Published in final edited form as:

Account Res. 2019 November ; 26(8): 498–511. doi:10.1080/08989621.2019.1694913.

The Risk-Benefit Ratio of Studying Psychiatric Symptoms via Daily Diary Methods

Caitlin Wolford-Clevenger, Ph.D.[1], **Mickeah J. Hugley, B.S.**[1], **James McNulty, PhD**[2], **Lawrence Christian Elledge, PhD**[3], **Karen Cropsey, PsyD**[1], **Gregory L. Stuart, PhD**[3]

[1]University of Alabama at Birmingham

[2]Florida State University

[3]University of Tennessee-Knoxville

Abstract

Ethics committee members and researchers have deliberated about the risk-benefit ratio of researching sensitive issues such as psychiatric symptoms. Although research has suggested that inquiring about psychiatric symptoms in research generally does not cause harm, these findings have primarily arisen from cross-sectional studies. We examined whether this generalized to repeated, daily assessments of psychiatric symptoms. We collected daily survey data on psychiatric symptoms over 90 days from a sample of 206 (150 women) college students. A subset of the sample (n = 80) provided reactions to study participation administered on the 90th day. Individuals who did not complete the 90th day survey reported higher levels of suicidal ideation and hopelessness than those who did. For individuals who completed the 90th, final assessment, reactions primarily fell within the neutral to positive range, with variation depending on their baseline levels of psychiatric symptoms and identification as religious. This study adds to past work by demonstrating that individuals who remained in the study had neutral to positive experiences. However, participants with greater suicidal ideation and hopelessness were likely to attrit, warranting caution in assuming a low risk-benefit ratio of these studies. Management of risks involved in repeated assessment studies may be informed by this work.

Keywords

repeated measures; harm; comfort; reactions; responses; beneficence

Introduction

Repeated measure designs such as ecological momentary assessment and daily diary methods are increasingly being used to understand the causes and variation of psychiatric symptoms within or across days. Although these methods afford the research benefits of increasing external validity, reducing recall bias, and establishing temporal associations

Correspondence concerning this article should be addressed to Caitlin Wolford-Clevenger, Volker Hall L107, 1670 University Blvd, Birmingham, AL 35233, cclevenger@uabmc.edu, 205-996-2452.

Declaration of Interest Statement

No potential conflict of interest was reported by the authors.

(Bolger, Davis, & Rafaeli, 2003; Shiffman, Stone, & Hufford, 2008), little is known about the possible risks and benefits these methods pose to participants. The Protection of Human Subjects Common Rule requires institutional review boards (IRBs) to determine whether a study's anticipated societal and direct benefits outweigh risks to participants and whether these risks are aptly minimized (United States Department of Health and Human Services [HHS]; 2009). Despite these provisions, evaluation of risks and benefits of psychiatric research remains challenging (Iltis et al., 2013; Lidz & Garverich, 2013). Possible reasons include the subjective nature of weighing risks and benefits and tendencies for IRB members to assign more risk to studies on psychiatric symptoms due to stigmatization (Lakeman & FitzGerald, 2009; Luebbert, Tait, Chibnall, & Deshields, 2008; Oldham, Haimowitz, & Delano, 1999). Research that delineates the risks and benefits of research involving repeated assessment of psychiatric symptoms would inform the review process of these studies.

Several studies have assessed risks and benefits of participation in research on psychiatric symptoms such as suicidal behaviors, suicidal ideation, nonsuicidal self-injury, depressive symptoms, hopelessness, posttraumatic stress disorder (PTSD) symptoms, and alcohol and drug use. Research on these psychiatric symptoms using various designs and samples has demonstrated that a majority of participants do not experience increased distress or symptoms (Biddle et al., 2013; de Beurs, Ghoncheh, Geraedts, & Kerkhof, 2016; Gibson, Boden, Benson, & Brand, 2014; Harris & Goh, 2017; Jacomb et al., 1999; Jorm, Kelly, & Morgan, 2007; Langhinrichsen-Rohling, Arata, O'brien, & Bower, 2006; Law et al., 2015; Muehlenkamp, Swenson, Batejan, & Jarvi, 2015; Mathias et al., 2012; Reynolds, Lindenboim, Comtois, Murray, & Linehan, 2006; Rivlin, Marzano, Hawton, & Fazel, 2012; Roth et al., 2017b; Smith, Cukrowicz, & Poindexter, 2010; Whitlock, Pietrusza, & Purington, 2013). For individuals who do experience distress, personal experience with psychiatric symptoms appears to increase this possibility (Jacomb et al., 1999; Jorm et al., 2007; Langhinrichsen-Rohling et al., 2006; Whitlock, Pietrusza, & Purington, 2013). However, participants have also reported experiencing perceived benefits from research on psychiatric symptoms, including improved mood (Biddle et al., 2013; Gibson et al., 2014) and motivation to increase healthy behaviors (Roth et al., 2017a).

Although research across various populations has shown that researching psychiatric symptoms generally does not cause harm, and may even provide some benefits, most studies assessed symptoms at only one time point. Given the increased usage of ecological momentary assessment and daily diary methods, additional tests of the generalizability of these findings to studies employing these methods are necessary, as they involve more opportunities to cause harm. Further, among studies on psychiatric symptoms employing repeated measures, assessments were infrequent (three to five time points; Mathias et al., 2012; Muehlenkamp et al., 2015; Smith, Cukrowicz, & Poindexter, 2010), with exception to one study that assessed symptoms up to five times a day for two weeks (Law et al., 2015). Finally, as explored in past work (e.g., Whitlock et al., 2013), identifying characteristics of people for whom repeated assessment may cause greater emotional harm could inform researchers about appropriate procedures to minimize harm. Whitlock and colleagues (2013) found no differences in reactions across racial, gender, or sexual orientation identities. An additional test of these findings and exploration of another demographic factor (i.e., identification as religious) will add to our understanding of who may experience greater

harm or benefit from repeated assessment of psychiatric symptoms. In particular, people who identify as religious may experience greater discomfort with a study repeatedly assessing mental health symptoms, as prior research has shown this population experiences greater stigma about mental health (Eisenberg, Downs, Golberstein, & Zivin, 2009).

The purpose of the present study was twofold: 1) to replicate prior work showing a low risk-benefit ratio for studying psychiatric symptoms in a study involving daily assessment of such symptoms over 90 days, and 2) to describe for whom such research might cause greater risk. First, we aimed to explore participants' reactions to the research by examining the percentages of participants having positive, neutral, and negative experiences in the study. We did not have a clear *a priori* hypothesis regarding these percentages due to lack of precise percentages identified in the literature. Second, we aimed to explore sociodemographic differences in these reactions, including age, gender, race, identification as religious, and baseline levels of suicidal behaviors, suicidal ideation, nonsuicidal self-injury, depressive symptoms, hopelessness, PTSD symptoms, and alcohol and drug use. Given the limited research regarding this question, this aim was also exploratory.

Materials and Methods

Participants

We screened 1,338 undergraduate students in introductory psychology courses at a large southeastern university to participate in a larger study on proximal correlates of suicidal thoughts and behaviors that collected daily surveys for 90 days. Thirty-three percent of these students were eligible and completed the baseline phase. Two-hundred and six participants completed at least three daily surveys ($n = 206$), the minimum needed for the larger study aims. A subset ($n = 80$) of this overall sample completed the 90th assessment day, during which the measure of reactions to study participation was administered. See Table 1 for descriptives of the overall sample and subsample that completed the 90th day.

Procedures

The research study received institutional review board approval. Participants were recruited via a southeastern university's Introduction to Psychology subject pool. Study participation was confidential (i.e., subject identities were known but not shared); however, identifiable information was never linked to their baseline and daily data (i.e., data were anonymous). A brief eligibility questionnaire assessed inclusion criteria (designed for the larger study aims): a) 18 years or older, b) consumed alcohol within the past month, and c) ever thought about/attempted suicide. Eligible individuals provided informed consent and completed a baseline survey online. The baseline survey assessed variables related to suicide risk including interpersonal needs, alcohol and drug use, capability for suicide, history of suicidal thoughts and behaviors, nonsuicidal self-injury, history of physical and sexual aggression victimization and perpetration, psychiatric symptoms. Participants who wanted to complete the second phase of the study entered their e-mail address in a separate survey (which was kept confidential). Data were made anonymous by requiring participants to create their own identification codes to link baseline and survey data—a 7-letter unique identifier (Yurek et al., 2008). The daily surveys were sent each day at 6 AM, with a reminder at 12 PM, for 90

days. Each survey asked participants to recall their previous day's behavior (i.e., from the time they awoke until the time they went to sleep).

The daily surveys assessed constructs similar to the baseline survey but in a briefer format, including interpersonal needs, alcohol and drug use, perceived capability for suicide, suicidal thoughts and behaviors, nonsuicidal self-injury, emotional states, and perpetration and victimization of aggression. Participants answered between 40 and 63 items, depending on whether they endorsed certain behaviors (e.g., if drug use was endorsed, type of drug was inquired about). Participants were compensated 75 cents to one dollar per survey, depending on time of study enrollment. Participants with 75% daily compliance were paid an additional $10, totaling to $77.50 to $100 in compensation. Incentive was increased five months into data collection to improve compliance.

Measures

Baseline measures related to the present study aims collected demographic information and history of alcohol use, drug use, depressive symptoms, hopelessness, PTSD symptoms, suicidal ideation, suicidal behaviors, and nonsuicidal self-injury. The 10-item Alcohol Use Disorders Identification Test (AUDIT) is a reliable and valid measure that assessed alcohol use and problems over the past year, with total scores ranging from 0 to 40 (Saunders, Aasland, Babor, De la Fuente, & Grant, 1993). The AUDIT had good internal consistency in the present study (α = .81). The Drug Use Disorders Identification (DUDIT; Stuart, Moore, Ramsey, Kahler, 2004), a reliable and valid 14-item questionnaire, assessed drug use with seven classes of drugs (e.g., stimulants) and problems in the previous year. Total possible scores range from 0 to 52. The internal consistency for the DUDIT in the present sample was acceptable (α = .74).

The 20-item Center for Epidemiological Studies-Depression Scale (CES-D) assessed past-week depressive symptoms. Total scores range from 0 to 60. The CES-D has been found to be reliable and valid (Radloff, 1977), and its internal consistency in the current sample was good (α = .90). Hopelessness was measured using the Brief Hopelessness Measure-Negative, which consists of two items assessing hopelessness on a 5-point scale (1 = absolutely disagree to 5 = absolutely agree). Total possible scores range from 2 to 10. The BHM has demonstrated convergent validity with the Beck Hopelessness Scale (Fraser et al., 2014). Good internal consistency was shown in the present sample (α = .81). PTSD symptom severity was assessed using the 20-item PTSD checklist for the DSM-5 (PCL-5; Blevins, Weathers, Davis, Witte, & Domino, 2015). The PCL-5 is on a 5-point scale (0 = "Not at all" to 4 = "Extremely"), with total possible scores ranging from 0 to 80. The scale has been found to be reliable and valid, and demonstrated good internal consistency in the present sample (α = .94).

Nonsuicidal self-injury was assessed using the 15-item Deliberate Self-Harm Inventory (DSHI; Gratz, 2001), a questionnaire that assesses deliberate injury of body tissue without intent to die. Items were on a yes-no scale, with total possible scores ranging from 0 to 15. The scale has demonstrated good validity and reliability (Gratz, 2001) and demonstrated adequate internal consistency in the present sample (α = .69). The severity of past-week suicidal ideation was assessed using the 4-item Hopelessness Depression Symptom

Questionnaire-Suicidality Subscale (HDSQ-SS; Metalsky & Joiner, 1997). Total possible scores range from 0 to 12. The scale has demonstrated reliability and validity in a prior sample of college students (Metalsky & Joiner, 1997) and in the present sample (α = .91). Regarding suicidal behavior, participants indicated whether they have never attempted suicide, attempted suicide once, or attempted suicide more than once. History of aborted, prevented, or interrupted suicide attempts was assessed via items modeled from the Columbia-Suicide Severity Rating Scale and Child and Adolescent Services Assessment (C-SSR; CASA; Brent et al., 2009; Burns, Angold, Magruder-Habib, Costello, & Patrick, 1997; Posner et al., 2011; Posner, Oquendo, Gould, Stanley, & Davies, 2007; e.g., "How many times have you ever started to do something to end your life but someone or something stopped you before you actually did anything?"). Items were on a 7-point Likert scale (0 = none to 6 = more than twenty times). The items were combined with the suicide attempt item, with total possible scores ranging from 0 – 25. The scale had marginally acceptable internal consistency (α = .69).

On the 90th day, participants completed the Reactions to Research Participation Questionnaire (RRPQ; Newman, Willard, Sinclair, & Kaloupek, 2001), a 23-item questionnaire that assesses five domains of reactions to participation: participation (i.e., willingness and voluntary nature of study; four items; total possible score 4 – 20), personal benefits (i.e., participants gained something meaningful from participation; four items; total possible score 4 – 20), emotional reactions (i.e., participation was emotional; four items; total possible score 4 – 20), perceived drawbacks (i.e., participation costs; six items; total possible score 6 – 30), and global evaluation (i.e., was the research useful and ethical; five items; total possible score 5–25). Participants were instructed to "complete the following items about your research experience." Each item was rated on a 5-point scale (1 = strongly disagree to 5 = strongly agree). Higher scores on the participation, personal benefits, and global evaluation subscales indicated more positive experiences. Higher scores on the perceived drawbacks and emotional reactions indicate more negative experiences (i.e., more drawbacks and greater emotional reactions). The subscales have shown acceptable to good internal consistency in prior samples (DePrince & Chu, 2008), as well as in the present sample (αs = .73–.96).

Data Analytic Strategy

Demographics and study variable descriptives were computed and compared across subsamples (those who did not complete the 90th day survey and those who did). We ran a *t*-test to examine differences in number of daily assessments completed between individuals who completed the 90th day assessment and those who did not. To explore our first aim to identify percentage of individuals having negative, neutral, and positive study experiences, we collapsed RRPQ item responses of strongly disagree and disagree into a "disagree" category and strongly agree and agree into an "agree category" and left the "unsure" category as is. We then ran frequencies on each item. To test our second aim, we examined correlations of reactions to research participation with the baseline continuous variables of PTSD symptoms, depressive symptoms, hopelessness, suicidal ideation, suicidal behaviors, nonsuicidal self-injury, alcohol use, and drug use. We used *t*–tests to examine differences in research reactions across the dichotomous variables of gender, race (non-White compared

with White, due to low sample size within non-White race/ethnicities), and identification as religious (yes/no).

Results

Descriptive analyses regarding how many daily assessments participants completed revealed the following: Overall sample: $M = 35.64$ days ($SD = 28.50$), Mode = 3; 90th day completers: $M = 59.12$ days ($SD = 25.98$), Mode = 90; 90th day non-completers: $M = 20.73$ days ($SD = 19.28$), Mode = 3. The mean differences in days completed between 90th day completers and non-completers were significant $t(204) = -12.27$, $p < .001$. Participants who did not complete the 90th day reported greater levels of hopelessness (Cohen's $d = 0.45$) and suicidal ideation (Cohen's $d = 0.56$) at baseline than those who completed the 90th day survey (See Table 1). Most participants who completed the 90th day survey agreed on items regarding voluntary and willing participation (85–95%; See Table 2). A smaller, but still majority, percentage of participants agreed that there were personal benefits to participating (62–75.9%). Similarly, most participants did not perceive drawbacks to the study (62.5–88.8%). Results varied more across the emotional reactions subscale, with 46.8 to 53.2% of the sample disagreeing that they had an emotional reaction; 11.3 to 27.8% stating they were unsure, and 19 to 41.5% agreeing that they experienced an emotional reaction. Finally, most participants found the research to be useful and ethical (92.5–95%).

Regarding our second research aim, correlations revealed that hopelessness was negatively associated with the perceived benefits subscale ($r = -.22$ $p = .049$), suggesting that people with higher levels of hopelessness perceived less benefits of the study. Suicidal ideation ($r = .25$, $p = .03$) and suicidal behaviors ($r = .29$, $p = .01$) were positively associated with the emotional reactions subscale, suggesting that people with more severe suicidal ideation and more frequent suicidal behavior experienced greater emotional reactions to the study. Alcohol use and problems were positively associated with the perceived drawbacks subscale ($r = .26$, $p = .02$), suggesting that people with greater alcohol use and problems perceived greater drawbacks to the study. Depressive symptoms were negatively associated with the participation subscale ($r = -.27$, $p = .03$), suggesting that people with depressive symptoms rated the voluntary nature of the study less positively. Suicidal behaviors were positively associated with the global evaluation subscale ($r = .28$, $p = .012$), suggesting that people with a greater frequency of suicidal behavior had a more positive, overall evaluation of the study. Individuals who identified as religious experienced greater emotional reactions in the study ($M = 11.87$, $SD = 3.57$) compared with individuals who did not identify as religious ($M = 9.84$, $SD = 3.33$); $t(75) = -2.52$, $p = .014$. PTSD symptoms, nonsuicidal self-injury, and drug use were not associated with any of the reactions subscales. There were no gender or race differences in the subscales. In short, hopelessness, suicidal ideation, alcohol use, depression, and identification as religious were associated with at least one of the reaction variables, with each association suggesting potential for a more negative experience for individuals high in these variables. However, individuals with greater frequency of past suicidal behaviors were more likely to view the study as potentially useful.

We wanted to gain a clearer picture of the severity of negative reactions among people high in hopelessness, suicidal ideation, suicidal behavior, alcohol use, depression, and

identification as religious. Put simply, we aimed to describe the severity of negative reactions among people with high levels of these symptoms. Therefore, post-hoc regression analyses were run to examine the mean levels of the relevant reaction variables for individuals highest in hopelessness, suicidal ideation, suicidal behaviors, alcohol use, and depression, as well as people who identified as religious. Using standardized versions of the variables, we recoded each variable such that "0" equaled high levels of the respective variable (e.g., standardized depressive symptoms variable – 1 = high depressive symptoms; 0 on religion variable = practices a religion). We recoded these variables in order to examine the degree of negative reactions among people with the *highest levels* of the respective correlates of such negative reactions. The variable (e.g., depressive symptoms) was then entered as a predictor, and the respective reactions variable (e.g., participation) was entered as the criterion variable. The intercept was used to determine the average levels of the criterion variable for individuals high in the predictor variable. The intercept was then divided by the number of subscale items (e.g., four items for the participation scale) to interpret the mean relative to the 5-point scale of the RRPQ (1 = strongly disagree to 5 = strongly agree).

For individuals high in hopelessness, the average of the personal benefits variable was 3.44, close to the "unsure" anchor. For individuals high in suicidal ideation, suicidal behavior, and who identified as religious, the respective averages of the emotional reactions variable were 3.21, 3.01, and 2.96—each close to the "unsure" anchor. For individuals high in alcohol use and problems, the average of the perceived drawbacks variable was 2.25, close to the "disagree" anchor. For individuals high in depressive symptoms, the average of the participation variable was 3.95, close to the "agree" scale.

Discussion

The present study is an important addition to the literature regarding the risks and benefits of participation in studies employing repeated assessments of psychiatric symptoms. For individuals who completed the 90th, final day, reactions to participation generally fell within the neutral to positive range, mirroring past work (e.g., Biddle et al., 2013; Gibson et al., 2014; Law et al., 2015; Whitlock et al., 2013). However, participants with higher levels of suicidal ideation and hopelessness did not complete the 90th survey, suggesting that our participants' reactions may not generalize to more distressed samples. This finding warrants caution in assuming a low risk-benefit ratio of these studies.

Also similar to past work, individuals high in hopelessness, suicidal ideation, alcohol use, and depressive symptoms were more likely to rate participation less positively (Jacomb et al., 1999; Jorm et al., 2007; Hesse et al., in press; Whitlock et al., 2013). Our work also added a novel finding that individuals who identified as religious may rate such participation less positively. This may be due to greater mental health stigma among individuals who identify as religious (Eisenberg et al., 2009). However, exploration of these data suggested that participants were "unsure" at worst regarding risks and benefits. Specifically, individuals high in hopelessness were unsure about experiencing personal benefits from participation, and individuals high in suicidal ideation/behavior and who identified as religious were unsure about their experience of emotional reactions during participation.

Thus, although participants who completed the final survey may have uncertainty about the extent to which study participation provided personal benefits or generated negative emotion, other aspects of participation were generally positive. However, the results of these exploratory analyses may be bound to this sample. Additionally, individuals who identify as religious are more dubious to receiving mental health treatment due to their beliefs. Further, from a cognitive dissonance standpoint (Festinger, 1957), participants who completed the 90th day may be less likely to evaluate the study negatively such that their attitudes are consistent with their behavior. However, the study's financial compensation might have reduced this inclination.

A similar concern is that the participants who completed the 90th survey had less hopelessness and suicidal ideation at baseline compared with participants who did not complete the 90th survey. The effect sizes for these differences ranged from small to medium, suggesting that these differences are not trivial. Participants with greater hopelessness and suicidal ideation may have experienced repeated assessments more negatively and hence did not complete the daily assessment phase and 90th survey. Indeed, among those who completed the final survey, participants high in suicidal ideation were less certain about their emotional reactions during participation, and individuals high in hopelessness were uncertain about personal benefits of participation. To address this possibility of a higher risk-benefit ratio among individuals with higher suicidal ideation or hopelessness, researchers could address this explicitly. During informed consent, investigators can emphasize that individuals experiencing current suicidal ideation and hopelessness may find participation more distressing and less personally beneficial and emphasize that initiating or continuing participation is not mandatory. Additionally, the present study conducted informed consent electronically. Future studies involving repeated assessment of sensitive issues may benefit from the use of in-person informed consent to ensure participants fully understand the procedures, risks, benefits, and voluntary nature of the study.

Limitations of the present study included the small sample of undergraduate students with a history of alcohol use and suicidal thoughts and behaviors who completed the 90th day and thus the reactions to participation survey. Future work could improve on this limitation by attempting to contact participants who attrit to complete an exit survey on their reactions to participation and reasons for discontinuing. Future repeated assessment studies of psychiatric symptoms could also use qualitative methods to clarify participants' experiences in an in-person or phone debriefing session. Additionally, the design of the present study did not allow for determining whether it was repeated assessment, assessment of sensitive issues, or both, that impacted reactions to participation. Future work could randomly assign participants to conditions, such as by repeated or cross-sectional assessment and by neutral or psychiatric issues, to clarify the unique impacts of these design aspects. Finally, although this sample consisted of a particularly vulnerable population including individuals with histories of suicidal thoughts or behaviors, it may not generalize to samples with greater variability in psychiatric symptoms, age, race, gender, and other sociodemographic variables.

Acknowledgements

We extend our gratitude to the participants who shared their experiences with us for this study.

Please note that this work was supported, in part, by grant F31AA024685 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Thomas Fellowship. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding sources.

References

Biddle Lucy, Cooper Jayne, Amanda Owen-Smith Emily Klineberg, Bennewith Olive, Keith Hawton, Kapur Nav, Donovan Jenny, and Gunnell David. 2013 “Qualitative, interviewing with vulnerable populations: Individuals’ experiences of participating in suicide and self-harm based research.” Journal of Affective Disorders 145, 3: 356–362. doi: 10.1016/j.jad.2012.08.024 [PubMed: 23021191]

Blevins Christy A., Weathers Frank W., Davis Margaret T., Witte Tracy K., and Domino Jessica L.. 2015 “The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation.” Journal of Traumatic Stress 28 (6): 489–498. doi: 10.1002/jts.22059 [PubMed: 26606250]

Bolger Niall, Davis Angelina, and Rafaeli Eshkol. 2003 “Diary methods: Capturing life as it is lived.” Annual Review of Psychology 54(1): 579–616. doi: 10.1146/annurev.psych.54.101601.145030

de Beurs DP, Ghoncheh R, Geraedts AS, and Kerkhof AJFM. 2016 “Psychological distress because of asking about suicidal thoughts: A randomized controlled trial among students.” Archives of Suicide Research 20 (2): 153–159. doi: 10.1080/13811118.2015.1004475 [PubMed: 25751130]

Eisenberg D, Downs MF, Golberstein E, & Zivin K (2009). Stigma and help seeking for mental health among college students. Medical Care Research and Review, 66, 522–541. doi: 10.1177/1077558709335173 [PubMed: 19454625]

Festinger Leon. 1962 A Theory of Cognitive Dissonance. Palo Alto: Stanford University Press.

Gibson Susanne, Zoe VR Boden Outi Benson, and Brand Sarah L.. 2014 “The impact of participating in suicide research online.” Suicide and Life-Threatening Behavior 44 (4): 372–383. doi: 10.1111/sltb.12082 [PubMed: 24527848]

Gratz Kim L. “Measurement of deliberate self-harm: Preliminary data on the Deliberate Self-Harm Inventory.” 2001 Journal of Psychopathology and Behavioral Assessment 23(4): 253–263. doi: 10.1037/t04163-000

Harris Keith M., and Ting-Ting Goh Melissa. 2017 “Is suicide assessment harmful to participants? Findings from a randomized controlled trial.” International Journal of Mental Health Nursing 26(2): 181–190. doi: 10.1111/inm.12223 [PubMed: 27083774]

Hesse Calvin A., Strauss Catherine, Shorey Ryan C., Stuart Gregory L., and Cornelius Tara L.. 2018 “Examination of the transient changes in affect resulting from participation in research addressing intimate partner violence.” Journal of Interpersonal Violence. Advance online publication. doi: 10.1177/0886260518795172

Iltis Ana S., Misra Sahana, Dunn Laura B., Brown Gregory K., Campbell Amy, Earll Sarah A., Glowinski Anne, Hadley Whitney B., Pies Ronald, and DuBois James M.. 2013 “Addressing risks to advance mental health research.” JAMA Psychiatry 70 (12): 1363–1371. doi: 10.1001/jamapsychiatry.2013.2105 [PubMed: 24173618]

Jacomb Patricia A.., Jorm Anthony F., Rodgers Bryan, Korten Ailsa E., Henderson Scott, and Christensen Helen. 1999 “Emotional response of participants to a mental health survey.” Social Psychiatry and Psychiatric Epidemiology 34(2): 80–84. doi: 10.1007/s001270050115 [PubMed: 10189813]

Jorm Anthony F., Kelly Claire M., and Morgan Amy J.. 2007 “Participant distress in psychiatric research: A systematic review.” Psychological Medicine 37 (7): 917–926. doi: 10.1017/s0033291706009779 [PubMed: 17224097]

Lakeman Richard, and FitzGerald Mary. 2009 “The ethics of suicide research: The views of ethics committee members.” Crisis 30(1): 13–19. doi: 10.1027/0227-5910.30.1.13 [PubMed: 19261563]

Langhinrichsen-Rohling Jennifer, Arata Catalina, O'Brien Natalie, and Bower David. 2006 "Just How Upsetting Are Self-Report Surveys Anyway?." Violence and Victims 21(4): 425–444. doi: 10.1891/0886-6708.21.4.425 [PubMed: 16897911]

Law Mary Kate, Michael Furr R, Mayfield Arnold Elizabeth, Mneimne Malek, Jaquett Caroline, and Fleeson William. 2015 "Does assessing suicidality frequently and repeatedly cause harm? A randomized control study." Psychological Assessment 27(4): 1171–1181. doi: 10.1037/pas0000118 [PubMed: 25894705]

Lidz Charles W., and Garverich Suzanne. 2013 "What the ANPRM missed: additional needs for IRB reform." 41(2): 390–396. doi: 10.1111/jlme.12050

Luebbert Rebecca, Tait Raymond C., Chibnall John T., and Deshields Teresa L.. 2008 "IRB member judgments of decisional capacity, coercion, and risk in medical and psychiatric studies." Journal of Empirical Research on Human Research Ethics 3(1): 15–24. doi: 10.1525/jer.2008.3.1.15 [PubMed: 19385779]

Mathias Charles W., Michael Furr R, Sheftall Arielle H., Hill-Kapturczak Nathalie, Crum Paige, and Dougherty Donald M.. 2012 "What's the harm in asking about suicidal ideation?" Suicide and Life-Threatening Behavior 42(3): 341–351. doi: 10.1111/j.1943-278x.2012.0095.x [PubMed: 22548324]

Muehlenkamp Jennifer J., Swenson Lance P., Batejan Kristen L., and Jarvi Stephanie M.. 2015 "Emotional and behavioral effects of participating in an online study of nonsuicidal self- injury: An experimental analysis." Clinical Psychological Science 3(1): 26–37. doi: 10.1177/2167702614531579

Newman Elana, Willard Traci, Sinclair Robert, and Kaloupek Danny. 2001 "Empirically supported ethical research practice: The costs and benefits of research from the participants' view." Accountability in Research 8 (4): 309–329. doi: 10.1080/08989620108573983 [PubMed: 12481796]

Oldham John M., Haimowitz Stephan, and Delano Susan J.. 1999 "Protection of persons with mental disorders from research risk: a response to the report of the National Bioethics Advisory Commission." Archives of General Psychiatry 56 (8): 688–693. doi: 10.1001/archpsyc.56.8.688 [PubMed: 10435600]

Reynolds Sarah K., Lindenboim Noam, Katherine Anne Comtois Angela Murray, and Linehan Marsha M.. 2006 "Risky assessments: Participant suicidality and distress associated with research assessments in a treatment study of suicidal behavior." Suicide and Life- threatening behavior 36 (1): 19–34. doi: 10.1521/suli.2006.36.1.19 [PubMed: 16676622]

Rivlin Adrienne, Marzano Lisa, Hawton Keith, and Fazel Seena. 2012 "Impact on prisoners of participating in research interviews related to near-lethal suicide attempts." Journal of Affective Disorders 136 (1–2): 54–62. doi: 10.1016/j.jad.2011.09.009 [PubMed: 21975135]

Roth Alexis M., Felsher Marisa, Reed Megan, Goldshear Jesse L., Truong Quan, Garfein Richard S., and Simmons Janie. 2017 "Potential benefits of using ecological momentary assessment to study high-risk polydrug use." mHealth 3 (10): 46–56. doi: 10.21037/mhealth.2017.10.01 [PubMed: 29184898]

Roth Alexis M., Rossi John, Goldshear Jesse L., Truong Quan, Armenta Richard F., Lankenau Stephen E., Garfein Richard S., and Simmons Janie. 2017 "Potential risks of ecological momentary assessment among persons who inject drugs." Substance Use & Misuse 52 (7): 840–847. doi: 10.1080/10826084.2016.1264969 [PubMed: 28426353]

Shiffman Saul, Stone Arthur A., and Hufford Michael R.. 2008 "Ecological momentary assessment." Annual Review of Clinical Psychology 4: 1–32. doi: 10.1093/oxfordhb/9780199381708.013.1

Smith Phillip, Cukrowicz Kelly, and Poindexter Erin. 2010 "The effect of participating in suicide research: Does participating in a research protocol on suicide and psychiatric symptoms increase suicide ideation and attempts?" Suicide and Life-Threatening Behavior 40 (6): 535–543. doi: 10.1521/suli.2010.40.6.535 [PubMed: 21198322]

Stuart Gregory L., Moore Todd M., Ramsey Susan E., and Kahler Christopher W.. 2004 "Hazardous drinking and relationship violence perpetration and victimization in women arrested for domestic violence." Journal of Studies on Alcohol 65 (1): 46–53. doi: 10.15288/jsa.2004.65.46 [PubMed: 15000503]

United States Department of Health and Human Services. 2009 Protection of human subjects. 45 CFR §46. Retreived from: http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html.

Whitlock Janis, Pietrusza Celeste, and Purington Amanda. 2013 "Young adult respondent experiences of disclosing self-injury, suicide-related behavior, and psychological distress in a web-based survey." Archives of Suicide Research 17 (1): 20–32. doi: 10.1080/13811118.2013.748405 [PubMed: 23387400]

Yurek Leo A., Vasey Joseph, and Sullivan Havens Donna. 2008 "The use of self-generated identification codes in longitudinal research." Evaluation Review 32 (5): 435–452. doi: 10.1177/0193841×08316676 [PubMed: 18477737]

Table 1.

Descriptives and comparisons of subsamples

	Did not Complete 90th Day (*n* = 126)		Completed 90th day (*n* = 80)			Effect Size
	n	**%**	***n***	**%**	***p***	***Cramer's V***
Gender					.19	.095
Women	89	70.6	61	79.2		
Men	37	29.4	16	20.8		
Race/ethnicity					.45	.065
Non-Latino White	106	84.8	63	78.8		
Latino or Non-White	19	15.2	17	21.2		
Sexual Orientation					.42	.069
Heterosexual	110	87.3	65	82.3		
Gay, Lesbian, Bisexual, Other	16	12.7	14	17.7		
Usual residence					.25	.12
With friends/roommates	85	69.1	48	60.8		
With spouse/family	22	17.9	22	27.8		
Alone	16	13	9	11.4		
Academic Level					.06	.21
Freshman	87	69	55	69.6		
Sophomore	23	18.3	14	17.7		
Junior	11	8.7	1	1.3		
Senior	5	4	8	10.1		
Yearly Household Income					.95	.06
Less than $50,000	43	38.4	30	39.5		
$50,000-$100,000	31	27.7	21	27.6		
$100,000-$150,000	19	17	11	14.5		
$150,000-$200,000	6	5.4	6	7.9		
Greater than $200,000	13	11.6	8	10.5		
Practice a religion?					.46	.07
Yes	79	64.2	46	59		
No	44	35.8	32	41		
	M	***SD***	***M***	***SD***	***p***	***Cohen's d***
Age	19.05	2.39	19.21	3.40	.43	.10
Depressive Symptoms	29.20	10.64	26.42	9.92	.11	.27
Hopelessness	3.38	2.58	2.33	2.11	**.003**	**.45**
PTSD Symptoms	28.52	17.81	25.20	17.55	.19	.19
Alcohol Use and Problems	9.85	6.27	8.38	5.74	.10	.24
Drug Use and Problems	4.41	5.41	5.18	7.08	.39	.12
Nonsuicidal Self-Injury	1.79	1.96	1.70	1.99	.74	.05
Suicidal Ideation (past week)	1.30	2.06	0.37	1.09	**<.001**	**.56**
Suicidal Behavior (past year)	1.76	2.23	1.68	1.96	.81	.04

Note: Chi-Square and Cramer's V analyses conducted on dichotomous and nominal variables, respectively. Cramer's V used for effect sizes for both dichotomous and nominal variables. *T*-tests conducted on continuous variables.

Table 2.

Descriptives of Reactions to Research Participation Questionnaire

	M	*SD*	% Disagreed	% Unsure	% Agreed
Participation Subscale	17.13	3.42			
1. I like the idea I contributed to science	4.21	0.96	5	7.5	87.5
2. I was glad to be asked to participate	4.10	0.91	5	10	85
3. I felt I could stop testing at any time	4.38	0.93	6.3	1.3	92.5
4. Participation was a choice I freely made	4.44	0.93	5	0	95
Personal Benefits Subscale	15.01	3.70			
5. I gained insight about my experiences through research participation	4.44	0.93	8.9	29.1	62
6. I gained something positive from participating	3.80	1.06	11.4	19	69.6
7. I found participating beneficial to me	3.84	0.97	10.1	13.9	75.9
8. I found participating in this study personally meaningful	3.57	1.01	15.2	24.1	60.8
Emotional Reactions Subscale	10.91	3.66			
9. The research raised emotional issues for me that I had not expected	2.82	1.17	46.8	17.7	35.4
10. I experienced intense emotions during the research session	2.54	1.05	53.2	27.8	19
11. I was emotional during the research session	2.72	1.15	49.4	17.7	32.9
12. The research made me think about things I didn't want to think about	2.84	1.25	47.5	11.3	41.5
Perceived Drawbacks Subscale	12.45	3.55			
13. The study procedures took too long	2.21	0.97	71.3	17.5	11.3
14. Participating in this study was inconvenient for me	1.91	0.72	88.8	6.3	5
15. I found participating boring	2.43	1.02	62.5	17.5	20
16. I found the questions too personal	1.97	0.63	88.8	10	1.3
17. Knowing what I know now, I would participate in this study again if given the opportunity	3.95	1.01	8.8	15	76.3
18. Had I known in advance what participating would be like, I still would have agreed to participate	4.01	1.00	10	7.5	82.5
Global Evaluation Subscale	21.43	3.91			
19. I think the research is for a good cause	4.29	0.86	3.8	3.8	92.5
20. I believe this study's results will be useful to others	4.19	0.83	3.8	3.8	92.5
21. I was treated with respect and dignity	4.27	0.83	3.8	1.3	94.9
22. I trust that my replies will be kept private	4.31	0.87	5	0	95
23. I understood the consent form	4.36	0.82	3.8	2.5	93.8