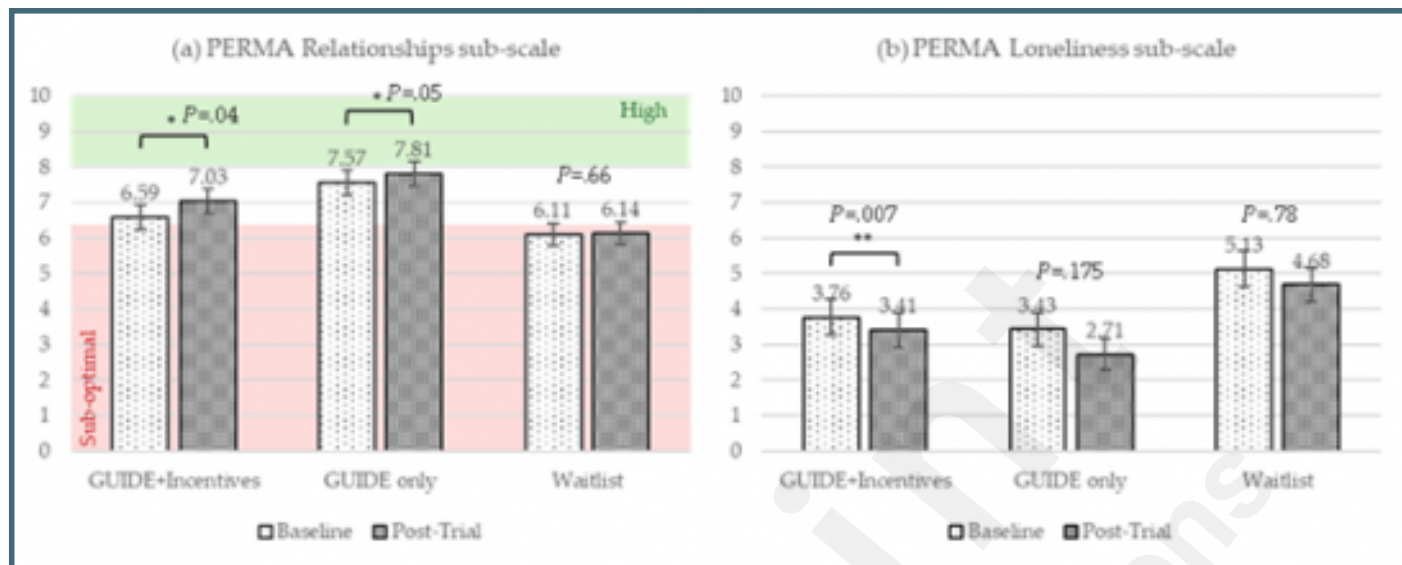
73. (PDF) Effect Size Estimates: Current Use, Calculations, and Interpretation. ResearchGate 2024 Oct 22; doi: 10.1037/a0024338
74. Boerma T, Hosseinpoor AR, Verdes E, Chatterji S. A global assessment of the gender gap in self-reported health with survey data from 59 countries. *BMC Public Health* 2016 Dec;16(1):675. doi: 10.1186/s12889-016-3352-y
75. Lipschitz JM, Pike CK, Hogan TP, Murphy SA, Burdick KE. The Engagement Problem: a Review of Engagement with Digital Mental Health Interventions and Recommendations for a Path Forward. *Curr Treat Options Psychiatry* 2023 Aug 25;10(3):119–135. doi: 10.1007/s40501-023-00297-3
76. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg* 2018 May;126(5):1763. doi: 10.1213/ANE.0000000000002864
77. Terlizzi E, Zablotsky, Benjamin B. Symptoms of Anxiety and Depression Among Adults: United States, 2019 and 2022. National Center for Health Statistics (U.S.); 2024 Nov. doi: 10.15620/cdc/164018
78. Jones S. Describing the Mental Health Profile of First Responders: A Systematic Review. *J Am Psychiatr Nurses Assoc* SAGE Publications Inc STM; 2017 May 1;23(3):200–214. doi: 10.1177/1078390317695266
79. Kligyte V, Connelly S, Thiel C, Devenport L. The Influence of Anger, Fear, and Emotion Regulation on Ethical Decision Making. *Hum Perform Routledge*; 2013 Sep 1;26(4):297–326. doi: 10.1080/08959285.2013.814655
80. Troy AS, Mauss IB. Resilience in the face of stress: emotion regulation as a protective factor. In: Southwick SM, Litz BT, Charney D, Friedman MJ, editors. *Resil Ment Health* 1st ed Cambridge University Press; 2011. p. 30–44. doi: 10.1017/CBO9780511994791.004ISBN:978-0-521-89839-3
81. Money can't buy happiness: A randomized controlled trial of a digital mental health app with versus without financial incentives - PMC. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10286537/> [accessed Dec 18, 2024]
82. Wild J, El-Salahi S, Esposti MD. The effectiveness of interventions aimed at improving wellbeing and resilience to stress in first responders: A systematic review.
83. Witarto BS, Visuddho V, Witarto AP, Bestari D, Sawitri B, Melapi TAS, Wungu CDK. Effectiveness of online mindfulness-based interventions in improving mental health during the COVID-19 pandemic: A systematic review and meta-analysis of randomized controlled trials. *PLOS ONE Public Library of Science*; 2022 Sep 21;17(9):e0274177. doi: 10.1371/journal.pone.0274177
84. Gould CE, Kok BC, Ma VK, Zapata AML, Owen JE, Kuhn E. Veterans Affairs and the Department of Defense mental health apps: A systematic literature review. *Psychol Serv*

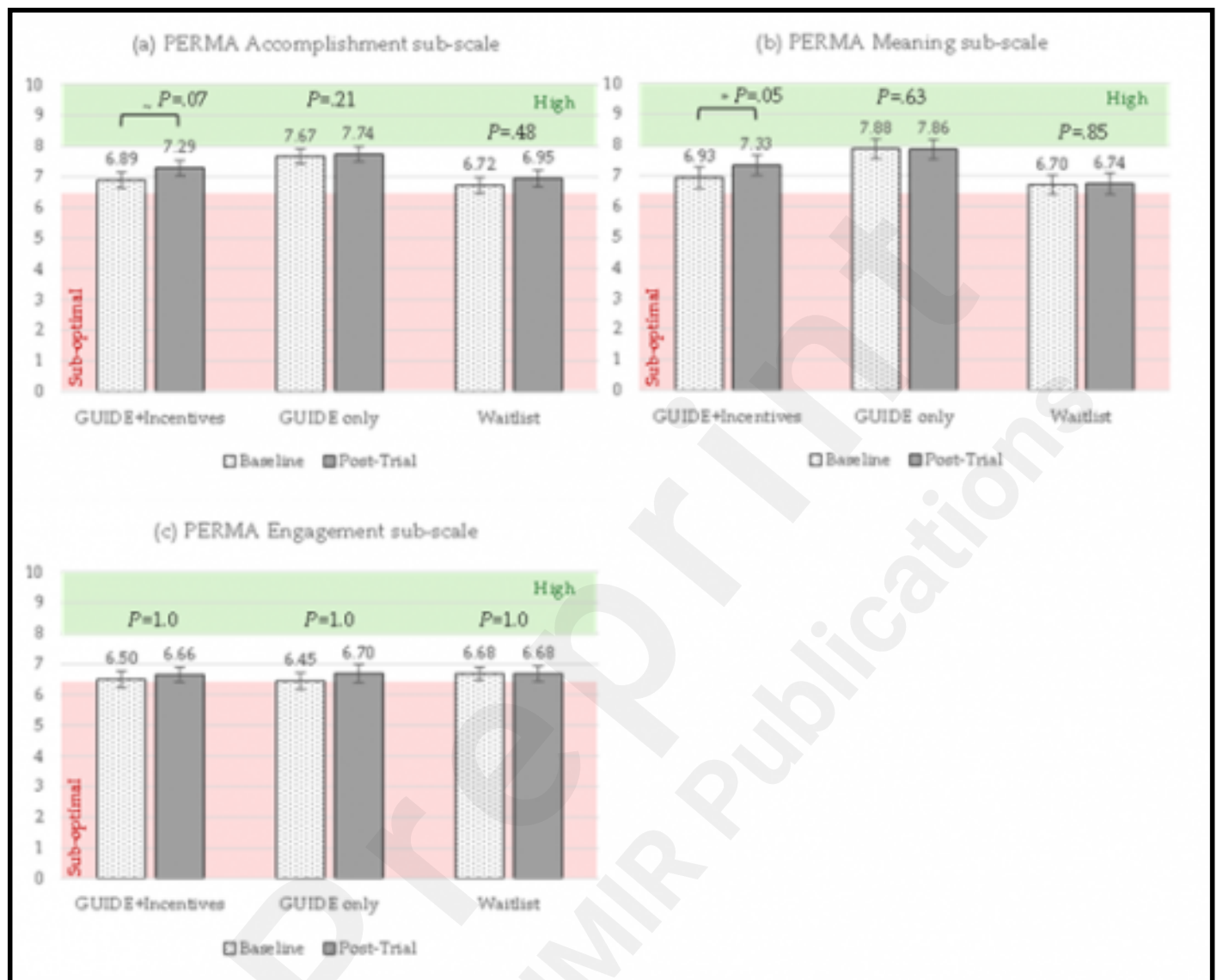
2019 May;16(2):196–207. doi: 10.1037/ser0000289

85. Men accounted for about 75 percent of workers in protective service occupations in 2020 : The Economics Daily: U.S. Bureau of Labor Statistics. Bur Labor Stat. Available from: <https://www.bls.gov/opub/ted/2021/men-accounted-for-about-75-percent-of-workers-in-protective-service-occupations-in-2020.htm> [accessed Dec 2, 2024]
86. 2022 Demographics Profile Active-Duty Members.
87. Schaeffer K. The changing face of America's veteran population. Pew Res Cent. 2023. Available from: <https://www.pewresearch.org/short-reads/2023/11/08/the-changing-face-of-americas-veteran-population/> [accessed Dec 2, 2024]
88. Carson LM, Marsh SM, Brown MM, Elkins KL, Tiesman HM. An analysis of suicides among first responders — Findings from the National Violent Death Reporting System, 2015–2017. *J Safety Res* 2023 Jun;85:361–370. doi: 10.1016/j.jsr.2023.04.003
89. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. *Am Psychol US: American Psychological Association*; 2003;58(1):5–14. doi: 10.1037/0003-066X.58.1.5
90. Czap NV, Czap HJ, Khachaturyan M, Burbach ME. Comparing Female and Male Response to Financial Incentives and Empathy Nudging in an Environmental Context. *Rev Behav Econ* 2018 Mar 26;5(1):61–84. doi: 10.1561/105.00000079
91. Linardon J. Rates of attrition and engagement in randomized controlled trials of mindfulness apps: Systematic review and meta-analysis. *Behav Res Ther* 2023 Nov 1;170:104421. doi: 10.1016/j.brat.2023.104421
92. Stone DN, Bryant SM, Wier B. Why Are Financial Incentive Effects Unreliable? An Extension of Self-Determination Theory. *Behav Res Account* 2010 Jan 1;22(2):105–132. doi: 10.2308/bria.2010.22.2.105
93. Gulliver A, Caelear AL, Sunderland M, Kay-Lambkin F, Farrer LM, Batterham PJ. Predictors of acceptability and engagement in a self-guided online program for depression and anxiety. *Internet Interv* 2021 Sep 1;25:100400. doi: 10.1016/j.invent.2021.100400
94. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. *Health Informatics J SAGE Publications Ltd*; 2020 Sep 1;26(3):2042–2066. doi: 10.1177/1460458219896492

Supplementary Files

Figures

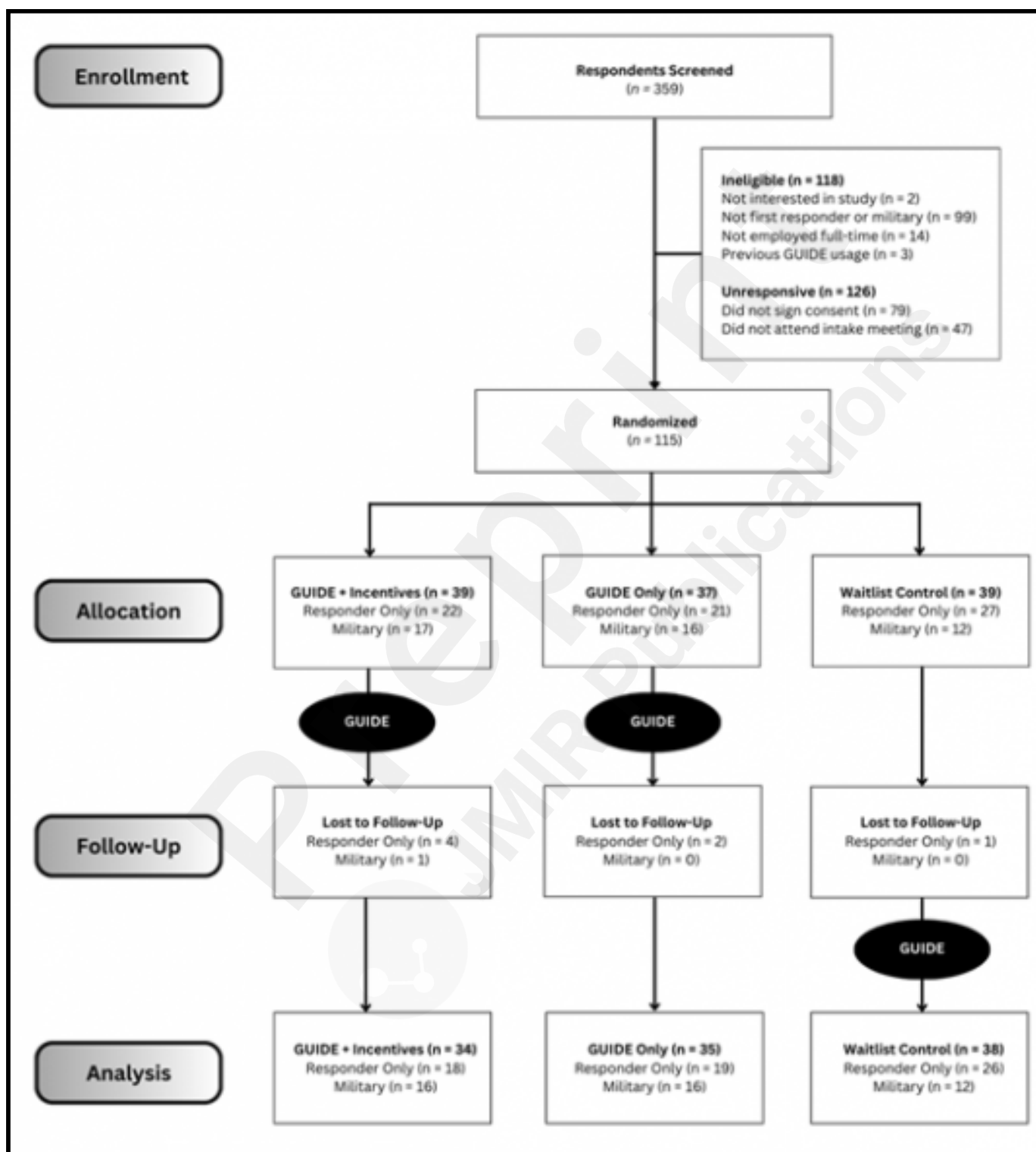


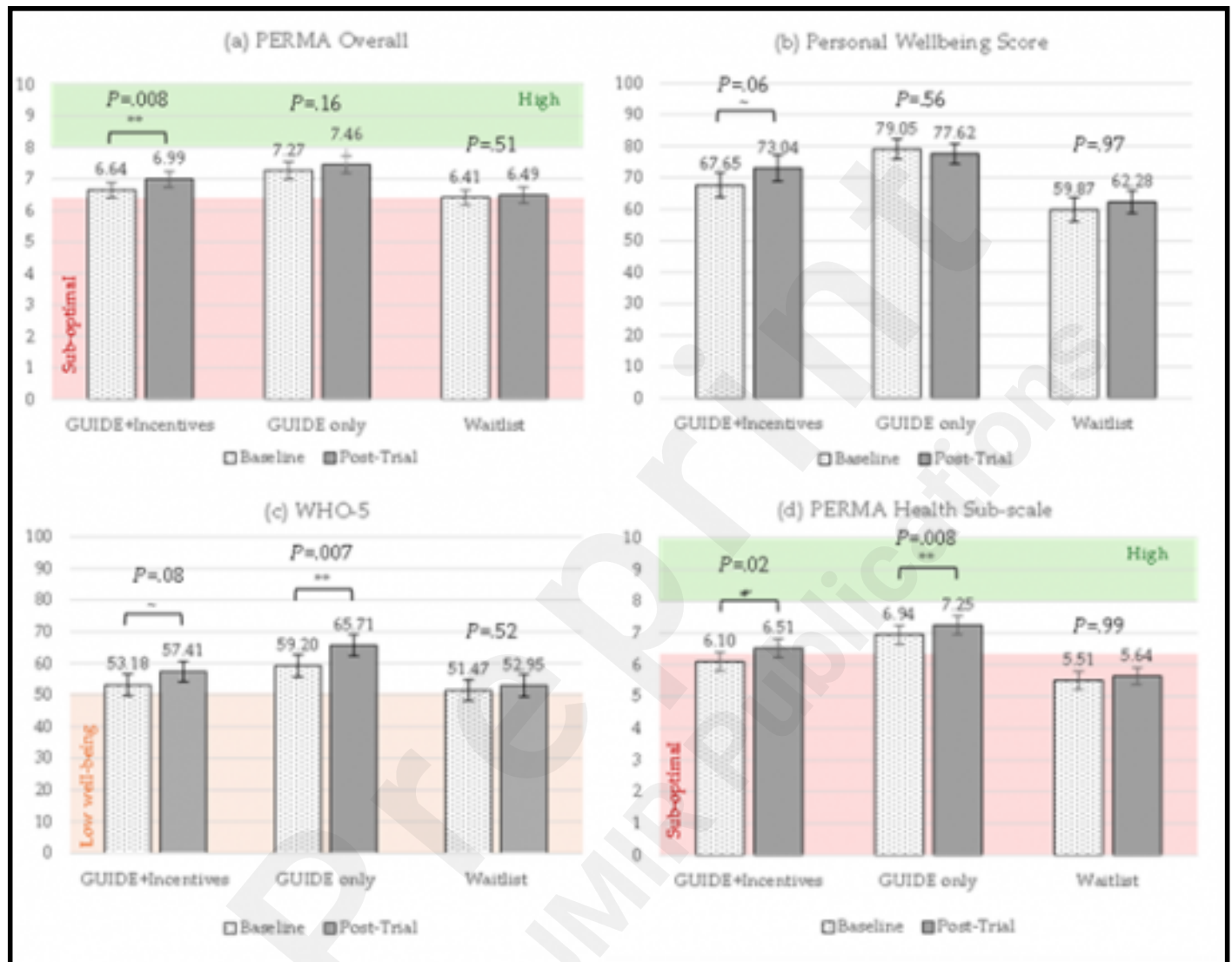


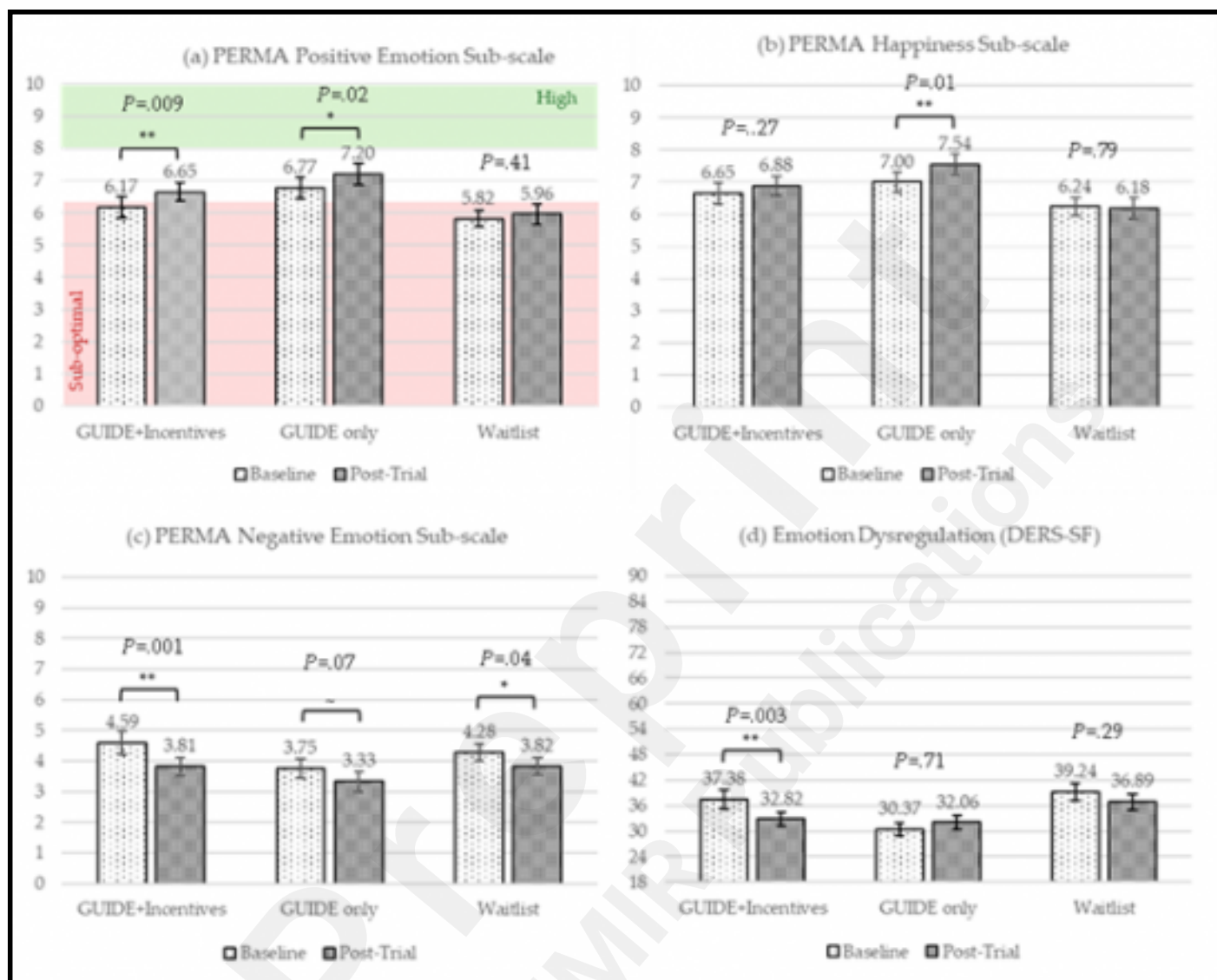
The GUIDE App Model.

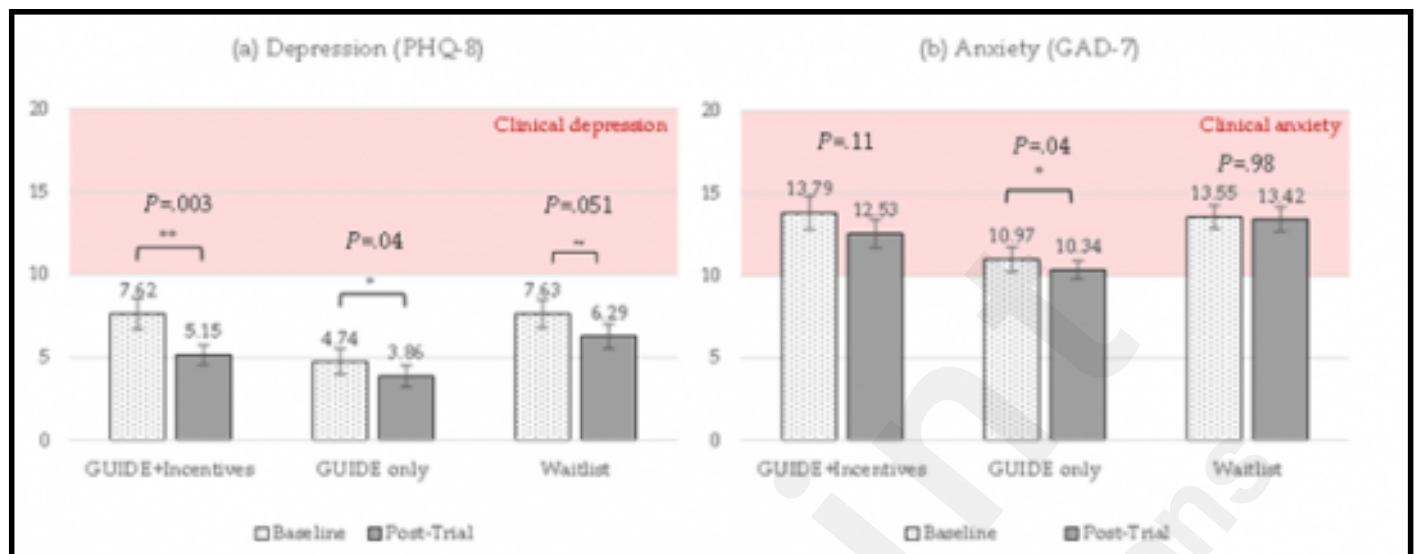


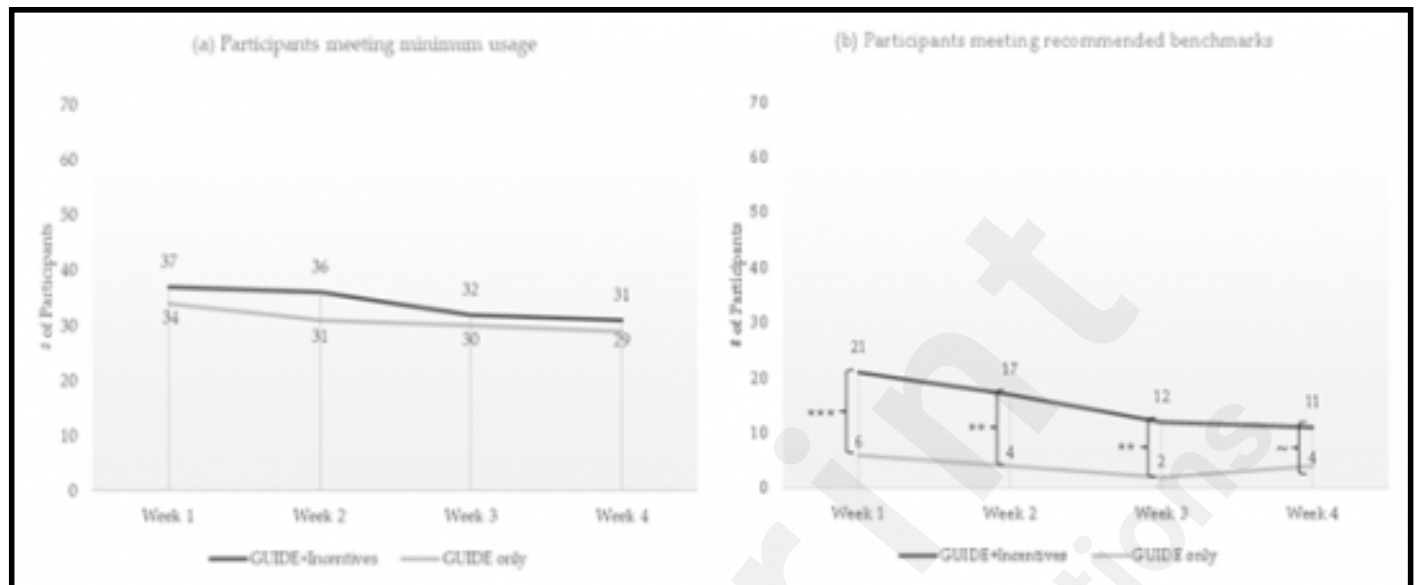
CONSORT flow diagram of the parallel randomized control trial. Note: Responder only=first responder without military affiliation. Military=veteran, reservist or active-duty. Lost to follow-up=allocated participant that did not complete post-trial assessment survey.











Multimedia Appendixes

Deidentified data from randomized controlled trial.

URL: <http://asset.jmir.pub/assets/8bd0af7879d8ff86cfc11adc614fed76.xlsx>

CONSORT-EHEALTH checklist.

URL: <http://asset.jmir.pub/assets/86a876912fccd76b03dce85a134c9d47.pdf>

Supplementary materials - tables and charts.

URL: <http://asset.jmir.pub/assets/4c242dc09ad5e978c6cf9afd6103b292.docx>

GUIDE Course Inventory.

URL: <http://asset.jmir.pub/assets/503cdab30d3f0d64a0600b5af95e89bd.pdf>