Behaviour Research and Therapy 117 (2019) 87–96

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Real-time monitoring technology in single-case experimental design research: Opportunities and challenges

Kate H. Bentley[a,*], Evan M. Kleiman[b], Grace Elliott[b], Jeffery C. Huffman[a], Matthew K. Nock[a,b]

[a] Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, USA
[b] Department of Psychology, Harvard University, USA

ARTICLE INFO

Keywords:
Single-case experimental design
Real-time monitoring
Ecological momentary assessment
Passive sensing
Technology

ABSTRACT

Single-case experimental design (SCED) is a rigorous method of studying behavior and behavior change. A key characteristic of SCED is repeated, systematic assessment of outcome variables, which is critical to achieving high internal validity, collecting a sufficient number of observations to conduct adequately powered statistical analyses, capturing dynamic and fine-grained changes in outcomes, and tailoring interventions at the individual level. Recent advances in real-time monitoring technology, such as digital ecological momentary assessment, passive smartphone-based behavioral tracking, and physiological assessment with wearable biosensors, are extremely well-suited to conducting these repeated, systematic measurements. Here, we discuss the rationale for incorporating real-time data collection technologies within SCED and highlight how recent studies have paired SCED with real-time monitoring. We also present original data illustrating how real-time digital monitoring can provide an idiographic and granular view of behavior (in this case, suicidal ideation). Last, we discuss the challenges of, and offer our recommendations for, using real-time monitoring technologies in SCED research.

Single-case experimental design (SCED) is a research methodology in which a single subject or small group of subjects is used to test causal relations among phenomena of interest (Barlow, Nock, & Hersen, 2009). SCED is well-suited for use in research on psychological and behavioral phenomena due to several key features, including: rigorous experimental manipulation of independent variables (usually a phase or condition), repeated and systematic measurement of dependent variables over time, potential to flexibly tailor treatment at the individual level, and plethora of SCED types to best match the research question at hand (Barlow et al., 2009; Kazdin, in press; Smith, 2012). The practical (e.g., time- and cost-efficiency) and scientific (e.g., high internal validity) advantages of SCED over more prominent between-subject designs in the study of behavior change have been discussed at length elsewhere (e.g., Barlow et al., 2009; Kazdin, in press; Kazdin 2010; Morgan & Morgan, 2001; Nock, Michel, & Photos, 2007). Here, we focus on the usefulness of one particular assessment strategy – *real-time monitoring* – in SCED research.

Despite its emerging popularity across all areas of psychological research, real-time monitoring, which can include both active assessments (e.g., ecological momentary assessment) and passive sensing, remains underutilized in SCED studies. In this article, we provide a rationale specifically for the use of real-time monitoring in SCED and highlight how researchers have applied real-time monitoring to facilitate examination of a variety of research questions in SCED with a level of rigor and accuracy that is not possible with more traditional assessments. We then present original idiographic data obtained through real-time monitoring from our team's research on suicidal ideation. Last, we discuss some of the key challenges associated with – and offer our recommendations for best practices in – the use of real-time monitoring in SCED.

1. Rationale for real-time monitoring in SCED

1.1. SCED requires repeated, systematic assessment

To provide a rationale for the use of real-time monitoring specifically in SCED research, it is first necessary to emphasize the importance of repeated, systematic assessment in SCED. First, repeated, systematic assessment throughout all phases is critical to obtaining the high internal validity characteristic of SCED. This is particularly relevant for SCED studies testing the efficacy of interventions, in which repeated assessments generally must be conducted before, during, and after interventions are applied to demonstrate that changes in the dependent variable occur *when and only when* the intervention is introduced and

* Corresponding author. Depression Clinical and Research Program, Massachusetts General Hospital/Harvard Medical School, One Bowdoin Square, 6th Floor, Boston, MA, 02114, USA.
E-mail address: kbentley@mgh.harvard.edu (K.H. Bentley).

https://doi.org/10.1016/j.brat.2018.11.017
Received 28 April 2018; Received in revised form 8 October 2018; Accepted 26 November 2018
Available online 07 December 2018
0005-7967/

rule out threats to internal validity (e.g., passage of time, life events, medication changes, etc.). Second, repeated assessment provides a sufficiently large number of observations to conduct well-powered statistical analyses with even the very small numbers of participants typical of SCED studies (Davis et al., 2013; Smith, 2012). Third, repeated, systematic measurement is critical in SCED studies aimed to answer questions about *how* and *why* behavior change processes unfold (i.e., mediators and mechanisms of change) (Kazdin 2007, 2010; Smith, 2012). Finally, repeated, systematic assessment allows SCED researchers to tailor how they manipulate their independent variables (in many cases, meaning customize treatment) for each subject in an ongoing, data-driven fashion (e.g., Bentley, Nock, Sauer-Zavala, Gorman, & Barlow, 2017).

SCED researchers typically ensure repeated, systematic assessment by either the administration of traditional self-report and clinician-rated measures at regular intervals (e.g., each week in the study) or direction observation of behavioral phenomena in the setting of interest (e.g., children's behavior in the classroom) (e.g., Barlow et al., 2009). Each of these strategies, however, has limitations. First, measures that ask participants to reflect on the intensity or frequency of a symptom or behavior over the past week, month, etc. are associated with retrospective recall biases and subject to experimenter demand characteristics. These traditional assessments also typically ask participants to provide an overall rating of their symptoms, which does not allow for a more nuanced understanding of dynamically and naturally occurring changes in symptoms or behaviors. Though some behavioral phenomena can be observed directly, more sensitive or stigmatized behaviors – such as suicidal and nonsuicidal self-injury (NSSI), the focus of our team's work – are more challenging and can even be unethical to observe directly. Additionally, measures that require administration or observation by a clinician can be resource- and time-intensive and may not be feasible for many researchers and clinicians interested in conducting small-scale SCED studies. Finally, assessments that take place in a laboratory or clinical setting may not generalize to the individual's natural environment (e.g., Trull & Ebner-Priemer, 2014).

1.2. Real-time monitoring facilitates repeated, systematic assessment in SCED

To circumvent the limitations of more traditional assessments, real-time monitoring (also known as ambulatory assessment; Shiffman, Stone, & Hufford, 2008), which refers to the assessment of constructs of interest (e.g., symptoms, cognitions, behaviors) outside of the laboratory, has increasingly been utilized in behavioral research. Detailed thoroughly elsewhere (e.g., Kleiman & Nock, 2017; Trull & Ebner-Priemer, 2014), the advantages of real-time monitoring over more traditional retrospective measures in psychological science include increased ecological validity, reduced retrospective recall and heuristic biases (e.g., Fredrickson & Kahneman, 1993), and ability to capture dynamic changes in psychological, cognitive, and behavioral phenomena over short periods of time. Real-time monitoring also facilitates improved understanding of sensitive and stigmatized behaviors that cannot be directly observed in a research or clinical setting. Though earlier real-time monitoring strategies were largely paper and pencil-based (such as written daily diaries) (e.g., Bolger, Davis, & Rafaeli, 2003; Thiele, Laireiter, & Baumann, 2002), recent advances in digital and mobile technologies have facilitated increased precision and temporal accuracy in self-report and objective outcome measurements (e.g., Smith, 2012; Vilardaga, Bricker, & McDonnell, 2014). Here, we describe the potential of real-time monitoring technologies to facilitate the repeated, systematic assessment needed in SCED.

1.2.1. Ecological momentary assessment

Currently, the most widely used real-time monitoring technology is ecological momentary assessment (EMA) or experience sampling, an *active* form of real-time monitoring (Kleiman & Nock, 2017). EMA typically involves participants responding to prompts, initiating responses to structured- or semi-structured questions, or entering self-report data into a mobile device (e.g., smartphone, tablet, personal digital assistants [PDAs]). Data can be collected via EMA at fixed or random time intervals (e.g., interval-contingent or signal-contingent) or initiated by the subject when an event occurs (i.e., event-contingent). As the name implies, EMA has the potential to provide *ecologically valid* information on behaviors or symptoms of interest as these behaviors or symptoms naturally occur, thus potentially strengthening SCED researchers' conclusions regarding the relations between independent and dependent variables.

For typical group-based designs (e.g., randomized clinical trials), unlike SCED, assessments conducted at relatively few time points do not sacrifice the strength of conclusions as data are collapsed across what is typically a much larger number of subjects. This is not the case with SCED, which tend to have fewer participants (or just one participant). With frequently repeated EMA entries over time-limited periods, researchers can overcome this limitation in SCED by acquiring a sufficiently large number of data points to conduct adequately powered (idiographic, within-individual) analyses, even with a single participant or very small number of participants. Information collected via technology-assisted EMA is also typically time-stamped, which helps ensure the temporal validity of these data, a key consideration in SCED studies aimed to elucidate functional or causal (Hamaker & Wichers, 2017) relations between independent and dependent variables (e.g., does change occur *when and only when* the intervention is applied?). Prompting brief (e.g., 5 or 10 item) EMA surveys through a mobile device that an individual can complete quickly at various points throughout their day also does not need to require a great deal of participant time or effort. This makes these remote assessments well-suited to SCED studies conducted by researchers or clinicians who lack the monetary resources to provide compensation for time-intensive in-person assessment visits.

Along these lines, the high (and rising) rates of smartphone ownership (Pew Research Center, 2017) support the feasibility of collecting EMA data through mobile apps on a large-scale. Research participants are increasingly likely to both already own a mobile device that can be used for data collection (nearly 90% of people in the United States ages 18 to 49 own a smartphone; Pew Research Center, 2017), which makes collecting EMA data via smartphone in SCED studies realistic for investigators who cannot provide participants with expensive digital devices. By using a smartphone for EMA data collection, participants can also enter information into their personal device unobtrusively and without the burden of carrying a separate study device. This may increase the feasibility of conducting SCED research with treatment-seeking (i.e., non-research-seeking) patients in clinical practice, a topic we elaborate upon further below.

1.2.2. Passive sensing

New technologies have also been harnessed for *passive* real-time monitoring (or sensing) of naturally occurring phenomena (e.g., Kleiman & Nock, 2017; Vilardaga, Bricker, & McDonell, 2014). Passive sensing refers to collecting data unobtrusively, without active data entry from the subject. Researchers are now able to passively collect data on smartphone usage patterns such as screen time, number of text messages, call duration, app usage, and social media activity (Kleiman & Nock, 2017). Smartphones and wearable devices such as wrist-worn biosensors can be used to capture a broad variety of behaviors and symptoms including patterns of activity and movement via geospatial activity (e.g., GPS) and accelerometer, sleep duration and quality, physiological indices such as heart rate and skin conductance, and voice characteristics (e.g., Ben-Zeev, Scherer, Wang, Xie, & Campbell, 2015; Torous, Onnela, & Keshavan, 2017). The rising ownership rates (and declining costs) of fitness trackers (e.g., Fitbits) (Statista, 2017) increase the feasibility of tracking objective indicators of common targets in behavioral treatments (e.g., exercise, sleep) and potential physiological

correlates or mechanisms of change (e.g., heart rate during an exercise- or mindfulness-based intervention). One attractive element of commercial wearables is that they output metrics (e.g., step count, sleep latency, average heart rate) that are easily understood and potentially actionable by clinicians and patients alike.

In SCED studies, passive real-time monitoring has great promise to facilitate the collection of *objective* behavioral phenomena. As passive sensing does not require the individual to enter any data directly, these methods are generally associated with lower participant burden than EMA and thus may be well-suited to clinicians seeking to collect clinically informative data on behavioral targets or treatment outcomes (e.g., sleep, exercise, social activity, etc.) from non-research-seeking patients in SCED studies. Further, these strategies offer the opportunity to collect data in a *dynamic* and *continuous* way, rather than sampling phenomena of interest at specific time intervals. This is a notable advantage for SCED research in which repeated, systematic measurement of outcomes is critical to elucidating functional or causal relations between independent and dependent variables. For example, using GPS and accelerometer to measure mobility could provide information on whether an exercise intervention (independent variable) leads to within-individual changes in physical activity (dependent variable) compared to a baseline phase in a multiple-baseline or alternating treatment SCED. Alternatively, the association between restricting teenagers' smartphone access after a certain time at night (independent variable) and sleep quality (dependent variable) might be assessed in SCED using a wearable sleep tracker without relying on subjects to subjectively report on sleep habits.

Passive sensing also can be combined with actively collected data (e.g., EMA) to increase the validity and reporting accuracy of outcome variables in SCED, particularly when the dependent variable is a more complex symptom or construct. For example, a researcher might use participants' responses to EMA surveys on affect in conjunction with passively collected indicators of social activity such as call and text logs and GPS to indicate severity of social anxiety and assess the impact of an intervention on this multi-faceted outcome in SCED. The benefits of objectively measuring behaviors such as physical activity and sleep are well-documented (e.g., Choi, Chen, Stein, Klimentidis, & Wang, 2018; Prince et al., 2008; van de Water, Holmes, & Hurley, 2011) and though not specific to SCED research, have the potential to significantly increase the accuracy and external validity of conclusions from studies using these assessment methods.

1.3. Real-time monitoring and SCED in clinical practice

One advantage of SCED over larger group-based designs is the potential for these methods to be readily deployed in clinical practice given the relatively low cost and resources needed (e.g., Barlow et al., 2009; Kazdin, in press; Nock et al., 2007). Clinicians can use SCED to rigorously examine whether (and how rapidly) an intervention (psychological, behavioral, pharmacological, etc.) is working to inform their clinical decision-making, as well as to elucidate the treatment components responsible for change (e.g., in multicomponent CBT, a therapist and patient might want to know whether changes in cognitive style or activity levels precede and drive changes in symptoms). Below we describe several ways in which active and passive real-time monitoring technologies may increase the feasibility of conducting clinically useful SCED studies in clinical settings, thereby promoting the integration of science and practice.

First, applying EMA or passive sensing in SCED studies in clinical practice may increase the timeliness with which clinicians receive information that can be used to inform tailoring of treatment, which is a common aim of SCED (e.g., Barlow et al., 2009). Most current mobile platforms upload and transfer data from the smartphone or other device in very short time frames (e.g., every hour, whenever WiFi is available). Thus, clinicians using SCED can view patients' data well in advance of their next scheduled session and possibly, use this information to inform treatment planning. Clinicians who are monitoring their patients' data remotely might also contact an individual between sessions if incoming data suggest that a change to intervention dosing, a sooner appointment, or phone check-in is needed. To the extent that real-time monitoring technologies facilitate clinician-patient communication, it is possible that clinicians applying these tools in SCED studies may enhance the therapeutic relationship (e.g., Hsin & Torous, 2018), shared decision-making (e.g., van Os et al., 2017), and ultimately, improve clinical outcomes.

Another way clinicians might use incoming idiographic real-time monitoring data to inform treatment planning is to identify the dynamic processes and specific symptoms that maintain or drive patients' conditions (e.g., Bringmann et al., 2016; Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017; Greene, Gelkopf, Eskamp, & Fried, 2018) and thus may be most critical to target in personalized interventions (e.g., Fisher & Boswell, 2016), potentially deployed within a SCED. Such applications of real-time monitoring in clinical practice are not without challenges, several of which we elaborate upon further later on in this paper; as just one example, determining how to synthesize and communicate real-time data in ways that are actionable to clinicians is not yet clear. Given the complex statistical approaches needed to make sense of vast amounts of intensive longitudinal data, automated algorithms that determine the interventions most likely to produce the best (and most rapid) response for an individual patient based on real-time data (e.g., Fernandez, Fisher, & Chi, 2017; van Os et al., 2018) are promising to ensure the feasibility of these assessments in SCED studies conducted in clinical settings.

Perhaps the most common criticism of SCED is the issue of generality, or the degree to which findings from a single participant or small number of participants generalize to other settings and populations. This concern is usually mitigated by emphasizing the feasibility (in terms of cost and time) of replication in SCED, as well as the high degree of experimental control and potential to identify relevant controlling variables, as these factors increase the likelihood that other researchers can replicate a SCED study almost exactly as it has been conducted previously (Morgan & Morgan, 2001; Normand, 2016; Tate et al., 2016). Further, it is also possible that the increased ecological validity associated with real-time monitoring may further enhance the generality of SCED research. Completing assessments in the "real world" may decrease the potential for clinic- or lab-specific characteristics to influence responding and thus result in discrepant findings across different settings. Given the well-documented discrepancies between EMA and retrospective assessments of behaviors such as substance use (for a review see: Morgenstern, Kuerbis, & Muench, 2014) and internal experiences such as depressive symptoms and suicidal ideation (e.g., Torous et al., 2015), research that systematically evaluates whether the real-time monitoring versus standard measures makes results from a SCED study more likely to be replicated would be informative.

1.4. Systematic literature review of SCED studies with real-time monitoring

Given the applicability of real-time monitoring technologies to SCED, we sought to identify studies that have used these assessment strategies to measure dependent variables in the context of a SCED. We conducted a search for published English-language journal articles in PubMed and PsycInfo databases using the following terms: "single-case"/"single-case experimental design"/"single case"/"SCED"/"SCD" *crossed with* "ecological momentary assessment"/"EMA"/"real-time"/"real time"/"ambulatory assessment"/"experience sampling"/"ESM"/"smartphone"/"palm pilot"/"sensor"/"digital"/"technology"/"daily diary." The titles and abstracts of all journal articles were searched up to September 2018. A total of 33 abstracts were initially identified and the titles and abstracts were screened identify potentially relevant studies. We also identified an additional 10 studies through review of reference lists and review articles. We used these criteria for article

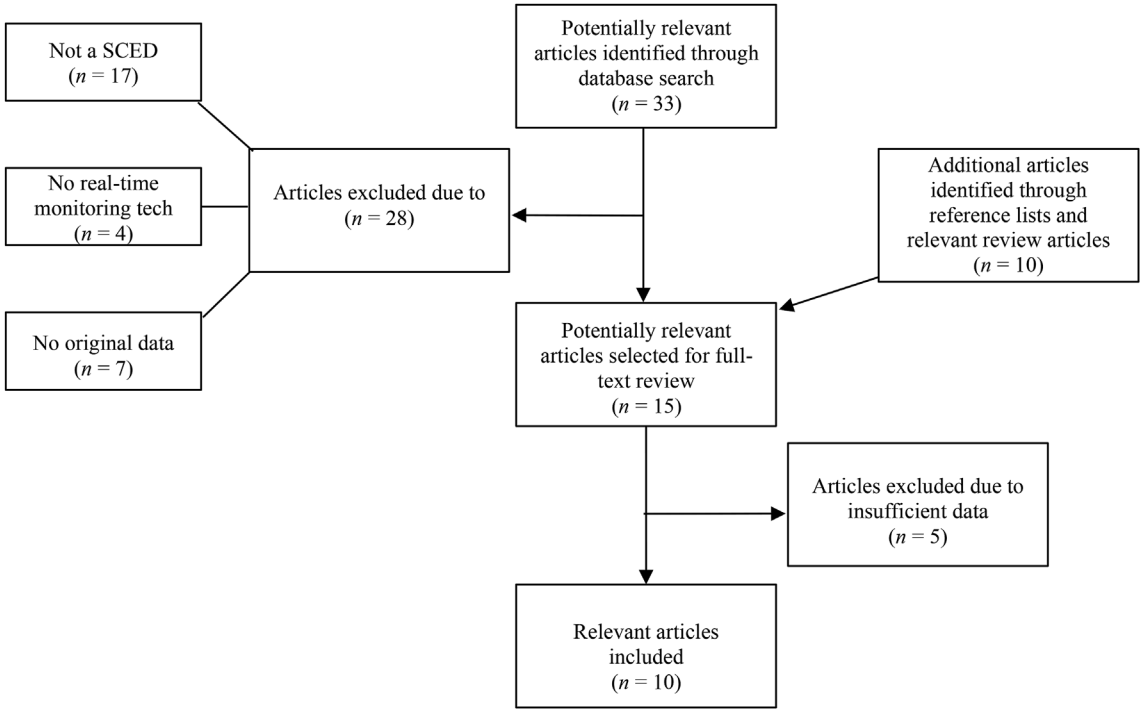

Fig. 1. Selection flow chart of studies using single-case experimental design (SCED) and real-time monitoring technology to assess outcome variables.

selection: (1) the study employed a SCED (e.g., alternating treatment, multiple baseline, combined series, changing criterion, reversal, mixed design), not a quasi-experimental design, case study, naturalistic design, "N-of-1" without rigorous experimental manipulation, etc., (2) the study used real-time monitoring technology (e.g., palm pilot, tablet, smartphone, wearable device) to assess primary outcome variables *in situ* and in (or close to in) real time (studies that exclusively used paper-and-pencil daily diaries were excluded given our emphasis on technology), (3) the study focused on a behavioral or psychological phenomenon, and (4) the study presented original data. When it was unclear based on title and abstract alone whether these criteria were met, the full-text was reviewed. The numbers of articles excluded (and reasons for exclusion) are presented in Fig. 1. After applying these criteria and examining the references lists of relevant review articles from the initial literature search, 10 articles were identified to meet these criteria for article selection.

Relevant information of the 10 included articles by authors and year is shown in Table 1. Here, we provide a summary of the real-time monitoring methods and themes observed across these studies, each of which used active or passive real-time monitoring to evaluate the impact of an intervention on an outcome. First, a broad variety of SCED types are represented in only 10 studies: alternating treatment, multiple baseline, reversal, and phase change. This highlights the flexibility with which real-time monitoring can be applied in the context of diverse SCEDs. Second, all studies used a relatively small number of participants: between one (Wichers & Groot, 2016) and 17 (Daskalova et al., 2016). Even with these small "sample sizes," the use of real-time monitoring resulted in large numbers of observations (for the six studies that included this exact information, the range of analyzed data points was 32 [from 3 participants in Cushing, Jensen, & Steele, 2010] to 1474 [from one participant in Wichers & Groot, 2016], with a median of 423). Seven of the 10 studies presented both visual and statistical analyses (defined here as any inferential statistics; for example, statistical significance of change between phases or correlations between independent and dependent variables), in line with best practice guidelines for SCED (Barlow et al., 2009; Tate et al., 2016).

Third, six of 10 studies used active forms of real-time monitoring, with three studies that used EMA (Bentley et al., 2017; Shingleton et al., 2012; Wichers & Groot, 2016). Other active real-time monitoring methods used included self-reported adherence to target behaviors (e.g., sleep) and daily subjective ratings of outcomes (Daskalova et al., 2016; Taylor et al., 2016), adherence to diet self-monitoring (Cushing et al., 2010), and video recordings of engagement in a behavior (e.g., carbon monoxide monitor use) via a web camera (Raiff & Dallery, 2010; Reynolds, Dallery, Shroff, Patak, & Leraas, 2008). Though Raiff and Dallery (2010) and Reynolds et al. (2008) used real-time monitoring that required active subject participation, we classify these assessments as objective as subjects were required to record themselves using a health monitoring device rather than reporting retrospectively on whether they used it. Only three of 10 studies measured a primary outcome or target behavior with (at least in part) passive assessments: daily step count with a pedometer (Nyman, Goodwin, Kwasnicka, & Callaway, 2016; Sniehotta, Pressau, Hobbs, & Arujo-Soares, 2012), physical activity with a wearable (Taylor et al., 2016), and sleep metrics (Daskalova et al., 2016).

Six of 10 studies required subjects to use a separate study device for data collection (e.g., personal digital device, pedometer, web camera), with the four most recent studies (Bentley et al., 2017; Daskalova et al., 2016; Shingleton et al., 2016; Taylor et al., 2016) using the participants' own smartphone; Taylor et al. (2016) used both the participant's smartphone and a wearable sensor. Regarding frequency of assessments, all studies used much shorter time intervals between assessments than is typical in most SCED studies, ranging from daily (Daskalova et al., 2016; Shingleton et al., 2016) to up to three times per day (e.g., Cushing et al., 2010) to up to 10 times per day (e.g., Wichers & Groot, 2016), whereas two studies employed continuous monitoring of at least one target behavior or outcome (Nyman et al., 2016; Sniehotta et al., 2012; Taylor et al., 2016). For the three SCED studies that employed EMA, two used interval-contingent monitoring only (Shingleton et al., 2017; Wichers & Groot, 2016) and one (Bentley et al., 2017) used a combination of interval- and event-contingent monitoring.

These studies exemplify how real-time monitoring can facilitate assessment of dependent variables with a level of reporting accuracy (e.g., Bentley et al., 2017 [self-injurious urges and acts]; Daskalova

Table 1
Studies using real-time monitoring technology to assess outcomes in a single-case experimental design.

Article	SCED	Subjects	Data Points	Intervention Tested	Real-time Monitoring Technology	Dependent Variable(s) Assessed with Real-time Monitoring Technology
Bentley et al. (2017)	Multiple baseline/ combined series	n = 10 (adults with nonsuicidal self-injury)	n = 855	Mindful emotion awareness training and cognitive flexibility	EMA at least daily (smartphone)	Nonsuicidal self-injury urges and acts
Cushing et al. (2010)	Multiple baseline	n = 3 (overweight adolescents)	n = 32	Personal electronic device for self-monitoring adherence in pediatric weight management program	Electronic self-monitoring up to three times per day (personal digital device)	Diet self-monitoring adherence
Daskalova et al., 2016	Phase change (ABAB)	n = 17 (undergraduate students)	n = at least 571 days with sleep data (exact n not reported)	System integrating passive and active data from a sleep app and individualized text messages	Self-report survey and sensor data (smartphone)	Perceived restfulness, sleep onset latency, frequency of awakenings
Nyman et al. (2016)	Alternating treatment	n = 8 (older adults)	n = 496	Goal-setting (text message prompted) or self-monitoring	Passive and continuous sensing (piezoelectric pedometer)	Daily step count
Raiff and Dallery (2010)	Reversal	n = 4 (adolescents with type 1 diabetes)	n = 60	Web-based contingency management program for blood glucose testing	Video recordings of at-home blood glucose monitor use up to eight per day (web cam)	Adherence to blood glucose testing
Reynolds et al. (2008)	Reversal	n = 4 (daily-smoking adolescents)	n = 350	Web-based contingency management program for smoking abstinence	Video recordings of at-home carbon monoxide monitor use three times per day (web cam)	Carbon monoxide levels
Shingleton et al. (2016)	Alternating treatment	n = 12 (adults with eating disorders)	Not reported	Personalized motivational text message intervention	EMA daily (smartphone)	Dietary restraint, restriction, motivation
Sniehotta et al. (2012)	Alternating treatment	n = 10 (normal or overweight adults)	Not reported	Goal-setting (text message prompted) or self-monitoring	Passive and continuous sensing (piezoelectric pedometer)	Daily step count
Taylor et al., 2016	Phase change ($AB_1B_2B_3$)	n = 13 (university students)	Not reported	Mobile platform for conducting single-case self-experiments	Mobile app use (smartphone) and passive sensing (wearable fitness tracker)	App adherence (e.g., daily use of app), self-reported adherence to target behaviors (e.g., sleep, physical activity) and daily self-reported outcomes (e.g., stress, happiness, productivity); passive activity tracking
Wichers and Groot (2016)	Phase change (AB)	n = 1 (adult with major depression)	n = 1474	Discontinuation of antidepressant (venlafaxine)	EMA up to 10 times per day (personal digital device)	Momentary affective states

Note. SCED = single-case experimental design. EMA = ecological momentary assessment.

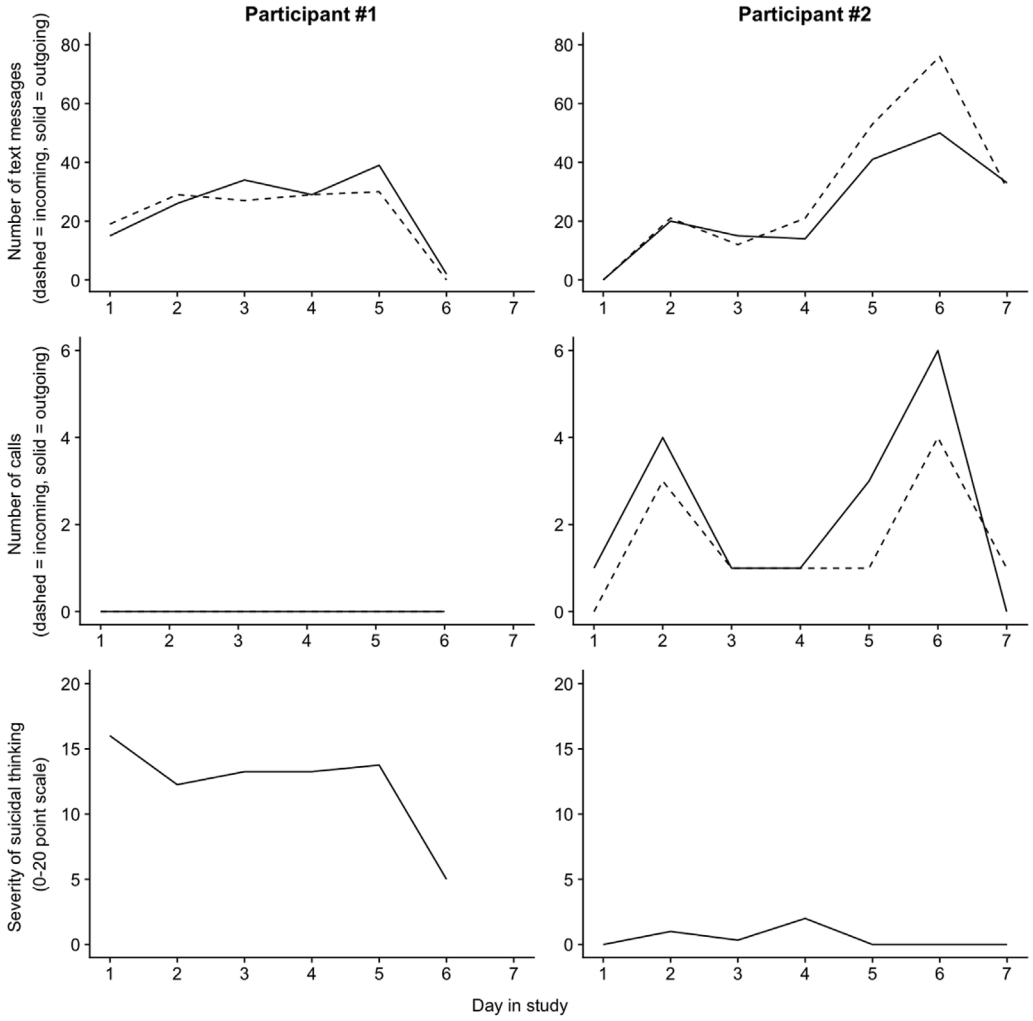

Fig. 2. Example of real-time monitoring data from two participants (Participant 1 in the left panel and Participant 2 in right panel). *Note:* No smartphone data were available for Participant 1 on Day 7 in the study.

et al., 2016 [subjective sleep ratings]; Shingleton et al., 2016 [dietary restraint, restriction, motivation]; Taylor et al., 2016 [happiness, stress, and productivity]), granularity (e.g., Wichers & Groot, 2016 [momentary affect]), or ecological validity (e.g., Daskalova et al., 2016 [sleep metrics]; Nyman et al., 2016 and Sniehotta et al., 2012 [step count]) that would be more difficult (if not impossible) to obtain with traditional assessments. Given the potential value of incorporating real-time monitoring technologies in SCED research, we were surprised to find relatively few published studies that have combined these methods to date (and only two SCED studies that used a combination of active and passive real-time monitoring).

1.5. *Original data: real-time active and passive monitoring in an idiographic study*

Here, we present original, idiographic data from our research team that illustrates how active and passive real-time monitoring technologies can provide a more granular view of behavior at the individual level and demonstrate functional associations between variables of interest. Fig. 2 shows one week of data from two participants who participated in a larger real-time monitoring study during their psychiatric inpatient stay. The findings presented here are original, as prior publications from this study (Kleiman et al., 2017, 2018) have not included passive sensing data. These new data illustrate how real-time monitoring makes it possible to observe individual-level relations between real-world behaviors (social interaction via phone calls and text messages) and a highly clinically-relevant outcome variable (self-reported suicidal thoughts). Rather than asking participants to actively (and retrospectively) self-report on their number of social interactions, these passive data provide an objective (and thus, potentially more accurate and less biased) indicator of interpersonal communication while undergoing inpatient treatment. As this study did not manipulate an independent variable, it would not be considered SCED, but the data granularity at the level of the individual is in line with SCED research.

As seen in Fig. 2, Participant 1 (in the left panel) demonstrated concurrent decreases in suicidal ideation as the number of text messages decreased, whereas Participant 2 (in the right panel) had a less discernible relation between passive and active data. For Participant 2, it might be inferred that increased social contact (measured via number of incoming and outgoing texts and phone calls) is negatively associated with suicidal thinking, whereas the opposite was observed for Participant 1. Thus, these data could be especially useful from an idiographic point of view as they may help tease apart why an intervention is effective for some individuals but not others. For example, it may be that social contact is not inherently always positive (e.g., Participant 1) and simply encouraging a patient to reach out to others more frequently may lead to *increases* in suicidal thinking if the social contact is stressful in nature. Moreover, these data illustrate individual-level differences in *how* people communicate (e.g., Participant 1 never made a phone call, but texted regularly whereas Participant 2 regularly made calls and sent

Table 2
Key challenges and recommendations for real-time monitoring technology in SCED research.

Challenge	Recommendations
Participant compliance and retention	• Balance the desire for maximum numbers of data points with participant burden • Consider the effects of sensor sampling rates on device battery life • Compensate participants for compliance (while considering the potential for reduced generality)
Privacy and confidentiality concerns	• Select platforms that meet best practice security guidelines • Consult experts in information technology, professional organizations, and online resources • Discuss privacy safeguards, how data will be used, and potential utility of real-time monitoring during the informed consent process
Real-time data management/analysis	• Ensure qualified staff, expertise, and resources for data cleaning, integration of multiple data streams, and data analysis • Exercise caution when drawing conclusions about complex psychological and behavioral phenomena from passive sensing data
Potential for reactivity to real-time monitoring	• Consider factors that may influence the potential for reactivity, such as the target behavior and complexity of monitoring protocol • Closely monitor for reactivity during non-intervention phases and extend baseline phases until stability is reached as needed
Lack of universal device ownership	• Provide a low-cost study smartphone/wearable device if feasible
Responding to risk in real-time data	• Establish protocol for monitoring and responding to incoming real-time data prior to study start (likely in collaboration with IRB) • Be transparent with participants about protocols for responding to indicators of risk during the informed consent process

text messages). Beyond the clinical implications, such data also lend themselves well to inferential tests of hypotheses (e.g., using hierarchal linear modeling; Davis et al., 2013).

2. Challenges and recommendations for real-time monitoring in SCED

Despite the advantages of using real-time monitoring technology in SCED, there are also many challenges associated with implementing these assessment strategies. Here, we discuss several important issues to consider when using real-time monitoring in SCED and where appropriate, offer recommendations for addressing these challenges (summarized in Table 2). Given how frequently digital health platforms and apps change, we do not list (or recommend) any one specific tool over another; rather, we address more general considerations in this rapidly advancing field.

First, compliance with real-time assessment tools is key, and thus it is crucial that SCED using real-time monitoring technology keep participants engaged. Response fatigue is a potential issue in studies using these methods, as responding to multiple prompts per day over extended periods of time can be burdensome and lead to large amounts of missing data or dropout. When incorporating these methods in SCED studies, researchers must strike a balance between the desire for large numbers of data points and reasonable participant burden. Passive sensing may be less burdensome than active monitoring; however, even passive sensing requires some level of engagement such as remembering to wear and charge a wearable or ensuring that a smartphone app is running in the background. Relatedly, mobile sensing apps – especially those that use very high or continuous sampling rates of sensor data – can have a noticeable effect on battery life (e.g., Boonstra et al., 2018). Researchers must weigh the sampling rates that provide the most accurate, fine-grained sensor data against the effects on battery consumption. Notably, some apps designed for research offer customizable sampling rates for GPS and accelerometer sensors (e.g., Torous, Kiang, Lorme, & Onnela, 2016).

To address issues related to compliance in real-time monitoring research, we (and many others) have found that compensating participants for completing real-time assessments (e.g., Bentley et al., 2017; Kleiman et al., 2018, 2017; Nock, Prinstein, & Sterba, 2009) or wearing sensors helps incentivize compliance. For example, our team usually pays participants per each completed assessment and day with complete sensor data as well as a bonus payment for high overall compliance (e.g., responding to at least 75% of EMA prompts). This may be less realistic, however, for clinicians using SCED methods in the context of an ongoing practice due to a lack of research funds and the potential ethical issues associated with paying one's own patients for providing clinical data. Like any treatment outcome research, results from SCED studies that compensate patients for their participation may not generalize to routine clinical settings. The trade-offs between maximizing compliance with monetary incentives and generality of SCED studies using real-time monitoring must be considered carefully.

Second, participants may be hesitant to engage with these digital tools for other reasons, including a lack of trust in the technology to securely capture active or passive data and to keep this informational confidential (e.g., Torous & Roberts, 2017; Trull & Ebner-Priemer, 2014). These concerns may be especially relevant for researchers hoping to collect data that may be perceived as more personal, such as recording the content of text messages, phone calls, or specific locations from GPS (e.g., Di Matteo. Fine, Fotinos, Rose, & Katzman, 2018). Participants may also question whether the information gained from passive sensing is accurate and useful (e.g., Dennison, Morrison, Conway, & Yardley, 2013), thus potentially reducing willingness to provide these data as part of a SCED study. It may be possible to increase engagement by having a discussion with participants about privacy, what information can and cannot be accessed, how their data will be used, and the potential utility of using these devices (e.g., more accurate determinations of whether or not an intervention is working, reduced need for time-intensive, costly trips to the lab or clinic for in-person assessments). Researchers and clinicians are encouraged to use platforms that employ best practices for ensuring data security and confidentiality, including on the device itself and during transmission to secure servers. Given that technology is rapidly evolving, consulting with experts in information technology at a local hospital or university to ensure that the apps and devices meet best practice security guidelines has been recommended (Torous, Chan, Yellowlees, & Boland, 2016). Organizations such as the American Psychiatric Association the Anxiety and Depression Association of America, as well as PsyberGuide, provide resources to assist with selection of mobile apps.

Third, real-time monitoring data can be cumbersome to analyze (especially in an ongoing fashion to monitor the effects of an intervention and tailor treatment, which is often a goal of SCED). Even on the low end (e.g., active monitoring with only a few data points per day), datasets can become extremely large even with a small number of participants. For example, the study by Reynolds et al. (2008) included above collected approximately 88 data points per participant. Even a small SCED study (e.g., two or three active participants at any one time) can involve time-intensive management of incoming data streams. This issue is especially relevant for devices that continually collect passive data, which can create up to several hundred data points per minute (e.g., devices like the Empatica E4 sample heart rate 64 times *per second*).

In addition to the sheer volume of data points, real-time monitoring

data (especially from passive sensing) can be very messy. The collection of physiological and behavioral data for extended periods of time as individuals go about their daily lives introduces countless potential confounding factors (e.g., exercise, erratic device wear, substance use) and artefacts that are not present in more constrained laboratory-based experiments. Real-time monitoring data also may not always be as accurate as assumed by the user/researcher. For example, though wrist-worn sensors tend to more accurately estimate some variables, like number of steps, they can over-estimate others, such as sleep time and efficiency, and heart rate (e.g., Benedetto et al., 2018; Evenson, Goto, & Furberg, 2015). Along these lines, an increase in heart rate, skin conductance, or motion captured by a sensor, for example, may influenced by a variety of factors – an increase in psychological distress, agitation, excitement, exercise, or medication side effect, just to name a few – which can make it difficult to map these objective data onto specific psychological constructs or outcomes of interest and thus operationalize specific independent and dependent variables in SCED. Though time-matched self-report surveys can help provide contextual information about objective data from specific time frames of interest, EMA of course cannot be conducted continuously. Accordingly, it is important to exercise caution when drawing conclusions from real-time monitoring data (especially regarding *causal* relations between variables) and consider the complexity of data that will come from using real-time monitoring in a SCED study. Ensuring the presence of appropriate staff, expertise, and resources for processinf these data, including handling the missing data that is inevitable when conducting repeated, frequent assessments over extended time periods, is critical.

Fourth, repeatedly asking participants about their thoughts, feelings, and behaviors or other mental health symptoms in EMA protocols may increase their awareness of or insight into these phenomena (e.g., van Ballegooijen et al., 2016; Wichers et al., 2011) and thus influence outcomes and internal validity in SCED. Indeed, the therapeutic effects of self-monitoring as an intervention in and of itself for many mental health conditions and populations is well-established (e.g., Burke, Wang, & Sevick, 2011; Kazdin 1974; Runyan et al., 2013). The potential for reactivity to EMA (and even passive sensing, which can lead to participants feeling “observed”) is especially a concern when using these assessments over longer periods of time or to assess outcomes in SCED studies aimed to test interventions. It can become difficult to decipher whether changes in outcomes are due to the experimental intervention, real-time monitoring, or a combination. This issue can be at least partially addressed with a baseline (monitoring-only) phase (e.g., Bentley et al., 2017; Shingleton et al., 2016), which can be extended as needed until stability (or worsening) in outcome variables is observed (e.g., Barlow et al., 2009) and thus critical to isolating any potential independent effect of self-monitoring.

Overall, the evidence for reactivity to EMA is mixed; for example, in the substance use literature, whereas some studies have found little to no evidence of reactivity (e.g., Hufford, Shields, Shiffman, Paty, & Balabanis, 2002; Stone et al., 2003), others have observed some reactivity to EMA for certain symptoms and experiences related to smoking (e.g., McCarthy, Minami, Yeh, & Bold, 2015; Rowan et al., 2007). SCED researchers should also consider the many factors likely to influence the potential for monitoring reactivity, including the complexity of assessment protocols (e.g., a simple EMA protocol assessing momentary affect may present less risk of reactivity than more intensive surveys assessing the thoughts, feelings, and contextual events leading up to engagement in a high-risk behavior), temporal ordering of smartphone surveys and behaviors of interest (e.g., will participants be asked about cravings *before* drinking or drug use has occurred, or only afterward in a daily report), and outcomes of interest (e.g., episodic behaviors versus subjective experiences) (e.g., Wray, Merrill, & Monti, 2014).

Fifth, real-time monitoring technology, despite becoming more accessible every year, is still disproportionally not accessible in some populations that may be of high clinical interest in SCED research. Though, as noted earlier, nearly 90% of people ages 18 to 49 own a smartphone (Pew Research Center, 2017), the same is true for only 46% of people 65 and older. Moreover, despite the fact that almost 80% of people of all ages in urban and suburban areas own a smartphone, the same is true for only 65% of people in rural areas. Thus, individuals who we may most want to target in our research (and are the least likely to be able to access traditional in-person delivery modes of assessment and care) may be the least likely to have access to the basic technology needed to engage in real-time monitoring. It is now possible for researchers to purchase and provide smartphones and wearable devices for under $100 each, but this still presents a challenge of added cost. Further, many of the more widely used programs designed for real-time monitoring are expensive themselves (e.g., several thousand dollars for an annual subscription), which may be cost-prohibitive for smaller-scale SCED studies, especially those conducted within a clinical practice. In a related vein, some apps are not compatible with iPhone and Android, and not all data streams are available on all devices. For example, the default security in Apple iPhone devices prevents any app from accessing call or text message logs.

Finally, receiving incoming data in real-time (or close to real-time) from patients can present important scientific, clinical, and ethical challenges to researchers. For those of us interested in collecting EMA data on self-injurious thoughts and behaviors, for example, there is not yet formal consensus about the best practices for responding to real-time indicators of risk (e.g., self-report EMA data on suicidal intent, potentially also in combination with a participant's GPS coordinates) or multiple missed surveys in a row, which might also suggest increased risk. Scientists conducting research on other high-risk behaviors, such as illicit drug use and other-directed violence, face similar issues. In our real-time monitoring studies, during informed consent, we are transparent with participants about our protocols for monitoring and responding to risk (including the delay between submitting an EMA survey and our team reviewing and responding to the data, etc.) and ensure their understanding through quiz-like questions. We also work closely with our IRBs to develop and update study risk management protocols as needed.

Clinicians conducting SCED studies with real-time monitoring in their practice may face similar ethical-social-legal issues. Even when high-risk behaviors are not explicitly assessed, does incoming real-time data that suggests a worsening in depressive or anxious symptoms necessitate the clinician following up with the patient before the next visit? If the clinician has access to these data but does not monitor or respond to them regularly, are they liable if an adverse outcome were to occur? For SCED research in a clinical setting, it may be less feasible to have a research assistant monitor or assist in responding to incoming data; for busy providers, this may be a practical barrier to the deployment of real-time monitoring. As these technologies continue to become more advanced, however, we anticipate that the development of streamlined, automated systems that leverage patients’ digital devices to detect increases in risk will facilitate the widespread use of real-time assessment in SCED and clinical practice.

3. Conclusions

SCED is a practical, cost-effective, and rigorous research methodology with great applicability to the study of behavior and behavior change. A hallmark feature of SCED is the repeated, systematic assessment of one's dependent variables across all study phases. Real-time monitoring technology, which includes both active (e.g., smartphone-based EMA) and passive assessment (e.g., wearable sensors, smartphone activity tracking), is increasingly utilized in psychological research to capture fine-grained, continuous data outside of the laboratory or clinic. Despite the potential for active and passive real-time monitoring to facilitate the repeated, systematic measurement needed in SCED studies, we found only a small number of published studies that have combined real-time monitoring and SCED thus far. Though real-time

monitoring technology is associated with challenges (e.g., participant burden, cost, privacy considerations, potential for reactivity, complex data analyses), when used appropriately, these strategies can significantly augment the granularity and ecological validity of individual-level data obtained in SCED research. These assessment methods may even increase the feasibility and potential value of SCED studies conducted in clinical settings. We anticipate that in the coming years, researchers and clinicians will increasingly combine these two powerful methods to produce new insights into human behavior and novel interventions for mental health conditions.

Funding

This work was supported by the Stuart T. Hauser Research Training Program in Biological and Social Psychiatry (T32MH016259) (KHB) and the Chet and Will Griswold Suicide Prevention Fund (MKN).

References

van Ballegooijen, W., Ruwaard, J., Karyotaki, E., Ebert, D. D., Smit, J. H., & Riper, H. (2016). Reactivity to smartphone-based ecological momentary assessment of depressive symptoms (MoodMonitor): Protocol for a randomized controlled trial. *BMC Psychiatry, 16*, 359. https://doi.org/10.1186/s12888-016-1065-5d.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case research designs: Strategies for studying behavior change.* New York, NY: Allyn and Bacon.

Ben-Zeev, D., Scherer, E. A., Wang, R., Xie, H., & Campbell, A. T. (2015). Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric Rehabilitation Journal, 38*(3), 218–226. https://doi.org/10.1037/prj0000130.

Benedetto, S., Caldato, C., Bazzan, E., Greenwood, D. C., Pensabene, V., & Actis, P. (2018). Assessment of the Fitbit Charge 2 for monitoring heart rate. *PloS One, 13*(2), e0192691. https://doi.org/10.1371/journal.pone.0192691.

Bentley, K. H., Nock, M. K., Sauer-Zavala, S., Gorman, B. S., & Barlow, D. H. (2017). A functional analysis of two transdiagnostic, emotion- focused interventions on non-suicidal self- injury. *Journal of Consulting and Clinical Psychology, 85*(6), 632–646. https://doi.org/10.1037/ccp0000205.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030.

Boonstra, T. W., Nicholas, J., Wong, Q. J. J., Shaw, F., Townsend, S., & Christensen, H. (2018). Using mobile phone sensor technology for mental health research: Integrated analysis to identify hidden challenges and potential solutions. *JMIR, 20*(7), e10131.

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., et al. (2016). Assessing temporal emotion dynamics using networks. *Assessment, 23*(4), 425–435.

Burke, L. E., Wang, J., & Sevick, M. A. (2011). Self-monitoring in weight loss: A systematic review of the literature. *Journal of the American Dietetic Association, 111*(1), 92–102.

Choi, K. W., Chen, C.-Y., Stein, M. B., Klimentidis, Y. C., & Wang, M.-J. (2018). Major depressive disorder working group of the psychiatric genomics consortium. In J. W. Smoller (Ed.). *Testing causal bidirectional influences between physical activity and depression using mendelian randomization*https://doi.org/10.1101/364232 bioRXiv.

Cushing, C. C., Jensen, C. D., & Steele, R. G. (2010). An evaluation of a personal electronic device to enhance self-monitoring adherence in a pediatric weight management program using a multiple baseline design. *Journal of Pediatric Psychology, 36*(3), 301–307.

Daskalova, N., Metaxa-Kakavouli, D., Tran, A., Nugent, N., Boergers, J., McGeary, J., & Huang, J. (2016). SleepCoacher: A personalized self-experimentation system for sleep recommendations. *UIST Proceedings of the 29th Annual Symposium on User Interface Software and Technology*https://doi.org/10.1145/2984511.2984534.

Davis, D. H., Gagné, P., Fredrick, L. D., Alberto, P. A., Waugh, R. E., & Haardörfer, R. (2013). Augmenting visual analysis in single-case research with hierarchical linear modeling. *Behavior Modification, 37*, 62–89. https://doi.org/10.1177/0145445512453734.

Dennison, L., Morrison, L., Conway, G., & Yardley, L. (2013). Opportunities and challenges for smartphone applications in supporting health behavior change: Qualitative study. *Journal of Medical Internet Research, 15*(4)https://doi.org/10.2196/jmir.2583.

Di Matteo, D., Fine, A., Fotinos, K., Rose, J., & Katzman, M. (2018). Patient willingness to consent to mobile phone data collection for mental health apps: Structured questionnaire. *JMIR Mental Health, 5*(3), e56.

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity, 12*, 159. https://doi.org/10.1186/s12966-015-0314-1.

Fernandez, K. C., Fisher, A. J., & Chi, C. (2017). Development and initial implementation of the dynamic assessment treatment algorithm (DATA). *PloS One, 12*(6), e0178806.

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment, 23*(4), 496–506.

Fisher, A. J., Reevez, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 126*(8), 1044–1056.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65*, 45–55.

Greene, T., Gelkopf, M., Epskamp, S., & Fried, E. I. (2018). Dynamic networks of PTSD symptoms during conflict. *Psychological Medicine, 28*, 1–9. https://doi.org/10.1017/S0033291718000351.

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics. *Intensive longitudinal data. Current Directions in Psychological Science: Vol. 26*, (pp. 10–15). . https://doi.org/10.1177/0963721416666518 (1).

Hsin, H., & Torous, J. (2018). *Creating boundaries to empower digital health technology BJPsych Open, Vol. 4*, 235–237 (4).

Hufford, M. R., Shields, A. L., Shiffman, S., Paty, J., & Balabanis, M. (2002). *Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. Psychology of Addictive Behaviors, Vol. 16*, 205 (3).

Kazdin, A. E. (1974). Reactive self-monitoring: The effects of response desirability, goal setting, and feedback. *Journal of Consulting and Clinical Psychology, 42*(5), 704.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432.

Kazdin, A. E. (2010). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.

Kazdin, A.E. (in press). Single-case experimental designs. Evaluating interventions in research and clinical practice. Behaviour Research and Therapy.

Kleiman, E. M., & Nock, M. K. (2017). Advances in scientific possibilities offered by real-time monitoring technology. *Psychiatry, 80*(2), 118–124. https://doi.org/10.1080/00332747.2017.1325661.

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.

Kleiman, E. M., Turner, B. J., Fedor, S., Beale, E. E., Picard, R. W., Huffman, J. C., et al. (2018). Digital phenotyping of suicidal thoughts. *Depression and Anxiety, 35*(7), 601–608.

McCarthy, D. E., Minami, H., Yeh, V. M., & Bold, K. W. (2015). An experimental investigation of reactivity to ecological momentary assessment frequency among adults trying to quit smoking. *Addiction, 110*(10), 1549–1560. https://doi.org/10.1111/add.12996.

Morgan, D. L., & Morgan, R. K. (2001). Single-participant research design: Bringing science to managed care. *American Psychologist, 56*(2), 119–127. https://doi.org/10.1037/0003-066X.56.2.119.

Morgenstern, J., Kuerbis, A., & Muench, F. (2014). Ecological momentary assessment and alcohol use disorder treatment. *Alcohol Research, 36*(1), 101–109.

Nock, M. K., Michel, B. D., & Photos, V. (2007). Single-case research designs. In D. McKay (Ed.). *Handbook of research methods in abnormal and clinical psychology* (pp. 337–350). Thousand Oaks, CA: Sage Publications.

Nock, M. K., Prinstein, M. J., & Sterba, S. K. (2009). Revealing the form and function of self- injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. *Journal of Abnormal Psychology, 118*(4), 816–827.

Normand, M. P. (2016). Less is more: Psychologists can learn more by studying fewer people. *Frontiers in Psychology, 7*, 934. https://doi.org/10.3389/fpsyg.2016.00934.

Nyman, S. R., Goodwin, K., Kwasnicka, D., & Callaway, A. (2016). Increasing walking among older people: A test of behavior change techniques using factorial randomized *N*-of-1 trials. *Psychology and Health, 31*(3), 313–330. https://doi.org/10.1080/08870446.2015.1088014.

van Os, J., Verhagen, S., Marsman, A., Peeters, F., Bak, M., Marcelis, M., et al. (2017). The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. *Depression and Anxiety, 34*(6), 481–493. https://doi.org/10.1002/da.22647.

Pew Research Center (2017). *Mobile phone ownership over time.* Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/fact-sheet/mobile/.

Prince, S. A., Adamo, K. B., Hamel, M., Hardt, J., Connor Gorber, S., & Tremblay, M. A. (2008). Comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 5*(1), 56.

Raiff, B., & Dallery, J. (2010). Internet-based contingency management to improve adherence with blood glucose testing recommendations for teens with type 1 diabetes. *Journal of Applied Behavior Analysis, 43*(3), 487–491.

Reynolds, B., Dallery, J., Shroff, P., Patak, M., & Leraas, K. (2008). A web-based contingency management program with adolescent smokers. *Journal of Applied Behavior Analysis, 41*, 597–601.

Rowan, P. J., Cofta-Woerpel, L., Mazas, C. A., Vidrine, J. I., Reitzel, L. R., Cinciripini, P. M., et al. (2007). Evaluating reactivity to ecological momentary assessment during smoking cessation. *Experimental and Clinical Psychopharmacology, 15*(4), 382.

Runyan, J. D., Steenbergh, T. A., Bainbridge, C., Daugherty, D. A., Oke, L., & Fry, B. N. (2013). A smartphone ecological momentary assessment/intervention "app" for collecting real-time data and promoting self-awareness. *PLoS One, 8*(8), e71325.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.

Shingleton, R. M., Pratt, E. M., Gorman, B., Barlow, D. H., Palfai, T. P., & Thompson-Brenner, H. (2016). Motivational text message intervention for eating disorders: A single-case alternating treatment design using ecological momentary assessment. *Behavior Therapy, 47*, 325–338.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*(4), 1–70. https://doi.org/

10.1037/a0029312.
Sniehotta, F. F., Pressau, J., Hobbs, N., & Arujo-Soares, V. (2012). Testing self-regulation interventions to increase walking using factorial randomized *N*-of-1 trials. *Health Psychology, 31*(6), 733–737. https://doi.org/10.1037/a0027337.
Statista. Wearable user penetration rate in the United States, in 2017, by age. (2017). Retrieved from https://www.statista.com/statistics/739398/us-wearable-penetration-by-age/ Accessed April 2, 2018.
Stone, A. A., Broderick, J. E., Schwartz, J. E., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. *Pain, 104*, 343–351.
Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., & Wilson, B. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology, 4*(1), 1–9. https://doi.org/10.1037/arc0000026.
Taylor, S., Sano, A., Ferguson, C., Mohan, A., & Picard, R. (2016). QuantifyMe: an open-source automated single-case experimental design platform. *Sensors, 18*, 1097. https://doi.org/10.3390/s18041097.
Thiele, C., Laireiter, A.-R., & Baumann, U. (2002). Diaries in clinical psychology and psychotherapy: A selective review. *Clinical Psychology & Psychotherapy, 9*(1), 1–37. https://doi.org/10.1002/cpp.302.
Torous, J.,B., Chan, S. R., Yellowlees, P. M., & Boland, R. (2016). To use or not? Evaluating ASPECTS of smartphone apps and mobile technology for clinical care in psychiatry. *Journal of Clinical Psychiatry, 77*(6), e734–e738.
Torous, J. B., Kiang, M. V., Lorme, J., & Onnela, J.-P. (2016). New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health, 3*(2), e16. https://doi.org/10.2196/mental.5165.
Torous, J. B., Onnela, J.-P., & Keshavan, M. (2017). New dimensions and new tools to realize the potential of RDoC: Digital phenotyping via smartphones and connected devices. *Translational Psychiatry, 7*, e1053. https://doi.org/10.1038/tp.2017.25.
Torous, J. B., & Roberts, L. W. (2017). Needed innovation in digital health and smartphone applications for mental health: Transparency and trust. *JAMA Psychiatry, 74*(5), 437–438. https://doi.org/10.1001/jamapsychiatry.2017.0262.
Torous, J. B., Staples, P., Shanahan, M., Lin, C., Peck, P., Keshavan, M., et al. (2015). Utilizing a personal smartphone custom app to assess Patient Health Questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Mental Health, 2*(1), e8. https://doi.org/10.2196/mental.3889.
Trull, T. J., & Ebner-Priemer, U. (2014). THe role of ambulatory assessment in psychological science. *Current Directions in Psychological Science, Vol. 23*, 466–470. https://doi.org/10.1177/0963721414550706.
Vilardaga, R., Bricker, J., & McDonell, M. (2014). The promise of mobile technologies and single case designs for the study of individuals in their natural environment. *Journal of Contextual Behavioral Science, 3*(2), 148–153. https://doi.org/10.1016/j.jcbs.2014.03.003.
van de Water, A. T., Holmes, A., & Hurley, D. A. (2011). Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography: A systematic review. *Journal of Sleep Research, 20*(1), 183–200. https://doi.org/10.1111/j.1365-2869.2009.00814.x.
Wichers, M., & Groot, P. C. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics, 85*, 114–116. https://doi.org/10.1159/000441458.
Wichers, M., Simons, C. J. P., Kramer, I. M. A., Hartmann, J. A., Lothmann, C., Myin-Germeys, I., et al. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica, 124*, 262–272.
Wray, T. B., Merrill, J. E., & Monti, P. M. (2014). Using ecological momentary assessment to assess situation level predictors of alcohol use and alcohol-related consequences. *Alcohol Research, 36*(1), 19–27.