Check for updates

A meta-analysis on the affect regulation function of real-time self-injurious thoughts and behaviours

Kevin S. Kuehn[®]¹[⊠], Jonas Dora¹, Melanie S. Harned[®]^{2,3}, Katherine T. Foster¹, Frank Song[®]¹, Michele R. Smith[®]¹ and Kevin M. King[®]¹

Prominent theories suggest that self-injurious thoughts and behaviours are negatively reinforced by decreased negative affect. The present meta-analysis quantifies effects from intensive longitudinal studies measuring negative affect and self-injurious thoughts and behaviours. We obtained data from 38 of the 79 studies (48%, 22 unique datasets) involving N = 1,644 participants (80% female, 75% white). Individual-participant data meta-analyses revealed changes in affect pre/post self-injurious thoughts and behaviours. In antecedent models, results supported increased negative affect before nonsuicidal self-injurious behaviour (k = 14, 95% Cl 0.09 to 0.31) and suicidal thoughts (k = 14, 95% Cl 0.03 to 0.19). For consequence models, negative affect was reduced following nonsuicidal self-injurious thoughts (k = 6, 95% Cl -0.79 to -0.44), nonsuicidal self-injurious behaviours (k = 14, 95% Cl -0.73 to -0.19) and suicidal thoughts (k = 13, 95% Cl -0.79 to -0.23). Findings, which were not moderated by sampling strategies or sample composition, support the affect regulation function of self-injurious thoughts and behaviours.

ith an age-standardized rate of 9.0 per 100,000 people, suicide was the 18th leading cause of death in 2019 worldwide¹. While many more individuals contemplate and/or attempt suicide, approximately 800,000 people die by suicide each year¹. In the United States, 4.8% of adults 18 years and older seriously considered suicide in 2018², while about 0.5% of US adults reported they attempted suicide². Nonsuicidal self-injury (NSSI), defined as deliberate damage to one's body tissue without the intention to die³, is a risk factor for future suicidal behaviour⁴. NSSI is suspected to increase one's tolerance for painful stimuli, removing barriers to attempting suicide. Worldwide, an estimated 17% of adolescents engage in NSSI⁵.

At the same time, researchers' ability to predict self-injurious thoughts and behaviours (SITBs) is poor. A recent meta-analysis of risk factors found that predictive models have not produced larger effect sizes over the past 50 years despite substantially more studies⁶. Critical innovation is needed to ensure that research on suicide-related phenomena is equipped to reliably predict who is at risk for death by suicide and when this risk is most imminent.

Leading theories propose that NSSI is maintained, at least in part, by an immediate reduction in negative affect^{7,8}. Although initially developed to explain the continuation of NSSI behaviours, emerging work suggests this affect regulation hypothesis may extend to other forms of SITBs, including suicidal cognitions9. Specifically, the affect regulation hypothesis proposes that negative affect is (1) increased before the occurrence of SITBs (the antecedent hypothesis) but (2) reduced following (the consequence hypothesis), which (3) increases the probability that someone experiences a SITB in the future in response to negative affect. Although the last component, that the relief from negative affect increases the probability of future SITBs, has not been tested empirically, treatments for a variety of psychological disorders target broadening and building emotion regulation skills¹⁰⁻¹³. Indeed, the development of effective emotion regulation strategies is hypothesized to be a mechanism of action accounting for reductions of SITBs in treatment¹⁴.

The affect regulation hypothesis is fundamentally a within-person process, positing that when people experience negative affect, SITBs then function to provide relief from distressing negative affect. However, the preponderance of evidence to date for this hypothesis has relied on retrospective reports in which people were asked to reflect about their affective states coinciding with their SITBs and the possible reasons they experienced them. For example, Nock and Prinstein⁷ asked 108 self-harming psychiatrically hospitalized youths how often they engaged in NSSI for 22 different reasons upon intake. Confirmatory factor analysis yielded a four-factor structure, with automatic negative reinforcement (that is, the affect regulation hypothesis) being the most frequently endorsed reason for NSSI (for example, self-harming to reduce emotional distress). Other aspects identified were automatic positive reinforcement (for example, engaging in NSSI to feel pain/a different emotion) as well as interpersonal negative and positive reinforcement. A recent meta-analysis confirmed this initial work as the affect regulation function was the most endorsed reason for NSSI across studies, with an estimated 63–78% of people reporting this function¹⁵.

However, research on retrospectively reported motives for behaviour does not always replicate when motivational processes are measured in the moment. For example, although affect regulation motives are frequently identified as a proximal predictor of alcohol use, ecological momentary assessment (EMA) research has failed to find consistent evidence that affect regulation plays an important role in drinking episodes when measured in real time¹⁶. In general, retrospective methods have the potential for inaccuracy due to recall bias17, which may be particularly pronounced for individuals already experiencing high levels of negative affect18. It is possible, for example, that individuals who experience elevated levels of negative affect are more likely to engage in SITBs. Additionally, retrospective methods do not allow for the establishment of clear temporal precedence between negative affect and SITBs, nor do they offer the ability to adequately test process theories, such as the affect regulation hypothesis, or model contextual factors that may alter these processes.

¹Department of Psychology, University of Washington, Seattle, WA, USA. ²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA. ³Veterans Affairs Puget Sound Health Care System, Seattle, WA, USA. ^{Sd}e-mail: kevinskuehn@gmail.com

ARTICLES

Missing the within-person nature of the affect regulation hypothesis could run the risk of committing Simpson's paradox¹⁹, an ecological fallacy in which between-person conclusions (for example, that those high in negative affect are more likely to experience SITBs) are expected to generalize to a within-person process (for example, if someone experiences elevated negative affect, they then are more likely to experience a SITB). Repeated observations over time enable one to tease apart between-person and within-person variance²⁰, and to more directly test process-based hypotheses by establishing temporal precedence. Thus, there are two types of research designs that are adequate in addressing the affect regulation hypothesis: experiments and intensive longitudinal methods.

A few notable experimental manipulations have provided evidence for the antecedent and consequence hypotheses in laboratory environments, with at least three studies reporting physiological and subjective changes pre and post NSSI proxies²¹⁻²³. For example, Reitz et al.²¹ experimentally incised participants' forearms and, for participants diagnosed with borderline personality disorder (BPD), detected relief from subjective distress following the incision. Similarly, Welch et al.²² used an imagery design in which participants diagnosed with BPD were provided personally tailored scripts describing their own accidental death, episode of NSSI or suicide attempt. Participants reported decreased negative affect and psychophysiological activity following NSSI and accidental death imagery. Finally, Franklin et al.23 reported that psychophysiological measures of negative affect, but not self-report measures, were reduced following an NSSI proxy. Although controlled laboratory studies were valuable for testing causal models of the affect regulation hypothesis in SITBs, the generalizability of their results to real-world occurrences across both individuals and time is unclear as they artificially induce single experiences that may not effectively model mechanisms triggering SITBs in daily life.

Intensive longitudinal methods can track SITBs in vivo alongside their proximal risks and consequences. These methods include ecological momentary assessment, in which participants are surveyed multiple times per day, as well as daily diary designs. The resulting data offer a rich opportunity to develop insights into what accounts for variance in SITBs over time (that is, within-person) and across people (that is, between-person). Although intensive longitudinal methods have been used to develop insights across many areas of clinical research, these methods are especially well suited to capture the specific phenomenology of SITBs^{24,25}, which appear to be short-lived and highly variable²⁶.

There has been a rapid increase in the use of intensive longitudinal methods to study the affect regulation hypothesis of SITBs. Despite several narrative reviews, a quantitative synthesis has been impossible due to significant heterogeneity in the way intensive longitudinal studies have been designed and executed, making it impossible to compute meta-analytic estimates of effect sizes from published reports. Studies have varied in terms of the number of surveys delivered per day, the amount of time in between surveys, the demographics and clinical presentations of the participants studied (for example, individuals diagnosed with BPD or those admitted to psychiatric inpatient units, etc.), the specific SITB variable tested, the proximal risk factors measured, the analytic strategy used to test hypotheses and the reporting practices of individual studies. Variations in designs, such as the frequency or time interval between observations, may influence the observed effect size²⁷, as the association between the experience of negative affect and the risk of SITB is likely to be strongest when both are measured close in time. Differences in analytic strategies makes an estimation of the cumulative effect from published articles impossible from published studies alone as coefficients reported in studies can reflect very different model assumptions (such as using a linear versus binary outcome, or a multilevel versus a structural equation model).

For example, Armey et al.²⁸ studied 36 college students with an ecological momentary design of six random samples per day over a 7 day period. Measuring negative affect on a 1-5 scale composed of nine items (for example, afraid, guilty and scornful), they²⁸ used growth curve modelling to test changes in negative affect in the five observations surrounding NSSI behaviours, comparing this trajectory in negative affect with five random observations of those who did not engage in NSSI behaviours. In line with the affect regulation hypothesis, the authors reported that negative affect increased before NSSI but decreased in the observations following, while negative affect exhibited no change when NSSI was not present. Bresin et al.²⁹ on the other hand, used a 14 day daily diary design with 67 college-aged participants and measured negative affect on a five-item 0-5 scale (for example, distressed, guilty and angry at self). Using multilevel models, they found that individuals were more likely to experience an NSSI urge on days they experienced higher levels of negative affect (odds ratio=8.00). In contrast, Kiekens et al.³⁰ used dynamic structural equation modelling with data (N=30) derived from a 12 day, eight time per day ecological momentary assessment design with negative affect measured on a six-item 0-6 scale (for example, stressed, irritated and anxious). The authors reported evidence that within-person changes in negative affect prospectively predicted NSSI thoughts and behaviour, however only NSSI thoughts, but not behaviour, remained significant after controlling for the occurrence of NSSI thoughts at the previous timepoint.

Although each study, at least partially, supports the affect regulation hypothesis, it is clearly impossible to synthesize these findings with a single pooled effect. Thus, the field is left with reports of single, isolated studies involving small numbers of participants and with limited ability to summarize results across studies. Without consistent coefficients or means and standard deviations of negative affect surrounding reports of SITBs, systematic reviews can only 'vote count' published effects, counting the proportion of significant to nonsignificant effects as evidence, or lack thereof, for the affect regulation hypothesis. Vote counting as a means of reviewing a body of literature has long been considered an inadequate method³¹ because statistical significance is only a function of sample size and the magnitude of the effect³². Null effects from underpowered studies (common in SITB research) cannot be taken as evidence for the absence of an effect, but merely the lack of evidence that an effect exists. Pooling effects across multiple studies improves the estimation of a common effect and increases power to detect even smaller effects than any single study³¹. Moreover, vote counting ignores variation in design across studies and makes it impossible to discern whether effects are stronger in certain designs or samples. Leveraging the full power of intensive longitudinal methods by pooling effects from individual studies to quantify the magnitude of the antecedent and consequence hypotheses is crucial to better understand the within-person process of affect regulation in SITBs.

The success of intensive longitudinal methods in SITB research hinges on the ability to catalogue converging evidence across diverging sampling and analytic methods. To date, six systematic reviews of intensive longitudinal studies on SITBs have been published³³⁻³⁸. Three of these reviews focused exclusively on NSSI³³⁻³⁵, and all used vote counting or qualitative summaries of existing studies. Specifically, Hamza and Willoughby³³ reviewed the 18 studies testing the association between emotion regulation and NSSI in either an experimental or intensive longitudinal design. The authors reported that the experimental studies reviewed suggest that negative affect is decreased following NSSI proxies and that, although the seven intensive longitudinal studies consistently offered support for the antecedent hypothesis of affect regulation, only two of these studies examined the consequence hypothesis. Of these two, there were mixed findings, with one reporting reduced negative affect and the other reporting increased positive, but not decreased,



Fig. 1 | CONSORT flow diagram. Flowchart from literature search to inclusion in quantitative synthesis.

negative affect. Rodríguez-Blanco et al³⁵. reviewed 23 studies examining NSSI in intensive longitudinal methods and reported that most of these studies focused on short-term affective changes in response to NSSI. In the studies reviewed, they³³ noted that affect regulation was the most commonly self-reported function of NSSI and highlighted the mixed findings of the consequence hypothesis, with one study detecting increases of negative affect following NSSI, some finding no change and a few reporting decreases.

Finally, Hepp et al.³⁴ used vote counting to summarize the literature on the four-function model in NSSI^{7,39}, of which the affect regulation hypothesis is one component. Their narrative summary concluded that there is evidence for the antecedent hypothesis, in that most studies reported increases in negative affect before NSSI but noted that studies reported both significant and nonsignificant effects regarding the consequence hypothesis. All three reviews suggest that negative affect is increased before NSSI but are inconclusive with regards to reduced negative affect post NSSI, and the magnitude of effect sizes for either hypothesis across studies is yet unknown.

The present study aims to solve methodological and substantive gaps in the current understanding of SITBs as they are experienced in daily life. We meta-analysed individual-participant data (IPD) from all available studies which measured negative affect and SITBs in intensive longitudinal data, including data from several unpublished studies, to calculate a standardized effect size estimate for both antecedent and consequence analyses for each of the various SITB outcomes. Using data provided by study authors, we calculated within-subject standardized coefficients. Specifically, we tested whether negative affect was elevated before SITBs (versus non-SITB occasions, the antecedent hypothesis), relative to an individual's average level of negative affect. We then examined whether negative affect was reduced following SITBs (as compared to the timepoint SITB was reported, the consequence hypothesis), relative to an individual's average level of affect.

Results

Description of the included studies. A total of 79 studies were initially selected for inclusion in the present review (Fig. 1 and Supplementary Table 1). These 79 studies included 5,888 participants. The average age was 28.21 years (weighted mean 31.71 years, median 28.02 years). Participants were primarily female (weighted mean 73.68%, median 76.70%) and white (weighted mean 70.60%, median 74.00%). Of the 79 studies, 22 (27.85%) measured NSSI thoughts, 28 (35.44%) measured NSSI behaviours, 34 (43.04%) measured suicidal thoughts and only 3 (3.80%) observed at least one instance of a suicidal behaviour.

In the 22 unique datasets included in the quantitative synthesis, which included 1,644 participants, the average age was 24.42 years (median 23.28 years). Participants were also primarily female (mean 80.00%, median 78.00%) and white (mean 74.73%, median 75.00%). The frequency with which an SITB was observed varied significantly across datasets and SITB outcomes (ranging from 0.05% for NSSI

Table 1 | Average effect sizes by SITB variable for antecedent and consequence models

	Antecedent models			
	k	β	SE	95% CI
NSSI thoughts	6	0.06	0.06	-0.07 to 0.19
NSSI behaviours	14	0.20	0.06	0.09 to 0.31
Suicidal thoughts	13	0.11	0.04	0.03 to 0.19
	Consequence models			
	k	β	SE	95% CI
NSSI thoughts	6	-0.63	0.09	–0.79 to –0.44
NSSI behaviours	14	-0.47	0.14	-0.73 to -0.19
Suicidal thoughts	13	-0.52	0.14	-0.79 to -0.23

behaviours in Kuehn et al. (unpublished manuscript) to 79.01% for suicidal thoughts in ref. ⁴⁰). Of the 22 studies, 6 (27%) used a daily diary design, while the rest applied an ecological momentary assessment design. Additionally, 8 of the 22 (36.36%) assessed current negative affect with an SITB measured since the last signal (that is, a retrospective effect), while 15 of the 22 studies queried negative affect and SITBs on the same timescale (that is, momentarily; one study assessed NSSI thoughts and NSSI behaviours on different time scales). Most of the studies (17/22) exclusively used a signal-contingent design, with a few using a combination of signal-and event-contingent prompts.

In the following sections, we report the results from the IPD meta-analyses in both antecedent and consequence analyses across the SITB variables measured. These 22 unique datasets corresponded to a little more than half of the studies initially eligible for inclusion. Characteristics of the studies included in the meta-analyses are presented in Supplementary Tables 2 and 3, and average effect sizes are presented in Table 1.

NSSI thoughts. Antecedent hypothesis. Six datasets measured negative affect and NSSI thoughts. Results are reported in Supplementary Fig. S1. One of these eight studies reported an increase in negative affect before thoughts of NSSI relative to when participants did not think about NSSI. The average effect size was $\beta = 0.06$ (k = 6, 95% CI -0.07 to 0.19), suggesting that, before NSSI thoughts, negative affect was estimated to be increased by 0.06 s.d. relative to moments not followed by NSSI thoughts. The range of the 95% credible interval indicates that the data were inconclusive with regards to the degree of negative affect experienced before NSSI thoughts. Small effects in either direction as well as a null result retained posterior plausibility.

Consequence hypothesis. These six datasets were again used to test for decreased negative affect following thoughts of NSSI relative to the timepoint in which NSSI thoughts were reported. Results are reported in Supplementary Fig. S1. All six datasets showed evidence of reduced negative affect with an overall average effect size of $\beta = -0.63$ (k = 6, 95% CI -0.79 to -0.44). This suggests that, following NSSI thoughts, participants' negative affect decreased by 0.63 s.d. relative to moments in which a participant continued to think about NSSI. The 95% credible interval did not include zero, suggesting effects are consistent with medium to large effects in the expected direction.

NSSI behaviours. Antecedent hypothesis. Fourteen datasets were included in the antecedent hypothesis of NSSI behaviours^{28–30,41–50}. Results from this test are reported in Fig. 2. Of the 14 studies,

7 (50.00%) detected a significant effect with an average effect size of $\beta = 0.20$ (k = 14, 95% CI 0.09–0.31). This indicates that affect was estimated to be increased by 0.20 s.d. before NSSI behaviours relative to moments not followed by self-harm. The 95% credible interval did not include zero, suggesting that the results are consistent with a small effect in the anticipated direction.

Consequence hypothesis. Of the 14 datasets, 10 (71.43%) detected evidence of decreased negative affect in the consequence hypothesis of NSSI behaviours. Results from these models are reported in Fig. 2. The effect size from the three-level model was $\beta = -0.47$ (k = 14, 95% CI -0.73 to -0.19), suggesting that negative affect was reduced 0.47 s.d. following NSSI behaviour relative to moments in which a participant continued to self-injure. The 95% credible interval did not include zero, suggesting effects are consistent with medium to large effects in the expected direction.

Suicidal thoughts. Antecedent hypothesis. Thirteen datasets measured negative affect and suicidal thoughts^{29,30,40,44,47,51-57}. Results from the antecedent hypothesis are reported in Fig. 3. Of the 13 datasets, 7 (53.85%) detected evidence for increased negative affect before suicidal thoughts. The average effect size across these datasets was $\beta = 0.11$ (k = 14, 95% CI 0.03–0.19), suggesting that negative affect was estimated to be increased by 0.11 s.d. before suicidal thoughts relative to moments not followed by suicidal thinking. The 95% credible interval did not include zero, suggesting that the results are consistent with a small effect in the anticipated direction.

Consequence hypothesis. In the consequence hypothesis (Fig. 3), all 13 datasets (100%) found that negative affect was reduced following suicidal thoughts. The average effect size was $\beta = -0.52$ (k = 13, 95% CI -0.79 to -0.23), suggesting that negative affect was reduced 0.52 s.d. following suicidal thoughts relative to moments in which a participant continued to think about suicide. The 95% credible interval did not include zero, suggesting effects are consistent with small to large effects in the expected direction.

Moderators. Results of moderation analyses are presented in Tables 2 and 3 and Supplementary Table 4.

Antecedent NSSI behaviour

In NSSI behaviour antecedent models, whether a study recruited participants diagnosed with BPD moderated the effect size (β = 0.29, s.e.m. 0.09, 95% CI 0.10–0.48). Studies that included participants diagnosed with BPD observed larger effects between negative affect and NSSI behaviour.

Consequence NSSI behaviour

In consequence analyses of NSSI behaviour, we found no evidence of moderation.

Antecedent suicidal thoughts

We did not detect any significant moderators in antecedent suicidal thought models.

Consequence suicidal thoughts

In consequence analyses of suicidal thoughts, we again found no evidence of moderation.

Sensitivity analyses. Due to the variability in study designs, we standardized raw data to calculate effect sizes. In treating all datasets the same, there were a few decisions we made that differed from the published articles. As these decisions may have inadvertently biased our results, we conducted some sensitivity analyses to compare findings under different scenarios.

First, in two datasets^{9,56} in which suicidal thoughts were measured continuously, there was not a clear way to dichotomize the variable to indicate the presence or absence of suicidal thoughts. These two datasets also had a relatively high number of nonzero suicidal thoughts. Therefore, in both datasets, we centred this

ARTICLES



Fig. 2 | Forest plots of NSSI behaviours in both antecedent and consequence models. Effect sizes from antecedent (top) and consequence (bottom) models of the negative affect, NSSI behaviour association (n = 897 participants nested in 14 studies). Data presented as random-effects estimates from a three-level model ±95% credible interval. Average effect size (fixed effect from three-level model and 95% credible interval) is shown at the bottom of the respective plots.

variable within person and considered suicidal thoughts to be present when an individual's observation was more than 1 s.d. greater than their own average level of suicidal thoughts. Results in which we loosened (from 1 s.d. to 0.5 s.d.) and tightened (1.5 s.d. and 2 s.d.) this operationalization are presented in Supplementary Figs. S2–S4. Average effect size in antecedent analyses increased up to 0.01 (average β ranged from 0.11 to 0.12) under different conditions, while consequence analyses ranged ±0.02 (average β ranged from -0.50 to -0.54) in various scenarios.

Second, one dataset^{46,58} combined NSSI thoughts and NSSI behaviours. We included this datapoint in both analyses but report effects with this study excluded in Supplementary Table 5. There are minimal differences in NSSI thought and behaviour models when excluding this study.

Third, one study⁵⁷ used a measure of depressive symptoms in the past 24 h instead of depressed affect. Since the items (depressed, hopeless, interested and worried) overlapped with other conceptualizations of negative affect, we included this datapoint in the pooled analysis. Results with this study excluded are presented in Supplementary Table 6, and whether this study is included does not affect any of our findings.

Fourth, NSSI thought models with urges removed are presented in Supplementary Table 7. Removing NSSI urges and analysing only studies that asked specifically about NSSI thoughts did not affect results.

Finally, as done in prior studies^{30,42}, we tested whether controlling for SITB at t - 1 influenced antecedent estimates and whether controlling for negative affect at t - 1 influenced consequence estimates. These results are reported in Supplementary Table 8. Controlling for dependent variables at t - 1 did not affect any of the main analyses.

Publication bias. We tested for publication bias by removing unpublished studies and reanalysing results based exclusively on published articles. These analyses are reported in Supplementary Table 9. Results were nearly identical when excluding unpublished articles.

Results summary. Overall, results from the meta-analyses suggest that negative affect is consistently increased before the occurrence



Affect reported following (versus together with) suicidal thoughts

Fig. 3 | **Forest plots of suicidal thoughts in both antecedent and consequence models.** Effect sizes in antecedent (top) and consequence (bottom) models of the association of negative affect with suicidal thoughts (n = 1,108 participants nested in 13 studies). Data presented as random-effects estimates from a three-level model \pm 95% credible interval. Average effect size (fixed effect from three-level model and 95% credible interval) is shown at the bottom of the respective plots.

of various SITBs (95% CI $\beta = 0.09-0.31$ for NSSI behaviour and 95% CI $\beta = 0.03-0.19$ for suicidal thoughts). There was also evidence for reduced negative affect following SITBs (95% CI $\beta = -0.79$ to -0.44 for NSSI thoughts, 95% CI $\beta = -0.73$ to -0.19 for NSSI behaviours and 95% CI $\beta = -0.79$ to -0.23 for suicidal thoughts). There did not appear to be publication bias in the literature, although most of the intensive longitudinal studies of SITBs published to date include small samples. Studies that included participants diagnosed with BPD detected stronger effects, but only in NSSI behaviour antecedent analyses. We did not find evidence of moderation for any of the other factors tested.

Discussion

The primary aim of this meta-analysis was to estimate the strength of the evidence for the affect regulation hypothesis of SITBs from studies using intensive longitudinal methods. Although the affect regulation hypothesis is the predominant theory regarding the maintenance of NSSI behaviours, the model had primarily been tested in single studies using relatively small sample sizes across a wide variety of methods, and systematic reviews had been unable to estimate pooled effect sizes, limiting our understanding of the magnitude and precision of these effects. Additionally, although there was preliminary evidence to suggest that affect regulation may similarly maintain suicidal thoughts^{9,59}, most of the evidence to date narrowly focused on NSSI behaviours. Results from this meta-analysis suggest that there is broad support for the affect regulation hypothesis in maintaining SITBs, such that negative affect is generally higher before NSSI behaviours and suicidal thoughts and exhibits moderate to large reductions following all SITBs.

Broadening prior systematic reviews, we included studies that measured both NSSI and suicidal forms of SITBs, in addition to analysing observation-level data across studies. In both antecedent and consequence analyses, effects were largely consistent across the range of SITBs. Pre-STIB affect was weakly to moderately elevated across SITBs, and there were only small differences in effect sizes across all forms of SITBs. This implies that negative affect alone is not likely to accurately discriminate NSSI thoughts or behaviours from suicidal thoughts or behaviours. Future research should

Table 2 | Moderation results in antecedent and consequence NSSI behaviour analyses

NSSL	behaviours	antecedent

Madauatau	
Winderator	

Sampling (1, EMA; 0,

daily diary)

Moderator			
	β (s.e.m.)	95% Cl lower bound	95% CI upper bound
No. of prompts	0.03 (0.01)	0.01	0.06
No. of hours	-0.01 (0.01)	-0.02	0.01
Frequency of SITB	-0.08 (0.49)	-1.02	0.89
Compliance	-0.34 (0.25)	-0.81	0.19
No. of women	0.34 (0.35)	-0.39	1.00
No. white	0.28 (0.29)	-0.34	0.82
Mean age	0.02 (0.01)	-0.01	0.04
BPD (1, yes; 0, no)	0.29 (0.09)	0.10	0.48
Sampling (1, EMA; 0, daily diary)	0.10 (0.13)	-0.17	0.35
Timing (1, concurrent; 0, lagged)	0.10 (0.12)	-0.13	0.33
NSSI behaviours conseq	uence		
	β (s.e.m.)	95% Cl lower bound	95% Cl upper bound
No. of prompts	0.03 (0.04)	-0.04	0.11
No. of hours	-0.00 (0.01)	-0.03	0.02
Frequency of SITB	-0.02 (0.49)	-0.98	0.94
Compliance	-0.48 (0.40)	-1.26	0.34
No. of women	0.44 (0.43)	-0.40	1.27
No. white	0.50 (0.48)	-0.46	1.39
Mean age	-0.00 (0.03)	-0.06	0.06
BPD (1, yes; 0, no)	0.15 (0.24)	-0.33	0.61

Timing (1, concurrent; -0.07 (0.23) -0.53 0.38 0, lagged) investigate between- and within-person differences which might better discriminate who is most likely to engage in NSSI and sui-

-0.03 (0.25)

0.43

-0.52

NATURE HUMAN BEHAVIOUR

Table 3 | Moderation results in antecedent and consequence suicidal thought analyses

Suicidal thoughts antecedent				
Moderator				
	β (s.e.m.)	95% CI lower bound	95% CI upper bound	
No. of prompts	0.02 (0.01)	-0.01	0.05	
No. of hours	-0.00 (0.01)	-0.02	0.01	
Frequency of SITB	-0.03 (0.14)	-0.31	0.25	
Compliance	0.04 (0.24)	-0.42	0.53	
No. of women	-0.02 (0.35)	-0.71	0.67	
No. white	0.37 (0.36)	-0.38	1.06	
Mean age	0.00 (0.01)	-0.01	0.01	
BPD (1, yes; 0, no)	0.01 (0.09)	-0.16	0.19	
Sampling (1, EMA; 0, daily diary)	0.11 (0.09)	-0.05	0.29	
Timing (1, concurrent; 0, lagged)	0.12 (0.16)	-0.19	0.46	
Suicidal thoughts consequence				
	β (s.e.m.)	95% CI lower bound	95% CI upper bound	
No. of prompts	-0.00 (0.05)	-0.11	0.11	
No. of hours	-0.01 (0.02)	-0.05	0.03	
Frequency of SITB	0.20 (0.31)	-0.41	0.79	
Compliance	-0.06 (0.35)	-0.73	0.62	
No. of women	0.41 (0.40)	-0.38	1.19	
No. white	-0.21 (0.41)	-1.01	0.59	
Mean age	0.01 (0.02)	-0.03	0.05	
BPD (1, yes; 0, no)	-0.08 (0.23)	-0.54	0.38	
Sampling (1, EMA; 0, daily diary)	-0.21 (0.22)	-0.63	0.23	
Timing (1, concurrent; 0, lagged)	-0.08 (0.32)	-0.70	0.54	

investigate between- and within-person differences which might better discriminate who is most likely to engage in NSSI and suicidal forms of SITBs and indicate when suicide risk is most acute. For example, within-person fluctuations in reflexive emotion regulation strategies⁶⁰ or momentary improvements in self-efficacy to avoid self-injurious behaviour^{26,61} may discriminate episodes of NSSI from suicidal forms of SITBs⁶².

Effect sizes were substantially larger in consequence models across all forms of SITBs. These findings suggest that relief from distressing negative affect is crucial in reinforcing SITBs and indicate that behavioural interventions should continue to teach replacement behaviours, such as emotion regulation skills, to obtain relief from negative affect in lieu of SITBs. Low-cost and scalable treatments, such as ecological momentary interventions⁶³, could be designed to increase an individual's ability to downregulate from increased negative affect immediately before an SITB and/or block the relief experienced from SITBs.

It is notable that there were large differences in magnitude between the antecedent and consequence models of negative affect and SITBs. Although the evidence provides clear support that SITBs function to reduce negative affect, the evidence that negative affect could be used to predict when SITBs would occur was only modest. This is critical for understanding the contexts and conditions in which SITBs are likely to occur and should dampen enthusiasm for the field's ability to predict and intervene on SITBs until we can collectively gain greater precision of prediction. Although it is vital to align sampling in longitudinal data collection with the time frame in which a hypothesized process is expected to occur²⁵, there is no systematic research that has examined when SITBs may occur following peaks in negative affect. Thus, it may be that antecedent effects, on average, were diluted by observing affect far in time from the occurrence of SITBs (although it is important to note that the estimated effect did not vary as a function of the number of EMA prompts). Future research should seek to understand the temporal dynamics of how and when peaks in negative affect lead to SITBs.

Intensive longitudinal designs offer a clear improvement over prior methods in better understanding SITBs, and there is a growing number of studies using this method. Despite this promise, there was considerable between-study heterogeneity. A variety of factors could explain this cross-study variation in results. Various definitions of negative affect, negative emotions and types of SITBs have been studied using differing statistical tests and methods. For example, study sample sizes ranged from 11 to 1,709 (mean 103.30), the number of prompts per day ranged from 1 to 12 per day (mean 4.98) and the duration of the study period ranged from 3 to 196 days (mean 19.37 days, median 14 days). Analyses also varied widely between projects, with some examining mean levels of general negative affect (broadly defined), others focusing on variance or

ARTICLES

activation and some looking at specific forms of negative affect. Even when similar SITB and negative affect variables were measured, the scale of the negative affect variable differed and standardized coefficients were not reported. To account for some of this heterogeneity and compare inferences across studies, we standardized the negative affect variables, but this meant that the average effect sizes derived from these meta-analyses included many different operationalizations and measures.

Although the methods used to test hypotheses were diverse, the participants enroled in these studies certainly were not. Young adult, white women were overrepresented across studies, suggesting that these findings may not adequately characterize the affect regulation hypothesis of SITBs in men, children, older adults or Black, Indigenous and people of colour. To ensure that findings from such studies apply to everyone at risk, research studies need to make every effort to recruit more diverse samples.

Beyond the between-study heterogeneity within this body of literature, there were a few more notable limitations to the present study. First, due to the various measurements across studies, we made decisions necessary to standardize analyses across studies. This meant that our methods inevitability had to diverge from a few published studies. Sensitivity analyses that consider how different decisions affected the results are included in the Supplementary information. Second, we were only able to obtain data from about half of the published articles. Every effort to obtain as much data as possible was made, but more data would certainly be preferable. Diverging research designs, analytic techniques and reporting practices meant we were unable to extract information necessary to calculate effect size from any of the published articles, relying on individual authors to provide raw data. Finally, our moderation analyses were exploratory and examined study- as opposed to individual-level characteristics. We tested any variable we could operationalize from the published articles, but it is likely that testing some of the moderators (for example, compliance and frequency of SITB) on the individual level would most probably lead to more precise estimates.

Although intensive longitudinal methods offer a promising solution for stagnation in suicide research, studies to date do not yet offer clear, well-converged inferences about SITBs. Instead, the close review presented here has highlighted critical opportunities for improving this work to catalyse advances in the field. By incorporating principles of open science⁶⁴ (that is, preregistering hypotheses and making data/analysis scripts publicly available), meta-analyses could be conducted more efficiently, possibly leading to more reliable findings to assist in the prevention of SITBs. Additionally, the field should value replication studies from intensive longitudinal designs that verify findings to increase confidence in the stability of our discoveries and, if replication is not supported, may assist in identifying potential moderators and mediators.

Furthermore, exploration of reliable effects will require larger periods of data collection to ensure that tests of effects are based on ample observations of SITBs. Fourteen days, the average duration of intensive longitudinal studies, is an especially narrow window to observe processes that, although prevalent at the between-person level, are of short duration and rare at the within-person level. Increasing the window of observation for SITBs will increase the number of observations, thereby increasing power to explain effects. The use of a longitudinal burst design or a multiple period of intensive longitudinal methods windows separated by months and years could also increase power while balancing participant response fatigue⁶⁵.

Finally, all statistical tests reviewed in the present report commonly estimate average effects across individuals, relying on the assumption that effects are homogeneous between people and over time. However, the degree of between-study heterogeneity observed in this review and recent findings suggesting that SITBs are probably a multiply determined complex process that vary widely across individuals^{26,66} suggest that any assumption of homogeneity in the link between negative affect and SITBs on the person level may be incorrect^{67,68}. It is possible that this variability may be so great both between and within individuals that links between SITBs and candidate risk factors could be nonergodic (that is, effects are not consistent across people or within the same person over time). Rather than using a nomothetic, group-level approach to render uniform inferences across diverse individuals, an idiographic approach^{67–70} may be especially beneficial for identifying consistency in relations between negative affect (or other antecedents/consequences) and SITBs across individuals and time.

Despite all these challenges, the present meta-analysis was able to extend the three prior systematic reviews by quantifying the association between negative affect and SITBs across studies. The prior reviews noted mixed evidence for some aspects of the affect regulation hypothesis. By pooling studies together, we were able to provide a much more refined understanding of the affective dynamics in the hours and minutes before and after SITBs. Results demonstrated a small effect for increased within-person affect before NSSI behaviours and suicidal thoughts (but not NSSI thoughts) and medium to large effects for decreased negative affect following all forms of SITBs.

Methods

Article search and study selection. We conducted a thorough search for intensive longitudinal studies measuring both negative affect and SITBs following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines⁷¹. The search was performed using PsycINFO, PubMed, Google Scholar and PsyArxiv with the following search terms: 'suicid* AND ecological momentary assessment', 'suicid* AND experience sampling', 'suicid* AND daily diary', 'suicid* AND ambulatory assessment'. For published articles, we included studies circulated before 15 January 2022. We also included preprints, studies that included a measure of negative affect in an intensive longitudinal study of at least one SITB but never published on the negative affect-SITB association and one dataset in which data collection was complete but the main manuscript was not yet published. Inclusion criteria for the current review were (1) an empirical study (for example, not literature reviews, theoretical articles, case studies, etc.) that measured (2) an SITB variable (that is, NSSI urges, thoughts, or behaviours and/or suicidal urges, thoughts or behaviours) in intensive longitudinal methods and (3) negative affect continuously in intensive longitudinal methods.

This initial search identified 168 unique possible studies for inclusion (163 published articles, 2 preprints, 2 unpublished dissertations and 1 unpublished dataset). Consistent with PRISMA guidelines, the titles, abstracts and keywords of these 168 studies and datasets were initially reviewed independently by two people (K.S.K. and F.S.) to assess whether they appeared to meet inclusion criteria. Thirty studies were removed in this initial search as they were determined to be nonempirical articles. The two raters reviewed 138 full-text articles and descriptions of datasets. Interrater agreement for the initial inclusion/exclusion determination was substantial (κ =0.86). The two raters met following their independent searches to resolve disagreements on nine of the articles. Of the original 168 studies, 89 were excluded based on initial inclusion/exclusion criteria, leaving 79 articles or datasets eligible for inclusion.

Compilation of data. Because of the variation in analytic procedures and reporting practices, none of the 79 studies reported the necessary information to calculate standardized effect sizes from the published articles alone. We attempted to contact the corresponding author of each study individually to provide us with the raw data. We obtained the raw data for 22 of the 57 unique datasets (N=1,644), which corresponded to 38 of the 79 articles (48.10%). We made three total attempts to contact corresponding authors before excluding studies. To obtain unpublished data, we also emailed the list-serv for the American Association of Suicidology and asked all contacted authors if they had any unpublished data relevant to the present study.

Participants in the studies for which we had data were significantly younger (t(df=51.78) = 2.56, P=0.01, Cohen's d=0.72, 95% CI t 1.19-9.77). There were no other differences in terms of publication year, sample size, duration of the observation period, number of prompts per day, percentage of female participants or percentage of white participants in the sample.

Meta-analyses. We conducted six separate IPD meta-analyses⁷². Antecedent models examined changes in pre-SITB negative affect, while consequence models detected differences in negative affect post SITB. Both antecedent and consequence models were run separately for NSSI thoughts, NSSI behaviours, and suicidal thoughts. As there were a small number of studies that measured NSSI urges (k=1), and suicidal urges (k=1), thoughts and urges were combined and labelled NSSI

NATURE HUMAN BEHAVIOUR

thoughts and suicidal thoughts. Only one study reported any instances of suicidal behaviour (that is, suicide attempts), limiting our ability to include this outcome.

We standardized negative affect in each study to account for measurement differences. We then aggregated individual studies into a combined dataset before separating within- and between-person variance of the standardized affect variable. To tease apart within- and between-person variance, we first averaged each participant's EMA responses across the study period to create person-level averages. We then centred each observation of negative affect within-person by subtracting the participant-level mean from each EMA observation, and then grand-mean-centred those participants' averages⁷³. Thus, by centreing level 1 variables within-person, any one observation reflects an individual's deviation from their own average across all timepoints.

For each of the six analyses, we used a three-level Bayesian multilevel model (observations nested within individuals nested within studies). Besides accounting for the nested structure, we also nested these participant-level intercepts within studies to account for differences in SITBs across studies. We further included random slopes to account for variability in the negative affect–SITB association between participants and studies. For all models, we lagged the data so that we could make comparisons in negative affect before and following incidents of SITB.

We used the following R syntax to calculate antecedent effects:

NA.standard.CWP.lag ~ 0 + Intercept + NSSI_thgts + (0 + Intercept | Study) + (0 + Intercept | Study:PID) + (0 + NSSI thgts | Study) + (0 + NSSI_thgts | Study:PID)

Although somewhat counterintuitive, we chose to use within-person negative affect at t - 1 as the dependent variable instead of SITBs. As we were concerned with the relative difference in affect between an SITB versus a non-SITB report, the coefficient from the above model provided a clean and interpretable effect size of interest because it estimated the differences in within-person negative affect at observations before observing SITB versus not. Conversely, a model with a dichotomous SITB as the dependent variable would produce a coefficient that, when exponentiated, represents a change in the odds ratio of an SITB for increasing levels of negative affect. We felt that a model with negative affect as the dependent variable allowed for effect sizes to be comparable between antecedent and consequence models and produce a fect sizes that were easy to interpret.

Consequence models were based on Kleiman et al.⁹ and reflect the relative difference in negative affect at timepoint t when an SITB was reported as compared with mean levels of within-person affect when an SITB was not reported at t + 1. As such, here, we compare a participant's reported level of negative affect together with the report of SITB versus the report of negative affect at the timepoint following the report of SITB. The following R script was used for analyses:

NA.standard.CWP ~ 0 + Intercept + NSSI_thgts + (0 + Intercept + NSSI_thgts | Study/PID/pair)

Priors. To specify prior information, in line with recommendations from Gelman et al.⁷⁴, we chose a weakly informative prior. The purpose of these priors is to regularize parameters while having minimal influence on the results, provided sufficient amounts of data are available. The following priors were used:

Intercept $\sim N(0, 1)$

B (SITB) $\sim N(0, 1)$

s.d. ~ Student t (3, 0, 2.5)

 $\sigma \sim$ Student *t* (3, 0, 2.5)

We used the 'brms' package 75 in the R statistical environment 76 to conduct all analyses.

Moderation analyses. We tested for moderation using the interaction between our hypothesized moderators and SITB variables in a three-level model. Moderation was tested in both antecedent and consequence models. We did not test for moderation with NSSI thought models because of the small number of studies. We examined sources of heterogeneity between studies and used the following variables as moderators (all moderators were tested as study-level characteristics):

Number of prompts per day. We extracted the number of intensive longitudinal prompts sent to participants per day (mean 5.04, s.d. 3.31, range 1–12).

Number of hours between prompts. The average amount of time (in hours) between study prompts was either stated in the publication or calculated by dividing the average duration of the observation period by the number of prompts (mean 10.43, s.d. 10.41, range 0.5–24).

Frequency of SITB. The frequency with which each SITB variable was observed over the course of the study was calculated by dividing the number of observations in which an SITB was endorsed by the total number of observations (NSSI thoughts: mean 12.44%, s.d. 7.37%, range 3.15–26.60%; NSSI behaviours: mean

3.53%, s.d. 2.85%, range 0.05-8.92%; suicidal thoughts: mean 22.82%, s.d. 27.28%, range 3.73-79.01%).

Compliance rate. The proportion of surveys completed by participants was either reported in the study or calculated from the raw data (mean 69.15%, s.d. 17.46%, range 36–100%).

Percentage of sample who identify as female. We extracted the percentage of the sample that identified as female from each study (mean 74.73%, s.d. 36.66%, range 38.30–91.00%).

Percentage of sample who identify as white. We derived the percentage of the sample reporting a white identity from each study (mean 80.00%, s.d. 11.07%, range 57.00–100%).

Mean age of the sample. We also extracted the mean age of the sample from each study (mean 24.42 years, s.d. 7.02 years, range 15.50–37.90 years).

BPD. We created a binary variable that indicated whether studies included participants diagnosed with BPD. Seven studies (31.82%) mentioned BPD as an inclusion criterion and/or reported enough diagnostic information to determine whether this was assessed.

Sampling strategy (EMA versus daily diary). We categorized studies as using an EMA (n=16, 72.72%) or daily diary (n=6, 27.27%) design.

Momentary versus retrospective prompts. We then examined the wording of the time frame assessed in prompts and categorized them as momentary (n = 15, 68.18%) or retrospective (n = 7, 31.82%).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Although this study was not preregistered, the raw data are publicly available at https://github.com/kskuehn/NA-SITB_meta.

Code availability

Analysis scripts are also publicly available at https://github.com/kskuehn/ NA-SITB_meta.

Received: 11 April 2021; Accepted: 17 March 2022; Published online: 28 April 2022

References

- Suicide rate estimates, age-standardized estimates by WHO region. Global Health Observatory Data Repository, World Health Organization https://apps. who.int/gho/data/node.main.MHSUICIDEASDR?lang=en (2021).
- Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: results from the 2019 national survey on drug use and health (HHS publication no. PEP20-07-01-001, NSDUH Series H-55). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration https://www. samhsa.gov/data/ (2020).
- 3. Nock, M. K. Self-injury. Annu. Rev. Clin. Psychol. 6, 339-363 (2010).
- Whitlock, J. et al. Nonsuicidal self-injury as a gateway to suicide in young adults. J. Adolesc. Health 52, 486–492 (2010).
- Gillies, D. et al. Prevalence and characteristics of self-harm in adolescents: meta-analyses of community-based studies 1990–2015. J. Am. Acad. Child Adolesc. Psychiatry 57, 733–741 (2018).
- Franklin, J. C. et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol. Bull.* 143, 187–232 (2017).
- Nock, M. K. & Prinstein, M. J. A functional approach to the assessment of self-mutilative behavior. J. Consult. Clin. Psychol. 72, 885–890 (2004).
- Chapman, A. L., Gratz, K. L. & Brown, M. Z. Solving the puzzle of deliberate self-harm: the experiential avoidance model. *Behav. Res. Ther.* 44, 371–394 (2006).
- Kleiman, E. M. et al. Are suicidal thoughts reinforcing? A preliminary real-time monitoring study on the potential affect regulation function of suicidal thinking. *J. Affect. Disord.* 232, 122–126 (2018).
- Berking, M. & Lukas, C. A. The affect regulation training (ART): a transdiagnostic approach to the prevention and treatment of mental disorders. *Curr. Opin. Psychol.* 3, 64–69 (2015).
- Levitt, J. T. & Cloitre, M. A clinician's guide to STAIR/MPE: treatment for PTSD related to childhood abuse. *Cogn. Behav. Pract.* 12, 40–52 (2005).
- Linehan, M. M. Cognitive-Behavioral Treatment of Borderline Personality Disorder (Guilford Press, 1993).

- Stasiewicz, P. R. et al. Affect regulation training (ART) for alcohol use disorders: development of a novel intervention for negative affect drinkers. J. Subst. Abus. Treat. 45, 433–443 (2013).
- Neacsiu, A. D., Bohus, M. & Linehan, M. M. in *Handbook of Emotion Regulation* (ed. Gross, J. J.) 491–507 (Guilford Press, 2014).
- Taylor, P. J. et al. A meta-analysis of the prevalence of different functions of non-suicidal self-injury. J. Affect. Disord. 227, 759–769 (2018).
- Littlefield, A. K., Talley, A. E. & Jackson, K. M. Coping motives, negative moods, and time-to-drink: exploring alternative analytic models of coping motives as a moderator of daily mood-drinking covariation. *Addict. Behav.* 37, 1371–1376 (2012).
- Bradburn, N. M., Rips, L. J. & Shevell, S. K. Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 236, 157–161 (1987).
- Eysenck, M. W. & Mogg, K. in *The Handbook of Emotion and Memory: Research and Theory* (ed. Christianson, S. A.) 429–450 (Psychology Press, 1992).
 Blyth, C. R. On Simpson's paradox and the sure-thing principle. *J. Am. Stat.*
- Assoc. **67**, 364–366 (1972). 20. Curran, P. J. & Bauer, D. J. The disaggregation of within-person and
- Curran, P. J. & Bauer, D. J. The disaggregation of within-person and between-person effects in longitudinal models of change. *Annu. Rev. Psychol.* 62, 583–619 (2011).
- Reitz, S. et al. Incision and stress regulation in borderline personality disorder: neurobiological mechanisms of self-injurious behaviour. Br. J. Psychiatry 207, 165–172 (2015).
- Welch, S. S., Linehan, M. M., Sylvers, P., Chittams, J. & Rizvi, S. L. Emotional responses to self-injury imagery among adults with borderline personality disorder. J. Consult. Clin. Psychol. 76, 45–51 (2008).
- Franklin, J. C. et al. The functions of nonsuicidal self-injury: support for cognitive–affective regulation and opponent processes from a novel psychophysiological paradigm. J. Abnorm. Psychol. 119, 850–862 (2008).
- Kleiman, E. M. & Nock, M. K. Real-time assessment of suicidal thoughts and behaviors. *Curr. Opin. Psychol.* 22, 33–37 (2018).
- Kiekens, G., Robinson, K., Tatnell, R. & Kirtley, O. J. Opening the black box of daily life in nonsuicidal self-injury research: with great opportunity comes great responsibility. *JMIR Ment. Health* 8, e30915 (2021).
- Kleiman, E. M. et al. Digital phenotyping of suicidal thoughts. *Depress.* Anxiety 35, 601–608 (2018).
- King, K. M. et al. Longitudinal modeling in developmental neuroimaging research: common challenges, and solutions from developmental psychology. *Dev. Cogn. Neurosci.* 33, 54–72 (2018).
- Armey, M. F., Crowther, J. H. & Miller, I. W. Changes in ecological momentary assessment reported affect associated with episodes of nonsuicidal self-injury. *Behav. Ther.* 42, 579–588 (2011).
- Bresin, K., Carter, D. L. & Gordon, K. H. The relationship between trait impulsivity, negative affective states, and urge for nonsuicidal self-injury: a daily diary study. *Psychiatry Res.* 205, 227–231 (2013).
- Kiekens, G. et al. Fluctuations in affective states and self-efficacy to resist non-suicidal self-injury as real-time predictors of non-suicidal self-injurious thoughts and behaviors. *Front. Psychiatry* 11, 214 (2020).
- Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. Introduction to Meta-Analysis (Wiley, 2021).
- Friedman, L. Why vote-count reviews don't count. *Biol. Psychiatry* 49, 161–162 (2001).
- Hamza, C. A. & Willoughby, T. Nonsuicidal self-injury and affect regulation: recent findings from experimental and ecological momentary assessment studies and future directions. J. Clin. Psychol. 71, 561–574 (2015).
- Hepp, J. et al. A systematic review of daily life studies on non-suicidal selfinjury based on the four-function model. *Clin. Psychol. Rev.* 82, 101888 (2020).
- Rodríguez-Blanco, L., Carballo, J. J. & Baca-Garcia, E. Use of ecological momentary assessment (EMA) in non-suicidal self-injury (NSSI): a systematic review. *Psychiatry Res.* 263, 212–219 (2018).
- 36. Sedano-Capdevila, A., Porras-Segovia, A., Bello, H. J., Baca-García, E. & Barrigon, M. L. Use of ecological momentary assessment to study suicidal thoughts and behavior: a systematic review. *Curr. Psychiatry Rep.* 23, 1–17 (2021).
- Gee, B. L., Han, J., Benassi, H. & Batterham, P. J. Suicidal thoughts, suicidal behaviours and self-harm in daily life: a systematic review of ecological momentary assessment studies. *Digital Health* 6, 2055207620963958 (2020).
- Rabasco, A. & Sheehan, K. The use of intensive longitudinal methods in research on suicidal thoughts and behaviors: a systematic review. *Arch. Suicide Res.* 1–15 (2021).
- 39. Nock, M. K. Why do people hurt themselves? New insights into the nature and functions of self-injury. *Curr. Dir. Psychol. Sci.* 18, 78-83 (2009).
- Peters, E. M. et al. Instability of suicidal ideation in patients hospitalized for depression: an exploratory study using smartphone ecological momentary assessment. *Arch. Suicide Res.* 26, 1–14 (2020).
- Vansteelandt, K. et al. The affect stabilization function of nonsuicidal self-injury in borderline personality disorder: an ecological momentary assessment study. *Behav. Res. Ther.* 92, 41–50 (2017).

- Houben, M. et al. The emotion regulation function of nonsuicidal self-injury: a momentary assessment study in inpatients with borderline personality disorder features. J. Abnorm. Psychol. 126, 89 (2017).
- Santangelo, P. S. et al. Ecological momentary assessment of affective and interpersonal instability in adolescent non-suicidal self-injury. J. Abnorm. Child. Psych. 45, 1429–1438 (2017).
- Wolford-Clevenger, C. A daily diary study of intrusive PTSD symptoms and suicidal ideation among transgender and gender diverse adults. *Psychol. Trauma-US* 13, 768–771 (2017).
- Muehlenkamp, J. J. et al. Emotional states preceding and following acts of non-suicidal self-injury in bulimia nervosa patients. *Behav. Res. Ther.* 47, 83–87 (2009).
- Hochard, K. D., Heym, N. & Townsend, E. The unidirectional relationship of nightmares on self-harmful thoughts and behaviors. *Dreaming* 25, 44–58 (2015).
- Czyz, E. K., King, C. A. & Nahum-Shani, I. Ecological assessment of daily suicidal thoughts and attempts among suicidal teens after psychiatric hospitalization: lessons about feasibility and acceptability. *Psychiat. Res.* 267, 566–574 (2018).
- Selby, E. A., Franklin, J., Carson-Wong, A. & Rizvi, S. L. Emotional cascades and self-injury: investigating instability of rumination and negative emotion. J. Clin. Psychol. 69, 1213–1227 (2013).
- Selby, E. A. et al. The dynamics of pain during nonsuicidal self-injury. *Clin. Psychol. Sci.* 7, 302–320 (2019).
- Lear, M. K., Wilkowski, B. M. & Pepper, C. M. A daily diary investigation of the defective self-model among college students with recent self-injury. *Behav. Ther.* 50, 1002–1012 (2019).
- Husky, M. et al. Predictors of daily life suicidal ideation in adults recently discharged after a serious suicide attempt: a pilot study. *Psychiat. Res.* 256, 79–84 (2017).
- Kaurin, A., Dombrovski, A.Y., Hallquit, M.N. & Wright, A.G. Momentary interpersonal processes of suicidal surges in borderline personality disorder. *Psychol. Med.* 1–11 (2020).
- Kaurin, A., Dombrovski, A., Hallquist, M. & Wright, A. G. Daily suicidal surge and attempted suicide in borderline personality disorder. Preprint at https://doi.org/10.31234/osf.io/hswea
- 54. Forkmann, T. et al. Assessing suicidality in real time: a psychometric evaluation of self-report items for the assessment of suicidal ideation and its proximal risk factors using ecological momentary assessments. J. Abnorm. Psych. 127, 758–769 (2018).
- Kleiman, E. M. et al. Examination of real-time fluctuations in suicidal ideation and its risk factors: results from two ecological momentary assessment studies. J. Abnorm. Psych. 126, 726–738 (2017).
- Kleiman, E. M. et al. Examination of real-time fluctuations in suicidal ideation and its risk factors: results from two ecological momentary assessment studies. J. Abnorm. Psychol. 126, 726–738 (2017).
- Salim, S., Robinson, M. & Flanders, C. E. Bisexual women's experiences of microaggressions and micro affirmations and their relation to mental health. *Psychol. Sex. Orientat. Gend. Divers.* 6, 336 (2019).
- Hochard, K. D., Ashcroft, S., Carroll, J., Heym, N. & Townsend, E. Exploring thematic nightmare content and associated self-harm risk. *Suicide Life Threat. Behav.* 49, 64–75 (2019).
- 59. Mou, D. et al. Negative affect is more strongly associated with suicidal thinking among suicidal patients with borderline personality disorder than those without. *J. Psychiatr. Res.* **104**, 198–201 (2018).
- King, K. M., Feil, M. C. & Halvorson, M. A. Negative urgency is correlated with the use of reflexive and disengagement emotion regulation strategies. *Clin. Psychol. Sci.* 6, 822–834 (2018).
- Czyz, E. K. et al. Self-efficacy to avoid suicidal action: factor structure and convergent validity among adults in substance use disorder treatment. *Suicide Life Threat. Behav.* 44, 698–709 (2014).
- Anestis, M. D., Pennings, S. M., Lavender, J. M., Tull, M. T. & Gratz, K. L. Low distress tolerance as an indirect risk factor for suicidal behavior: considering the explanatory role of non-suicidal self-injury. *Compr. Psychiatry* 54, 996–1002 (2013).
- 63. Versluis, A., Verkuil, B., Spinhoven, P., van der Ploeg, M. M. & Brosschot, J. F. Changing mental health and positive psychological well-being using ecological momentary interventions: a systematic review and meta-analysis. *J. Med. Internet Res.* 18, e152 (2016).
- Kirtley, O. J., Lafit, G., Achterhof, R., Hiekkaranta, A. P. & Myin-Germeys, I. Making the black box transparent: a template and tutorial for registration of studies using experience-sampling methods. *Adv. Methods Pract. Psychol. Sci.* 4, 2515245920924686 (2021).
- 65. Sliwinski, M. J. Measurement-burst designs for social health research. Soc. Personal. Psychol. Compass 2, 245–261 (2008).
- 66. King, C. A. et al. Five profiles of adolescents at elevated risk for suicide attempts: differences in mental health service use. J. Am. Acad. Child Adolesc. Psychiatry 59, 1058–1068 (2020).

ARTICLES

NATURE HUMAN BEHAVIOUR

- 67. Kaurin, A., Dombrovski, A. Y., Hallquist, M. N. & Wright, A. G. Integrating a functional view on suicide risk into idiographic statistical models. *Behav. Res. Ther.* **150**, 104012 (2021).
- Kuehn, K. S., Foster, K. T., Czyz, E. & King, C. A. Identifying person-specific coping responses to suicidal urges: a case series analysis and illustration of the idiographic method. Suicide Life Threat. Behav. https://doi.org/10.1111/ sltb.12841 (2022).
- 69. Wright, A. G. & Woods, W. C. Personalized models of psychopathology. Annu. Rev. Clin. Psychol. 16, 49-74 (2020).
- Foster, K. T. & Beltz, A. M. Advancing statistical analysis of ambulatory assessment data in the study of addictive behavior: a primer on three person-oriented techniques. *Addict. Behav.* 83, 25–34 (2018).
- 71. Moher, D. et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6, e1000097 (2009).
- Riley, R. D., Lambert, P. C. & Abo-Zaid, G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *Br. Med. J.* 340, c221-c227 (2010).
- 73. Enders, C. K. & Tofighi, D. Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychol. Methods* **12**, 121 (2007).
- 74. Gelman, A., Simpson, D. & Betancourt, M. The prior can often only be understood in the context of the likelihood. *Entropy* **19**, 555–567 (2017).
- 75. Bürkner, P. C. brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw. https://www.jstatsoft.org/v080/i01 (2017).
- 76. R Core Team: *R: a language and environment for statistical computing* (R Foundation for Statistical Computing, 2013).

Acknowledgements

This research was supported by a National Research Service Award from the National Institute of Mental Health (F31MH117827, PI K.S.K.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

The authors thank the many individuals who contributed to this manuscript by sharing their data. We are grateful for their generosity and hope this work assists in our collective efforts to reduce the global burden of self-injury.

Author contributions

K.S.K: project planning and conceptualization, data curation, literature search, data analysis, software, writing (drafting and editing) and final approval. J.D.: data analysis, software, writing (drafting and editing) and final approval. M.S.H.: writing (drafting and editing) and final approval. K.A.F.: writing (drafting and editing) and final approval. F.S.: literature search, writing (drafting and editing) and final approval. M.R.S.: validation, writing (editing) and final approval. K.M.K.: project planning and conceptualization, data analysis, methodology, supervision, writing (drafting and editing) and final approval.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-022-01340-8.

Correspondence and requests for materials should be addressed to Kevin S. Kuehn.

Peer review information *Nature Human Behaviour* thanks Glenn Kiekens, Raffaella Calati and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

nature portfolio

Corresponding author(s): Kevin Kuehn

Last updated by author(s): 22-03-15

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

Statistics

For	all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.
n/a	Confirmed
	\boxtimes The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
	The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.
	A description of all covariates tested
	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
	For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
	Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated
	Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.
So	ftware and code

Policy information about <u>availability of computer code</u> Data collection No software was used for data collection Data analysis R Statistical Environment, 'brms'

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Although this study was not preregistered, the raw data are publicly available (https://github.com/kskuehn/NA-SITB_meta).

Field-specific reporting

Life sciences

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Individual participant data (IPD) meta-analysis (quantitative) from intensive longitudinal studies of self-injurious thoughts and behaviors.
Research sample	We obtained the raw data for 22 of the 57 unique data sets (N = 1,644; 80% female; 75% white), which corresponded to 38 of the 79 articles (48.10%). We made three total attempts to contact corresponding authors prior to excluding studies. To obtain unpublished data, we also emailed the list-serv for the American Association of Suicidology and asked all contacted authors if they had any unpublished data relevant to the present study. Participants in the studies for which we had data were significantly younger t (df = 51.78) = 2.56 , p = .01; Cohen's d = 0.72 , 95% CI t = $1.19 - 9.77$. There were no other differences in terms of publication year, sample size, duration of the observation period, number of prompts per day, percentage of female participants, or the percentage of white participants in the sample.
Sampling strategy	Sampling strategies differed between studies. Most used a convenience sample of individuals at high-risk for suicide. We attempted to obtain all available data
Data collection	We conducted a thorough search for intensive longitudinal studies measuring both negative affect and SITBs following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines58. The search was performed using PsycINFO, PubMed, Google Scholar, and PsyArxiv with the following search terms "suicid* AND ecological momentary assessment", "suicid* AND experience sampling", "suicid* AND daily diary", "suicid* AND ambulatory assessment". For published articles, we included studies circulated before January 15, 2022. We also included pre-prints, studies that included a measure of negative affect in an intensive longitudinal study of at least one SITB but never published on the negative affect-SITB association, and one dataset in which data collection was complete but the main manuscript was not yet published. Inclusion criteria for the current review were: (a) an empirical study (e.g., not literature reviews, theoretical articles, case studies etc.) (b) measured a SITB variable (i.e., NSSI urges, thoughts, or behaviors and/or suicidal urges, thoughts, or behaviors) in intensive longitudinal methods. We obtained the raw data for 22 of the 57 unique data sets (N = 1,644; 80% female; 75% white), which corresponded to 38 of the 79 articles (48.10%
Timing	2009 - 2022. For published articles, we included studies circulated before January 15, 2022.
Data exclusions	This initial search identified 168 unique possible studies for inclusion (163 published articles, two pre-prints, two unpublished dissertations, and one unpublished dataset). Consistent with PRISMA guidelines, the titles, abstracts, and keywords of these 168 studies and datasets were initially reviewed independently by two people (the first and fourth authors) to assess whether they appeared to meet inclusion criteria. Thirty studies were removed in this initial search as they were determined to be non-empirical articles. The two raters reviewed 138 full-text articles and descriptions of datasets to determine whether they met inclusion criteria. Inter-rater agreement for the initial inclusion/exclusion determination was substantial ($\varkappa = .86$.). The two raters met following their independent searches to resolve disagreements on 9 of the articles. Of the original 168 studies, 89 were excluded based on initial inclusion/exclusion criteria leaving 79 articles or datasets eligible for inclusion.
	We followed the exact same procedures as the published articles for excluding cases with missing data (e.g., excluded if missing <30% of observations).
Non-participation	We did not obtain data for 35 unique data sets. This was either because corresponding authors declined to provide raw data or did not respond to our inquiries.
Randomization	Observational data. We controlled for auto-regressive effects, tested for moderators, and completed sensitivity analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
\boxtimes	Antibodies	\boxtimes	ChIP-seq
\boxtimes	Eukaryotic cell lines	\boxtimes	Flow cytometry
\ge	Palaeontology and archaeology	\boxtimes	MRI-based neuroimaging
\ge	Animals and other organisms		
	Human research participants		
\boxtimes	Clinical data		
\boxtimes	Dual use research of concern		

Human research participants

Policy information about studies involving human research participants				
Population characteristics	s (N = 1,644; 80% female; 75% white			
Recruitment	Individual studies handled recruitment of research participants and were all approved by ethics boards at various universities/research organizations.			
Ethics oversight	N/A - present study did not include data collection			

Note that full information on the approval of the study protocol must also be provided in the manuscript.

nature portfolio | reporting summary