AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Assessment

ISSN: 1040-3590

2024, Vol. 36, No. 1, 66–80
https://doi.org/10.1037/pas0001288

Intervening on High-Risk Responses During Ecological Momentary Assessment of Suicidal Thoughts: Is There an Effect on Study Data?

Kate H. Bentley[1, 2, 3], Alexander J. Millner[3, 4], Adam Bear[3], Lia Follet[3], Rebecca G. Fortgang[1, 2, 3], Kelly L. Zuromski[3, 4], Evan M. Kleiman[5], Daniel D. L. Coppersmith[3], Franchesca Castro-Ramirez[3], Yael Millgram[3], Adam Haim[6], Suzanne A. Bird[1, 2], and Matthew K. Nock[1, 3, 4]

[1] Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, United States
[2] Department of Psychiatry, Harvard Medical School
[3] Department of Psychology, Harvard University
[4] Franciscan Children's Hospital, Boston, Massachusetts, United States
[5] Department of Psychology, Rutgers, The State University of New Jersey
[6] National Institute of Mental Health, Washington, DC, United States

Ecological momentary assessment (EMA) is increasingly used to study suicidal thoughts and behaviors (STBs). There is a potential ethical obligation for researchers to intervene when receiving information about suicidal thoughts in real time. A possible concern, however, is that intervening when receiving responses that indicate high risk for suicide during EMA research may impact how participants respond to questions about suicidal thoughts and thus affect the validity and integrity of collected data. We leveraged data from a study of adults and adolescents ($N = 434$) recruited during a hospital visit for STBs to examine whether monitoring and intervening on high-risk responses affects subsequent participant responding. Overall, we found mixed support for the notion that intervening on high-risk responses influences participants' ratings. Although we observed some evidence of discontinuity in subsequent responses at the threshold used to trigger response-contingent interventions, it was not clear that such discontinuity was caused by the interventions; lower subsequent responses could be due to effective intervention, participant desire to not be contacted again, or regression to the mean. Importantly, the likelihood of completing surveys did not change from before to after response-contingent intervention. Adolescents were significantly more likely than adults, however, to change their initial suicidal intent ratings from above to below the high-risk threshold after viewing automated response-contingent pop-up messages. Studies explicitly designed to assess the potential impact of intervening on high-risk responses in real-time monitoring research are needed, as this will inform effective, scalable strategies for intervening during moments of high suicide risk.

Public Significance Statement
There may be an ethical obligation to intervene when participants in real-time monitoring studies report being at high risk for suicide; however, one possible concern is that intervening could impact how (and whether) participants respond to questions about suicide. Overall, we found some evidence that contacting participants when they report high suicidal intent may impact their responses, although the mechanism through which the lowering of responses occurs is not clear.

Keywords: ecological momentary assessment, suicide, suicidal thoughts, reactivity

Supplemental materials: https://doi.org/10.1037/pas0001288.supp

This article was published Online First November 2, 2023.

Kate H. Bentley https://orcid.org/0000-0002-6875-6440

This study received funding from the National Institute of Mental Health under Grants U01MH116928 (to Matthew K. Nock), K23MH120436 (to Kate H. Bentley), K23MH120439 (to Kelly L. Zuromski), K23MH132766 (to Rebecca G. Fortgang), and F31MH130055 (to Daniel D. L. Coppersmith).

The authors would like to thank Alexis Christie, Merryn Daniel, Dylan DeMarco, Victoria Joyce, Flynn Kelly, Onyinye Obi-Obasi, Narise Ramlal, and Tida Tambedou for their invaluable work conducting recruitment and research procedures with participants in this study. The authors would also like to thank all the study participants for their time and willingness to contribute to this research.

Kate H. Bentley played a lead role in conceptualization and writing–original draft, a supporting role in data curation, and an equal role in formal analysis, investigation, methodology, and supervision. Alexander J. Millner played a lead role in writing–review and editing, a supporting role in formal analysis, and an equal role in conceptualization, data curation, investigation, methodology, and supervision. Adam Bear played a lead role in data curation, formal analysis, and visualization and a supporting role in conceptualization, methodology, and writing–review and editing. Lia Follet played a supporting role in formal analysis, investigation, and writing–review and editing and an equal role in data curation. Rebecca G. Fortgang played a supporting role in writing–review and editing and an equal role in investigation, methodology, and supervision. Kelly L. Zuromski played a

continued

Suicidal thoughts and behaviors (STBs) are highly prevalent public health problems. In 2020, nearly 46,000 Americans died by suicide, making it the 12th leading cause of death in the United States (Centers for Disease Control and Prevention, 2020). Suicidal thoughts and nonfatal suicidal behaviors are also common. For example, it is estimated that 1.2 million suicide attempts take place in the United States each year (Centers for Disease Control and Prevention, 2020)—and are associated with considerable public health costs in their own right (Shepard et al., 2016). Unfortunately, our ability to predict and prevent STBs remains poor (Franklin et al., 2017; Woodford et al., 2019). This may be in part due to the field lacking a strong understanding of fundamental properties of STBs, including how suicidal thoughts unfold naturalistically (e.g., how quickly suicidal thoughts occur, how long they persist) and when and why people transition from thinking about suicide to engaging in suicidal behavior (Millner et al., 2020).

Methods of real-time monitoring—namely, ecological momentary assessment (EMA)—are powerful tools to improve our understanding of STBs as these phenomena naturally occur, as well as the proximal, time-varying risk factors that predict their occurrence (Kleiman et al., 2023). Real-time monitoring has several advantages over traditional assessment methods, including circumventing the well-established limitations of retrospective recall (Schacter, 1999; thus potentially promoting more accurate and ecologically valid reporting) and facilitating the collection of many—hundreds or even thousands of—data points within individuals over time. Self-report data on STBs, as well as related affective, cognitive, and behavioral experiences, gleaned from EMA have the potential to be combined with other real-time data streams, such as continuous physiological data from wearable biosensors or passive smartphone sensors, to comprehensively characterize and predict short-term increases in suicide risk (Kleiman et al., 2023). One long-term goal of much of this work is to inform the development of just-in-time adaptive interventions that deploy the right amount and type of support at the right time to individuals at risk (Coppersmith, Dempsey, et al., 2022).

Safety and Ethical Considerations in Real-Time Research on STBs

Given the promise of using real-time monitoring methods to study STBs, as well as the increasing ubiquity of smartphones (Pew Research Center, 2021) and accessibility of EMA data collection platforms, it is no surprise that recent years have witnessed a proliferation of EMA research on STBs (Ammerman & Law, 2022; Gee et al., 2020; Rabasco & Sheehan, 2022; Sedano-Capdevila et al., 2021). This relatively new area, however, brings notable safety and ethical considerations, some of which suicide researchers have been grappling with for decades (Fisher et al., 2002; Hom, Podlogar, et al., 2017; Pearson et al., 2001) and others that are unique to EMA. One of the key issues is whether (and how) to monitor and intervene in response to incoming real-time data reflecting high current risk of suicide. Given that EMA typically consists of asking participants to report on the severity of their current suicidal thoughts at the present moment, researchers may receive information in real time that a study participant is experiencing strong thoughts of suicide or high intent to act *right now*. This begs the question of whether researchers conducting this work have an ethical, clinical, or regulatory obligation (National Institute of Mental Health, n.d.) to monitor incoming real-time data from suicidal individuals and intervene upon indication that a person may make a suicide attempt in the very near future. Furthermore, as some research teams are now passively and continuously collecting global positioning system data alongside EMA, it may be possible to pinpoint a participant's exact location while they report being at high risk (or a suicide attempt is in progress), posing the possibility to send what could be lifesaving help right away.

A recent consensus statement on these issues documented strong (about 90%) agreement among a panel of experts (e.g., researchers, clinicians) and individuals with lived experience that investigators using intensive longitudinal methods to study STBs should review incoming data on an ongoing (at least every weekday) basis and deploy interventions based on self-reported suicide risk (Nock et al., 2021). Such interventions generally include contacting participants to conduct a suicide risk assessment or review a safety plan, or notifying an emergency contact. In contrast to this strong consensus, however, a recent systematic review revealed that nearly 40% of studies do *not* actually monitor and intervene upon incoming EMA data on suicide risk (Bentley et al., 2021). Certain study characteristics appeared to influence whether researchers do or do not intervene on high-risk responses; for example, studies conducted with adolescents (vs. adults) were more likely to monitor data and intervene in real time. Studies were also fairly evenly split (48% vs. 52%) between using and not using automated "pop-up" messages (that generally provide crisis resources) when survey responses met predefined risk thresholds (Bentley et al., 2021).

Why might researchers choose *not* to monitor and intervene on incoming real-time data on suicide risk? One reason is that doing so often requires significant person power and resources (e.g., trained clinicians to contact participants and conduct risk assessments). Certain data collection platforms may not have the capacity to send real-time notifications to study staff (e.g., some only store data locally on participants' phones during the study). Monitoring incoming data in real-time may also not be feasible for studies that enroll very large numbers of individuals nationwide (or even

supporting role in writing–review and editing and an equal role in investigation, methodology, and supervision. Evan M. Kleiman played a supporting role in conceptualization, funding acquisition, and writing–review and editing and an equal role in investigation, methodology, and supervision. Daniel D. L. Coppersmith played a supporting role in conceptualization, investigation, methodology, and writing–review and editing. Franchesca Castro-Ramirez played a supporting role in investigation, methodology, and writing–review and editing. Yael Millgram played a supporting role in investigation, methodology, and writing–review and editing. Adam Haim played a supporting role in funding acquisition, investigation, methodology, and writing–review and editing. Suzanne A. Bird played a supporting role in funding acquisition, investigation, methodology, resources, and writing–review and editing. Matthew K. Nock played a lead role in funding acquisition, investigation, methodology, project administration, resources and supervision, a supporting role in data curation, and an equal role in conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Kate H. Bentley, Department of Psychiatry, Massachusetts General Hospital, 165 Cambridge Street, Boston, MA 02114, United States. Email: kbentley@mgh.harvard.edu

globally) or are not well funded. Researchers may also be concerned that responding promptly to incoming data may be outside their roles as investigators (and more in line with, e.g., crisis responders). Another factor is that studies in which individuals are recruited to participate anonymously—for example, via online methods—are necessarily limited in the potential to intervene (at least via outreach from study staff) in response to incoming data. Beyond these practical considerations, some have raised another potential, not trivial concern: that monitoring and intervening on high-risk responses in observational, real-time monitoring studies of STBs could meaningfully impact the resultant data.

Potential Impacts of Intervening on High-Risk Responses in Observational Research

It is possible that intervening in response to incoming data on STBs may influence how (or whether) participants respond to study assessments and thus impact the validity and integrity of collected data. As the goal of most real-time monitoring research is to observe phenomena unfold naturalistically, this is an important consideration, as researchers generally strive to minimize the impact of their study procedures on the experiences they seek to observe. There are several possible ways (or mechanisms through which) monitoring and intervening on high-risk responses in real-time research on STBs could influence study data, two of which we describe below.

Treatment Effects

One possible way that monitoring and intervening on incoming high-risk responses may influence study data is if such interventions effectively reduce the likelihood of experiencing STBs. There is now much evidence showing that brief suicide-focused interventions—namely, conducting follow-up contacts and developing/reviewing safety plans, both of which are included in many studies' safety protocols (Bentley et al., 2021; Nock et al., 2021)—can prevent suicide attempts among those at elevated risk (Doupnik et al., 2020; Nuij et al., 2021). Indeed, delivering follow-up contacts and safety planning, as well as conducting suicide risk assessments, are all recommended as part of routine clinical care for those identified as at risk for suicide by leading national organizations (e.g., National Action Alliance for Suicide Prevention: Transforming Health Systems Initiative Work Group, 2018). Thus, it is with good reason that some combination of follow-up contact, safety planning, and suicide risk assessment are often included in response-contingent intervention protocols aimed to help protect individuals participating in real-time monitoring research (Bentley et al., 2021). To our knowledge, however, no research has yet been conducted to test whether deploying such interventions *in real time* and *in response to suicidal thoughts reported on momentary surveys* reduces the risk of suicide attempt. Furthermore, as noted above, if response-contingent interventions effectively reduce participants' risk of STBs in EMA research, such interventions could be perceived as undermining the main goal of these studies, which is to observe how STBs unfold naturalistically and ultimately inform the much-needed development of better ways to detect and intervene with those at risk. On the other hand, discovering that such real-time monitoring and intervention methods can decrease the short-term risk of STBs and thus result in direct benefits to participants would be a significant advance for current suicide prevention efforts.

Reactivity

Another way that monitoring and response-contingent interventions may impact study data is through *reactivity*, or people altering their behavior due to awareness of being observed. Although reactivity has long been considered in psychological science (Nelson, 1977), it is especially relevant in EMA research (Carr et al., 2020; Shiffman, 2015), as participants' perceptions of being watched by researchers may be amplified with frequent assessments, especially if coupled with newer passive monitoring strategies (e.g., Mohr et al., 2017). In EMA studies in which participants are aware that their survey responses about STBs are not only being closely monitored but also may trigger interventions, reactivity may be especially salient. One possibility is that the awareness that one's responses are being monitored and may lead to potentially protective, helpful interventions could contribute to participants being *more* willing to disclose suicidal thoughts; this could have direct benefits for participants. On the other hand, participants may also be *less* likely to report STBs (or complete surveys entirely) if they know that study staff will soon see their responses or that doing so may lead to an intervention (e.g., Deming et al., 2021). Hesitancy to disclose STBs to avoid unwanted consequences of doing so (e.g., involuntary hospitalization) or because of other stigma-related barriers has long been acknowledged in research and clinical realms (Blanchard & Farber, 2020; Hom, Stanley, et al., 2017; Høyen et al., 2022). Being made aware that reporting STBs via EMA could result in response-contingent interventions may discourage disclosure and thus influence participants' *overall* response patterns; for example, studies that monitor and respond to incoming data could end up with more suppressed ratings of suicidal thoughts than studies that do not. Deploying response-contingent interventions could also lead to specific changes in the responses *following* intervention receipt (such as either reduced willingness to report STBs again or reinforced reporting of STBs in the future to elicit desired interventions). Perhaps the greatest potential risk of reactivity is that suppressed STB ratings could adversely affect the potential to deliver timely interventions and thus prevent suicidal behavior; that is, if the possibility of receiving response-contingent interventions deters participants from disclosing suicidal thoughts, it could preclude the chance to deploy what may ultimately be life-saving interventions. It is also possible that, in rare cases, response-contingent interventions could be iatrogenic (e.g., if they result in unnecessary hospitalization or police involvement; Ward-Ciesielski & Rizvi, 2021).

Importantly, both treatment effects and reactivity to response-contingent interventions in EMA research may result in reduced reporting of STBs, which would make it difficult to tease them apart. Thus, thoughtful control conditions that are likely to elicit one effect while ruling out the other would be needed to distinguish between treatment effects and reactivity. Here, we take an initial step toward examining such potential effects in a purely observational real-time monitoring study of STBs, as we know of no other research aimed to determine the effects of response-contingent interventions in EMA studies to date. Results have the potential to inform the refinement of best practice guidelines for this rapidly expanding area of research (e.g., Nock et al., 2021) and assist investigators in considering the benefits and (potential) risks of intervening when participants report current thoughts of suicide.

The Present Study

To improve understanding of the potential effects of response-contingent interventions in real-time monitoring research, we leverage data from a large, recently completed EMA study of adults and adolescents recruited during an emergency department (ED) visit or psychiatric inpatient stay for STBs. In this study, incoming real-time self-report data on suicidal intent were closely monitored, and responses above an a priori "high-risk" threshold triggered a series of response-contingent automated pop-up messages and phone-based outreach from the study team. Given that all incoming responses that exceed the high-risk threshold triggered response-contingent interventions (i.e., there was no comparison condition in which high-risk responses do not trigger the interventions), we focus on examining whether receiving interventions is associated with specific within-person differences in responses. We sought to address three specific questions.

Question 1: Do people report lower suicidal intent after receiving a high-risk response-contingent intervention (relative to responses for which there is no intervention)? If the suicidal intent ratings that follow response-contingent interventions are different from those that do not follow response-contingent interventions, this may suggest that the interventions are influencing participants' responses (either through treatment effects or reactivity). We hypothesized that suicidal intent ratings would show a larger decrease following response-contingent interventions than not following interventions, when controlling for the continuous effect of previous levels of suicidal intent on future suicidal intent. We were also interested in exploring whether response-contingent interventions appear to influence constructs beyond that used to trigger the interventions (suicidal intent), and thus, with no a priori hypotheses, also tested for differences in subsequent ratings of suicide urge, negative affect, and positive affect.

Question 2: Are participants less likely to respond to survey prompts after receiving response-contingent interventions? If response-contingent interventions are perceived as unwanted or unhelpful, participants might be less likely to respond to surveys entirely after receiving a response-contingent intervention (although we note that decreased responding could also be due to the initiation of a higher level of care in response to a suicidal crisis, which may present other barriers to completing surveys). We hypothesized that participants would overall be *less* likely to respond to survey prompts after receiving response-contingent interventions, when controlling for any decrease in survey completion rates over time. We also explored, with no a priori hypothesis, whether any such differences vary by sample type (adults vs. adolescents).

Question 3: How often do participants lower initially high ratings of suicidal intent after being presented with automated response-contingent pop-up messages? If response-contingent interventions are perceived as unwanted or unhelpful, we might expect participants to alter their initial suicidal intent rating shortly after being made aware that their response has triggered an imminent intervention in an effort to evade being contacted by the study team. Thus, with no a priori hypotheses, we explore how often participants lower an initially high suicidal intent rating immediately after receiving automated response-contingent pop-up messages that include a notification that they may soon be contacted by researchers. We also explore whether this behavior of lowering responses varies by sample type (adults vs. adolescents).

Transparency and Openness

This study, which involved an analysis of data from a larger existing project, was not formally preregistered. Data and materials are available from the first author upon request. Access to anonymized data will be available through the National Institute of Mental Health Data Archive at the completion of the larger project of which this study is a component. All procedures were approved by the governing hospital/university institutional review boards in accordance with the provision of the World Medical Association Declaration of Helsinki.

Method

Participants

Participants were 198 adults and 236 adolescents ($N = 434$) enrolled between May 2019 and August 2022 in a large intensive longitudinal monitoring study aimed to improve understanding of STBs during the posthospital period. Adults were recruited in the United States at a large urban hospital ED and adolescents from an inpatient psychiatry unit. Primary inclusion criteria were: (a) age 18+ years old (adults) or 12–19 years old (adolescents); (b) English fluency; (c) presentation at hospital (ED or psychiatric inpatient stay) with suicidal thoughts; (d) smartphone ownership; and (e) willingness to provide at least one collateral contact person. All participants provided written informed consent (adults) and/or assent (adolescents).

Procedure

Consent and Enrollment

Adults were enrolled in the study either during an in-person meeting with a research assistant in the emergency department (ED; or, if discharged to the same hospital's inpatient unit, the inpatient unit) or (at various points since the COVID-19 pandemic) via a phone meeting up to 4 weeks after ED discharge. Adolescents were enrolled via in-person meeting during inpatient hospitalization. During the informed consent process, participants were made aware that, although their responses to the EMA surveys may take up to 24 hr to review, if at any point the study team became concerned about their safety, the study team would reach out to them (and, for the adolescent sample, their parent or guardian) via phone to ensure they are safe. Of relevance to the present investigation, participants were *not* made aware at the start of the study of the specific high-risk response threshold used to trigger outreach (described in detail below). Participants were told that if they could not be reached, study staff would then reach out to their collateral contact(s). Last, it was noted that, though extremely rare, if research staff had reason to believe the participant was at imminent risk of suicide and neither they nor their collateral contact(s) responded to outreach, confidentiality may be broken by calling emergency services. Adults also developed a personalized safety plan (using a modified, briefer version of the template developed by Stanley & Brown, 2012) collaboratively with the research assistant during enrollment. For adolescents, research staff (with participants' permission) obtained the personalized safety plan documented in their health record during the index inpatient hospitalization.

EMA

Adult participants began receiving prompts to complete EMA surveys (via LifeData software installed on their smartphones) immediately following their enrollment; for adolescents, EMA prompts began after inpatient discharge (as they did not have access to their smartphone on the psychiatric unit). EMA prompts occurred six times daily for the first 3 months (84 days) of the study. The first and last surveys of the day were at scheduled times, with other prompts sent randomly within predefined intervals. For 3 months after the EMA period, participants received one prompt per day. EMA surveys took under 5 min to complete and included questions assessing facets of suicidal thoughts and a range of affect states and contextual factors. Participants were compensated $1 for each completed survey. After 3 consecutive days of no data, study staff attempted to reach participants to troubleshoot issues with the technology and encourage them to reengage. After five unanswered contacts over 26 days, participants were considered lost to follow-up (though still able to reengage at any point up until 6 months).

Risk Monitoring and Response-Contingent Interventions

Briefly, the monitoring and response-contingent intervention protocol consisted of a series of automated pop-up messages and phone-based outreach from study staff. First, regarding the pop-up messages, when a participant selected a response of ≥8 out of 10 on the EMA item assessing suicidal intent (see the Measures section), they were immediately presented with a pop-up message that stated "We want you to stay safe" and notified them that they (a) would soon be shown their safety plan (within the app), (b) would be sent an automated message in about 15 min to check in, and (c) may soon be contacted by phone. They were then presented with a follow-up question that asked, "Are you going to kill yourself today?" with the following response options: "Yes," "No," and "Don't know." When "No" was selected, a pop-up message appeared notifying the participant that they would be shown their personalized safety plan from the enrollment visit on the next screen and that the research team may be in touch with them soon to check-in. For "Yes" or "Don't know" responses, participants were presented with a recommendation to call their clinician, call 911, or go to the nearest emergency room, and then asked to select which of these they would do (or "None of the above"). During this series of pop-up messages, participants were able to select a "back" button to return to (and potentially modify) their suicidal intent rating before submitting the survey. If they lowered their initial rating to below the high-risk threshold (<8), they did not receive the automated pop-up messages again. Fifteen minutes after selecting ≥8 on suicidal intent, participants received another prompted, briefer follow-up survey assessing suicidal urges and intent, as well as any resources they may have used.

Regarding phone-based outreach, selecting ≥8 on the suicidal intent item triggered a real-time notification to the study risk monitoring team, comprised of masters- or doctoral-level clinical psychologists with training in suicide risk assessment. Within 24 hr (but usually much sooner; the mean time elapsed from high-risk response to outreach was 3.0 hr [$SD = 4.8$ hr]), between the hours of 9 a.m. and 9 p.m. and 7 days/week, a risk monitoring team member contacted the participant via phone call (and potentially also text message if the participant did not answer or respond promptly to the call). For adolescents, the parent or legal guardian was contacted first to notify them of the high-intent response, before contacting the adolescent directly. Once the risk monitoring team member reached the participant, they followed an established, detailed phone-based risk assessment protocol. Based primarily on participants' verbal responses to questions about current suicidal intent and plan to kill oneself within the next week, as well as real-time consultation with other risk team members as needed, a risk monitoring team member primarily either (a) reviewed, updated, and encouraged the participant to use their safety plan and go to the ED if they became unable to stay safe (if not at imminent risk) or (b) in the rare case of imminent risk, enacted a structured protocol for connecting them with emergency services immediately. The risk monitoring team member could also develop a plan to call the participant back later that day (often, 1 hr later) for a reassessment. The study used an established, detailed protocol for cases when participants did not respond promptly to such outreach, including repeat calls and texts, contacting collaterals, and, as a very last resort, calling emergency services.

Measures

A subset of specific EMA items was analyzed for the present investigation.

Suicidal Intent

Suicidal intent was assessed with the following item: "Right now, how strong is your intention to kill yourself today?" Participants responded on a 0 (*I am definitely not going to kill myself today*) to 10 (*I am definitely going to kill myself today*) scale. Ratings of ≥8 on this item triggered the study's response-contingent intervention protocol, as detailed above.

Suicide Urge

Suicide urge was assessed with: "Right now, how strong is your urge to kill yourself?" Participants responded on a 0 (*not at all*) to 10 (*very strong*) scale.

Negative Affect

Overall negative affect was assessed with the following item: "Right now, how much do you feel negative?" with an accompanying definition for negative ("generally, how negative you feel right now"). Responses were on a 0 (*not at all*) to 10 (*very much*) scale.

Positive Affect

Overall positive affect was assessed with an item (and definition) analogous to the above item for negative affect, but with "positive" substituted for "negative."

Analytic Strategy

All data processing and analysis were conducted in R (R Core Team, 2022) and used the *tidyverse* package (Wickham et al., 2019). For the present investigation, data were only used from the first 3 months of the study, when EMA surveys were prompted six times per day.

Response Types

Our definitions for response types to the suicidal intent item are below.

High-Risk Response. We defined high-risk response as a *submitted* EMA survey with the suicidal intent item rated ≥8 (and for which the response was not determined to be a mistake, per below).

Mistake Response. We defined *mistake responses* as observations when a participant initially selected (and may have submitted) a rating of ≥8 on the EMA item assessing suicidal intent, but upon potentially being contacted by the research team, confirmed that they meant to record a different (non-high-risk) rating *and* there were no indicators (e.g., response ≥5 on suicide urge or verbal indications of distress) that the initial selection of ≥8 for intent was not a mistake.

Lowered Response. We defined *lowered response* as non-mistake observations when a participant initially selected ≥8 on suicidal intent, thus triggering a web-based notification to the study team, but after viewing the automated pop-up messages and prior to submitting the survey, lowered their response to <8, meaning their final submitted response fell below the high-risk threshold.

Question 1

We approached the question of whether the suicidal ratings that follow a high-risk response-contingent intervention differ from the ratings that do not follow a response-contingent intervention from a regression discontinuity analysis framework. In brief, regression discontinuity allows researchers to probe the effects of an intervention in observational studies in which a specific criterion is used to assign participants to receive or not receive the intervention (e.g., there is no randomization; Gelman et al., 2021). In other words, regression discontinuity can be used to investigate the effects of interventions that are delivered based on a given assignment variable: in this case, reaching a threshold of ≥8 on the suicidal intent EMA item. One conceptual assumption driving our discontinuity analysis was that suicidal intent ratings that fall just above the high-risk intervention threshold should reflect a similar "true" or underlying level of suicidal intent as the ratings that fall just below the threshold; thus, in the absence of any interventions occurring at the threshold, we would expect the responses that follow ratings falling just above the threshold (e.g., intent = 8) to be similar or perhaps, due to autocorrelation, slightly higher overall than those following ratings that fall just below the threshold (e.g., intent = 7).

To test for discontinuity at the response-contingent intervention threshold, we built a series of multilevel regression models using the R package *lme4* (Bates et al., 2015), in which previous suicidal intent ratings were used to predict future suicidal intent ratings, with random effects for participants. Our primary model used suicidal intent at time t to predict suicidal intent at $t + 1$ (the next completed EMA) in those participants who provided *at least one* high-risk response during the study. Our secondary model used only those participants with exactly *one* high-risk response; this offered a potentially more stringent test of discontinuity as the intervention threshold could not have been "learned" prior to the first (and for these participants, only) time a high-risk rating was selected (as they were not informed about the specific threshold during enrollment). All discontinuity models were restricted to observations for which previous suicidal intent ratings were ≥5 to reduce the imbalance between those that triggered and did not trigger interventions.

Key to regression discontinuity analysis, we centered suicidal intent at t on the intervention threshold by subtracting 7.5 from each rating and included a variable indicating whether suicidal intent at t was ≥8 (with 0 = suicidal intent was 5, 6, or 7 so response-contingent interventions were not triggered and 1 = suicidal intent was 8, 9, or 10 so interventions were triggered) as our *discontinuity term*. Examining the significance (using a threshold of $p < .05$) of the discontinuity term coefficient allowed us to determine if crossing the threshold used to trigger the interventions at t predicts suicidal intent at $t + 1$ over and above the continuous effect of increasing levels of intent at t (e.g., from 5 to 6 to 7), thus potentially suggesting that the response-contingent interventions have a specific association with subsequent responses. The direction of the discontinuity term would indicate whether crossing the intervention threshold predicts higher (i.e., *positive discontinuity*) or lower (i.e., *negative discontinuity*) future responses over and above the continuous effect of increasing intent. We also built exploratory models using other constructs (suicide urge, negative and positive affect) as outcomes at $t + 1$ to assess for discontinuity for items beyond only that used to trigger the interventions. We also built two "placebo" models testing hypothetical intervention thresholds (of ≥7 and ≥9) to determine if discontinuity exists only at the actual high-risk threshold (≥8). All models controlled for each participant's mean previous intent and included hours elapsed (centered at a 2-hr time gap) between observations (t and $t + 1$) as an interaction term with both the previous intent rating and discontinuity term and excluded mistakes that were not corrected before submitting the survey.

Importantly, we also sought to determine whether any observed discontinuity is clearly attributable to the response-contingent interventions exerting an effect on future survey responses and not, for example, other phenomena such as regression to the mean—meaning, relatively high-intent ratings to be followed by lower intent ratings and relatively low ratings to be followed by higher ratings. Thus, as a robustness check, we ran a second series of multilevel regression models in which the temporal order of predictor and outcome variables were switched. These analogous models tested whether *future* suicidal intent (at t) predicts *previous* suicidal intent (e.g., at $t - 1$) rather than, as detailed above, previous intent (at t) predicts future intent (e.g., at $t + 1$). As it is impossible for interventions triggered by a response at t to influence responses prior to t, we would interpret significant discontinuity coefficients in these "backward-in-time" models to suggest that discontinuity at the high-risk threshold cannot explicitly be attributed to intervention effects.

Question 2

To determine whether participants are less likely to complete surveys after receiving response-contingent interventions, we tested whether a survey prompt occurred before or after a response-contingent interventions predicts the likelihood of survey completion. We ran a series of multilevel binomial regression models in which whether a survey prompt occurred before or after a participant first received response-contingent interventions was the (binary) independent variable and whether or not a survey prompt was completed was the (binary) dependent variable, with random effects for participants. These models also controlled for day in study (given

the overall observed reduction in EMA response rates over time). The first model used all participants who received a response-contingent intervention (for any high-risk, lowered, or mistake response), and the second and third models were run only among adults and then adolescents, respectively, who received an intervention. For all models, we also required that participants' first response-contingent interventions occurred at least 4 (but less than 81) days into the study to ensure that there were at least 3 days of prompts both before and after receipt of interventions for valid comparisons.

Question 3

To explore how often participants lower an initially high suicidal intent rating after receiving automated response-contingent pop-up messages (as this may reflect an attempt to evade receiving a response-contingent intervention and thus suggest reactivity), we first calculated the percentage of participants (adults and adolescents) to lower an initially selected high-risk rating during the study, as well as the percentage of initially selected high-risk ratings that were lowered after viewing the automated pop-up interventions but prior to submitting the survey, excluding mistake responses. We then fit a multilevel logistic regression model in which sample (adults vs. adolescents) was the (binary) independent variable and whether a survey for which the response to the intent item was initially rated ≥8 was lowered to <8 prior to submitting (i.e., lowered vs. submitted high-risk response) was the (binary) dependent variable.

Results

Descriptive Statistics

Sample Demographics and STB History

The average age of the adult sample (n = 180 who provided demographic information) was 31.20 years (SD = 11.27, range = 18–69). Among adults, 51.98% identified as female, 38.98% as male, 3.39% nonbinary/gender nonconforming, 2.26% transgender men, 1.69% transgender women, and 1.69% other. Of adults, 76.70% identified as White, 13.07% Black, 5.11% Asian, 2.27% American Indian/Alaskan Native, and 8.52% other (participants could select multiple); 11.9% of adults identified as Hispanic/Latinx. For adolescents (n = 227 with demographic data), the average age was 14.87 years (SD = 1.65, range = 12–18). Of adolescents, 64.44% identified as female, 16.44% nonbinary/gender nonconfirming, 11.11% male, 3.11% transgender men, and 4.89% other. Among adolescents, 83.48% identified as White, 12.05% Asian, 5.36% Black, 0.89% Native Hawaiian/Pacific Islander, 0.45% American Indian/Alaskan Native, and 4.91% other; 9.63% were Hispanic/Latinx.[1]

Regarding self-reported history of STBs (Fox et al., 2020) at baseline, 53.3% of adults had made a lifetime suicide attempt (of these, 30.2% made an attempt in the past month), and the mean number of days in the past week with suicidal thoughts was 3.53 (SD = 2.51). Among adolescents, 62.6% reported a lifetime suicide attempt (of these, 62.0% made an attempt in the past month), and the mean number of days in the past week with suicidal thoughts was 3.9 (SD = 2.47).[2]

EMA Compliance

Participants enrolled up until September 1, 2022, were included in this investigation. Over the first 3-month EMA period, across all enrolled participants, a total of 62,066 prompted EMA surveys (M = 143.01 per participant, SD = 133.86) were completed, for an overall compliance rate of 29.04% per participant (47.81% of days, on average, with at least one completed survey). Adults completed 30,334 prompted surveys (M = 153.20, SD = 150.72), and adolescents completed 31,732 surveys (M = 134.46 per participant, SD = 117.53).[3] Regarding responses included in our primary discontinuity models (Question 1), the median duration between completed surveys was 3.17 hr (M = 7.90, SD = 30.04); between high-risk responses and the next completed survey, the median duration was 3.75 hr (M = 18.99 hr, SD = 67.67). Of all included index survey responses, 96% were followed by another completed survey within 24 hr, and for high-risk responses only, 88% had another completed survey within 24 hr.

Question 1: Suicidal Intent Following Response-Contingent Interventions

Figure 1 presents a joint distribution of previous (at t) suicidal intent and subsequent (at $t + 1$) suicidal intent ratings. All participants (N = 434) are included in Figure 1, though only those participants who provided the specific suicidal intent rating indicated on each axis are reflected in the dark red solid circles (see note below the figure). First, the zero-inflated nature of the suicidal intent item and relative sparsity of responses at or above 8 (the high-risk intervention threshold) are evident. This plot also indicates an overall positive linear association between previous and subsequent intent ratings through previous intent of 7 (and most clearly from 0 through 5). There is an observable reduction in the association between previous and future suicidal intent at previous intent ratings of 8, suggesting the potential for negative discontinuity at the high-risk threshold.

A total of 62 participants (14.29% of the overall sample) provided at least one high-risk (neither lowered nor mistake) response during the study (n = 35 adults [17.68% of all adults] and n = 27 [11.44%] adolescents[4]); this was the subsample used for the primary discontinuity model. Results from the primary and secondary multilevel linear *regression* models (suicidal intent at t predicting suicidal intent at $t + 1$, with a discontinuity term set at the response-contingent intervention threshold and controlling for both time between observations and mean suicidal intent at t) are presented in Table 1. In the primary hypothesized model (n = 62 with at least one high-risk response), the discontinuity term (−0.44) was negative and nonsignificant (p = .051).[5] In our

[1] Demographics for the primary subsamples in Question 1 (n = 62) did not differ significantly from the overall samples.

[2] For the primary subsamples in Question 1 (n = 62), 81.2% of adults reported a lifetime attempt (38.5% past month), and mean number of days with suicidal thoughts in the past week was 3.77 (SD = 3.01); for adolescents, 74.1% reported a prior attempt (45.0% past month); the mean number of past-week days with suicidal thoughts was 4.29 (SD = 2.74).

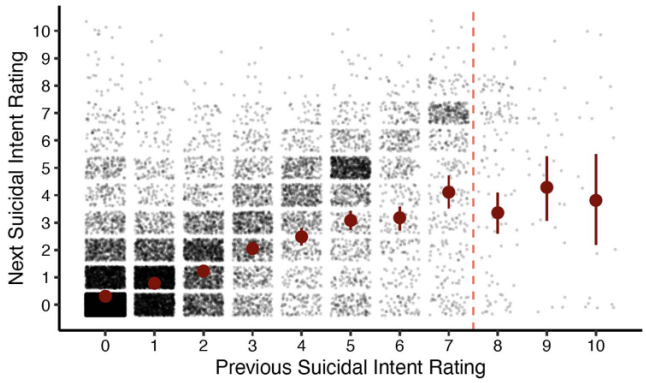
[3] Compliance for the primary Question 1 subsamples (n = 62) was higher (M = 197.13 EMAs) than the overall sample.

[4] There was not a statistically significant difference between the proportions of adults versus adolescents with at least one high-risk response (p = .064).

[5] As a robustness check, we also ran this primary model when excluding lowered responses; without lowered responses, the negative discontinuity term was similar in magnitude (−0.39) and nonsignificant (p = .09).

Figure 1

Joint Distribution of Previous (at t) Suicidal Intent and Subsequent (at t + 1) Suicidal Intent Ratings in the Full Sample (N = 434)



Note. Small black dots represent raw observations across all participants. Large, dark red solid circles show the overall mean across participant-level mean responses for each suicidal intent rating; each large solid circle only includes participants who provided the specific suicidal intent rating at any point during the study. The error bars are bootstrapped confidence intervals (CIs) at the level of participant means. The dashed vertical line indicates the high-risk response-contingent intervention threshold (suicidal intent ratings of ≥8). See the online article for the color version of this figure.

secondary model, among only those participants with *exactly one* high-risk response ($n = 29$), the resultant discontinuity term was negative (−1.42) and significant ($p < .05$). Results from the "placebo" discontinuity models using hypothetical thresholds of 7 and 9 are presented in Supplemental Tables S1 and S2; the resultant discontinuity terms were positive and small (0.01 and 0.08, respectively), and not significant ($p = .94$ and $p = .86$, respectively).

Results from the exploratory models using other constructs (suicide urge, negative affect, and positive affect) at $t + 1$ as outcomes are presented in Supplemental Table S3. Negative, statistically significant discontinuity terms were observed for both future suicide urge ($p < .05$) and negative affect ($p < .01$), suggesting the presence of negative discontinuity at the high-risk threshold for these other constructs distinct from but closely related to suicidal intent. The discontinuity term for positive affect, however, did not reach statistical significance.

Regarding our robustness checks, Figure 2 presents discontinuity terms and slope coefficients from a series of models using *previous* intent to predict *future* intent (through $t + 10$; i.e., the tenth next completed survey) and an analogous series of backward-in-time models using *future* intent to predict *previous* intent (through $t - 10$; i.e., the tenth previous survey). The discontinuity and slope coefficients for these two sets of models generally mirror each other, with discontinuity terms and slopes decreasing overall in magnitude when moving both forward and backward in time from the index survey, forming concave shapes. This suggests an overall similarity in both discontinuity and slope when moving forward and backward in time. Results from the model using future intent (at t) to predict previous intent (at $t - 1$) in all participants with at least one high-risk response ($n = 62$) are shown in Table 2. For this model, the discontinuity term (−0.47) was negative and significant ($p < .05$), indicating that *future* suicidal intent predicted *previous* intent at the high-risk threshold over and above increasing future intent. In the subset of participants with exactly one high-risk response ($n = 29$), the discontinuity term from the model using future intent to predict previous intent was large and negative (−0.75) but not statistically significant ($p = .28$; Table 2). We also found significant, negative discontinuity coefficients in the analogous models using *future* suicidal intent (at t) to predict both *previous* suicide urge and negative affect (at $t - 1$) ratings (Supplemental Table S4).

Table 1

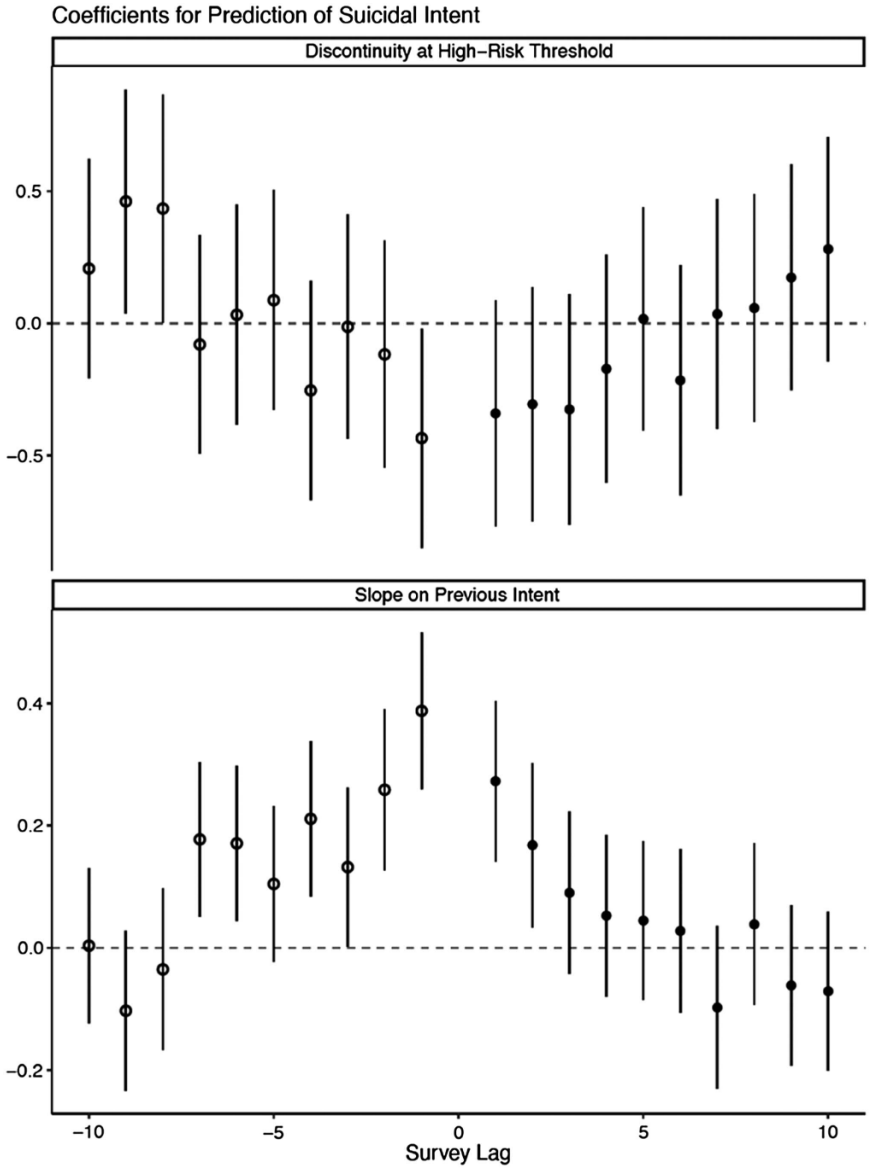
Multilevel Models Testing Discontinuity at High-Risk Threshold: Previous Suicidal Intent Predicting Future Suicidal Intent

	Future suicidal intent					
	Participants with at least one high-risk response ($n = 62$)			Participants with exactly one high-risk response ($n = 29$)		
Predictor	Estimate	95% CI	p	Estimate	95% CI	p
(Intercept)	2.88	[2.39, 3.37]	**<.001**	3.43	[2.56, 4.30]	**<.001**
Time between surveys	−0.02	[−0.02, −0.01]	**<.001**	−0.11	[−0.18, −0.05]	**.001**
Previous suicidal intent (value)	0.32	[0.19, 0.46]	**<.001**	0.10	[−0.25, 0.44]	.583
Previous suicidal intent (mean)	−0.69	[0.53, 0.84]	**<.001**	0.35	[0.15, 0.55]	**.001**
Previous high-risk response (discontinuity term)	−0.44	[−0.89, 0.00]	.051	−1.42	[−2.79, −0.05]	**.042**
Time Between Surveys × Previous Intent (value)	−0.00	[−0.01, −0.00]	**.007**	−0.04	[−0.07, −0.02]	**.001**
Time Between Surveys × High-Risk Response	0.02	[0.00, 0.03]	**.023**	0.13	[0.04, 0.21]	**.003**
Random effects						
σ^2	3.35			4.10		
$\tau_{00\ participant}$	0.61			0.09		
ICC	0.15			0.02		
$N_{participant}$	62			29		
Observations	1880			361		
Marginal R^2/conditional R^2	0.267/0.380			0.133/0.152		

Note. Previous observations at t and future observations at $t + 1$. Time between surveys refers to the hours elapsed between observations at t and $t + 1$, centered at a 2-hr gap. Previous suicidal intent (value) centered on the high-risk threshold by subtracting 7.5 from the actual response. The discontinuity term indicates whether the response at t was high-risk or not. Bolded p values are significant at <.05. CI = confidence interval; ICC = intraclass correlation coefficient.

Figure 2

Discontinuity Coefficients for Prediction of Suicidal Intent Forward and Backward in Time



Note. Discontinuity term and slope coefficients from a series of multilevel models using previous suicidal intent to predict future intent (forward in time from the index survey, shown in solid circles) and future intent to predict previous intent (backward in time from the index survey, in outlined circles) The *x*-axis indicates the number of completed EMA surveys since the index survey (at 0), where −10 = the 10th survey before the index survey and 10 = the 10th survey after the index survey. Discontinuity coefficients correspond to model discontinuity terms (binary indicators of whether or not the index survey exceeded the high-risk threshold). Slope coefficients represent the relationship between previous and future (solid circles) or future and previous (outlined circles) suicidal intent ratings.

In summary, we observed some evidence of negative discontinuity for both subsequent *and* previous suicidal intent, suicide urge, and negative affect ratings at the threshold used to trigger response-contingent interventions among participants with a high-risk response. Results for the primary hypothesized model for future suicidal intent ratings, however, did not reach statistical significance ($p = .051$), ultimately yielding mixed support for our hypothesized effects.

Table 2
Multilevel Models Testing Discontinuity at High-Risk Threshold: Future Suicidal Intent Predicting Previous Suicidal Intent

	Future suicidal intent					
	Participants with at least one high-risk response (n = 60)			Participants with exactly one high-risk response (n = 27)		
Predictor	Estimate	95% CI	p	Estimate	95% CI	p
(Intercept)	2.95	[2.41, 3.49]	**<.001**	2.99	[2.09, 3.89]	**<.001**
Time between surveys	−0.02	[−0.04, −0.00]	**.037**	−0.05	[−0.12, 0.02]	.143
Future suicidal intent (value)	0.42	[0.29, 0.56]	**<.001**	0.18	[−0.15, 0.52]	.284
Future suicidal intent (mean)	0.66	[0.49, 0.84]	**<.001**	0.47	[0.23, 0.71]	**<.001**
Future high-risk response (discontinuity term)	−0.47	[−0.91, −0.02]	**.039**	−0.75	[−2.12, 0.61]	.280
Time Between Surveys × Future Intent (value)	−0.01	[−0.02, 0.00]	.080	−0.02	[−0.05, 0.01]	.278
Time Between Surveys × High-Risk Response	0.02	[−0.01, 0.04]	.298	−0.04	[−0.25, 0.16]	.685
Random effects						
σ^2	3.21			3.77		
$\tau_{00\ participant}$	0.92			0.39		
ICC	0.22			0.09		
$N_{participant}$	60			27		
Observations	1890			364		
Marginal R^2/conditional R^2	0.251/0.417			0.121/0.203		

Note. Future observations at t and previous observations at $t - 1$. Time between surveys refers to hours elapsed between observations at $t - 1$ and t, centered at a 2-hr gap. Future suicidal intent (value) centered on the high-risk threshold by subtracting 7.5 from the actual response. The discontinuity term indicates whether the response at t was high-risk or not. Bolded p values are significant at <.05. Two fewer participants included in these models (n = 60) than those in Table 1 because for two participants, the first and only time they provided a suicidal intent rating of ≥5 was their first completed survey, so there were no relevant previous observations at $t - 1$. CI = confidence interval; ICC = intraclass correlation coefficient.

Question 2: Survey Completion Following Response-Contingent Interventions

Results from three multilevel logistic regression models (for all participants who received response-contingent interventions [for a high-risk, lowered, or mistake response] during the study, then broken down by adults and adolescents) indicated that, when controlling for day in study, EMA completion did not differ from before to after response-contingent interventions (Table 3).[6]

Question 3: Lowered Responses Following Automated Pop-Up Messages

A total of 31 (7.14%) participants (8 [4.04%] adults and 23 [9.75%] adolescents; $\chi^2(1, N = 434) = 5.284, p < .05$) had at least one lowered response during the study (i.e., they initially selected a high-risk rating on suicidal intent but after viewing the automated response-contingent pop-up messages, lowered their rating to below the high-risk threshold before submitting the survey). After excluding mistake responses, over one fifth (22.13%) of all suicidal intent ratings that were initially rated ≥8 were lowered to below the threshold after viewing the automated pop-up messages (9.80% of initially high-risk ratings for adults compared to 42.86% for adolescents). For lowered responses, the most common final submitted response was 7 (the maximum rating that still falls below the high-risk threshold; 27.78% of lowered responses), followed by 6 (16.67% of lowered responses) and 0 (also 16.67%). The multilevel logistic regression model showed that adolescents were significantly more likely than adults to lower initially high-risk ratings to below the high-risk threshold after automated pop-up messages (Table 4).[7]

Discussion

Real-time monitoring is increasingly used in suicide research with the goals of better understanding how STBs unfold in real time, improving suicide risk prediction, and informing intervention development. Receiving information in real-time or very close to it, however, about participants' current suicidal thoughts poses critical questions related to whether (and how) researchers should intervene with the goal of preventing suicidal behavior from occurring during these studies. One largely unaddressed consideration is that monitoring of incoming data and delivery of response-contingent interventions could affect participant responding, perhaps either due to interventions effectively reducing STBs or reactivity. In the present study, we sought to understand whether monitoring and intervening on high-risk responses in a large EMA study of suicidal adults and adolescents has an impact on collected data. Overall, we found some evidence that response-contingent interventions involving outreach from the research team may exert effects on study data. First, although we observed some evidence of discontinuity in responses at the high-risk intervention threshold, discontinuity could not specifically be attributed to the response-contingent interventions. Second, response-contingent interventions did not impact the likelihood of completing surveys. Third, lowering one's initial suicidal intent ratings to below the high-risk threshold after viewing response-contingent pop-up messages notifying them to expect contact from the study team was common, and adolescents were more likely to do so than adults. Below, we discuss each of these findings in detail.

First, we found some evidence of negative discontinuity in subsequent suicidal intent ratings at the high-risk intervention threshold. Specifically, in the secondary model (using only those

[6] Models run using the 62 participants from Question 1 who received response-contingent interventions specifically for a high-risk response (not only for lowered responses or mistakes) produced the same nonsignificant results.

[7] We additionally tested if other demographic (gender and race) or clinical characteristics (history of suicide attempt) were associated with lowered responses; each of these other multilevel models was not significant.

Table 3
Multilevel Models Comparing Survey Completion Before Versus After Response-Contingent Intervention

	Survey completed								
	Participants who received interventions (n = 89)			Adults who received interventions (n = 41)			Adolescents who received interventions (n = 48)		
Predictor	*OR*	95% CI	*p*	*OR*	95% CI	*p*	*OR*	95% CI	*p*
(Intercept)	0.56	[0.48, 0.66]	**<.001**	0.56	[0.44, 0.70]	**<.001**	0.57	[0.46, 0.70]	**<.001**
Day in study	0.99	[0.99, 0.99]	**<.001**	0.99	[0.99, 1.00]	**<.001**	0.99	[0.99, 0.99]	**<.001**
Before versus after first intervention	0.96	[0.91, 1.02]	.184	1.01	[0.93, 1.10]	.803	0.94	[0.87, 1.02]	.133
Random effects									
σ^2		3.29			3.29			3.29	
$\tau_{00\ participant}$		0.56			0.57			0.54	
ICC		0.15			0.15			0.14	
$N_{participant}$		89			41			48	
Observations		7,488			3,449			4,039	
Marginal R^2/conditional R^2		0.014/0.158			0.006/0.152			0.026/0.162	

Note. Included participants are those who received response-contingent interventions (for a high-risk, lowered, or mistake response) at least four but fewer than 81 days into the study. Before versus after intervention referred to whether a prompted EMA survey occurred before or after a participant's first response-contingent intervention. The outcome variable was whether or not a prompted EMA was completed. Bolded p values are significant at <.05. CI = confidence interval; EMA = ecological momentary assessment; ICC = intraclass correlation coefficient.

participants with exactly one high-risk response) but not the primary model (all participants with at least one high-risk response), when a response to the suicidal intent item exceeded the high-risk threshold (≥8 out of 10) and thus resulted in response-contingent interventions, there was a stronger "downward pull" for the next rating than when a response to the suicidal intent item was still elevated but did not exceed the high-risk intervention threshold (≥5 but <8). Given that we might expect ratings following a response of 8 on suicidal intent, for example, to be very similar to those following a rating of 7, this result was noteworthy and poses the possibility that response-contingent interventions may have impacted subsequent responses. However, the *backward-in-time* model corresponding to the primary hypothesized model also showed significant negative discontinuity (with an estimate of similar magnitude to that from the forward-in-time model) at the high-risk threshold, whereas the backward-in-time model corresponding to the secondary model did not (although the discontinuity term was still large and negative). In other words, for the larger sample of participants with at least one high-risk response, when a response to the suicidal intent item was high risk, there was also a stronger downward pull for the previous rating than when a response fell just below the high-risk threshold. As it is impossible for a response-contingent intervention to causally influence ratings *prior to* the intervention, the fact that we saw negative discontinuity at the high-risk threshold in this backward-in-time model suggests that discontinuity may not necessarily be due to the *interventions* exerting an effect on responses.

It is also noteworthy that we saw evidence of discontinuity from both forward-in-time and backward-in-time models for suicide urge and negative affect, not only the item used to trigger interventions (suicidal intent). This could be interpreted to suggest that participants may have suppressed their ratings on multiple items to avoid intervention because, as noted earlier, they were not made aware of the specific triggering item (or threshold) at the start of the study. It is also possible, however, that participants quickly "learned" that only their responses to suicidal intent triggered interventions. Along these lines, the fact that we saw discontinuity for these other related items (albeit in both forward- *and* backward-in-time models) may also be interpreted to indicate that participants did *not* alter their subsequent responses specifically to avoid intervention.

Although it is still unclear why we observed some evidence of discontinuity only at the high-risk threshold (7 vs. 8) and not at other "hypothetical" thresholds (e.g., 6 vs. 7 or 8 vs. 9), there are other explanations that could help explain discontinuity in ratings surrounding (both before and after) high-risk responses. First, regression to the mean would result in relatively high ratings to generally be followed by lower ratings and relatively low ratings to be followed by higher ratings. Second, elevated suicidal intent may be a highly unstable psychological state, and therefore, such states tend to be preceded and followed by fluctuations of greater magnitudes than more moderate suicidal intent states. In other words, people may be less likely to stay at elevated levels of suicidal thinking across multiple momentary observations over short time periods. Recent descriptive work observing high instability of

Table 4
Multilevel Model Testing Lowered Responses by Age Group

	Lowered response		
Predictor	*OR*	95% CI	*p*
(Intercept)	0.00	[0.00, 0.11]	**.001**
Age (adults vs. adolescents)	70.88	[3.39, 1481.04]	**.006**
Random effects			
σ^2		3.29	
$\tau_{00\ participant}$		27.79	
ICC		0.89	
$N_{participant}$		84	
Observations		242	
Marginal R^2/conditional R^2		0.121/0.907	

Note. Results from a multilevel logistic regression model with random effects for participants. The outcome was whether or not an initially selected high-risk rating for suicidal intent was immediately and explicitly lowered to below the high-risk threshold after viewing the response-contingent automated intervention but prior to submitting the survey. Bolded p values are significant at .05. CI = confidence interval; ICC = intraclass correlation coefficient.

suicidal thoughts—both intent and urges—over minutes/hours (Coppersmith, Ryan, et al., 2022; Kleiman et al., 2017) supports this possibility.

It is also possible that relatively low number of high-risk responses in the study overall may have rendered our discontinuity model results unreliable. Although we enrolled a high-risk clinical sample and monitored participants during the period of highest known risk for suicide (posthospital visit for STBs), the sparsity of high-risk responses may not be unexpected given that the use of a suicidal intent item assessing the intention to kill oneself *today* (a specific, short-term timeframe) on a scale with fairly extreme anchors (e.g., 10 = "I am definitely going to kill myself today"). It is also possible that participants avoided using the high end of the suicidal intent rating scale due to being made aware at the start of the study that the researchers would intervene if they were at immediate risk of suicide; this may represent a more general form of reactivity to the awareness that one's responses may trigger interventions. Along these lines, the relative sparsity of high-risk responses may also partially reflect that the actual deployment of interventions either effectively reduced the risk of STBs among those who provided at least one high-risk response (i.e., treatment effects) or reactivity following receipt of a response-contingent intervention. Future studies with relevant control conditions are needed to conclude definitively whether people are less likely to report high suicidal intent when made aware that their responses may trigger interventions or report high suicidal intent again after response-contingent interventions.

Second, the likelihood of responding to survey prompts did not change significantly from before to after receiving response-contingent interventions when controlling for the general trend of reduced compliance over time. We had hypothesized that participants would be *less* likely to respond to survey prompts after receiving response-contingent interventions, perhaps either due to reactivity to the interventions (e.g., reduced willingness to respond to surveys and thus potentially receive another intervention from the research team) or reports of high suicidal intent being followed by initiation of higher levels of care (e.g., inpatient hospital stays) that could present barriers to participants continuing to complete surveys regularly. We also considered the opposite trend (compliance *increasing* after receiving response-contingent interventions), perhaps due to participants feeling more cared for by the study team or connected to the research and thus *more* motivated to complete surveys after interventions. Overall, however, our finding that compliance did not change significantly from before to after the response-contingent interventions suggests that intervening on high-risk responses with pop-up messages and direct outreach from the study team may be unlikely to impact whether or not participants respond to future prompts.

Our third aim was to explore the phenomenon of lowered responses, which may suggest reactivity to response-contingent interventions. We found that just over 20% of the time that participants selected a high-risk rating, they lowered their initial rating to below the high-risk threshold after viewing the response-contingent pop-up messages but before submitting the survey. Due to the close temporal proximity of viewing the pop-up messages and lowering the initial rating, we presume that participants tended to change their initial responses in an attempt to evade outreach from the study team. Additionally, the fact that 7 (the highest possible rating that still falls below the high-risk threshold) was the most common final submitted rating for lowered responses may suggest that evasion of outreach best explains this phenomenon. It is also possible, however, that lowered responses were due to suicidal intent dropping naturally over the course of seconds/minutes or other content included in the pop-up messages (e.g., reminder to contact clinical resources) having an immediate therapeutic benefit and participants deciding to "update" their intent rating before submitting the survey. Because we were only aware of initially *high-risk* ratings that were changed (due to the selection of ≥8 triggering the automated notification to study staff), not *non*-high-risk ratings that may have been changed prior to submitting the survey, we were unable to test whether participants were more likely to change initial (high-risk) ratings that triggered response-contingent interventions than those (non-high-risk) ratings that did not.

We also found that lowered responses were more common in the adolescent than the adult sample. If such lowered responses are indeed reflective of attempts to evade outreach from the study team, this finding would suggest that adolescents may be more reactive to the possibility of receiving outreach from the study team when experiencing high suicidal intent. The study's response-contingent intervention protocol included first contacting adolescents' parents or guardians prior to contacting the adolescent, which could be perceived as aversive, for example due to the potential for parents to initiate unwanted conversations, alert the adolescent's providers, bring them to the hospital, or take other undesirable actions. This procedural difference may have contributed to the higher rate of lowered responses among adolescents, as well as a lower (although not statistically different) percentage of adolescents who provided at least one "real" (not lowered) high-risk response in general during the study. Adolescents may also find other aspects of the outreach protocol more aversive than adults; for example, speaking to an adult member of the study team during moments of acute distress may be especially unappealing. It is possible that interventions that do not involve call-based outreach (e.g., those delivered over text or that are fully automated) or do not directly involve parents/guardians may be viewed as more acceptable. It is also possible that the higher likelihood of lowered responses for adolescents may be due to other factors (e.g., rapid changes in suicidal intent over the course of seconds/minutes and a tendency to "update" one's ratings before survey submission, developmental considerations) that may differ between adolescents and adults.

In summary, in this observational study, we found mixed support overall for high-risk response-contingent interventions having an impact on participants' responses. It is especially noteworthy, however, that adolescents were significantly more likely than adults to lower initially elevated suicidal intent ratings after viewing response-contingent pop-up messages. Whether response-contingent interventions influence participants' responses (and the nature of effects) may, of course, be different for different people: an issue we return to in the Future Directions section.

Limitations

These findings must be considered in the context of several limitations. First, as stated above, data were drawn from a study that was not designed to experimentally evaluate the effects of response-contingent interventions; as such, we lacked control conditions that would permit between-person (or matched between-observation) comparisons of response-contingent intervention effects on study

data. Second, as noted above, high-risk responses were relatively rare overall, which may have presented power issues that could have impacted the stability of our multilevel models or, for the compliance analysis, contributed to the null group-level findings. We also lacked the statistical power to assess participant-level heterogeneity in response-contingent intervention effects. Third, and along these lines, overall EMA compliance was relatively low; although not entirely unexpected given the clinically severe nature of the sample and long monitoring period (e.g., Burke et al., 2017; Jacobucci et al., 2023), this may have biased the current findings and their generalizability. Our EMA data may not be missing at random; for example, participants may have been less likely to complete surveys when experiencing higher suicidal intent, and response behavior may differ following high-risk periods during which individuals chose to respond (vs. did not respond) to survey prompts. Fourth, even in this severe clinical sample, suicide attempts were (fortunately) relatively rare, which prevented us from examining the impact of response-contingent interventions on suicidal behavior. Fifth, we did not collect qualitative data from participants on *why* they lowered their response to the suicidal intent item after viewing the pop-up intervention, which means we do not know whether lowered responses were attempts to evade outreach from the study team or other reasons. Last, participants were primarily White, non-Hispanic/Latinx, and female, which limits the degree to which our results may generalize to more diverse individuals.

Future Directions

The present research lays a foundation for several future directions of research aimed to optimize intervention strategies that are maximally effective at reducing the risk of proximal suicidal behavior while also minimizing the potential for (unwanted forms of) reactivity. First and foremost, research that leverages randomized between-group or rigorous within-person study designs, such as the recent microrandomized trial (Klasnja et al., 2015; Qian et al., 2022) to experimentally evaluate the effects of different types of response-contingent interventions across a range of momentary suicide risk levels is needed. This line of research has the potential to provide more definitive conclusions as to whether there or not there are positive or negative effects of response-contingent interventions on participants (or study data) and also elucidate which types of interventions (e.g., human outreach, automated pop-up messages, interactive automated tools; e.g., Jaroszewski et al., 2019) are most accessible and effective for whom (and when). The latter would have implications far beyond the context of real-time monitoring research but rather for detecting and intervening with those at risk for suicide on a broader scale. As noted above, it is possible that heterogeneity in response-contingent intervention effects exists across individuals. Our analyses all consisted of group-level effects due to power limitations; however, it is reasonable to suppose that for some individuals, certain response-contingent interventions may have minimal overall impact, whereas for others, the same intervention(s) may be quite effective at reducing STBs, lead to increased hesitancy (or more willingness) to disclose STBs again, or even be harmful. Studies that take an idiographic approach to assessing the effects of different real-time interventions aimed to reduce short-term risk of suicidal behavior are needed. Last, more studies that collect feedback from those with lived experience about their experiences with and recommendations for response-contingent interventions (e.g., Hom et al., 2021; Nock et al., 2021), both within the context of EMA research and beyond, are needed. This will directly inform the development and refinement of person-centered intervention strategies for high-risk moments.

Conclusions

Overall, these results offer mixed support for the notion that intervening on high-risk responses in real-time monitoring research on STBs directly impacts study data. Adolescents do appear more likely than adults, however, to modify their initial responses in attempt to evade receiving phone-based outreach from researchers. Future research that is explicitly designed to probe the effects of monitoring and response-contingent interventions is needed to identify and optimize effective, scalable intervention strategies for moments of high suicide risk.

References

Ammerman, B. A., & Law, K. C. (2022). Using intensive time sampling methods to capture daily suicidal ideation: A systematic review. *Journal of Affective Disorders*, *299*, 108–117. https://doi.org/10.1016/j.jad.2021.10.121

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bentley, K. H., Maimone, J. S., Kilbury, E. N., Tate, M. S., Wisniewski, H., Levine, M. T., Roberg, R., Torous, J. B., Nock, M. K., & Kleiman, E. M. (2021). Practices for monitoring and responding to incoming data on self-injurious thoughts and behaviors in intensive longitudinal studies: A systematic review. *Clinical Psychology Review*, *90*, Article 102098. https://doi.org/10.1016/j.cpr.2021.102098

Blanchard, M., & Farber, B. A. (2020). "It is never okay to talk about suicide": Patients' reasons for concealing suicidal ideation in psychotherapy. *Psychotherapy Research*, *30*(1), 124–136. https://doi.org/10.1080/10503307.2018.1543977

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., French, B., Mancino, J., Mendez, D., Strollo, P., & Rathbun, S. L. (2017). Ecological momentary assessment in behavioral research: Addressing technological and human participant challenges. *Journal of Medical Internet Research*, *19*(3), Article e77. https://doi.org/10.2196/jmir.7138

Carr, D. J., Adia, A. C., Wray, T. B., Celio, M. A., Pérez, A. E., & Monti, P. M. (2020). Using the internet to access key populations in ecological momentary assessment research: Comparing adherence, reactivity, and erratic responding across those enrolled remotely versus in-person. *Psychological Assessment*, *32*(8), 768–779. https://doi.org/10.1037/pas0000847

Centers for Disease Control and Prevention. (2020). *Web-based Injury Statistics Query and Reporting System (WISQARS): Fatal injury reports*. https://webappa.cdc.gov/sasweb/ncipc/mortrate.html

Coppersmith, D. D. L., Dempsey, W., Kleiman, E. M., Bentley, K. H., Murphy, S. A., & Nock, M. K. (2022). Just-in-time adaptive interventions for suicide prevention: Promise, challenges, and future directions. *Psychiatry*, *85*(4), 317–333. https://doi.org/10.1080/00332747.2022.2092828

Coppersmith, D. D. L., Ryan, O., Fortgang, R., Millner, A., Kleiman, E., & Nock, M. (2022). *Mapping the timescale of suicidal thinking*. PsyArXiv. https://doi.org/10.31234/osf.io/eus2q

Deming, C. A., Harris, J. A., Castro-Ramirez, F., Glenn, J. J., Cha, C. B., Millner, A. J., & Nock, M. K. (2021). Inconsistencies in self-reports of suicidal ideation and attempts across assessment methods. *Psychological Assessment*, *33*(3), 218–229. https://doi.org/10.1037/pas0000976

Doupnik, S. K., Rudd, B., Schmutte, T., Worsley, D., Bowden, C. F., McCarthy, E., Eggan, E., Bridge, J. A., & Marcus, S. C. (2020).

Association of suicide prevention interventions with subsequent suicide attempts, linkage to follow-up care, and depression symptoms for acute care settings: A systematic review and meta-analysis. *JAMA Psychiatry*, *77*(10), 1021–1030. https://doi.org/10.1001/jamapsychiatry.2020.1586

Fisher, C. B., Pearson, J. L., Kim, S., & Reynolds, C. F. (2002). Ethical issues in including suicidal individuals in clinical research. *IRB: Ethics & Human Research*, *24*(5), 9–14. https://doi.org/10.2307/3563804

Fox, K. R., Harris, J. A., Wang, S. B., Millner, A. J., Deming, C. A., & Nock, M. K. (2020). Self-injurious thoughts and behaviors interview-revised: Development, reliability, and validity. *Psychological Assessment*, *32*(7), 677–689. https://doi.org/10.1037/pas0000819

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin*, *143*(2), 187–232. https://doi.org/10.1037/bul0000084

Gee, B. L., Han, J., Benassi, H., & Batterham, P. J. (2020). Suicidal thoughts, suicidal behaviours and self-harm in daily life: A systematic review of ecological momentary assessment studies. *Digital Health*, *6*. https://doi.org/10.1177/2055207620963958

Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and other stories*. Cambridge University Press.

Hom, M. A., Bauer, B. W., Stanley, I. H., Boffa, J. W., Stage, D. L., Capron, D. W., Schmidt, N. B., & Joiner, T. E. (2021). Suicide attempt survivors' recommendations for improving mental health treatment for attempt survivors. *Psychological Services*, *18*(3), 365–376. https://doi.org/10.1037/ser0000415

Hom, M. A., Podlogar, M. C., Stanley, I. H., & Joiner, T. E. (2017). Ethical issues and practical challenges in suicide research. *Crisis*, *38*(2), 107–114. https://doi.org/10.1027/0227-5910/a000415

Hom, M. A., Stanley, I. H., Podlogar, M. C., & Joiner, T. E., Jr. (2017). "Are you having thoughts of suicide?" Examining experiences with disclosing and denying suicidal ideation. *Journal of Clinical Psychology*, *73*(10), 1382–1392. https://doi.org/10.1002/jclp.22440

Høyen, K. S., Solem, S., Cohen, L. J., Prestmo, A., Hjemdal, O., Vaaler, A. E., Galynker, I., & Torgersen, T. (2022). Non-disclosure of suicidal ideation in psychiatric inpatients: Rates and correlates. *Death Studies*, *46*(8), 1823–1831. https://doi.org/10.1080/07481187.2021.1879317

Jacobucci, R., Ammerman, B. A., & McClure, K. (2023). *Examining missingness at the momentary level in clinical research using ecological momentary assessment: Implications for suicide research*. PsyArXiv. https://doi.org/10.31234/osf.io/7yfhp

Jaroszewski, A. C., Morris, R. R., & Nock, M. K. (2019). Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. *Journal of Consulting and Clinical Psychology*, *87*(4), 370–379. https://doi.org/10.1037/ccp0000389

Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychology*, *34*(Suppl.), 1220–1228. https://doi.org/10.1037/hea0000305

Kleiman, E. M., Glenn, C. R., & Liu, R. T. (2023). The use of advanced technology and statistical methods to predict and prevent suicide. *Nature Reviews Psychology*, *2*(6), 347–359. https://doi.org/10.1038/s44159-023-00175-y

Kleiman, E. M., Turner, B. J., Fedor, S., Beale, E. E., Huffman, J. C., & Nock, M. K. (2017). Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *Journal of Abnormal Psychology*, *126*(6), 726–738. https://doi.org/10.1037/abn0000273

Millner, A. J., Robinaugh, D. J., & Nock, M. K. (2020). Advancing the understanding of suicide: The need for formal theory and rigorous descriptive research. *Trends in Cognitive Sciences*, *24*(9), 704–716. https://doi.org/10.1016/j.tics.2020.06.007

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, *13*(1), 23–47. https://doi.org/10.1146/annurev-clinpsy-032816-044949

National Action Alliance for Suicide Prevention: Transforming Health Systems Initiative Work Group. (2018). *Recommended standard care for people with suicide risk: Making health care suicide safe*. Education Development Center.

National Institute of Mental Health. (n.d.). *Conducting research with participants at elevated risk for suicide: Considerations for researchers*. Retrieved April 4, 2023, from https://www.nimh.nih.gov/funding/clinical-research/conducting-research-with-participants-at-elevated-risk-for-suicide-considerations-for-researchers.shtml#regulatory

Nelson, R. O. (1977). Assessment and therapeutic functions of self-monitoring. *Progress in Behavior Modification*, *5*, 263–308. https://doi.org/10.1016/B978-0-12-535605-3.50012-1

Nock, M. K., Kleiman, E. M., Abraham, M., Bentley, K. H., Brent, D. A., Buonopane, R. J., Castro-Ramirez, F., Cha, C. B., Dempsey, W., Draper, J., Glenn, C. R., Harkavy-Friedman, J., Hollander, M. R., Huffman, J. C., Lee, H. I. S., Millner, A. J., Mou, D., Onnela, J., Picard, R. W., … Pearson, J. L. (2021). Consensus statement on ethical & safety practices for conducting digital monitoring studies with people at risk of suicide and related behaviors. *Psychiatric Research and Clinical Practice*, *3*(2), 57–66. https://doi.org/10.1176/appi.prcp.20200029

Nuij, C., van Ballegooijen, W., de Beurs, D., Juniar, D., Erlangsen, A., Portzky, G., O'Connor, R. C., Smit, J. H., Kerkhof, A., & Riper, H. (2021). Safety planning-type interventions for suicide prevention: Meta-analysis. *The British Journal of Psychiatry*, *219*(2), 419–426. https://doi.org/10.1192/bjp.2021.50

Pearson, J. L., Stanley, B., King, C. A., & Fisher, C. B. (2001). Intervention research with persons at high risk for suicidality: Safety and ethical considerations. *The Journal of Clinical Psychiatry*, *62*(Suppl. 25), 17–26.

Pew Research Center. (2021). *Mobile fact sheet*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Qian, T., Walton, A. E., Collins, L. M., Klasnja, P., Lanza, S. T., Nahum-Shani, I., Rabbi, M., Russell, M. A., Walton, M. A., Yoo, H., & Murphy, S. A. (2022). The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations. *Psychological Methods*, *27*(5), 874–894. https://doi.org/10.1037/met0000283

Rabasco, A., & Sheehan, K. (2022). The use of intensive longitudinal methods in research on suicidal thoughts and behaviors: A systematic review. *Archives of Suicide Research*, *26*(3), 1007–1021. https://doi.org/10.1080/13811118.2021.1903635

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*(3), 182–203. https://doi.org/10.1037/0003-066X.54.3.182

Sedano-Capdevila, A., Porras-Segovia, A., Bello, H. J., Baca-García, E., & Barrigon, M. L. (2021). Use of ecological momentary assessment to study suicidal thoughts and behavior: A systematic review. *Current Psychiatry Reports*, *23*(7), Article 41. https://doi.org/10.1007/s11920-021-01255-7

Shepard, D. S., Gurewich, D., Lwin, A. K., Reed, G. A., Jr., & Silverman, M. M. (2016). Suicide and suicidal attempts in the United States: Costs and policy implications. *Suicide & Life-Threatening Behavior*, *46*(3), 352–362. https://doi.org/10.1111/sltb.12225

Shiffman, S. (2015). Commentary on McCarthy et al. (2015): Ecological momentary assessment—Reactivity? Intervention? *Addiction*, *110*(10), 1561–1562. https://doi.org/10.1111/add.13050

Stanley, B., & Brown, G. K. (2012). Safety planning intervention: A brief intervention to mitigate suicide risk. *Cognitive and Behavioral Practice*, *19*(2), 256–264. https://doi.org/10.1016/j.cbpra.2011.01.001

Ward-Ciesielski, E. F., & Rizvi, S. L. (2021). The potential iatrogenic effects of psychiatric hospitalization for suicidal behavior: A critical review and recommendations for research. *Clinical Psychology: Science and Practice*, *28*(1), 60–71. https://doi.org/10.1111/cpsp.12332

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Woodford, R., Spittal, M. J., Milner, A., McGill, K., Kapur, N., Pirkis, J., Mitchell, A., & Carter, G. (2019). Accuracy of clinician predictions of future self-harm: A systematic review and meta-analysis of predictive studies. *Suicide & Life-Threatening Behavior*, *49*(1), 23–40. https://doi.org/10.1111/sltb.12395

Received April 26, 2023
Revision received September 9, 2023
Accepted September 14, 2023 ■