AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Bulletin

ISSN: 0033-2909

2020, Vol. 146, No. 12, 1117–1145
http://dx.doi.org/10.1037/bul0000305

Interventions for Suicide and Self-Injury: A Meta-Analysis of Randomized Controlled Trials Across Nearly 50 Years of Research

Kathryn R. Fox
University of Denver

Xieying Huang
Florida State University

Eleonora M. Guzmán
Columbia University

Kensie M. Funsch
Florida State University

Christine B. Cha
Columbia University

Jessica D. Ribeiro and Joseph C. Franklin
Florida State University

Self-injurious thoughts and behaviors (SITBs) are major public health concerns impacting a wide range of individuals and communities. Despite major efforts to develop and refine treatments to reduce SITBs, the efficacy of SITB interventions remains unclear. To provide a comprehensive summary of SITB treatment efficacy, we conducted a meta-analysis of published randomized controlled trials (RCTs) that have attempted to reduce SITBs. A total of 591 published articles from 1,125 unique RCTs with 3,458 effect sizes from the past 50 years were included. The random-effects meta-analysis yielded surprising findings: The overall intervention effects were small across all SITB outcomes; despite a near-exponential increase in the number of RCTs across five decades, intervention efficacy has not improved; all SITB interventions produced similarly small effects, and no intervention appeared significantly and consistently stronger than others; the overall small intervention effects were largely maintained at follow-up assessments; efficacy was similar across age groups, though effects were slightly weaker for child/adolescent populations and few studies focused on older adults; and major sample and study characteristics (e.g., control group type, treatment target, sample size, intervention length) did not consistently moderate treatment efficacy. This meta-analysis suggests that fundamental changes are needed to facilitate progress in SITB intervention efficacy. In particular, powerful interventions target the necessary causes of pathology, but little is known about SITB causes (vs. SITB correlates and risk factors). The field would accordingly benefit from the prioritization of research that aims to identify and target common necessary causes of SITBs.

Public Significance Statement

This meta-analysis found that many interventions produce small reductions in SITBs and that these effects endure across a range of sample and study characteristics. No interventions consistently produce large or moderate SITB reductions, and intervention efficacy has not improved across 50 years of research.

Keywords: meta-analysis, randomized controlled trials, self-injury, suicide, treatment

Supplemental materials: http://dx.doi.org/10.1037/bul0000305.supp

This article was published Online First October 29, 2020.

Kathryn R. Fox, Department of Psychology, University of Denver; Xieying Huang, Department of Psychology, Florida State University; Eleonora M. Guzmán, Teachers College, Columbia University; Kensie M. Funsch, Department of Psychology, Florida State University; Christine B. Cha, Teachers College, Columbia University; Jessica D. Ribeiro and Joseph C. Franklin, Department of Psychology, Florida State University.

Kathryn R. Fox and Xieying Huang contributed equally and share first authorship. Joseph C. Franklin served as the senior author of this article. We thank researchers who responded to our data request for the current meta-analysis.

Data and codes can be accessed using the following link: https://osf.io/r45bn/?view_only=d0aff7aa15ea487dbd889e565b69f371. The data analyzed during the current study are available from the corresponding authors on reasonable request.

Correspondence concerning this article should be addressed to Xieying Huang, Department of Psychology, Florida State University, 1107 West Call Street, Tallahassee, FL 32306, or to Kathryn R. Fox, Department of Psychology, University of Denver, 2155 South Race Street, Denver, CO 80210. E-mail: huang@psy.fsu.edu or kathryn.fox@du.edu

Suicide has plagued humanity throughout recorded history. Dating to around 2,000 BCE, the Egyptian article *Dispute of a Man with His Ba* is believed by many to be the first recorded reference to suicide (Pritchard, 1950). According to the Bible, Judas hanged himself after betraying Jesus. After the fall of the Third Reich in 1945, there were reports of widespread suicide in Germany, including the mass suicide of nearly 1,000 people in Demmin (Goeschel, 2006). Suicide and related self-injurious thoughts and behaviors (SITBs) continue to plague humanity today. Approximately 1.5% of all worldwide deaths are suicide deaths (Naghavi, 2019; World Health Organization, 2014). Across 17 countries, Nock et al. (2008) found pooled lifetime prevalence rates of 9.2% for suicide ideation, 3.1% for suicide plans, and 2.7% for nonfatal suicide attempts. Nonsuicidal self-injury (NSSI; e.g., self-cutting in the absence of suicidal intent) is estimated to affect approximately 5.5% of adults and up to 17.2% of adolescents (Swannell, Martin, Page, Hasking, & St John, 2014). With a current world population near 7.7 billion, these rates mean hundreds of millions of people around the globe engage in SITBs. In addition to this high prevalence, SITBs have tremendous social, emotional, and economic costs, with estimates near $94 billion per year in the United States alone (Shepard, Gurewich, Lwin, Reed, & Silverman, 2016).

In recent decades, there have been major institutional and scientific efforts directed at alleviating this major public health crisis, but these efforts have not yet achieved many of their central objectives. In the hopes of facilitating these efforts, the present study aims to meta-analyze all published randomized controlled trials (RCTs) of interventions that have attempted to reduce SITBs. We expect that the results will provide direction on how to systematically and sustainably reduce SITBs and will highlight where to best focus future institutional and scientific SITB intervention efforts. Below, we outline the historical context for SITB intervention efforts and our major meta-analytic questions of interest.

Brief Historical Overview of SITB Intervention Efforts

Notable Institutional Efforts and Milestones

Although SITBs have been a major problem throughout history, organized SITB intervention efforts were rare before the 1950s, with most efforts occurring since 1990. The first recorded effort may have been Harry Warren's Save-A-Life League, which was founded in 1906 in New York City (Fitzpatrick, 1983; Miller & Gould, 2013). Warren, a Baptist minister, was deeply affected after encountering suicide attempt survivors and hearing the stories of suicide decedents from bereaved loved ones. He placed newspaper ads urging those considering suicide to meet with him. Warren—and later a team of volunteers—would support and counsel a group of nearly 100 suicide ideators who would visit them each week (Miller & Gould, 2013).

The next notable organized SITB intervention effort did not occur until the 1950s. In 1953, the Samaritans were formed in London by Chad Varah, a vicar in the Church of England, and quickly spread across the United Kingdom and into other countries (Bagley, 1968). The Samaritans operated the first telephone helpline for people who were considering suicide or in an active suicidal crisis. In 1958, the Los Angeles Suicide Prevention Center was founded by Edwin Shneidman, Norman Farberow, and Robert Litman. It began by offering a range of services (e.g., in-clinic interventions, mobile crisis teams) but soon transitioned primarily into a call center, similar to the Samaritans (Litman, Farberow, Shneidman, Heilig, & Kramer, 1965). Other centers and helplines soon spread throughout the United States. Over the next few decades, several major suicide prevention organizations were formed (see Spencer-Thomas & Jahn, 2012), including the International Association for Suicide Prevention (1960), the Center for Studies of Suicide Prevention at the National Institutes of Mental Health (1967), the American Association of Suicidology (1968), and the American Foundation for Suicide Prevention (1987).

Institutional suicide intervention efforts began to accelerate in the 1990s, particularly in the United States. In 1990, the U.S. Department of Health and Human services called for a national reduction in fatal and nonfatal suicide attempts by the year 2000. In 1996, the United Nations published its guidelines for the implementation of national suicide intervention strategies. A year later, the United States Congress passed a resolution recognizing suicide as a national problem. In that same year, the Centers for Disease Control established its Center for Injury Prevention and Control, which included a focus on SITBs. And in 1999, the U.S. Surgeon General published a call to action for suicide prevention, and the National Hopeline Network (1-800-SUICIDE) was established.

These efforts continued through the 2000s. The National Suicide Prevention Lifeline (1-800-273-TALK) was launched in 2001 and the U.S. Department of Health and Human Services released its *National Strategy for Suicide Prevention* that same year. In 2004, the Garrett Lee Smith Memorial Act was approved, establishing the first federal grant program for youth suicide intervention. Motivated in part by suicides among current and former military service members during this time period, in 2007 the Joshua Omvig Veterans Suicide Prevention Act was approved. It established the first comprehensive program aimed at reducing suicide among veterans. Soon after, in 2011, the Defense Suicide Prevention Office—housed within the United States Department of Defense—was founded. In 2012, the U.S. Department of Health and Human Services released a revision of its 2001 *National Strategy for Suicide Prevention*, and the World Health Organization's *Preventing Suicide: A Global Imperative* followed soon thereafter in 2013.

The aforementioned examples of institutional efforts demonstrate that suicide prevention and intervention has been a major area of focus in recent decades across major public and private institutions. These efforts have often been accompanied by aspirational calls for reductions in suicide rates over specific time periods. For example, in 1990 the U.S. Department of Health and Human Services put out the Healthy People 2000 report (U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 1990), which called for a drop in the national suicide rate from 11.7 to 10.5 per 100,000 by the year 2000; the Healthy People 2010 report published in 2000 (U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 2000) called for a drop from 10.5 to 4.8 by the year 2010; and the Healthy People 2020 report published in 2010 (U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 2010) called for a drop from 11.3 to 10.2 by 2020. Similarly, the World Health Organization (2013) called for a 10% reduction in the global suicide rate by 2020; in 2015, the American Foundation for Suicide Prevention (2015) and the National Institute of Mental Health called for a 20% reduction in suicide by 2025; and countries from Japan to Scotland to Bhutan to New Zealand have called for similar reductions over the coming decade (Dorji et al., 2017; Ministry of Health, 2019; The Scottish Government, 2018; World Health Organization, 2015). With few exceptions (e.g., Japan in recent years, the U.S. in the 1990s), these targets have not been met. When they have been met, it is often unclear why a particular decline occurred or how to sustain it. For example, it is unclear why the United States suicide rate declined in the late 1990s and why this rate has steadily increased since the year 2000.

Notable SITB Intervention Techniques

A wide range of intervention techniques have been applied to SITB prevention. Some of the more idiosyncratic efforts have included techniques such as presenting suicide attempt survivors with videotapes of loved ones' reactions to their attempt (Resnik, Davison, Schuyler, & Christopher, 1973), poetry (Alexander, 1990), movement therapy (Chapman, 1971), and exercise (Rosen & Thomas, 1984). Below, we provide a brief and nonexhaustive overview of some of the most common SITB intervention approaches since 1900.

Social outreach and large-scale crisis intervention (1906–today). As described above, this approach began with Henry Warren's Save-A-Life League in 1906 and extended to the Samaritans and the Los Angeles Suicide Prevention Center in the 1950s. It remains popular today in the form of national suicide helplines. The development of large-scale computer networks in the 1980s provided another avenue for this approach. In the first reported use of the Internet to prevent SITBs, Barnett (1982) used the PLATO system in an effort to resolve another user's suicidality. This approach has evolved into a much more sophisticated form today, with recent efforts employing machine learning techniques to identify at-risk users on popular social media platforms and to direct them to crisis services (e.g., Jaroszewski, Morris, & Nock, 2019).

Psychodynamic therapy (1910s–today). Psychodynamic therapy was the dominant form of psychotherapy for all psychopathologies, including SITBs, during the first half of the 20th century. Menninger (1938) and his book *Man Against Himself* strongly influenced the way that many therapists and laypeople conceptualized and treated SITBs. Owing to the rise of psychiatric drugs in the 1950s and cognitive–behavioral approaches in the 1970s, however, psychodynamic therapy is less common today. Nevertheless, many different forms of psychodynamic therapy are still applied to SITBs (Levy, Yeomans, & Diamond, 2007). Notably, treatment targets for several of these interventions resemble those of other psychotherapies, including psychological pain (Draper & Margolis, 1976) and hopelessness (Hendin, 1991).

Prefrontal lobotomy (1930s–1950s). A 36-year-old woman with a history of suicide attempts was among the first prefrontal lobotomy patients of Egas Moniz (Moniz, 1937), the Portuguese neurologist who invented the technique and later won a Nobel Prize for it. This approach continued into the 1940s and 1950s. In a discussion about when to use this technique for severely depressed patients, Kalinowsky and Scarff (1948) reasoned: "These patients suffer continuously and they also make life unbearable for their families. They are often suicidal and always complete invalids. Here, no hesitation [for the lobotomy procedure] is justified" (p. 84). In a report extolling the advantages of transorbital lobotomies compared with traditional lobotomies, Freeman (1953) said, "Suicide accounted for 9 deaths total in the series of 1,819 patients surviving operation. These represented obvious failures at relieving depression. It is worthy of note that no suicides have occurred in more than 1,000 patients treated by application of the deep frontal cut in transorbital lobotomy" (p. 488). Because of the dangers of lobotomy procedures, the severe sequalae of these procedures, and the introduction of psychiatric drugs, lobotomies became rare by the 1960s.

Electroconvulsive therapy and related approaches (1940s–today). Electroconvulsive therapy became a popular intervention in the 1940s, primarily for depression. By the 1950s, several studies had examined the effect of electroconvulsive therapy on SITBs (e.g., Huston & Locher, 1948; Levy & Southcombe, 1953; Ziskind, Somerfeld-Ziskind, & Ziskind, 1945). It remains a popular intervention today, and similar approaches involving deep brain stimulation and transcranial magnetic stimulation have also been applied to SITBs in recent years (e.g., Weintraub et al., 2013).

Gatekeeper training, peer support, and institutional programs (1950s–today). As part of the Los Angeles Suicide Prevention Center's efforts, gatekeepers and other community members were trained to understand, recognize, and mitigate suicide risk (Farberow, 1969). Similar strategies were employed by many later groups and institutions, with this general approach later extending to schools (e.g., Ross, 1980) and the military (e.g., Knox, Litts, Talcott, Feig, & Caine, 2003). Some have been doubtful of its efficacy (e.g., Kutcher, Wei, & Behzadi, 2017), but others suggest that this approach may hold promise (Calear et al., 2016).

Pharmacotherapy (1950s–today). Psychiatric drugs for schizophrenia emerged in the mid-1950s, with drugs for anxiety and depression following soon thereafter. The first RCTs using pharmacotherapy to prevent SITBs were published in the early 1970s (Kay, Fahy, & Garside, 1970; Prien, Klett, & Caffey, 1973). Pharmacotherapy for SITBs became increasingly popular throughout the 1980s, 1990s, and 2000s, with drugs ranging from selective serotonin reuptake inhibitors to antipsychotics to ketamine being tested for potential effects on several SITB outcomes. But early on in the history of pharmacotherapy, some became concerned by emerging evidence that a large proportion of suicide attempts involved overdoses of psychiatric medications (Brophy, 1967). And in the 2000s, controversy arose over mixed evidence that certain antidepressants may cause suicidality (see Friedman & Leon, 2007) and evidence that antidepressant effects on depression may be largely indistinguishable from placebo effects (Kirsch, 2008). Despite these controversies, pharmacotherapy continues to be among the most popular suicide intervention technique today.

Acute psychiatric hospitalization (1960s–today). Prior to the 1960s, a person could be involuntarily and indefinitely hospitalized in most jurisdictions if he or she had been diagnosed with any mental illness that required treatment. Beginning in the 1960s, the legal requirement for involuntary hospitalization shifted toward the *dangerousness principle* in many jurisdictions in the United States and around the world (Simpson, 1984; Testa & West, 2010). This principle states that involuntary hospitalization should happen only if there is evidence that the individual is an imminent danger to themselves or other people. Consequently, people deemed at risk of imminent suicidal behavior are required to be hospitalized in many jurisdictions. Around this same time, legislation in many jurisdictions limited involuntary admissions to 72 hr or less and required additional documentation of dangerousness to extend admissions to longer periods. Among the best known of these are California's 1967 Lanterman-Petris-Short Act (i.e., the 5150 hold) and Florida's 1971 Baker Act. As a result of these policies, thousands of people deemed "high-risk" are hospitalized each day, most for 72 hr or less. To many, this common approach seems eminently reasonable, as it provides oversight and protection for people who might otherwise try to kill themselves. But in recent years many have grown concerned by evidence that suicide rates are extremely high immediately after discharge from psychiatric facilities, and evidence that rates remain elevated for years thereafter (Chung et al., 2017). Some have even argued that hospitalization has a causal effect on SITBs (Large, Chung, Davidson, Weiser, & Ryan, 2017).

Checking-in programs (1960s–today). In 1969, Jerome Motto began a program that involved sending caring letters to high risk individuals recently discharged from a psychiatric facility. This RCT of more than 3,000 participants revealed a decreased suicide rate among participants in the caring letters group (Motto & Bostrom, 2001; see also Motto, 1976). These results inspired several similar studies involving caring postcards, texts, and phone calls. But across the 15 years of Motto and Bostrom's (2001) study, effects on suicide were only significant across the first two years ($p = .043$), with effects weakening across later years. Later studies have obtained similar effects (see Milner, Carter, Pirkis, Robinson, & Spittal, 2015).

Cognitive and behavioral approaches (1960s–today). Throughout the 1960s, Aaron T. Beck developed a cognitive theory of depression and an intervention approach based on this theory (see Beck, 1976). By the 1970s, this theory and intervention had been extended to suicide (Rush & Beck, 1978). Around this same time, similar approaches centered on problem-solving, behavior therapy, social skills, and distress tolerance were developed (e.g., dialectical behavior therapy (DBT); Linehan, 1987). In recent decades, these approaches have continued to evolve (e.g., CBT-SP, Brown et al., 2005) and are among the most popular interventions for SITBs.

Means safety and restriction (1970s–today). Given their potential to affect millions of people over short periods of time, means safety and restriction techniques may have the most potential for instantiating large-scale reductions in SITBs. Means restriction programs have been around since at least the 1970s. Based on an analysis of suicide occurring in hospitals, Farberow, Ganzler, Cutter, and Reynolds (1971) proposed measures to restrict the ability to jump from a high place or to hang oneself in a hospital. Danto (1971) argued for restriction of firearms to prevent both suicide and homicide. Drawing on a public health model, Browning (1974) similarly argued for the restriction of popular means of suicide, especially firearms. Based on trends in Australia, Oliver and Hetzel (1972) proposed a restriction on sedative medications. Kreitman (1976) famously found that the switch from coal gas to natural gas was associated with decreases in suicide across England, Wales, and Scotland throughout the 1960s. In another well-known example, Hawton et al. (2004) obtained evidence that fatal and nonfatal SITBs with paracetamol and salicylates declined in the United Kingdom after legislation was enacted to alter the packaging of these medications.

Although these methods are promising, they are often difficult to evaluate, leaving ambiguity about the effects of these techniques. For example, in an extended analysis of data on paracetamol and salicylates in the United Kingdom, Bateman (2009) found that earlier interpretations were confounded with SITB changes unrelated to the restriction of these medications. Bateman (2009) concluded that these restrictions had no effect on SITBs or other deaths related to paracetamol. There is also mixed evidence on the degree to which means safety and restriction may lead to means substitution on a population level (Daigle, 2005; Kreitman, 1976; Yip et al., 2012).

Multilevel eclectic approaches (2000s–today). In recent years, several groups have proposed and enacted multilevel eclectic approaches to SITB intervention (see Baker, Nicholas, Shand, Green, & Christensen, 2018 for a review). These approaches span multiple levels (e.g., government, media, community, primary care, schools, families, gatekeepers, mental health care professionals) and multiple SITB intervention techniques (e.g., means restriction, safety planning, evidence-based therapy, social support). These broad, resource-intensive efforts have been adopted by an increasing number of hospital systems.

The Present Meta-Analysis

Existing Summaries of Knowledge About SITB Intervention Efficacy

Given several decades of SITB intervention efforts, researchers have long sought to identify and build on the most efficacious approaches. So far, however, firm knowledge about the efficacy of SITB intervention techniques has remained elusive. In one of the earliest reviews of this literature, Hawton et al. (1998) found that "There remains considerable uncertainty about which forms of psychosocial and physical treatments of patients who harm themselves are most effective" (p. 441). In another major review, Mann et al. (2005) located 18 RCTs and several relevant non-RCT studies. They concluded that "physician education in depression recognition and treatments and restricting access to lethal means reduce suicide rates. Other interventions need more evidence of efficacy" (p. 2064). In a similar review that same year, Goldney (2005) noted that "because of the almost complete absence of randomized controlled trials demonstrating the effectiveness of specific treatments, there is sometimes a degree of pessimism about our ability to prevent suicidal behaviors" (p. 128).

More recent work and reviews have echoed these earlier sentiments. Bolton, Gunnell, and Turecki (2015) summarized their review of the literature by noting, "Although several drug based and psychotherapy based treatments exist, the best approaches to reducing the risk of suicide are still unclear" (p. 10). In a review of SITB interventions for youths, Glenn, Franklin, and Nock (2015) concluded that "although research on interventions for SITBs has increased over the past 10 years, there are currently no well-established treatments for suicidal or nonsuicidal SITBs in youth" (p. 26). In discussing an estimated age-standardized decline in global mortality attributable to suicide since 1990, Naghavi (2019) noted that this decline was nearly identical to the estimated age-standardized decline in mortality attributable to all other causes. This prompted Naghavi (2019) to question "whether the decline in suicide mortality is due to suicide prevention activities, or whether it reflects general improvements to population health" (p. 9).

The Utility of a Broad Meta-Analysis of SITB RCTs

The overarching aim of the present meta-analysis is to advance knowledge about the efficacy of SITB intervention techniques. To accomplish this, we will conduct a meta-analysis of all qualifying published studies with an RCT design and a SITB outcome. We will include published dissertations and published studies, because our aim is to summarize what the most rigorously evaluated and publicly available information on SITBs indicates. This is the same approach taken by many recent meta-analyses of interventions for SITBs and psychopathology more broadly (e.g., Calati & Courtet, 2016; Murray, Quintana, Loeb, Griffiths, & Le Grange, 2019; van Bronswijk, Moopen, Beijers, Ruhe, & Peeters, 2019). We will focus on RCTs because they represent the most rigorous tests of whether a given intervention causes reductions in SITBs. RCTs are not perfect, but they do eliminate far more alternative explanations for an intervention effect than uncontrolled, naturalistic, quasi-experimental, and nonrandomized controlled studies. As such, RCTs provide the best available estimations of intervention efficacy.

To our knowledge, there are no broad meta-analyses of SITB intervention efforts. Extant broad reviews of this literature are either qualitative reviews or systematic reviews, and most of these conclude that there are few RCTs on this topic (e.g., Goldney, 2005). Extant meta-analyses are relatively narrow in focus, centering on specific SITB intervention techniques, populations, SITB outcomes, or some specific combinations of each of these (e.g., DeCou, Comtois, & Landes, 2019; Leavey & Hawkins, 2017; Riblet, Shiner, Young-Xu, & Watts, 2017). They also tend to include designs other than RCTs. As described in the Method section below, the present meta-analysis spans RCTs across all SITB intervention techniques, populations, and outcomes. It includes nearly 600 studies and more than 3,000 effect sizes from these RCTs. This broad meta-analysis will permit us to address several broad questions about SITB interventions.

Major Questions of Interest

There are perhaps hundreds of interesting and consequential questions about SITB intervention techniques, their efficacy, and moderators of their efficacy. Because of space limitations, the present article will focus on addressing some of the biggest outstanding questions about this literature. For example, we will focus on questions such as *What is the general trend for SITB intervention efficacy over time?*, instead of more specific questions such as *Over time, have antidepressants become more efficacious than CBT at preventing suicide attempts among adolescents?* Below, we note the major questions that we will endeavor to address in the present article.

What Does the Published SITB RCT Literature Look Like?

We sought to examine several characteristics of the SITB RCT literature, including the following: (a) the number of relevant articles and effect sizes across time, (b) the number of effect sizes for each SITB outcome, and (c) the types of SITB interventions that have been evaluated in RCTs and the number of effect sizes for each intervention.

What Are the Major Characteristics of SITB RCTs?

There are many RCT design features that are important to consider when evaluating findings. We were primarily curious about the following: (a) control group characteristics, (b) sample size, (c) pretreatment equivalence of SITBs across treatment and control groups, and (d) attrition rates by group.

What Is the Overall Efficacy of SITB Interventions, and Is This Moderated by SITB Outcome Type?

The overall efficacy of SITB interventions is unclear because no broad meta-analyses of these interventions has been conducted. We hypothesize that active interventions will be moderately efficacious overall. It is possible that existing interventions are more proficient at preventing certain SITBs. For example, it may be that existing interventions are better able to prevent more severe SITBs (e.g., suicide death, attempt), but it may also be that existing interventions are better able to prevent less severe SITBs (e.g., NSSI suicide ideation). Still another possibility is that existing interventions are equally efficacious at preventing all (or most) types of SITBs. The results of these analyses are expected to provide important information about which outcomes are most difficult to prevent, and where to concentrate intervention improvement efforts.

Have Our Interventions Become More Efficacious Over Time—Why or Why Not?

Ideally, clinical science builds on prior advances to produce substantial progress over long periods of time. We hypothesize that the SITB intervention literature has been progressive, with effect sizes gradually accruing across decades of research. If this pattern is supported, we will explore potential drivers of this pattern. For example, it may be that cognitive and dialectical behavior therapies are responsible for improved intervention effect sizes over time. Such a finding would suggest that future interventions should concentrate on building from these approaches. However, it is also possible that these analyses will indicate that the SITB intervention literature has not been progressive. In this scenario, we would attempt to determine why there has been stagnation and how to stimulate progress.

Are Some Interventions Better Than Others?

A wide range of interventions have been applied to SITBs. They vary greatly in terms of targets, time spent in treatment, number of therapists and/or psychiatrists involved, and the amount of training required. We hypothesized that certain treatments, especially multifaceted treatments originally designed for individuals at high risk for SITBs (e.g., dialectical behavior therapy), would outperform others. Such results would provide clear direction for future research, dissemination, and implementation efforts. But it is also possible that results will indicate that no intervention or set of interventions is substantially more efficacious than others. Such findings may indicate the need for fundamental changes in how we approach SITB interventions.

How Does the Ultimate Target Outcome of the Intervention (i.e., SITBs Versus Non-SITBs) Impact Treatment Efficacy?

Many interventions primarily target outcomes other than SITBs (e.g., depression) but still measure SITBs as an outcome. Some have suggested that targeting psychopathology in general may be sufficient to reduce SITBs (e.g., Blumenthal & Kupfer, 1988), whereas others suggest that it may be necessary to directly target SITBs (e.g., Mann et al., 2005). The answer to this question remains unclear. One possibility is that interventions are similarly efficacious regardless of their primary target outcome. If results support this view, it would suggest that there is no apparent benefit to specifically targeting SITBs—targeting related outcomes and conditions (e.g., depression) may be sufficient to reduce SITBs. But if intervention efficacy is moderated by ultimate target outcome, it would indicate that there is a benefit to directly targeting SITBs and unique causes of SITBs.

Are SITB Interventions More Efficacious for Adults or Youths?

Older adult, adult, and developmental psychopathology have become increasingly separate fields over the past few decades. SITB interventions likewise often differ based on age. We aimed to broadly determine whether SITB intervention efforts have been more successful for older adults, adults, or youths.

Do SITB Intervention Effects Last Beyond the Immediate Treatment Period?

It is possible that SITB intervention effects are enduring, such that they last beyond the immediate treatment period (i.e., the time during which the intervention is actively applied). Yet it is also possible that effects weaken substantially once the intervention ceases. Knowledge about this pattern would help to determine whether interventions should be continually applied to SITBs or whether SITB interventions might be efficacious on a more limited basis.

How Do Sample and Study Characteristics Influence SITB Intervention Effects?

As with all RCTs, each intervention effect size in the present meta-analysis will be relative to the control intervention. We hypothesize that more active control interventions will be associated with *smaller* effect sizes. For example, we expect that studies that employed cognitive–behavioral therapy as a control intervention will, on average, produce *smaller* effect sizes than studies that employed waitlists as the control intervention. Such findings would suggest that nonspecific or overlapping aspects of interventions (e.g., weekly meetings with a therapist; general behavioral principles) account for a notable portion of intervention effects.

SITBs occur in many different populations—including the general population and a range of clinical populations. It is unclear whether SITB interventions are particularly efficacious (or inefficacious) for any specific populations. For example, it could be that SITB interventions are most efficacious for general population samples, less efficacious for clinical population samples, and least efficacious for samples with a history of SITBs. Another possibility is that SITB interventions are similarly efficacious across all sample types. The results of these analyses will help to determine—broadly—for whom interventions are most efficacious, and for whom we must focus on developing more efficacious interventions.

Intervention length may also influence SITB intervention efficacy. It is possible that longer interventions deliver higher dosage and therefore might be associated with stronger effects. It is also possible that longer interventions are necessitated by higher SITB severity and therefore might be associated with weaker effects. It is similarly possible the intervention length does not influence results. Likewise, it is unclear how sample size might influence effects. We hypothesize that larger sample sizes will be associated with more precise effect estimates, but not necessarily larger or smaller estimates.

We hope that addressing these broad questions will advance knowledge about SITB intervention techniques and facilitate institutional and scientific efforts aimed at resolving this enduring public health crisis.

Method

Given the broad aims of the present meta-analysis, we conducted a comprehensive literature search to identify randomized control trials (RCTs) published in print or online before January 1, 2018 in PubMed, PsycINFO, and Google Scholar. We also identified published articles through ClinicalTrials.gov. Searches included permutations of the following terms: *treatment*, *intervention*, *therapy*, *suicide*, *self-injury*, *self-directed violence*, *self-harm*, *self-mutilation*, *self-cutting*, *self-burning*, and *self-poisoning*. To help ensure that we did not miss any relevant studies, we also searched the reference sections of reviews and meta-analyses that emerged from these search terms. As shown in Figure 1, these searches ultimately yielded 591 qualifying articles from 1,125 separate RCTs. See Supplement 1 in the online supplemental materials for a list of included studies. Of note, because the current study only involved the collection and study of existing data that are deidentified and publicly available, it did not require approval from the Institutional Review Board.

Inclusion and Exclusion Criteria

Inclusion criteria required random assignment to a treatment or control condition (including nontreatment, waitlist, and active controls) and assessment of some form of SITB post treatment. Given that we were primarily interested in the effects of intervention on the occurrence, frequency, and severity of SITBs, we excluded studies that only examined outcomes relevant to SITBs (e.g., attitudes toward SITBs, confidence in helping or treating individuals with SITBs). We excluded stud-

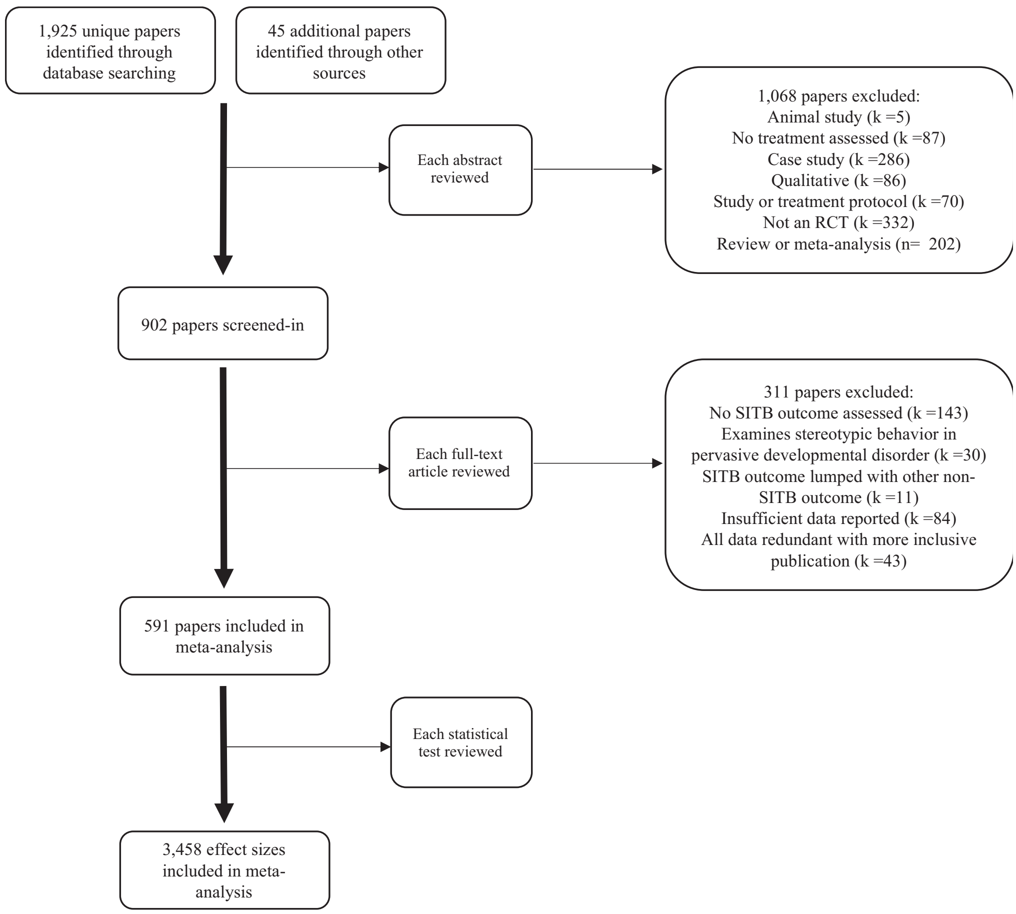

Figure 1. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow chart. RCT = randomized controlled trials; SITB = self-injurious thoughts and behaviors.

ies that were not written in English, studies that did not assess SITB(s) posttreatment, and studies that assessed composite outcomes including but not limited to SITBs posttreatment (e.g., risky behaviors, involving alcohol use and NSSI, death attributable to suicide and unintentional overdose). Studies that adopted an RCT design but did not assess the effects of interventions were also excluded. In cases in which studies did not include necessary statistical information (e.g., studies that listed only SITB engagement regardless of treatment group, studies that provided mean SITB engagement but without standard deviations or standard errors), we contacted the corresponding and/or first authors to obtain necessary data. In cases in which study authors no longer had access to those data and/or we did not hear back by August 1, 2020, studies (or specific treatment effects that lacked necessary information) were excluded.

The present meta-analytic strategy assumes that all effect sizes are independent. However, the majority of included studies reported on multiple forms of SITBs posttreatment (i.e., treatment effects) and/or assessed SITB treatment outcomes across multiple time points. It is likely that many of these outcomes correlate with each other, resulting in moderate dependence in included effect sizes. Simulation studies suggest that ignoring this type of dependence can result in slight underestimation of the effect sizes (e.g., Thompson & Becker, 2014). To reduce this possibility, we excluded redundant effect sizes for the same SITB treatment outcome published across multiple studies. Additionally, to obtain an idea of the impact of the upper limit of this dependence, we repeated all primary analyses assuming *complete* dependence across treatment effects within studies. Specifically, all effect sizes were averaged within each study (Scammacca, Roberts, & Stuebing, 2014). Results were consistent (see Supplement 2 in the online supplemental materials).

Data Extraction and Coding

Each treatment effect meeting our inclusion/exclusion criteria was coded on all dimensions described below. When a study reported multiple RCTs, statistics were extracted from each RCT whenever possible. If this was not possible, the aggregated statistics were extracted. Codes were developed by the lead authors in consultation with coauthors and based on several factors related to RCT methodologic quality (e.g., Schleider, Abel, & Weisz, 2015; Weisz et al., 2017). Studies were coded by authors and trained research assistants; 20% of articles were checked by KRF, and discrepancies were discussed and resolved. All codes were then checked again by XH and KMF, and discrepancies were discussed and resolved with KRF.

Author, year, and era codes. Article authors and year of publication were recorded for each treatment effect. Era of research was coded both based on the actual year of publication and via 10-year intervals starting with 1970s.

SITB treatment effects. Studies assessed a wide range of SITB outcomes. We coded treatment effects reflecting: (a) NSSI (i.e., intentional self-harm enacted without suicidal intent)[1]; (b) self-harm (i.e., intentional self-harm where suicidal intent was not assessed or required); (c) suicide ideation/plans (any form of suicidal thought and/or plan); (d) suicide attempt (i.e., any intentional self-harm with nonzero intent to die); (e) suicide death (i.e., any intentional self-harm resulting in death); (f) hospital visits and hospitalizations due to SITBs; and (g) other/combined SITBs. Suicide plans were lumped with suicidal thoughts because there were only a small number of effect sizes examining suicide plans. Outcomes that could not be placed into the above categories but involved some form of SITBs were classified as *other/combined SITBs* (e.g., suicidal gestures, outcomes combining both suicide attempt and death). Finally, as stated above, outcomes that both included a SITB and a non-SITB related thought or behavior (e.g., unintentional overdose, risky sexual behavior) were excluded.

Assessment time points. We extracted the following assessment time points: (a) pretreatment (i.e., baseline), (b) posttreatment, (c) follow-ups. Given that the primary interest of this study was to examine the treatment effects on SITBs after treatment completion, we reported analytic results from post treatment time points in the main text, and reported results from other time points in Supplement 3 in the online supplemental materials.

Sample severity. Sample severity was coded based on the study inclusion criteria. In studies in which participants were recruited from schools or the community and no psychopathology or SITB histories were required, samples were labeled as *general*. In studies in which participants were recruited based on exhibiting some form of psychopathology, samples were labeled *clinical*. In studies in which participants were recruited based on having a history of some form of SITB, samples were labeled *self-injurious*.

Sample age. Sample age was coded using the average age and/or range of ages included in a study, depending on what was reported. Sample age was labeled as: (a) child/adolescent if the study included only participants who were under 18 years of age at the start of the study; (b) adolescent/adult if the study included both participants under and over 18 at the start of the study; (c) adult if the study included participants over 18 years of age, but under 65 years of age, at the start of the study; (d) older adult if the study only included participants at least 65 years of age at the start of the study; (e) adult/older adult if the study included participants both at least 65 years of age and participants between 18 and 65 years of age; and (f) all age groups when participants were included regardless of age. We also coded the mean (or median, depending on if the mean were available) age of participants in the study.

Intervention target type. Treatments were coded based on their primary treatment target: psychopathology treatment, SITB treatment, or other (e.g., reducing physical illnesses).

Specific intervention type. Given the popularity of certain treatments for reducing SITBs, we also coded treatments based on the specific type of treatments administered, when possible. These included Dialectical Behavior Therapy (DBT), Cognitive Therapy/Cognitive Behavioral Therapy (CT/CBT), eclectic psychotherapy (i.e., therapies using a range of modalities), psychiatric medication, concurrent therapy and psychiatric medication, checking-in programs, safety planning/means restriction, and inpatient hospitalization, and so forth.

Intervention characteristics. Treatments were also coded based on their length (i.e., weeks of treatment) and treatment components (e.g., individual, group, family).

Control treatment characteristics. Control treatments were coded as (a) no treatment/waitlist, (b) placebo, and (c) active treatment.

Study quality. As recommended by the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Green, 2011), we utilized the Quality Assessment Tool for Quantitative Studies (Effective Public Health Practice Project, 2007) to categorize study quality into *weak*, *moderate*, or *strong*. The final categorization is based on seven main study aspects: selection bias, study design, confounders, blinding, data collection methods, and withdrawals. KMF completed the ratings based on the protocol, with 20% of the studies also independently coded by XH. Interrater agreement (κ = 0.72) was moderate and consistent with previous studies using the same tool (e.g., Armijo-Olivo et al., 2014; Hartling et al., 2009). In addition, we also coded several factors related to RCT methodological quality as recommended by prior research (e.g., Schleider et al., 2015; Weisz et al., 2017). These included the number of participants randomized to each condition, the number of participants in each condition *at the start* and *at the end* of treatment, therapist training, therapist supervision, and blinding status.

Study randomization and design. Randomization procedures were coded as *individual* if randomization occurred on an individual participant level. They were coded as *cluster* for studies in which randomization occurred on a higher level such as households, providers, schools, hospitals, and catchment areas. Each study was also coded based on whether they adopted a parallel, crossover, or dynamic waitlist design.

Meta-Analytic Methods

We conducted the meta-analysis using *R* (Version 3.6.2; R Core Team, 2019) with *metafor* package (Version 2.4–0; Viechtbauer, 2020). We first conducted a pooled meta-analysis examining treatment effects on *all* SITB outcomes. We then repeated the analyses systematically for each specific SITB outcome (e.g., suicide ideation, suicide attempt). Random Effects models were adopted to account for heterogeneity across studies, and I^2 tests were used to examine between-study heterogeneity. Of note, results were statistically equivalent when Fixed Effects models were utilized (see Supplement 4 in the online supplemental materials). To assess for publication bias, the following indices were calculated: Classic Fail-Safe *N*, Orwin's Fail-Safe *N*, Begg and Mazumdar Rank Correlation Test, Egger's Regression Test, funnel plot symmetry, and Duval and Tweedie's Trim and Fill Test.

[1] In cases where authors clearly measured self-harm enacted without suicidal intent, these behaviors were categorized as NSSI regardless of the term authors used.

Studies reported intervention effects in various ways, including the presence or absence of SITB engagement, the frequency of SITB engagement, and scores on items or scales assessing for SITBs. In pooling effect sizes across studies, it is important that effect sizes share the same scale and meaning (Higgins & Green, 2011). In this case, the presence and absence of SITB engagement often does not have the same meaning as the frequency of SITB engagement or means and standard deviations from scales. The former typically captures the number of occurrences of SITBs, whereas the latter two typically capture severity or intensity of SITBs. As such, they should not be combined in the same analysis. Hedges's *g* effect sizes were used to examine continuous treatment effects (e.g., means and standard deviations of the frequency of SITB engagement or scores on SITB scales). This effect is the standardized mean difference (Cohen's *d*) with a correction factor (*J*) for bias associated with small samples (Hedges, 1981). Therefore, Hedges's *g* is considered a robust measure of effect. These effect sizes may then be conservatively interpreted using the following conventions: small (0.2), medium (0.5), and large (0.8) effects (Cohen, 1988).

For binary outcomes (i.e., presence or absence of SITB engagement), risk ratios (RRs) were used to summarize effect sizes. Given the overall low prevalence rates of SITB outcomes, particularly suicide death, it was not uncommon for one treatment arm of the clinical trial to have zero events. Several statistical approaches can be employed to overcome this difficulty. Traditional methods such as inverse variance methods employ continuity corrections (i.e., adding a 0.5 correction to zero cells) to eliminate zero cells and to avoid computational problems when calculating the effect for each individual study. When trials have zero events in both treatment arms, these methods exclude them in the analyses. In addition to the inverse variance methods, we also considered other less commonly used methods that were deemed to be well suited for outcomes with rare events, such as the Peto method (Bradburn, Deeks, Berlin, & Russell Localio, 2007) and the Mantel Haenszel method (Lane, 2013; Mantel & Haenszel, 1959). Given that the results were statistically equivalent regardless of the methods (see Supplement 5 in the online supplemental materials), we elected to report the more commonly used risk ratios with inverse variance methods in the main text.[2]

We also sought to test whether existing treatments are better at reducing some SITBs compared with others. We thus assessed treatment effects for each SITB outcome separately. When 95% confidence intervals (CIs) were overlapping, effects were not significantly different. Finally, we tested moderators of treatment effects across SITB outcomes. To examine broad moderator effects, each moderator analysis was first conducted for *all* the effect sizes combined regardless of specific SITB outcomes. To assess whether certain moderator effects were unique to specific outcomes, moderator analyses were then systematically repeated for each SITB outcome category. In terms of power, moderator analyses would be most strongly powered to detect significant effects when all the effect sizes were analyzed in aggregate. For completeness, we report moderator analytic results for both the overall SITB outcome category and each specific SITB outcome (Supplement 6 in the online supplemental materials). Metaregressions were conducted for continuous moderators, including publication year, sample size, sample age, and treatment length. For binary or categorical moderators, separate effect estimates were obtained. These moderators include publication era, age group, intervention type, control treatment type, intervention target, sample severity, therapist training, supervision/adherence check, study quality, study randomization, and study design.

Results

Descriptive Statistics

Number of articles across time. A total of 591 articles comprising 1,125 unique RCTs met inclusion criteria and were included in analyses (see Figure 1). Articles were published as early as 1970 (Baastrup et al., 1970; Kay et al., 1970). However, more recent studies accounted for the majority of the articles included, with 87.99% published since 2000 and 60.74% published since 2010. From these qualifying articles, we identified 3,458 unique effect sizes meeting our inclusion criteria. The distribution of effect sizes across time followed a similar pattern, with 92.02% of effect sizes reported from articles published since 2000, and 67.21% since 2010 (see Figure 2). Of note, a total of 620 (17.93%) of the effect sizes reported zero events in both treatment arms (pretreatment: 22; posttreatment: 529; follow-ups: 69), and six (0.17%) of the effect sizes noted that every participant in both arms reported SITB events (pretreatment: five; posttreatment: one). In addition, two effect sizes (0.06%) reported means and standard deviations of zero in both arms. These effect sizes could not be meta-analyzed due to insufficient variance to estimate the treatment effect.

SITB outcomes. The most commonly reported treatment effects focused on suicide ideation, accounting for 30.86% of the effect sizes (see Figure 3). Other/Combined SITBs was the second most common outcome (26.72%), followed by suicide attempts (17.73%), suicide death (11.63%), self-harm regardless of intent (6.07%), NSSI (4.57%), and hospitalization attributable to SITBs (2.43%).

Assessment time points. Most effect sizes were obtained posttreatment (59.05%), with the rest obtained at follow-ups (19.40%) and baseline (21.54%).

Sample severity. Approximately 60.32% of effect sizes used clinical samples that required participants to experience some form of psychopathology to participate in the study, whereas 28.17% of the effect sizes used self-injurious samples (i.e., prior SITBs were required as an inclusion criterion). The rest of the effect sizes (11.51%) used general samples. As it relates to participants' prior SITBs, studies varied in their reporting. Approximately 26.60% of effect sizes provided information on the percentage of participants endorsing any type of SITBs, 17.09% on prior suicide ideation, 3.18% on suicide plan, 31.17% on prior suicide attempt, and 7.66% on NSSI. Among effect sizes providing these statistics, on average 82.83% of participants endorsed some form of prior SITBs, with 50.09% of participants endorsing prior suicide ideation, 27.46% endorsing prior suicide plan, 50.16% endorsing at least one prior suicide attempt, and 54.26% endorsing prior NSSI.

Sample characteristics. Sample sizes of the included articles ranged widely. Articles that conducted a pooled analysis of mul-

[2] Seventeen effect sizes (0.49%; one chi square statistic, two hazard ratios, seven rate ratios, seven odds ratios) could not be transformed to yield risk ratio estimates. They were excluded in subsequent analyses.

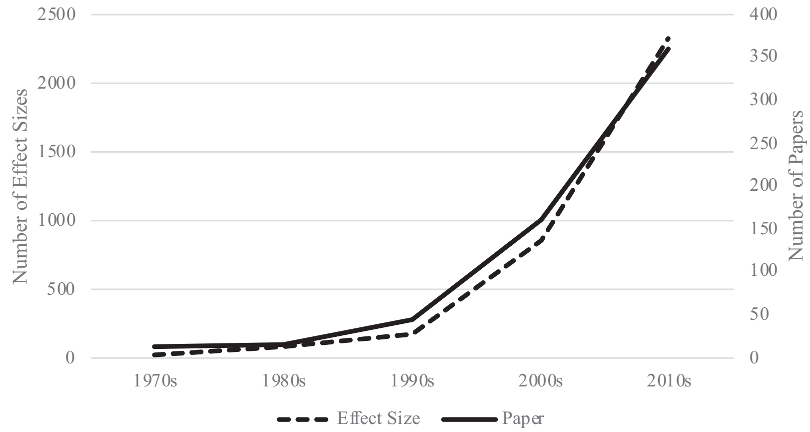

Figure 2. Number of articles and effect sizes across each era of research.

tiple RCTs reported sample sizes ranging from 268 to 48,277, with 2,194 as the median. Sample sizes of studies adopting cluster randomization ranged from 70 to 223,861, with 1,896 as the median. The rest of the articles reported sample sizes ranging from 10 to 18,154, with 200 as the median. The majority (78.89%) of the effect sizes were obtained from studies with sample sizes smaller than 500. Sample age was reported in 92.60% of the effect sizes. The mean age was 33.30 ($SD = 13.37$). Slightly fewer effect sizes (84.41%) provided sufficient information to determine the age *group* of the sample (vs. the specific age). Most effect sizes examined adult populations (38.09%). Much fewer focused exclusively on children and/or adolescents (21.11%) and older adults (0.78%). Approximately one fourth of effect sizes (24.90%) examined a mixed age population (mixed older adults and adults: 17.18%; mixed adolescents and adults: 6.68%; all age groups: 1.04%). Most effect sizes (92.94%) reported information on gender. On average, females constituted 61.99% of the samples.

Intervention target type. Most effect sizes were drawn from interventions primarily targeting psychopathology (63.48%), with a smaller percentage primarily targeting SITBs (29.76%). A small proportion of interventions targeted other outcomes (e.g., physical health; 6.77%).

Specific intervention type. Medication alone was the most common active treatment examined, accounting for 46.72% of the effect sizes. The rest of the effect sizes studied CT and CBT (11.68%), DBT (6.99%), and combinations of therapy and medication (6.25%). None of the other treatments accounted for more than 5% of the effect sizes, but notable additional interventions included psychotherapy employing a variety of modalities (4.42%). checking-in programs (2.54%), problem-solving therapy (1.62%), safety planning and/or means restriction (1.47%), psychoanalysis (0.93%), and inpatient hospitalization (0.12%).

Intervention characteristics. We were able to extract treatment length associated with 97.05% of the effect sizes. The distribution of treatment lengths was positively skewed (Shapiro-Wilk Normality test of normality: $p < .001$). The median treatment length was 12 weeks ($M = 23.83$, $SD = 31.49$). Most effect sizes were associated with active treatments that only included individual components (76.92%), with a much smaller percentage including both individual and group components (8.36%), individual and

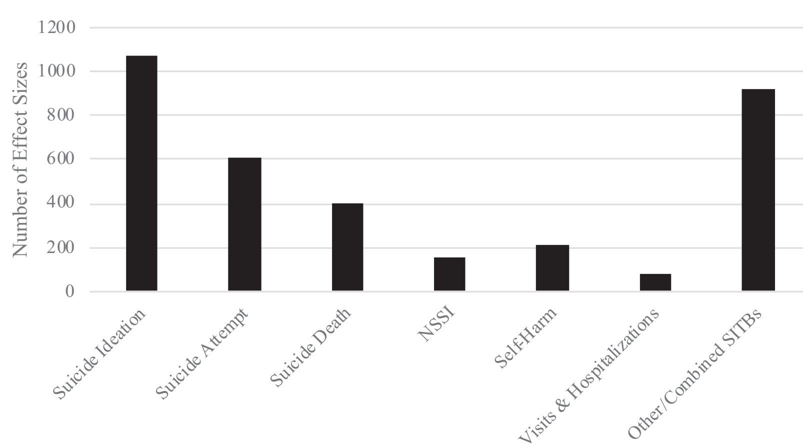

Figure 3. Number of effects sizes across outcomes. NSSI = nonsuicidal self-injury; SITBs = self-injurious thoughts and behaviors.

family components (5.41%), and group components only (4.54%). None of the other treatment components constituted more than 5% of the effect sizes, such as school-based programs (1.82%), family interventions (1.21%), or parent/therapist/primary care provider training (0.06%).

Control treatment characteristics. Most active treatment conditions were compared against an active control condition (60.12%), such as active psychotherapy and pharmacotherapy. The rest of the active conditions were compared against placebo conditions (34.59%), and no-treatment conditions (5.29%).

Methodological quality.

Quality ratings. More than half of the effect sizes (57.58%) were associated with weak study quality. More than a quarter of the effect sizes (36.21%) were yielded from studies with moderate quality, with only 6.22% of the effect sizes from studies with strong quality. Because of the elaborate scoring system of the Quality Assessment Tool for Quantitative Studies (Effective Public Health Practice Project, 2007), it is hard to induce which study aspects primarily contributed to the high proportion of weak study quality. However, it is likely that the difficulty in selecting a representative sample for RCTs (i.e., selection bias) might have played a role as only 0.90% of the effect sizes were drawn from samples rated as *very likely* to be representative of the target population (e.g., two large grade cohorts across schools in a city school system to test universal preventive interventions for schoolers; Wilcox et al., 2008). Moreover, more than half of the effect sizes (62.00%) were drawn from studies that did not report what percentage of participants approached or selected to participate agreed to participate. The challenges in blinding within RCT designs might have also contributed to the overall weak ratings. Approximately 38.37% of the effect sizes were from studies that did not report whether the SITB outcome assessors were blind to intervention condition. In terms of participants, only 0.40% of the effect sizes were drawn from studies that explicitly stated that participants were unaware of the research question. In short, because all included studies were RCTs, all included studies were methodologically strong relative to most studies. But, owing to a range of limitations (e.g., selection bias, blinding), the widely used Quality Assessment Tool for Quantitative Studies classifies over half of the SITB RCT literature as methodologically weak.

Retention rates. Because of differences in reporting, we considered the following three retention indices: participants retained from treatment assignment to the start of treatment, from treatment assignment to completion, and from the start of treatment to completion. A total of 2,693 effect sizes (77.88%) provided sufficient information to calculate retention at least one way. Retention rates were negatively skewed (Shapiro-Wilk Normality test of normality: all $ps < .001$). Within active treatment groups, the median percentage of participants retained from recruitment to the beginning of treatment was 98.68%. The median rates of participants retained from treatment assignment to completion, and the beginning of treatment to treatment completion were 75.76% and 79.77%, respectively. Within control comparison groups, the median percentage of participants retained from treatment assignment to the beginning of treatment was 99.74%. The median rates of participants retained from treatment assignment to completion, and the beginning of treatment to completion were 76.09% and 82.88%, respectively. Mann–Whitney-Wilcoxon's Tests showed that all three retention indices were significantly different by treatment conditions ($p < .001$, $p = .02$, $p = .001$, respectively), with slightly higher retention rates in control groups across all indices.

Therapists. When therapists were involved in the active treatment conditions (38.17%), 73.03% of active treatment conditions required specific pretreatment training for therapists and 82.12% provided supervision or conducted adherence checks throughout treatment to ensure treatment integrity.

Blinding. When blind status was not reported, we assumed the RCT was not blind. Approximately half of the effect sizes (49.86%) were *not* blind, 44.99% were double-blind, and 5.15% were single blind.

Study randomization and design. Most effect sizes (97.46%) were from studies adopting an individual-level randomization procedure, with the rest (2.54%) from studies adopting cluster randomization. Almost all included effect sizes (99.33%) were from studies using a parallel design, with very few using crossover design (0.64%) or dynamic waitlist design (0.03%).

Effect Estimates and Publication Bias

This meta-analysis set out to examine the extent to which existing interventions reduce SITBs after treatment administration; we focused on reporting effect estimates assessed at the posttreatment time point below. This decision is based on evidence from prior meta-analyses showing that mental health treatments tend to show the strongest effects immediately posttreatment (e.g., Carl et al., 2019; Eckshtain et al., 2020; Grenon et al., 2019; Oud et al., 2019). Primarily presenting posttreatment effect sizes will likely provide more optimistic estimates than aggregating effects assessed at follow-ups. However, it is important to note that there may be a discrepancy between the amount of time postintervention until maximum treatment potency is obtained compared with the amount of time postintervention until treatment effects can be observed (particularly for upstream interventions [e.g., school-based interventions, couples therapy] and low base-rate behaviors, like suicide deaths). Larger-scale and particularly longer-term treatment studies are needed to better ensure that significant treatment effects can be detected. This may be particularly true when trying to understand and test the ability of prevention programs to reduce SITBs across the general population.

Given that whether treatment groups exhibited statistically similar level of SITBs prior to beginning treatment provides important context for interpreting the post treatment effects, we report the detailed meta-analytic results of effect sizes obtained prior to treatment in Supplement 3 in the online supplemental materials. Briefly, active treatment groups and control groups did not differ on any of the SITB outcomes prior to treatment. Similarly, understanding whether and to what extent treatment effects were maintained at follow-ups might be of interest for some researchers. The detailed results are provided in Supplement 3 in the online supplemental materials. The effect estimates were generally statistically equivalent to those measured immediately after treatment.

Overall analyses. We first examined the overall effects of treatment on all types of SITB outcomes measured immediately after treatment completion (Figure 4 & 5). Analyses on binary SITB outcomes/rates of SITBs included 1,186 effect sizes in total, yielding a significant yet small treatment effect with a RR of 0.91

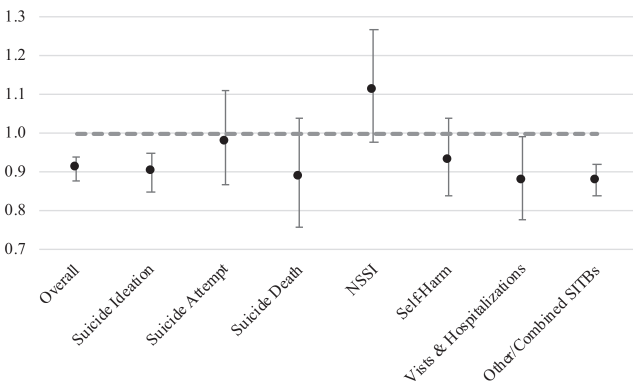

Figure 4. Posttreatment weighted risk ratios by outcome type. Dashed line indicates null effect (i.e., risk ratio = 1); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. NSSI = nonsuicidal self-injury; SITBs = self-injurious thoughts and behaviors.

(95% CIs [0.88, 0.94]; Figure 4 and Table 1). That is, on average, treatments reduced SITBs by approximately 9% compared with control group immediately post intervention. Heterogeneity across studies was low (see Table 1). Publication bias was minimal, and the funnel plot was symmetrical (Table 1 and Figure 6a). Analyses on outcomes reported in continuous form such as means and standard deviations included 314 effect sizes, which also produced a small but significant treatment effect (g = −0.17 [−0.22, −0.12]; Table 1 and Figure 5). Moderate heterogeneity across studies was detected (see Table 1). Results of publication bias tests indicated moderate bias, and the funnel plot appeared asymmetrical (Table 1 and Figure 6b).

Suicide ideation. A total of 395 effect sizes were included in the binary analyses. Results showed that treatment on average significantly reduced the occurrence of suicide ideation (RR = 0.90 [0.85, 0.95]). Results indicated low heterogeneity across studies and low publication bias (Table 1 and Supplement 7a in the online supplemental materials). Regarding continuous outcomes, a total of 114 effect sizes were included, resulting in a small but significant reduction in suicide ideation frequency and/or intensity (g = −0.09 [−0.15, −0.02]). Between-study heterogeneity was high, but publication bias was low (Table 1 and Supplement 7b in the online supplemental materials).

Suicide attempt. Binary analyses included 209 effect sizes. Results suggested that treatment on average did not significantly reduce the occurrence of suicide attempt (RR = 0.98 [0.87, 1.11]). Low heterogeneity and minimal publication bias were observed (Table 1 and Supplement 7c in the online supplemental materials). Only 12 effect sizes involved the means and standard deviations of the number of suicide attempt episodes in each condition. The pooled effect was nonsignificant (g = −0.10 [−0.27, 0.06]). Results showed moderate between-study heterogeneity, although publication bias appeared minimal (Table 1 and Supplement 7d in the online supplemental materials).

Suicide death. A total of 159 effect sizes were included in the binary analyses. The results indicated that interventions on average did not significantly reduce suicide death (RR = 0.89 [0.76, 1.04]). Between-study heterogeneity and publication bias were low (Table 1 and Supplement 7e in the online supplemental materials).

Nonsuicidal self-injury. Forty-six effect sizes examined NSSI as a discrete outcome. Results indicate that treatment on average did not significantly reduce the occurrence of NSSI (RR = 1.11 [0.98, 1.27]). Between-study heterogeneity and publication bias was not detected (Table 1 and Supplement 7f in the online supplemental materials). In terms of continuous outcomes, results across 20 effect sizes show that treatment did not significantly decrease the frequency and/or intensity of NSSI (g = −0.07 [−0.18, 0.04]). Again, heterogeneity between studies and publication bias were not detected (Table 1 and Supplement 7g in the online supplemental materials).

Self-harm. Regarding binary outcomes, 73 effect sizes examined self-harm regardless of suicidal intent. Treatment effects were nonsignificant (RR = 0.93 [0.84, 1.04]). Results indicated low heterogeneity and minimal publication bias (Table 1 and Supplement 7h in the online supplemental materials). For continuous outcomes, a total of 24 effect sizes were included for analysis. Treatment on average did not result in a significant reduction in the frequency and/or intensity of self-harm (g = −0.15 [−0.47, 0.10]). Between-study heterogeneity was moderate, though minimal publication bias was detected (Table 1 and Supplement 7i in the online supplemental materials).

Visits/Hospitalizations attributable to SITB. Thirty-six effect sizes examined emergency room visits and/or hospitalizations resulting from engagement in SITB as a binary outcome. Treatment on average showed efficacy in reducing hospitalization (RR = 0.88 [0.78, 0.99]). Results indicated minimal between-study heterogeneity and low publication bias (Table 1 and Supplement 7j in the online supplemental materials). Only five effect sizes examined hospitalization as a continuous outcome. The pooled effect was nonsignificant (g = −0.19 [−0.47, 0.10]). Results detected high between-study heterogeneity and publication bias (Table 1 and Supplement 7k in the online supplemental materials).

Other/Combined SITBs. A total of 268 effect sizes examined other types of SITBs (e.g., suicide preparations) and/or combined multiple types of SITBs (e.g., suicide preparations, suicide attempt

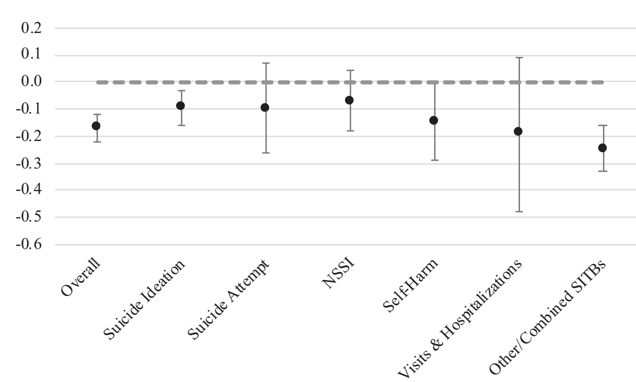

Figure 5. Posttreatment weighted Hedges's g by outcome type. Dashed line indicates null effect (i.e., Hedges's g = 0); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. NSSI = nonsuicidal self-injury; SITBs = self-injurious thoughts and behaviors.

Table 1
Posttreatment Effect Sizes and Publication Biases Across Outcomes

Binary	n	RR [95% CI]	p	I^2	Fail-safe N Classic	Fail-safe N Orwin's	Begg and Mazumdar Rank Correlation	Egger's Test of Intercept	Duval and Tweedie's Trim and Fill: Missing effect sizes	Duval and Tweedie's Trim and Fill: Adjusted RR
Overall	1,186	0.91 [0.88, 0.94]	<.001	22.13	8,497	335	$\tau = -0.01$, $p = .53$	$z = 1.34$, $p = .18$	0	0.91 [0.88, 0.94]
Suicide ideation	395	0.90 [0.85, 0.95]	<.001	29.68	1,938	456	$\tau = -0.04$, $p = .28$	$z = 0.36$, $p = .72$	0	0.90 [0.85, 0.95]
Suicide attempt	209	0.98 [0.87, 1.11]	.75	39.33	0	0	$\tau = 0.07$, $p = .12$	$z = 1.90$, $p = .06$	0	0.98 [0.87, 1.11]
Suicide death	159	0.89 [0.76, 1.04]	.14	0.00	0	173	$\tau = -0.09$, $p = .09$	$z = 0.47$, $p = .64$	0	0.89 [0.76, 1.04]
NSSI	46	1.11 [0.98, 1.27]	.11	0.00	0	0	$\tau = -0.01$, $p = .96$	$z = 0.19$, $p = .85$	0	1.11 [0.98, 1.27]
Self-harm	73	0.93 [0.84, 1.04]	.20	21.99	0	41	$\tau = 0.02$, $p = .84$	$z = -0.48$, $p = .63$	0	0.93 [0.84, 1.04]
Visits and hospitalizations	36	0.88 [0.78, 0.99]	.03	1.58	39	141	$\tau = -0.09$, $p = .45$	$z = -1.15$, $p = .25$	5	0.91 [0.80, 1.04]
Other/combined SITBs	268	0.88 [0.84, 0.92]	<.001	0.00	830	316	$\tau = -0.06$, $p = .14$	$z = 0.66$, $p = .51$	0	0.88 [0.84, 0.92]

Note. n = number of effect sizes; NSSI = nonsuicidal self-injury; RR = weighted mean risk ratio; SITBs = self-injurious thoughts and behaviors; 95% CI = 95% confidence interval. Dashes indicate unavailable information, I^2 indicates the percentage of variances attributable to heterogeneity between studies. Begg and Mazumdar Rank Correlation and Egger's Test of the Intercept test whether there is any evidence for the existence of publication bias. Classic fail-safe N values represent the number of studies needed to nullify the observed effects statistically. Orwin's Fail-Safe N represents the number of studies needed to nullify the observed effects clinically. Begg and Mazumdar Rank Correlation Test computes the rank order correlation between effect estimates and standard error; Egger's Test of the Intercept uses precision (i.e., the inverse of the standard error) to predict the standardized effect (i.e., effect size divided by the standard error). The size of the effect is reflected in the slope and bias is reflected in the intercept. Duval and Tweedie's Trim & Fill estimates the unbiased effect sizes after accounting for publication bias. Missing cases are the number of cases estimated as missing below the mean.
[a] Estimates are not reported for analyses involving fewer than five effect sizes, because a small number of effect sizes compromises the accuracy of estimates.

and death) and examined them as binary outcomes in their investigation. The pooled results showed a significant treatment effect in reducing the occurrence of SITBs (RR = 0.88 [0.84, 0.92]). Little between-study heterogeneity and publication bias were detected (Table 1 and Supplement 7l in the online supplemental materials). With respect to continuous outcomes, the pooled treatment effect across 139 effect sizes was significant ($g = -0.25$ [−0.34, −0.17]). High heterogeneity existed between studies, and moderate publication bias was observed (Table 1 and Supplement 7m in the online supplemental materials).

Moderator Analyses

Below, we report the moderator effects of treatment efficacy across outcomes (i.e., the overall aggregated effects) and any significant moderator effects for each type of SITB outcomes. Please see Supplement 6 in the online supplemental materials for the full results of moderator analyses associated with each outcome. We note here that outcome-specific effects were highly consistent with overall effects on aggregated SITBs.

Publication year and era. Considering all outcomes, publication year and era did not significantly moderate findings (Tables 2 and 3, Figure 7). In terms of specific SITB outcomes, treatment effects for other/combined SITBs were significantly weaker in 2000s (RR = 1.12 [0.97, 1.30]; Supplement 6 in the online supplemental materials). No other significant moderator effects were detected for any specific SITB outcomes (Supplement 6 in the online supplemental materials).

Sample size. Metaregression analyses were conducted to examine the effects of sample size on treatment effects. The effects were nonsignificant for SITBs analyzed as an aggregate or as specific SITB outcomes (see Table 2).

Sample age. To assess for the moderator effects of sample age on the overall meta-analytic results, we first conducted metaregressions of the mean sample age. For pooled binary outcomes, metaregression analyses indicated that treatment effects were significantly larger for older participants, although this moderator effect was small ($b = -0.004$; Table 2). For pooled continuous outcomes, treatment effects were consistent regardless of sample age. We also analyzed the effects of sample age by conducting separate analyses for different sample age groups. For binary outcomes, significant moderator effects were detected. Treatment effects were the smallest when samples only consisted of children and/or adolescents (see Table 3). No significant moderator effects were detected for continuous outcomes when analyzed as an aggregate.

In terms of specific SITB outcomes, metaregression analyses indicated that older sample age was associated with significantly larger treatment effects for suicide ideation when analyzed as a binary outcome and for self-harm when analyzed as a continuous outcome, although the effects were small (bs = −0.005, −0.02; Table 2). For suicide death, older sample age was associated with significantly weaker treatment effects, but again the effect was small ($b = 0.04$; Table 2). With respect to sample age groups, treatment effects were significantly larger for suicide attempt when

Table 1 (*continued*)

Continuous	n	Hedges's g [95% CI]	p	I^2	Fail-safe N Classic	Orwin's	Begg and Mazumdar Rank Correlation	Egger's Test of Intercept	Duval and Tweedie's Trim and Fill Missing effect sizes	Adjusted g
Overall	314	−0.17 [−0.22, −0.12]	<.001	79.56	20,237	899	$\tau = -0.13$, $p = .001$	$z = -2.90$, $p = .004$	0	−0.17 [−0.22, −0.12]
Suicide ideation	114	−0.09 [−0.15, −0.02]	.01	60.59	528	134	$\tau = -0.05$, $p = .42$	$z = -1.56$, $p = .12$	0	−0.09 [−0.15, −0.02]
Suicide attempt	12	−0.10 [−0.27, 0.06]	.23	54.25	3	26	$\tau = -0.03$, $p = .95$	$z = -0.75$, $p = .45$	2	−0.02 [−0.22, 0.19]
Suicide death	0[a]	—	—	—	—	—	—	—	—	—
NSSI	20	−0.07 [−0.18, 0.04]	.21	0.00	0	0	$\tau = -0.15$, $p = .39$	$z = 0.02$, $p = .98$	0	−0.07 [−0.18, 0.04]
Self-harm	24	−0.15 [−0.30, −0.01]	.04	55.71	72	75	$\tau = -0.27$, $p = .06$	$z = -1.24$, $p = .21$	1	−0.13 [−0.28, 0.02]
Visits and hospitalizations	5	−0.19 [−0.47, 0.10]	.20	58.14	3	17	$\tau = -0.40$, $p = .48$	$z = -0.87$, $p = .39$	1	−0.12 [−0.40, 0.16]
Other/combined SITBs	139	−0.25 [−0.34, −0.17]	<.001	87.62	9591	650	$\tau = -0.21$, $p < .001$	$z = -2.61$, $p = .01$	0	−0.25 [−0.34, −0.17]

the sample consisted of mixed adolescents and adults (RR = 0.51 [0.35, 0.77]), and significantly smaller for self-harm when the sample consisted of mixed older adults and adults (RR = 1.78 [1.14, 2.79]). Treatment effects were also significantly weaker for other/combined SITBs among children and adolescents (RR = 1.15 [1.00, 1.32]; Supplement 6 in the online supplemental materials). No other significant moderator effects were detected.

Treatment length. Metaregression analyses were conducted for both binary and continuous outcomes. Treatment length did not significantly moderate the treatment effects when all SITBs were analyzed as an aggregate (see Table 2). In terms of specific SITB outcomes, longer treatment length was associated with significantly but only slightly larger treatment effects for suicide attempt ($b = -0.002$) and other/combined SITBs ($b = -0.002$) reported in binary forms, and suicide attempt ($b = -0.02$), self-harm ($b = -0.01$), and visits and hospitalizations attributable to SITBs ($b = -0.01$) when they were reported in continuous forms (see Table 2).

Intervention target type. Overall, for pooled binary outcomes, the strength of treatment effects remained consistent regardless of whether treatment aimed to target SITBs or psychopathology but was significantly weaker when treatment aimed to target other outcomes (Table 3 and Figures 8 and 9). Treatment effects remained statistically equivalent for pooled continuous outcomes. Moderator effects of intervention target type were detected for other/combined SITBs where interventions with targets other than SITBs or psychopathology reported higher SITB events in the active condition (RR = 1.64 [1.08, 2.49]; Supplement 6 in the online supplemental materials). No other moderator effects were detected for specific SITB outcomes.

Specific intervention type. When all SITB outcomes reported in binary forms were considered as an aggregate, no moderator effects of specific intervention type were observed (Table 3 and Figure 10). When all SITB outcomes reported in continuous forms were analyzed in aggregate, the only intervention type that differed significantly from the pooled effect was checking-in programs, which exerted significantly weaker effect (Table 3 and Figure 11). In terms of specific SITB outcomes, the only detected moderator effects were the following: (a) the effects of combined psychotherapy and medication and checking-in programs for suicide ideation (RRs = 0.64 [0.58, 0.70], 0.70 [0.63, 0.78], respectively) were significantly stronger than the pooled effects (RR = 0.90 [0.85, 0.95]) when suicide ideation was reported in binary forms, whereas interventions not falling in the examined categories (i.e., Other) produced significantly weaker effects (RR = 1.01 [0.97, 1.05]; Supplement 6a in the online supplemental materials); (b) checking-in programs produced significantly larger effects for suicide attempt (RR = 0.52 [0.37, 0.73]) compared with the pooled effects (RR = 0.98 [0.87, 1.11]); Supplement 6b in the online supplemental materials); and (c) eclectic psychotherapy yielded larger effects for other/combined SITBs when measured in continuous forms ($g = -0.70$ [−0.95, −0.46]; Supplement 6g in the online supplemental materials).

Control treatment characteristics. When all SITB outcomes in binary forms were considered, studies where control groups received no treatment (e.g., waitlist control groups) yielded sig-

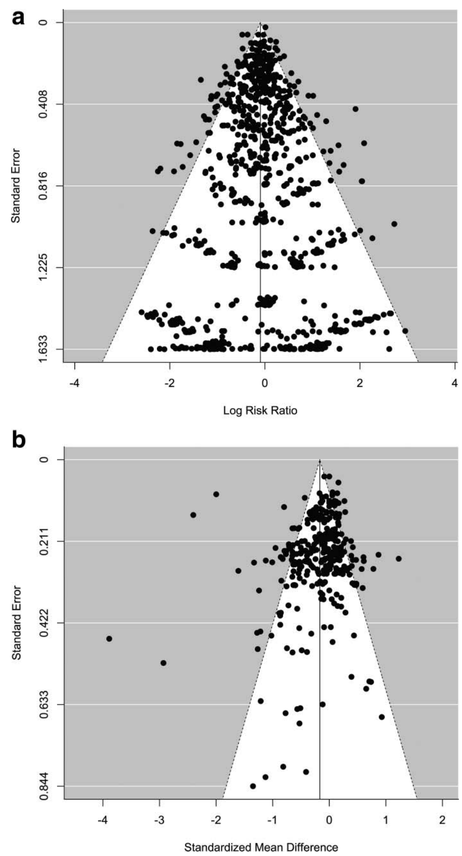

Figure 6. Funnel plots. (a) Funnel plot for binary outcomes. (b) Funnel plot for continuous outcomes. Shaded circles represent observed estimates; open circles represent imputed values estimated to be missing to the right of the mean; no open circles indicate no values were estimated to be missing.

nificantly weaker treatment effects (see Table 3). No other moderator effects were detected for either binary or continuous outcomes (see Table 3). When specific SITB outcomes were examined, the only two significant moderator effects were detected for suicide ideation and suicide attempt in binary forms: interventions on average did not reduce suicide ideation when the control condition involved no treatment (RR = 1.02 [0.97, 1.06]; Supplement 6a in the online supplemental materials); interventions reported more suicide attempts in the active condition compared with the placebo condition (RR = 1.57 [1.21, 2.03]; Supplement 6b in the online supplemental materials). No other significant moderator effects emerged (Supplement 6 in the online supplemental materials).

Sample severity. For both binary and continuous outcomes, treatment effects did not significantly differ depending on whether participants were required to exhibit SITBs or psychopathology for study inclusion (see Table 3). No significant moderator effects were detected for specific SITB outcomes either (Supplement 6 in the online supplemental materials).

Treatment components. Treatment effects remained statistically consistent regardless of treatment components (e.g., individual, group, family, school-based; Table 3). With respect to specific SITB outcomes, interventions involving both individual and group components yielded significantly larger reduction in self-harm (RR = 0.58 [0.44, 0.78]; $g = -0.56$ [−0.75, −0.37]) compared with the pooled effect (RR = 0.93 [0.84, 1.04], $g = -0.15$ [−0.30, −0.01]).

Therapist training, supervision/adherence check. For pooled analyses, the presence or lack of training of therapists did not significantly moderate the findings for binary outcomes (see Table 3). For continuous outcomes, interventions that did not report whether they incorporated training yielded significantly larger effects ($g = -0.37$ [−0.51, −0.23]; Table 3). The incorporation of adherence check of therapists did not moderate the results when continuous outcomes were analyzed in aggregate (see Table 3). Significant moderator effects were detected for suicide death: interventions that did not report adherence check status produced significantly larger effects (RR = 0.32 [0.17, 0.61]) compared with the pooled effects for suicide death (RR = 0.89 [0.76, 1.04]; Supplement 6c in the online supplemental materials).

Study quality. For pooled analyses, the study quality did not significantly moderate the findings for either binary or continuous outcomes (see Table 3). The same pattern applies to specific SITB outcomes (Supplement 6 in the online supplemental materials).

Study randomization and design. For pooled analyses, whether a study elected to randomize individual participants or larger units (e.g., schools, clinics, households) did not moderate the findings for the pooled analyses. Given that very few effect sizes were yielded from studies adopting a crossover or dynamic waitlist design, we were unable to generate separate effect estimates from those designs. No moderator effects emerged for specific SITB outcomes that involved sufficient effect sizes to meta-analyze (Supplement 6 in the online supplemental materials).

Discussion

SITBs have plagued humanity for millennia. Over the past several decades there have been major institutional and scientific efforts aimed at eliminating SITBs, with these efforts accelerating in recent years. These efforts have produced hundreds of interventions and hundreds of RCTs, but it has been unclear exactly which interventions are most efficacious and what conditions moderate their efficacy. The primary purpose of this meta-analysis was to shed light on these issues by addressing several broad questions about SITB intervention studies. Analyses revealed several surprising and notable findings, including the following: (a) overall intervention effects were small across all SITB outcomes; (b) the number of SITB RCTs increased exponentially across five decades, but intervention efficacy has not improved; (c) all SITB intervention types produced small effects, and no intervention was significantly and consistently stronger than any other; (d) efficacy was similar regardless of whether interventions primarily targeted SITBs or psychopathology; (e) efficacy was similar across age groups, though effects were slightly weaker for child/adolescent populations and very few studies focused on older adults; (f) the overall small intervention effects were generally maintained at follow-up assessments; and (g) intervention effects were consistently small regardless of sample and study characteristics (e.g., control group type, intervention length, sample size). Taken together, these findings suggest the need for fundamental changes in

Table 2
Meta-Regression Analyses

Measure	Publication year		Sample size		Sample age		Treatment length	
	b	*p*	*b*	*p*	*b*	*p*	*b*	*p*
Binary								
Overall	0.004	.14	−0.000	.13	**−0.004**	**.001**	−0.001	.14
Suicide ideation	0.007	.27	−0.000	.74	**−0.005**	**.01**	0.000	.84
Suicide attempt	−0.02	.06	−0.000	.07	−0.01	.05	**−0.002**	**.03**
Suicide death	−0.02	.16	0.000	.76	**0.04**	**.002**	−0.002	.14
NSSI	0.01	.78	0.000	.67	−0.005	.54	0.000	.94
Self-harm	0.01	.10	0.000	.54	−0.000	.98	−0.000	.96
Visits and hospitalizations	−0.002	.77	0.000	.35	0.01	.35	−0.004	.05
Other/combined SITBs	0.01	.09	−0.000	.76	−0.001	.61	**−0.002**	**.02**
Continuous								
Overall	0.004	.27	0.000	.88	−0.004	.07	−0.002	.20
Suicide ideation	−0.01	.37	0.000	.76	−0.01	.06	−0.001	.69
Suicide attempt	0.01	.63	0.000	.50	0.000	.99	**−0.02**	**<.001**
Suicide death	—	—	—	—	—	—	—	—
NSSI	−0.001	.95	0.000	.92	0.02	.20	0.003	.38
Self-harm	0.02	.12	0.000	.75	**−0.02**	**.047**	**−0.01**	**<.001**
Visits and hospitalizations	−0.06	.30	0.002	.44	−0.05	.32	**−0.01**	**.048**
Other/combined SITBs	0.001	.86	0.000	.74	0.004	.44	−0.000	.99

Note. NSSI = nonsuicidal self-injury; SITBs = self-injurious thoughts and behaviors. *b* indicates the regression coefficient, and dashes indicate unavailable information. Significant effects are bolded.

SITB interventions. We discuss each of these major findings and their implications below.

Overall Intervention Effects Were Small Across All Outcomes

A total of 1,125 RCTs (producing 591 published articles and 3,458 effect sizes) qualified for this meta-analysis. Given the difficulty of conducting RCTs with SITBs as an outcome, this is an impressive number of studies. Unfortunately, the overall effect of interventions on SITBs was small. In binary analyses, there was a 9% reduction (95% CI [6%, 12% reduction]) in the number of people reporting any SITBs in active groups compared with control groups. In continuous analyses, there was a small standardized mean difference (Hedges's $g = -0.17$; 95% CI [−0.22, −0.12]) in the frequency of SITBs in active groups compared with control groups. Directly analogous to Cohen's *d* effect size standards, Hedges's *g* effect sizes under 0.30 are considered small and are typically impossible to discern with the naked eye when viewing group distributions. Publication bias was minimal for binary analyses and minimal-to-moderate for continuous analyses. Both binary and continuous analyses converge on the conclusion that, overall, intervention effects on SITBs are small.

We note that these effects, and those subsequently discussed, do not refer to absolute reductions in SITBs across the course of studies. They instead refer to the relative SITB reductions in active intervention groups compared with control intervention groups. Psychiatric and medical symptoms often abate over time without intervention (Adams, Houle, Parker, & Burke, 2012; Arkowitz & Lilienfeld, 2006; Flett, Vredenburg, & Krames, 1995). This is why control groups—ideally with randomization and a similar protocol to the active interventions—are necessary to distinguish the causal effect of an intervention from passage-of-time effects and nonspecific effects. This means that, in the present meta-analysis, a 9% reduction refers to a 9% reduction in SITBs relative to the control group's SITBs. It does not refer to a 9% total reduction in SITBs from the beginning to the end of the study. However, we did not detect stronger effects when the control group involved no treatment or placebo compared with active treatment. This suggests that the relative reductions that we report are likely similar in magnitude to the causal effect of these interventions relative to the passage of time.

We explored whether these overall effects were moderated by SITB outcome type. The most common outcome examined was suicide ideation (30.86% of all effect sizes), followed by other/combined SITBs (26.72%), suicide attempts (17.73%), suicide death (11.63%), self-harm regardless of suicide intent (6.07%), NSSI (4.57%), and hospitalization attributable to SITBs (2.43%). These descriptive findings highlight that—despite being the most dangerous SITBs—suicide attempts and deaths account for a small proportion of studied outcomes. Additionally, outcome type did not significantly moderate the overall small intervention effects. As a whole, the results indicate that existing interventions have similarly small effects on all SITB outcomes.

Some might argue that small effects are all that should be expected from SITB interventions, especially in light of the complexity of these phenomena and in light of the fact that most interventions target few factors (e.g., Stanley & Mann, 2019). Others might argue that, given some interventions for other forms of psychopathology produce large effect sizes (e.g., Cohen's *d*s and Hedge's *g*s ~1.0 for CBT for depression and anxiety: Bandelow et al., 2018; Johnsen & Friborg, 2015), we should expect large effects from SITB interventions. We strongly agree with many SITB researchers that the causes of SITB are complex and that it is possible that this complexity places a low ceiling on potential effect sizes for SITB interventions. Indeed, the present meta-analysis was unable to reject this hypothesis. At the same time, we are reluctant to accept this hypothesis because we believe it is possible that fundamental changes to how we approach SITB

Table 3
Moderator Analyses

Moderator	Overall			
	Binary		Continuous	
	n	RR [95% CI]	*n*	*g* [95% CI]
Pooled effects	1,186	0.91 [0.88, 0.94]	314	−0.17 [−0.22, −0.12]
Decade				
1970s	8	0.83 [0.59, 1.18]	5	−0.16 [−0.31, −0.01]
1980s	10	0.82 [0.57, 1.17]	8	−0.13 [−0.41, 0.14]
1990s	80	0.78 [0.69, 0.88]	9	−0.10 [−0.31, 0.11]
2000s	370	1.02 [0.94, 1.11]	76	−0.24 [−0.36, −0.12]
2010s	718	0.91 [0.87, 0.94]	216	−0.14 [−0.19, −0.09]
Control group type				
No treatment	**23**	**1.01 [0.97, 1.06]**	29	−0.23 [−0.38, −0.07]
Placebo	544	0.97 [0.91, 1.04]	69	−0.15 [−0.27, −0.04]
Active	619	0.88 [0.84, 0.92]	216	−0.16 [−0.21, −0.10]
Intervention target type				
SITBs	206	0.89 [0.84, 0.93]	133	−0.20 [−0.28, −0.12]
Psychopathology	887	0.91 [0.87, 0.96]	179	−0.15 [−0.21, −0.08]
Other	**93**	**1.18 [0.98, 1.43]**	2[a]	—
Specific intervention type				
Medication only	816	0.94 [0.90, 0.99]	48	−0.22 [−0.37, −0.07]
CT/CBT	52	0.81 [0.70, 0.93]	75	−0.20 [−0.28, −0.11]
Eclectic psychotherapy	21	0.93 [0.78, 1.10]	22	−0.31 [−0.52, −0.10]
DBT	29	0.98 [0.83, 1.17]	41	−0.11 [−0.21, −0.01]
Psychotherapy and medication combined	80	0.80 [0.69, 0.92]	17	−0.25 [−0.40, −0.10]
Checking-in programs	29	0.87 [0.75, 1.00]	**10**	**−0.05 [−0.11, 0.01]**
Psychoanalysis/insight-based therapy	5	0.84 [0.63, 1.13]	3[a]	—
Problem solving therapy	6	0.66 [0.45, 0.97]	9	−0.30 [−0.45, −0.15]
Safety planning/means safety	3[a]	—	2[a]	—
Inpatient hospitalization	0[a]	—	0[a]	—
Other	145	0.94 [0.89, 1.00]	87	−0.03 [−0.12, 0.05]
Sample severity				
General	120	0.99 [0.86, 1.15]	13	−0.11 [−0.33, 0.12]
Clinical	901	0.93 [0.89, 0.97]	149	−0.13 [−0.21, −0.05]
Self-Injurious	165	0.86 [0.82, 0.92]	152	−0.20 [−0.26, −0.14]
Age group				
Adult	345	0.85 [0.81, 0.90]	152	−0.15 [−0.21, −0.10]
Children and/or adolescents	**307**	**1.08 [1.00, 1.17]**	37	−0.07 [−0.19, 0.04]
Older adults	9	0.68 [0.39, 1.21]	1[a]	—
Mixed adolescents and adults	54	0.87 [0.80, 0.95]	17	−0.15 [−0.28, −0.01]
Mixed older adults and adults	211	0.89 [0.83, 0.96]	46	−0.16 [−0.24, −0.09]
All age groups	27	0.84 [0.71, 1.00]	2[a]	—
Did not report	233	0.95 [0.86, 1.05]	59	−0.19 [−0.37, −0.01]
Treatment components				
Individual	1,052	0.92 [0.89, 0.96]	207	−0.15 [−0.21, −0.09]
Individual and family	56	0.92 [0.77, 1.09]	16	−0.12 [−0.21, −0.02]
Individual and group	44	0.76 [0.64, 0.90]	50	−0.22 [−0.33, −0.12]
Group	14	1.02 [0.81, 1.28]	29	−0.29 [−0.52, −0.07]
Family	6	1.05 [0.81, 1.36]	5	−0.20 [−0.60, 0.20]
School-based intervention	5	0.78 [0.61, 1.00]	1[a]	—
Parent training	0[a]	—	0[a]	—
Therapist/primary care provider training	1[a]	—	0[a]	—
Other	8	0.89 [0.76, 1.04]	6	0.07 [−0.28, 0.42]
Blind status				
Not blind	355	0.85 [0.80, 0.89]	236	−0.15 [−0.20, −0.10]
Single blind	45	0.96 [0.83, 1.10]	25	−0.16 [−0.31, −0.01]
Double blind	786	0.97 [0.92, 1.02]	53	−0.22 [−0.36, −0.08]
Therapist training				
No therapist involved	921	0.93 [0.89, 0.97]	132	−0.13 [−0.21, −0.05]
No training	1[a]	—	0[a]	—
Training	190	0.86 [0.78, 0.94]	142	−0.15 [−0.21, −0.08]
Did not report	74	0.88 [0.82, 0.96]	**40**	**−0.37 [−0.51, −0.23]**
Supervision/adherence check				
No therapist involved	921	0.93 [0.89, 0.97]	132	−0.13 [−0.21, −0.05]
No adherence check	0[a]	—	5	−0.49 [−0.91, −0.07]
Adherence check	209	0.82 [0.75, 0.90]	158	−0.16 [−0.22, −0.10]

(*table continues*)

Table 3 (*continued*)

Moderator	Overall			
	Binary		Continuous	
	n	RR [95% CI]	*n*	*g* [95% CI]
Did not report	56	0.92 [0.85, 1.00]	19	−0.51 [−0.92, −0.08]
Study quality				
Weak	706	0.92 [0.88, 0.96]	186	−0.19 [−0.26, −0.12]
Moderate	433	0.88 [0.83, 0.93]	101	−0.12 [−0.19, −0.06]
Strong	47	0.98 [0.89, 1.07]	27	−0.14 [−0.34, 0.06]
Study randomization				
Individual	1,167	0.91 [0.88, 0.94]	311	−0.17 [−0.22, −0.12]
Cluster	19	0.97 [0.90, 1.04]	3[a]	—
Study design				
Parallel	1,182	0.91 [0.88, 0.94]	312	−0.17 [−0.22, −0.12]
Crossover	3[a]	—	2[a]	—
Dynamic waitlist	1[a]	—	0[a]	—

Note. CBT = Cognitive Behavioral Therapy; CT = Cognitive Therapy; DBT = dialectical behavioral therapy; *n* = number of effect sizes; RR = weighted mean risk ratio; SITBs = self-injurious thoughts and behaviors; 95% CI = 95% confidence interval. Dashes indicate unavailable information. Effect estimates significantly different from the pooled effects are bolded.
[a] Estimates are not reported for analyses involving fewer than five effect sizes, because a small number of effect sizes compromises the accuracy of estimates.

interventions may yet produce moderate-to-large reductions in SITBs. We discuss some of these potential fundamental changes in the future directions section below. Should these fundamental changes fail to yield advances, it may be most prudent to accept the hypothesis that SITB interventions are limited to small effects. In this scenario, it may also be prudent to re-examine the ultimate goals and purpose of SITB research. That is, if we have already reached the low ceiling of potential SITB treatment effects, what is the main purpose of further SITB research?

These overall findings are disappointing, but we emphasize that small reductions in SITBs are still important. Any reduction in the harm and death caused by SITBs is meaningful and certainly better than the alternative. But we believe that we are in accord with most other researchers and clinicians when we conclude that, for the sake of public health, we as a field must do much better. Our hope was that moderation analyses would reveal more encouraging findings that may have been obscured by these overall analyses.

Intervention Effects Have Not Improved Over Time

We hypothesized that time would moderate intervention effect size magnitude, such that effects sizes would gradually increase across decades of research. The number of RCTs with a SITB outcome has increased at a near-exponential rate since the first qualifying study in 1970. More than half of qualifying RCTs were published during the past decade. Contrary to our expectations, results revealed that effect size magnitude has not increased across nearly five decades of research.

Several explanations could help to explain a lack of improvement over time. One possibility is that the publication of null effects has become more common over time, which may offset any effect size increases. However, we could not detect strong evidence for this potential explanation: Publication bias did not appear to decrease over time (see Supplement 8 in the online supplemental materials), and the odds of an article publishing null or negative findings did significantly increase over time but would only offset extremely small improvements in efficacy ($b = 0.03$, $p = .02$). For instance, approximately 6% of the effect sizes published in 2000s reported null or negative findings, and this rate increased to only approximately 9% for effect sizes published in 2010s. Another possibility is that aspects of study methods—for example, larger sample sizes, more severe populations, more stringent control groups, and improved trial quality—may have offset any gains in intervention efficacy. But we were unable to detect any significant moderation based on these methodological variables (see Tables 2 and 3, Figures 8–11, and Supplement 6 in the online supplemental materials), indicating that such factors do not appear to explain the lack of improvement over time. Yet another potential explanation is that more studies testing idiosyncratic-and-ineffective interventions have been published over time, decreasing potential gains made by more mainstream interventions (e.g., medications, CBT, DBT). But we could find no evidence that more mainstream interventions produced stronger effects than other interventions (see Table 3, Figure 11), and we could find no evidence that the efficacy of more mainstream interventions has increased over time—in fact, the effects of DBT decreased over time (see Supplement 9 in the online supplemental materials).

Another potential explanation is that the field has not made meaningful gains in its ability to treat and prevent SITBs. The present meta-analysis could find no evidence that we are identifying increasingly important treatment targets or increasingly effective methods of counteracting those targets. Notably, this pattern is consistent with meta-analytic evidence that SITB prediction also has not improved across several decades of research (Franklin et al., 2017) and evidence that the suicide rate has not decreased in many countries with the greatest access to SITB interventions (e.g., the United States; Centers for Disease Control and Prevention, 2018). It is similarly consistent with evidence that, despite large increases in treatment utilization, the rates of nonfatal suicidal thoughts and behaviors have not diminished in the United States (Kessler et al., 2005; Nock et al., 2013; Olfson et al., 2016). In other words, there is much evidence that SITB awareness, research, interventions, and intervention utilization have all dra-

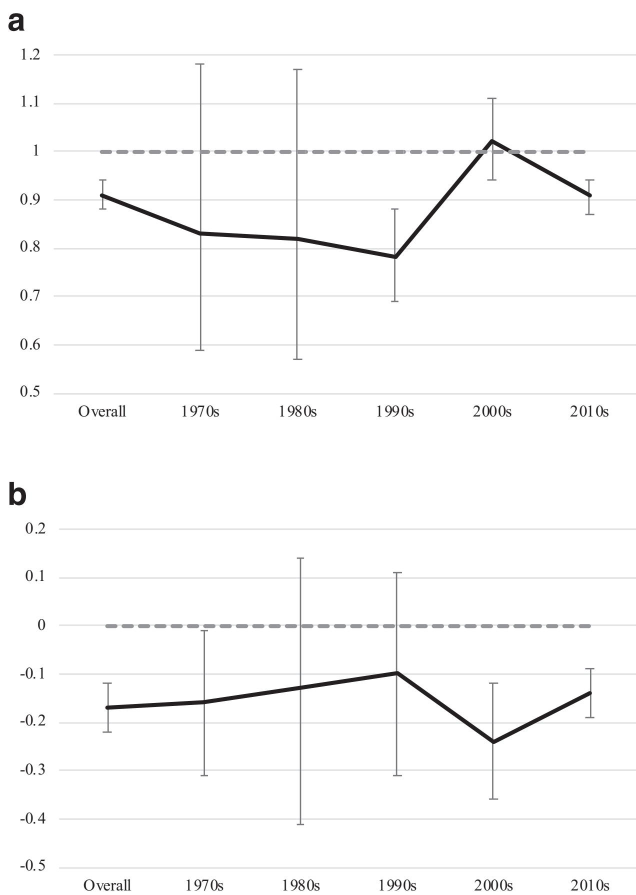

Figure 7. Posttreatment effects by decade. (a) Weighted odds ratios. (b) Weighted Hedges's *g*s. Dashed line indicates null effect (i.e., risk ratio = 1); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line.

matically increased across the past several decades. But because SITB intervention efficacy has not improved, these other gains have translated into limited effects on SITB rates. We conclude that this suggests the need for a fundamental change in how the field approaches SITB interventions.

All SITB Intervention Types Produce Similarly Small Effects

We hypothesized that certain SITB intervention types would produce much stronger effects than others. Although overall intervention effects were small, we believed that these overall analyses hid a few powerful effects. In particular, we believed that broad and intensive behaviorally based interventions designed for people with SITBs (e.g., dialectical behavioral therapy) would produce the strongest effects.

Reflecting the great diversity of intervention types applied to SITBs, a wide range of unique interventions were tested within qualifying RCTs. Psychoactive medications accounted for the largest proportion of intervention types (46.72%), CT and CBT (11.68%), DBT (6.99%), and combined medication and psychotherapy (6.25%), and several smaller types that accounted for less

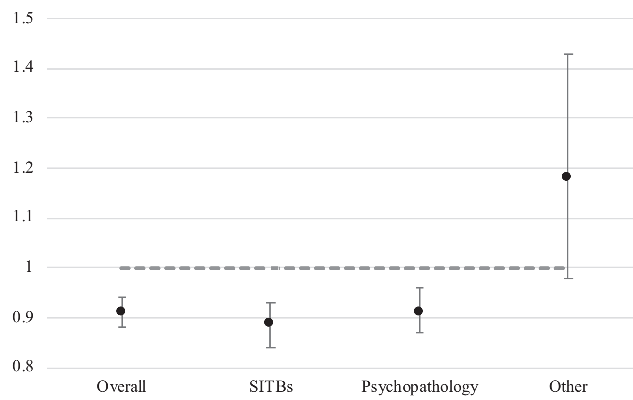

Figure 8. Posttreatment weighted risk ratios by intervention target type. Dashed line indicates null effect (i.e., risk ratio = 1); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. SITBs = self-injurious thoughts and behaviors.

than 5% each. Given how commonly they are applied to SITBs outside of studies, it is surprising that interventions such as acute psychiatric hospitalization, gatekeeper and peer support programs, electroconvulsive therapy, social outreach/helplines, safety planning and means safety/restriction, multilevel eclectic approaches, and psychoanalysis have rarely been evaluated within RCTs with SITBs as an outcome.

It is also notable that few interventions tested within RCTs were derived directly from SITB theories. Many SITB theory articles as well as articles purported to evaluate those theories often emphasize their treatment implications. For instance, a given theory may propose that constructs X, Y, and Z are the major causes of SITBs and recommend that interventions target X, Y, and Z. Surprisingly, we found few interventions developed on the basis of SITB theories. For example, only one article (Hill, 2015) examined interventions derived directly from the Interpersonal Theory of Suicide (Joiner, 2005; Van Orden et al., 2010). The most commonly tested interventions are modified versions of interventions originally

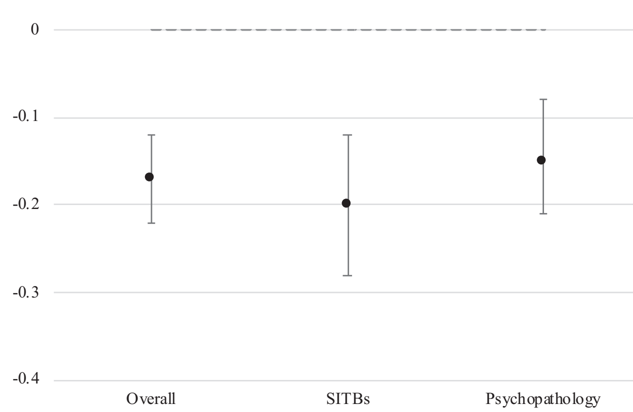

Figure 9. Posttreatment weighted Hedges's *g* by intervention target type. Dashed line indicates null effect (i.e., Hedges's $g = 0$); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. SITBs = self-injurious thoughts and behaviors.

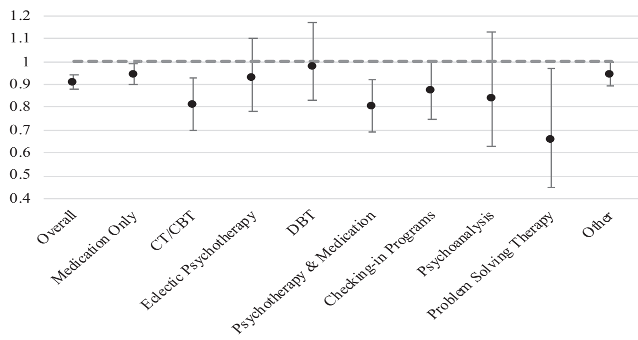

Figure 10. Posttreatment weighted risk ratios by specific intervention type. Dashed line indicates null effect (i.e., risk ratio = 1); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. CT/CBT = Cognitive Therapy/Cognitive Behavioral Therapy; DBT = dialectical behavior therapy.

intended for something else and are largely disconnected from the most popular SITB theories. Unlike many areas of medical and clinical science, the most popular SITB theories do not seem to correspond with the most popular SITB interventions, highlighting a key disconnect between theoretical and clinical research. Improvements in SITB intervention efficacy will likely require greater communication and collaboration among theoretical, experimental, and intervention researchers.

Once again contrary to our expectations, effect sizes were similarly small for all intervention types across both binary and continuous analyses. Regardless of whether interventions were simple (e.g., administration of a psychoactive drug), complicated (e.g., multifaceted approaches such as dialectical behavior therapy), developed primarily with SITBs in mind (e.g., dialectical behavior therapy), or developed primarily for some other purpose (e.g., antipsychotics), effect sizes were small. No intervention was significantly and consistently stronger than any other in terms of overall effects on SITBs. Several interventions may benefit from more study because they have been rarely tested within an RCT format (e.g., means safety/restriction, acute hospitalization, gatekeeper programs). However, the published RCTs on these interventions that do exist indicate that they are not significantly stronger than any other intervention type. In sum, a wide range of interventions produced similarly small reductions in SITBs, and no intervention type produced reliably moderate or large reductions. As with the above major findings, we reason that this general pattern suggests the need for fundamental changes in how we approach SITB interventions.

We intentionally focus on broader effects on overall SITBs because these effects are most reliable and space limitations prevent an exhaustive account of more specific effects. But we understand that many readers may be interested in more specific moderation analyses, particularly those concerning popular interventions. We note a few of these effects here, but strongly caution that these estimates may have limited reliability because of a relatively small number of effect sizes for such specific analyses. These results, along with many other more fine-grained analyses, are presented in Supplement 6 in the online supplemental materials.

We hypothesized that DBT and CT/CBT would be particularly effective SITB interventions. Overall analyses did not support this hypothesis (see above), and neither did finer-grained analyses. DBT only significantly reduced the severity/intensity of the combined SITB outcome (i.e., when studies included a wide range of SITBs in a single variable) and self-harm and marginally reduced the occurrence of SITB-related hospitalizations. DBT did not significantly reduce suicide ideation, suicide attempts, or NSSI. No qualifying study tested DBT's effects on suicide death. CT/CBT only reduced the combined SITB outcome (and only the continuous version, not the binary version) and suicide ideation (once again, only the continuous version, not the binary version). CT/CBT did not significantly reduce suicide attempts or suicide death. There were too few effect sizes to estimate CT/CBT's effects on self-harm, NSSI and SITB-related hospitalization. Checking-in programs produced a significant reduction in suicide ideation, although particular caution is needed for this analysis because it was only based on five effect sizes. Similarly, checking-in pro-

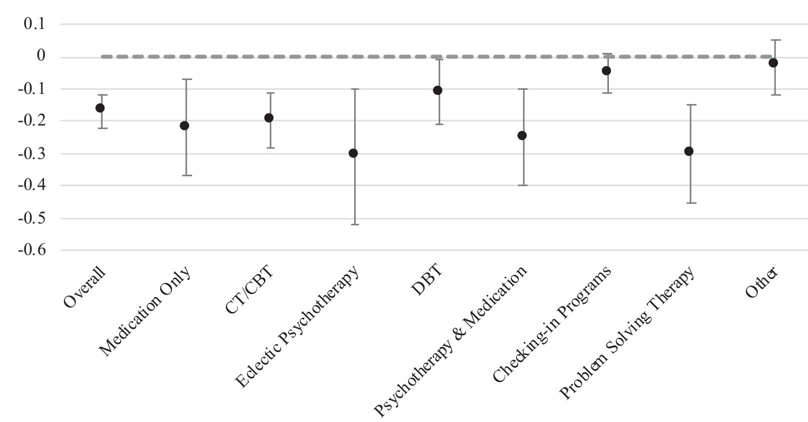

Figure 11. Posttreatment weighted Hedges's *g* by specific intervention type. Dashed line indicates null effect (i.e., Hedges's *g* = 0); error bars represent 95% confidence intervals. Effects were significant if confidence intervals do not overlap with the dashed line. CT/CBT = Cognitive Therapy/Cognitive Behavioral Therapy; DBT = dialectical behavior therapy.

grams also significantly reduced suicide attempts based on six effect sizes. These programs did not significantly reduce self-harm, and there were too few effect sizes to estimate effects on other outcomes. When pooling across outcomes, checking-in programs were associated with worse outcomes overall. In sum, outside of a few effects (e.g., DBT for self-harm; checking-in programs for suicide ideation and attempts), no specific intervention stood out as particularly helpful for any specific outcome.

Interventions Produce Similar Effects Regardless of the Primary Outcome Target

Some have proposed that interventions that primarily target SITBs (and SITB-specific causes) should be more efficacious than interventions that secondarily target SITBs. Others have proposed that it is not necessary to target SITBs to produce large reductions in SITBs—targeting phenomena such as depression should be sufficient. This disagreement stems from a larger disagreement about whether SITBs are simply an extension of psychopathology (in which case resolving psychopathology would resolve SITBs) or originate from unique causes (in which case resolving psychopathology would have little, or weaker, effect on SITBs). Most effect sizes were from treatments primarily targeting psychopathology (63.48%) and SITBs (29.76%), with miscellaneous targets accounting for the remaining RCTs. Results revealed that intervention effects were similarly small for interventions primarily targeting psychopathology or SITBs.

At first glance, this finding would seem to support the notion that, to reduce SITBs, it is not necessary to directly target SITBs or SITB-specific causes. But this conclusion would neglect a crucial aspect of this finding: *neither type of primary outcome target was associated with large or even moderate reductions in SITBs*. That is, neither approach has worked well. These results support the view of the SITB-specific proponents that primarily targeting psychopathology (or other outcomes) is not sufficient to produce strong reductions in SITBs because these interventions do not impact the central causes of SITBs. But these results also show that SITB-focused RCTs fared no better, indicating that—whatever the central causes of SITBs are—neither nonspecific nor SITB-specific interventions have successfully targeted them. As above, we conclude that these findings suggest the need for fundamental changes in our approach to SITB interventions.

Intervention Effects Were Small Across Age Groups

Self-injurious thoughts and behaviors occur across the life span, with different theories and intervention approaches often being applied depending on whether the patient is a child/adolescent, adult, or older adult. But it has been unclear whether SITB interventions are differentially efficacious across these groups. Most effect sizes were derived from RCTs including only adults (38.09%), with many fewer studies including solely children/adolescents (21.11%) or older adults (0.78%). Nearly one fourth of the effect sizes included a mix of these age groups (24.90%). Results revealed that intervention effects were small across all age groups. In binary analyses, effect sizes were particularly small for child/adolescent studies, though this moderation effect was not replicated in continuous analyses.

Paralleling other results, effects were small across all groups. It is important to note, however, that an age group with especially high rates of NSSI and suicide attempts (adolescents) was rarely studied. In addition to elevated prevalence, adolescence also offers a critical period for intervention (Glenn et al., 2017; Goldston et al., 2015). It is also notable that an age group with an especially high rate of suicide death (older adults) was also rarely the focus of RCTs. The field would benefit from more RCTs focused specifically on these two groups.

Small Intervention Effects Were Maintained at the Initial Follow-Up

Most RCTs are primarily concerned with effects at the end of the intervention (e.g., the final week of an 8-week course of psychoactive medication). Presumably, this is the point at which intervention effects are strongest, although such effects may be detected best at later time points for upstream interventions (e.g., school-based programs) and for rare outcomes like suicide death. Results revealed that effects for suicide ideation, suicide death, NSSI, visits and hospitalizations attributable to SITBs, and other/combined SITBs at the initial follow-up were not significantly different from effects immediately posttreatment. Effects for certain outcomes (e.g., suicide attempt in binary form, and self-harm in continuous form) were significantly larger at the initial follow-up than immediately posttreatment. Given that results indicated that overall intervention effects were generally small, this finding is surprising and promising. It suggests that the small benefits of SITB interventions tend to endure beyond the intervention itself.

Study and Sample Characteristics Did Not Moderate Intervention Effects

Study and sample characteristics can have large impacts on intervention effects. The present meta-analysis, however, did not detect any such consistent impacts on SITB intervention effects. Effects were largely statistically equivalent regardless of the following: sample type (i.e., prior SITB sample, psychopathology sample, other sample), sample size, intervention length, therapy components (e.g., individual, group, school-based, etc.), blindedness (i.e., nonblind, single-blind, double-blind), control group type (e.g., waitlist, treatment as usual, active control), or study quality.

We emphasize that the present findings should not be interpreted as indicating that these characteristics are unimportant. Rather, the present findings should be interpreted as showing that these characteristics are unimportant *when interventions effect sizes are small and contained within a restricted range*. If the literature had produced a wide range of effect sizes, we believe that many of these characteristics may emerge as important moderators. The constricted range of intervention effects again highlights the need for a fundamental change in how we approach SITB interventions.

Limitations

The findings of the present meta-analysis should be interpreted in light of its limitations, which fall into four major categories. First, several features of our meta-analytic method biased results toward larger intervention effects, meaning that the above analyses may reflect optimistic estimates of true effects. One such feature is that all effect sizes were treated as independent even though most RCTs produced multiple effect sizes (and thus, were likely partially dependent). We adopted this strategy because there was insufficient information to determine the degree of dependence among effect sizes. To estimate inflation from this strategy, we conducted a separate set of analyses assuming complete dependence among effect sizes from the same study (Supplement 2 in the online supplemental materials). These effects were similar to those assuming complete independence; true effects are likely between those reported above and these complete-dependence effect estimations. Another such feature of the present meta-analysis that may have inflated effects was the decision to only include published RCTs. We adopted this strategy because we aimed to summarize the most rigorously evaluated and publicly available information about SITB interventions. Unpublished studies tend to include smaller effects and more null effects than published studies; omitting these tends to inflate meta-analytic effect estimates. To adjust for this, we conducted publication bias analyses, which indicated minimal bias for binary analyses and low-to-moderate bias for continuous analyses. Despite these limitations, the overarching conclusion of the present meta-analysis remains the same: SITB intervention effects are small.

Second, as with most meta-analyses, we likely did not include all studies that met our general inclusion criteria (i.e., published RCT with a SITB outcome). We conducted extensive systematic searches across multiple databases and searched other articles for citations of relevant articles we may have missed. However, it is of course likely that some relevant RCTs were missed. Similarly, we located several qualifying studies that did not include sufficient statistical information to allow for inclusion. After contacting the authors of these articles, we were unable to obtain the necessary information and, correspondingly, unable to include those studies. Nevertheless, we included far more articles—591—than we had anticipated at the outset of this project, especially considering that prior systematic reviews of this literature concluded that there were very few RCTs on this topic (e.g., Goldney, 2005; 18 RCTs included in Mann et al., 2005). Given the consistency of effect sizes across time, outcome, intervention type, study sample, and several other moderators, we reason that it is unlikely that findings would have meaningfully changed if these omitted studies had been included.

Third, it is possible that some interventions do produce moderate or large effects, but they were not included in the present meta-analysis because of a lack of RCTs for these interventions. For example, there are very few RCTs of common interventions such as means safety and restriction, acute hospitalization, multilevel eclectic approaches, helplines and other social outreach approaches, and electroconvulsive therapy. It is difficult to evaluate some of these within traditional RCT designs. For instance, means safety and restriction interventions are often implemented on a population level by broad regulations (e.g., fences on bridges, firearm laws), precluding randomization and equivalent control groups. Similarly, interventions such as helplines are typically anonymous, making later assessments of intervention effects challenging. Likewise, given legal requirements in most jurisdictions, it is generally not possible or ethical to randomize some people to receive acute psychiatric hospitalization and some people to a control group. For certain interventions (e.g., electroconvulsive therapy), it is also difficult to create a sufficient control condition. These issues are challenging, but there are possible ways around these obstacles. Indeed, a small number of RCTs have examined means restriction/safety programs and acute psychiatric hospitalization. Notably, these interventions produced effects that are similar to those of more commonly studied interventions. But the present meta-analysis remains limited in that it cannot speak to the efficacy of interventions that have not been studied extensively within RCTs (e.g., firearm regulations).

Fourth, the present meta-analysis cannot account for potential idiographic effects of interventions. For example, it may be that cognitive–behavioral therapy works extremely well for 10% of people at-risk for SITBs but not at all for anyone else. That is, intervention effects may be consistently large for some people and consistently null for others. The aggregate effect of this, which is what all extant studies report, would be a small group effect. Although such idiographic effects cannot be ruled out by this meta-analysis, there is currently no direct evidence that certain interventions consistently work well for specific subsets of people and poorly for others. Such effects cannot be reliably determined from post hoc analyses (because such analyses would be circular); they must be determined from studies with specialized RCT designs and a priori identification of the individuals hypothesized to benefit or not benefit from the intervention. In any case, because of this limitation, the present meta-analysis can speak to nomothetic intervention effects but not potential idiographic intervention effects.

Major Future Directions

Although the present meta-analysis produced several disappointing findings, it also establishes a foundation that informs future research. In this section we outline six major future directions inspired by these meta-analytic results.

Test commonly utilized SITB interventions within RCT designs. Many of the most commonly applied SITB interventions in the real world (vs. academic studies) have rarely or never been tested within RCTs. For example, means safety and restriction, gatekeeper programs, helplines, and acute hospitalization are common SITB interventions in the real world, but few RCTs have evaluated these interventions. This is important because RCTs can control for many more alternative explanations for an intervention effect compared with naturalistic, interrupted time series, or quasi-experimental designs. Notably, many have argued that there are viable alternative explanations for the effects of non-RCT tests of these interventions (e.g., Bateman, 2009), and the few RCTs that exist on these interventions show small effects (e.g., Almeida et al., 2012; Pearson et al., 2017; Waterhouse & Platt, 1990). The field would greatly benefit from future work that employs creative approaches to safely, ethically, and effectively evaluate these hard-to-test interventions. Future research should also consider whether some of these interventions, like means safety and restriction, may be more viable in some countries than others. For example, interventions targeting means restriction may be more viable in countries where the most common method for attempting suicide is firearm or overdose/self-poisoning by a controllable substance, rather than death by hanging and overdose by commonly available/difficult to restrict substances.

Conduct studies that are sufficiently powered to reliably estimate intervention effects on suicide attempt and suicide death. Suicide attempts and death are the most serious forms of SITBs. They are also the hardest to study because they are the rarest and require the most resources to manage. As a result, most RCTs that include these behaviors as outcomes were underpowered to reliably estimate intervention effects. To reliably detect group differences in such a low base-rate outcome, studies would need to include thousands of participants for suicide attempt outcomes and tens of thousands of participants for suicide death outcomes. This would require extensive time, funding, and resources that are just not available for most RCTs. Unfortunately, even when summarizing across studies as we do in the present meta-analysis, we are likely *still* underpowered to detect significant differences across treatment and control groups in suicide deaths. This means that, after nearly 50 years of research, we still lack solid empirical information on the degree to which SITB interventions reduce suicide attempts and death. Future research would benefit from prioritizing studies sufficiently powered to accurately estimate effects on these two SITBs. Multiteam studies across multiple sites may be best suited for this endeavor.

Increase research on adolescent and elderly populations. Adolescents are particularly vulnerable to NSSI and suicide attempts (Glenn et al., 2017; Goldston et al., 2015), and elderly populations are at particular risk for suicide death (Hedegaard, Curtin, & Warner, 2020). In stark contrast to the prevalence rates in these populations, very few RCTs have been conducted with adolescent and elderly samples. Future studies are needed to fill this research gap.

Demonstrate intervention target engagement. Few studies have shown that their intervention effects are attributable to successful target engagement. For example, we are not aware of any studies that have demonstrated that selective serotonin reuptake inhibitors reduce SITBs by increasing serotonin levels. Including such tests in future studies would facilitate an understanding of *why* certain interventions work or do not work.

Prioritize research that can shed light on common necessary causes of SITBs. It is *possible* to develop an effective intervention for a pathology without understanding its causes. But this approach involves a lot of guessing and checking and can take a very long time. It is similar to attempting to sail to Australia from Portugal without a map or navigational equipment—it is possible, but extremely difficult and time-consuming. As the present results demonstrate, after nearly 50 years of trying, the SITB field has not hit upon a highly or even moderately efficacious intervention. With millions of people affected by SITBs each year, we cannot afford to continue guessing at SITB causes or proposing SITB causes based on correlational or longitudinal evidence. We reason that it is time to prioritize research that sheds light on the causes of SITBs. It is important to note that only experimental evidence allows for reasonable causal inferences—correlational and predictive evidence is not sufficient (see Franklin, Huang, Fox, & Ribeiro, 2018 for a discussion). It is likewise important to note that intervention RCTs—despite being experimental studies—also do not permit inferences about the causes of SITBs (i.e., *ex juvantibus* reasoning; cf. Lilienfeld, Smith, & Watts, 2016). This is because intervention studies test the causal effect of an intervention on SITBs; they do not test the potential causes of SITBs themselves (Franklin et al., 2018).

To test for the causes of SITBs, we need experimental studies that evaluate whether a particular manipulation makes SITBs more likely. It is obviously not feasible to conduct such studies on actual suicidal behaviors. But it is now possible to evaluate ideas about suicide causes within laboratory approximations of suicide (e.g., Franklin, Huang, & Bastidas, 2019). We note that, assuming proper risk management, it has always been possible to conduct experimental studies to cause suicide ideation in the lab (e.g., Chatard & Selimbegović, 2011), though such studies have been rare. It is similarly possible to conduct studies that make close approximations of NSSI more likely (e.g., Bresin & Gordon, 2013; Hooley, Ho, Slater, & Lockshin, 2010; Ludäscher et al., 2009). Prioritization of such work would advance knowledge about SITB causes, facilitating the identification of potentially important intervention targets.

We emphasize that experimental work should seek to identify common necessary causes of SITBs. Common necessary causes are causes that must be present for a large proportion of a given phenomenon to occur. For example, sunlight is a common necessary cause of plant growth for a large proportion of plants. Without sunlight, most plants would die. It is important to distinguish necessary causes from sufficient causes. Sufficient causes are causes that guarantee a phenomenon. Because sunlight does not guarantee plant growth (sunlight in the absence of water and other nutrients would not produce plant growth), it is not a sufficient cause of plant growth. It is also important to distinguish common necessary causes from what philosopher J. L. Mackie (1965) termed INUS conditions. These are Insufficient but Nonredundant parts of an Unnecessary but Sufficient (i.e., INUS) conditions for a phenomenon to occur. In other words, INUS conditions are causes that are small pieces of much larger causal webs that contribute to a small proportion of instances of a given phenomenon. For example, a lit cigarette is an INUS condition for a house fire. On its own, a lit cigarette cannot cause a house fire (i.e., it is an insufficient cause) and most house fires do not involve a lit cigarette (i.e., it is an unnecessary cause in most house fire instances). But lit cigarettes, in conjunction with many other factors, do play a small causal role in a small proportion of house fires.

Mackie (1965) argued that most of what we refer to as *causes* are in fact *INUS conditions*. It is similarly likely that most of what we consider SITB causes (e.g., the economy, unemployment, depressive symptoms, stress, neurotransmitter levels) are in fact INUS conditions for SITBs—they play small causal roles in a small proportion of SITB instances. Consistent with this view, hundreds of factors predict SITBs slightly better than chance, but none predict much more accurately than this (Fox et al., 2015; Franklin et al., 2017). This is because each predictor plays a small role in a small proportion of SITB instances. The present meta-analytic results are also consistent with this view. Because a large number of intervention targets play small causal roles in a small proportion of SITB instances, many interventions produce small SITB reductions, but none produce moderate or large reductions. We accordingly recommend that researchers go beyond identifying INUS conditions in their experimental SITB research—there are likely thousands of INUS conditions, each of which may generate small SITB reductions when targeted. We reason that researchers should instead seek to identify common necessary causes of SITBs. That is, researchers should attempt to identify causes that must be present for a large proportion of SITBs to occur. These are likely small in number but have the potential to generate moderate or large SITB reductions when targeted. We note a few candidate necessary SITB causes in the next section, but encourage researchers to go beyond this illustrative candidate set.

Develop interventions to target common necessary causes of SITBs. Over the past several decades, hundreds of different interventions have been applied to SITBs and all have produced similarly small effects. As noted above, we reason that this pattern is explained by existing interventions targeting INUS conditions. To change this pattern, we must fundamentally change how we approach SITB interventions—focusing on identifying and targeting common necessary causes rather than INUS conditions. Here we note three candidate common necessary causes for further study.

First and most familiar, the means of engaging in SITBs represent a potential type of common necessary SITB cause. For example, without a gun or ammunition, it is impossible to shoot oneself. The more common the means, the greater the proportion of SITB instances in which it plays a necessary causal role—for example, firearms play a necessary causal role in nearly half of suicide deaths in the United States (Centers for Disease Control and Prevention, 2018). Unfortunately, targeting this candidate presents multiple difficulties. For example, it is physically impossible to eliminate means in many situations (e.g., a high place for jumping, sharp object for cutting, a bag for suffocation, medications for overdoses, lighters for burning, etc.) and politically impossible to eliminate means in some situations (e.g., firearms in the United States). There is also mixed evidence on whether limiting one form of means leads to switching to alternative means (Daigle, 2005; Kreitman, 1976; Yip et al., 2012), and the few RCTs on this approach raise doubts about its potential (Almeida et al., 2012; Pearson et al., 2017; Waterhouse & Platt, 1990). Nonetheless, relatively little work has focused on this candidate, and many questions and possibilities remain. Studies on the topic, especially from a psychological perspective (vs. public health perspective), have been rare in part because of tradition and in part because of challenges associated with conducting this research. But we believe that these challenges are solvable and that the fruits of the research may be considerable.

Second, SITB concepts represent a potential type of common necessary SITB cause (Franklin, 2019). Concepts are collections of experiences tied together by language (i.e., a word or closely related words; Lindquist, MacCormack, & Shablack, 2015). We learn concepts throughout childhood and adolescence and continue to modify and add new concepts throughout adulthood (Hoemann, Xu, & Barrett, 2019). We automatically and effortlessly employ concepts to make sense of what we perceive, what we feel, and what we think, and to direct our behavior (Barrett, 2009, 2012, 2017; Hoemann & Feldman Barrett, 2019; Lindquist, 2013). For example, at this moment, you are effortlessly applying concepts to make sense of the marks on your screen as words with specific meanings. To experience a given psychological phenomenon, one must possess the corresponding concept for the phenomenon. For example, a technique called semantic satiation (e.g., repeating a word 30 times) temporarily disrupts a concept. If one repeats "anger" 30 times, it temporarily disrupts their ability to experience anger and perceive angry faces (e.g., Gendron, Lindquist, Barsalou, & Barrett, 2012; Lindquist, Barrett, Bliss-Moreau, & Russell, 2006). Concepts can be permanently disrupted by conditions such as semantic dementia. In such cases, individuals can no longer distinguish between sad, angry, fearful, and so forth, faces because they no longer possess the concepts for these emotions (e.g., Lindquist, Gendron, Barrett, & Dickerson, 2014).

This work suggests that to experience a SITB-related perception, emotion, cognition, or behavior, one must possess the corresponding SITB-related concept. In the absence of such concepts, it should not be possible for someone to experience SITBs. This is consistent with evidence on suicide concept maturation. Initial studies indicate that the suicide concept matures in late childhood and early adolescence (Cuddy-Casey & Orvaschel, 1997; Mishara, 1999). Rates of suicidality track with this concept maturation, with extremely low rates in early childhood, and adolescent rates approaching those of adults (Nock et al., 2013). We emphasize here that this perspective does not propose that a suicide concept is *sufficient* for suicidality (i.e., it is not the case that if you have a suicide concept, you will engage in suicidality), only that it is *necessary* for suicidality (i.e., only people with suicide concepts can engage in suicidality). Much more work is needed to evaluate this candidate, but given its grounding in basic psychological science and neuroscience (e.g., Barrett, 2017) and consistency with initial SITB work (e.g., Mishara, 1999; Nock et al., 2013), it could hold promise. If future work supports SITB-related concepts as common necessary causes of SITBs, there are several potential ways to target these concepts. For example, it may be possible to disrupt SITB concepts by disrupting their reconsolidation with ketamine, propranolol, or propofol. Recent studies indicate that these drugs disrupt the reconsolidation of long-held associations, producing long-term reductions in behaviors such as alcohol use (e.g., Das et al., 2019). Such procedures may similarly disrupt SITB concepts, producing long-term reductions in SITBs.

Third, a common necessary cause candidate specific to behavior is how people conceptualize the allostatic consequences of self-injurious behavior (Huang, Funsch, Park, & Franklin, in press). Work from basic psychological science, neuroscience, and physiology indicates that all behaviors are motivated by allostasis (Barrett, 2017; Sterling, 2012; Touroutoglou, Andreano, Adebayo, Lyons, & Barrett, 2019), which is the body's attempt to maximize energy efficiency by automatically anticipating the energy-related costs and benefits of behaviors. On a psychological level, energy inefficiency is represented as negative affect and energy efficiency is represented as positive affect (Barrett, 2017; Sterling, 2012). On a psychological level, this perspective proposes that people engage in behaviors that they anticipate will increase positive affect and decrease negative affect. Following this logic, a common necessary cause of self-injurious behavior may be the conceptualization that these behaviors will increase positive affect and decrease negative affect. This places the emphasis on the anticipated consequences of these behaviors rather than the antecedents. This emphasis has been a major part of NSSI research for decades (e.g., Nock & Prinstein, 2004), with experimental work demonstrating that NSSI does in fact increase positive affect and decrease negative affect (e.g., Franklin et al., 2013). However, NSSI interventions have primarily focused on reducing antecedents of NSSI (e.g., negative affect) rather than targeting its anticipated consequences.

Suicide research and interventions have seldom focused on the anticipated affective consequences of suicidal behavior. Recent experimental work indicates that these anticipated consequences may be a necessary cause in many instances of suicidal behavior. Using a VR suicide paradigm, Huang et al. (in press) found that negative affect induced by a stressful speech and social rejection did not cause an increase in VR suicide compared with the control group (i.e., VR suicide rates ~5% in the control group and stress groups). However, the anticipation that VR suicide would allow one to avoid a future stressful speech task caused a large increase in VR suicide (~30% VR suicide rate). Such findings suggest that interventions that cause individuals to conceptualize the consequences of suicidal behavior as less appealing may be effective. For example, an intervention might employ persuasive techniques to convince the individual that the consequences of suicidal behavior are likely to be extremely negative rather than beneficial (e.g., surviving but being maimed or disabled; terrible effects on family and friends). A recent experimental study is consistent with this possibility. Linthicum, Harris, and Ribeiro (2019) found that even small shifts in the anticipated probability of rewarding versus punishing consequences of engaging in VR suicide caused significant decreases in VR suicide rates.

Each of these three candidate common necessary SITB causes requires much further study, and we do not mean to imply that these are the only possible candidates. Our aim is for the above to illustrate examples of how we might refocus research in a way that could potentially generate more efficacious interventions. We encourage researchers to explore these candidates and to develop other candidates.

For Now, What Should Clinicians Do?

The present meta-analysis indicates that many existing interventions produce small reductions in SITBs. As noted above, this broad finding should prompt many new research directions. But where does this leave clinicians who are tasked with managing patients with SITBs *today*? Based on our read of the evidence, we have three interrelated recommendations for clinicians. First, recognize the limits of existing interventions. There is no robust evidence for highly efficacious intervention for SITBs and there is no consistent evidence that particular interventions work much better for certain SITB phenomena or for certain populations. Second, understand that there are many different interventions that can produce small group-level reductions in SITBs. It is important to connect at-risk individuals to one of these interventions. Third, consider applying the most scalable of these interventions to reach the most people in the most cost-effective manner. Nearly all existing interventions tested within RCT studies produce similar effects, meaning that short, cheap, and easily accessible interventions appear to be just as efficacious as long, expensive, and difficult-to-access interventions. To maximize SITB reductions, we accordingly recommend that—for now—clinicians disseminate the most scalable existing interventions.

Conclusion

After nearly 50 years of research and at least 1,125 unique RCTs, the literature of SITB RCTs indicates that existing interventions have small effects, this has not improved over time, and no known factor meaningfully moderates this pattern. These findings are disappointing, but they are consistent with recent meta-analyses showing poor SITB prediction (e.g., Fox et al., 2015; Franklin et al., 2017) and evidence that SITB rates have not meaningfully declined across the past several decades (e.g., Centers for Disease Control and Prevention, 2018). These findings also point the way for important future directions that may generate improvements in SITB intervention efficacy. We hope that this summary of the literature serves as a turning point, inspiring researchers and institutions to move in new and potentially more fruitful directions.

References

Adams, C. E., Houle, T. T., Parker, J. D., & Burke, R. S. (2012). Examining changes in depressive symptoms during substance abuse treatment in the context of regression to the mean. *Addictive Disorders & Their Treatment, 11,* 183–194. http://dx.doi.org/10.1097/ADT.0b013e31824709bb

Alexander, K. C. (1990). Communicating with potential adolescent suicides through poetry. *The Arts in Psychotherapy, 17,* 125–130. http://dx.doi.org/10.1016/0197-4556(90)90022-I

Almeida, O. P., Pirkis, J., Kerse, N., Sim, M., Flicker, L., Snowdon, J., . . . Pfaff, J. J. (2012). A randomized trial to reduce the prevalence of depression and self-harm behavior in older primary care patients. *Annals of Family Medicine, 10,* 347–356. http://dx.doi.org/10.1370/afm.1368

American Foundation for Suicide Prevention. (2015). *Project 2025*. Retrieved from. https://project2025.afsp.org/

Arkowitz, H., & Lilienfeld, S. O. (2006). Psychotherapy on trial. *Scientific American Mind, 17,* 42–49. http://dx.doi.org/10.1038/scientificamericanmind0406-42

Armijo-Olivo, S., Ospina, M., da Costa, B. R., Egger, M., Saltaji, H., Fuentes, J., . . . Cummings, G. G. (2014). Poor reliability between Cochrane reviewers and blinded external reviewers when applying the Cochrane risk of bias tool in physical therapy trials. *PLoS ONE, 9,* e96920. http://dx.doi.org/10.1371/journal.pone.0096920

Baastrup, P. C., Poulsen, J. C., Schou, M., Thomsen, K., & Amdisen, A. (1970). Prophylactic lithium: Double blind discontinuation in manic-depressive and recurrent-depressive disorders. *The Lancet, 266,* 326–330. http://dx.doi.org/10.1016/S0140-6736(70)92870-9

Bagley, C. (1968). The evaluation of a suicide prevention scheme by an ecological method. *Social Science & Medicine, 2,* 1–14. http://dx.doi.org/10.1016/0037-7856(68)90097-8

Baker, S. T. E., Nicholas, J., Shand, F., Green, R., & Christensen, H. (2018). A comparison of multi-component systems approaches to suicide prevention. *Australasian Psychiatry, 26,* 128–131. http://dx.doi.org/10.1177/1039856217743888

Bandelow, B., Sagebiel, A., Belz, M., Görlich, Y., Michaelis, S., & Wedekind, D. (2018). Enduring effects of psychological treatments for anxiety disorders: Meta-analysis of follow-up studies. *The British Journal of Psychiatry, 212,* 333–338. http://dx.doi.org/10.1192/bjp.2018.49

Barnett, D. C. (1982). A suicide prevention incident involving use of the computer. *Professional Psychology, 13,* 565–570. http://dx.doi.org/10.1037/0735-7028.13.4.565

Barrett, L. F. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition and Emotion, 23,* 1284–1306. http://dx.doi.org/10.1080/02699930902985894

Barrett, L. F. (2012). Emotions are real. *Emotion, 12,* 413–429. http://dx.doi.org/10.1037/a0027555

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12,* 1833–1844. http://dx.doi.org/10.1093/scan/nsx060

Bateman, D. N. (2009). Limiting paracetamol pack size: Has it worked in the U.K.? *Clinical Toxicology, 47,* 536–541. http://dx.doi.org/10.1080/15563650903093192

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York, NY: International Universities Press.

Blumenthal, S. J., & Kupfer, D. J. (1988). Overview of early detection and treatment strategies for suicidal behavior in young people. *Journal of Youth and Adolescence, 17,* 1–23. http://dx.doi.org/10.1007/BF01538721

Bolton, J. M., Gunnell, D., & Turecki, G. (2015). Suicide risk assessment and intervention in people with mental illness. *British Medical Journal, 351,* h4978. http://dx.doi.org/10.1136/bmj.h4978

Bradburn, M. J., Deeks, J. J., Berlin, J. A., & Russell Localio, A. (2007). Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. *Statistics in Medicine, 26,* 53–77. http://dx.doi.org/10.1002/sim.2528

Bresin, K., & Gordon, K. H. (2013). Changes in negative affect following pain (vs. nonpainful) stimulation in individuals with and without a history of nonsuicidal self-injury. *Personality Disorders, 4,* 62–66. http://dx.doi.org/10.1037/a0025736

Brophy, J. J. (1967). Suicide attempts with psychotherapeutic drugs. *Archives of General Psychiatry, 17,* 652–657. http://dx.doi.org/10.1001/archpsyc.1967.01730300012003

Brown, G. K., Ten Have, T., Henriques, G. R., Xie, S. X., Hollander, J. E., & Beck, A. T. (2005). Cognitive therapy for the prevention of suicide attempts: A randomized controlled trial. *Journal of the American Medical Association, 294,* 563–570. http://dx.doi.org/10.1001/jama.294.5.563

Browning, C. H. (1974). Suicide, firearms, and public health. *American Journal of Public Health, 64,* 313–317. http://dx.doi.org/10.2105/AJPH.64.4.313

Calati, R., & Courtet, P. (2016). Is psychotherapy effective for reducing suicide attempt and non-suicidal self-injury rates? Meta-analysis and meta-regression of literature data. *Journal of Psychiatric Research, 79,* 8–20. http://dx.doi.org/10.1016/j.jpsychires.2016.04.003

Calear, A. L., Christensen, H., Freeman, A., Fenton, K., Busby Grant, J., van Spijker, B., & Donker, T. (2016). A systematic review of psychosocial suicide prevention interventions for youth. *European Child & Adolescent Psychiatry, 25,* 467–482. http://dx.doi.org/10.1007/s00787-015-0783-4

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., . . . Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61,* 27–36. http://dx.doi.org/10.1016/j.janxdis.2018.08.003

Centers for Disease Control and Prevention. (2018). *Suicide rates rising across the U. S.* Retrieved from https://www.cdc.gov/media/releases/2018/p0607-suicide-prevention.html

Chapman, M. (1971). Movement therapy in the treatment of suicidal patients. *Perspectives in Psychiatric Care, 9,* 119–122. http://dx.doi.org/10.1111/j.1744-6163.1971.tb01083.x

Chatard, A., & Selimbegović, L. (2011). When self-destructive thoughts flash through the mind: Failure to meet standards affects the accessibility of suicide-related thoughts. *Journal of Personality and Social Psychology, 100,* 587–605. http://dx.doi.org/10.1037/a0022461

Chung, D. T., Ryan, C. J., Hadzi-Pavlovic, D., Singh, S. P., Stanton, C., & Large, M. M. (2017). Suicide rates after discharge from psychiatric facilities: A systematic review and meta-analysis. *Journal of the American Medical Association Psychiatry, 74,* 694–702. http://dx.doi.org/10.1001/jamapsychiatry.2017.1044

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Cuddy-Casey, M., & Orvaschel, H. (1997). Children's understanding of death in relation to child suicidality and homicidality. *Clinical Psychology Review, 17,* 33–45. http://dx.doi.org/10.1016/S0272-7358(96)00044-X

Daigle, M. S. (2005). Suicide prevention through means restriction: Assessing the risk of substitution. A critical review and synthesis. *Accident Analysis and Prevention, 37,* 625–632. http://dx.doi.org/10.1016/j.aap.2005.03.004

Danto, B. (1971). Firearms and their role in homicide and suicide. *Suicide and Life-Threatening Behavior, 1,* 10–17.

Das, R. K., Gale, G., Walsh, K., Hennessy, V. E., Iskandar, G., Mordecai, L. A., . . . Kamboj, S. K. (2019). Ketamine can reduce harmful drinking by pharmacologically rewriting drinking memories. *Nature Communications, 10,* 5187. http://dx.doi.org/10.1038/s41467-019-13162-w

DeCou, C. R., Comtois, K. A., & Landes, S. J. (2019). Dialectical behavior therapy is effective for the treatment of suicidal behavior: A meta-analysis. *Behavior Therapy, 50,* 60–72. http://dx.doi.org/10.1016/j.beth.2018.03.009

Dorji, G., Choki, S., Jamphel, K., Wangdi, Y., Chogyel, T., Dorji, C., & Nirola, D. K. (2017). Policy and governance to address depression and suicide in Bhutan: The national suicide-prevention strategy. *WHO South-East Asia Journal of Public Health, 6,* 39–44. http://dx.doi.org/10.4103/2224-3151.206163

Draper, D., & Margolis, P. (1976). A psychodynamic approach to suicide prevention. *Community Mental Health Journal, 12,* 376–382. http://dx.doi.org/10.1007/BF01411076

Eckshtain, D., Kuppens, S., Ugueto, A., Ng, M. Y., Vaughn-Coaxum, R., Corteselli, K., & Weisz, J. R. (2020). Meta-analysis: 13-year follow-up of psychotherapy effects on youth depression. *Journal of the American Academy of Child & Adolescent Psychiatry, 59,* 45–63. http://dx.doi.org/10.1016/j.jaac.2019.04.002

Effective Public Health Practice Project. (2007). *Quality assessment tool for quantitative studies*. Retrieved from https://www.ephpp.ca/quality-assessment-tool-for-quantitative-studies/

Farberow, N. L. (1969). Training in suicide prevention for professional and community agents. *The American Journal of Psychiatry, 125,* 1702–1705. http://dx.doi.org/10.1176/ajp.125.12.1702

Farberow, N. L., Ganzler, S., Cutter, F., & Reynolds, D. (1971). An eight-year survey of hospital suicides. *Suicide and Life-Threatening Behavior, 1,* 184–202.

Fitzpatrick, J. J. (1983). Suicidology and suicide prevention: Historical perspectives from the nursing literature. *Journal of Psychosocial Nursing and Mental Health Services, 21,* 20–28.

Flett, G. L., Vredenburg, K., & Krames, L. (1995). The stability of depressive symptoms in college students: An empirical demonstration of regression to the mean. *Journal of Psychopathology and Behavioral Assessment, 17,* 403–408. http://dx.doi.org/10.1007/BF02229059

Fox, K. R., Franklin, J. C., Ribeiro, J. D., Kleiman, E. M., Bentley, K. H., & Nock, M. K. (2015). Meta-analysis of risk factors for nonsuicidal self-injury. *Clinical Psychology Review, 42,* 156–167. http://dx.doi.org/10.1016/j.cpr.2015.09.002

Franklin, J. C. (2019). Psychological primitives can make sense of biopsychosocial factor complexity in psychopathology. *BMC Medicine, 17,* 187. http://dx.doi.org/10.1186/s12916-019-1435-1

Franklin, J. C., Huang, X., & Bastidas, D. (2019). Virtual reality suicide: Development of a translational approach for studying suicide causes. *Behaviour Research and Therapy, 120,* 103360. http://dx.doi.org/10.1016/j.brat.2018.12.013

Franklin, J. C., Huang, X., Fox, K. R., & Ribeiro, J. D. (2018). What suicide interventions should target. *Current Opinion in Psychology, 22,* 50–53. http://dx.doi.org/10.1016/j.copsyc.2017.08.002

Franklin, J. C., Puzia, M. E., Lee, K. M., Lee, G. E., Hanna, E. K., Spring, V. L., & Prinstein, M. J. (2013). The nature of pain offset relief in nonsuicidal self-injury: A laboratory study. *Clinical Psychological Science, 1,* 110–119. http://dx.doi.org/10.1177/2167702612474440

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

Freeman, W. (1953). Hazards of lobotomy: Study of two thousand operations. *Journal of the American Medical Association, 152,* 487–491. http://dx.doi.org/10.1001/jama.1953.03690060003002

Friedman, R. A., & Leon, A. C. (2007). Expanding the black box - depression, antidepressants, and the risk of suicide. *The New England Journal of Medicine, 356,* 2343–2346. http://dx.doi.org/10.1056/NEJMp078015

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion, 12,* 314–325. http://dx.doi.org/10.1037/a0026007

Glenn, C. R., Franklin, J. C., & Nock, M. K. (2015). Evidence-based psychosocial treatments for self-injurious thoughts and behaviors in youth. *Journal of Clinical Child and Adolescent Psychology, 44,* 1–29. http://dx.doi.org/10.1080/15374416.2014.945211

Glenn, C. R., Lanzillo, E. C., Esposito, E. C., Santee, A. C., Nock, M. K., & Auerbach, R. P. (2017). Examining the course of suicidal and nonsuicidal self-injurious thoughts and behaviors in outpatient and inpatient adolescents. *Journal of Abnormal Child Psychology, 45,* 971–983. http://dx.doi.org/10.1007/s10802-016-0214-0

Goeschel, C. (2006). Suicide at the end of the Third Reich. *Journal of Contemporary History, 41,* 153–173. http://dx.doi.org/10.1177/0022009406058692

Goldney, R. D. (2005). Suicide prevention: A pragmatic review of recent studies. *Crisis, 26,* 128–140. http://dx.doi.org/10.1027/0227-5910.26.3.128

Goldston, D. B., Daniel, S. S., Erkanli, A., Heilbron, N., Doyle, O., Weller, B., . . . Faulkner, M. (2015). Suicide attempts in a longitudinal sample of adolescents followed through adulthood: Evidence of escalation. *Journal of Consulting and Clinical Psychology, 83,* 253–264. http://dx.doi.org/10.1037/a0038657

Grenon, R., Carlucci, S., Brugnera, A., Schwartze, D., Hammond, N., Ivanova, I., . . . Tasca, G. A. (2019). Psychotherapy for eating disorders: A meta-analysis of direct comparisons. *Psychotherapy Research, 29,* 833–845. http://dx.doi.org/10.1080/10503307.2018.1489162

Hartling, L., Ospina, M., Liang, Y., Dryden, D. M., Hooton, N., Krebs Seida, J., & Klassen, T. P. (2009). Risk of bias versus quality assessment of randomised controlled trials: Cross sectional study. *British Medical Journal, 339,* b4012. http://dx.doi.org/10.1136/bmj.b4012

Hawton, K., Arensman, E., Townsend, E., Bremner, S., Feldman, E., Goldney, R., . . . Träskman-Bendz, L. (1998). Deliberate self harm: Systematic review of efficacy of psychosocial and pharmacological treatments in preventing repetition. *British Medical Journal, 317,* 441–447. http://dx.doi.org/10.1136/bmj.317.7156.441

Hawton, K., Simkin, S., Deeks, J., Cooper, J., Johnston, A., Waters, K., . . . Simpson, K. (2004). U. K. legislation on analgesic packs: Before and after study of long term effect on poisonings. *British Medical Journal, 329,* 1076–1079. http://dx.doi.org/10.1136/bmj.38253.572581.7C

Hedegaard, H., Curtin, S. C., & Warner, M. (2020). *Increase in Suicide Mortality in the United States, 1999–2018* (NCHS data brief (Issue 362)). Retrieved from https://www.cdc.gov/nchs/products/databriefs/db362.htm

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128. http://dx.doi.org/10.3102/10769986006002107

Hendin, H. (1991). Psychodynamics of suicide, with particular reference to the young. *The American Journal of Psychiatry, 148,* 1150–1158. http://dx.doi.org/10.1176/ajp.148.9.1150

Higgins, J. P. T., & Green, S. (2011). Assessment of study quality and risk of bias. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 5). London, UK: The Cochrane Collaboration.

Hill, R. M. (2015). *Open trial and pilot randomized controlled trial of a novel program to reduce perceived burdensomeness*. Miami: Florida International University. http://dx.doi.org/10.25148/etd.FIDC000057

Hoemann, K., & Feldman Barrett, L. (2019). Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. *Cognition and Emotion, 33,* 67–76. http://dx.doi.org/10.1080/02699931.2018.1535428

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology, 55,* 1830–1849. http://dx.doi.org/10.1037/dev0000686

Hooley, J. M., Ho, D. T., Slater, J., & Lockshin, A. (2010). Pain perception and nonsuicidal self-injury: A laboratory investigation. *Personality Disorders, 1,* 170–179. http://dx.doi.org/10.1037/a0020106

Huang, X., Funsch, K. M., Park, E. C., & Franklin, J. C. (in press). Anticipated consequences as the primary causes of suicidal behavior: evidence from a laboratory study. *Behaviour Research and Therapy*.

Huston, P. E., & Locher, L. M. (1948). Manic-depressive psychosis: Course when treated and untreated with electric shock. *Archives of Neurology and Psychiatry, 60,* 37–48. http://dx.doi.org/10.1001/archneurpsyc.1948.02310010043003

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

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin, 141,* 747–768. http://dx.doi.org/10.1037/bul0000015

Joiner, T. E. (2005). *Why people die by suicide*. Cambridge, MA: Harvard University Press.

Kalinowsky, L. B., & Scarff, J. E. (1948). The selection of psychiatric cases for prefrontal lobotomy. *The American Journal of Psychiatry, 105,* 81–85. http://dx.doi.org/10.1176/ajp.105.2.81

Kay, D. W. K., Fahy, T., & Garside, R. F. (1970). A seven-month double-blind trial of amitriptyline and diazepam in ECT-treated depressed patients. *The British Journal of Psychiatry, 117,* 667–671. http://dx.doi.org/10.1192/bjp.117.541.667

Kessler, R. C., Demler, O., Frank, R. G., Olfson, M., Pincus, H. A., Walters, E. E., . . . Zaslavsky, A. M. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *The New England Journal of Medicine, 352,* 2515–2523. http://dx.doi.org/10.1056/NEJMsa043266

Kirsch, I. (2008). Challenging received wisdom: Antidepressants and the placebo effect. *McGill Journal of Medicine, 11,* 219–222.

Knox, K. L., Litts, D. A., Talcott, G. W., Feig, J. C., & Caine, E. D. (2003). Risk of suicide and related adverse outcomes after exposure to a suicide prevention programme in the U.S. Air Force: Cohort study. *British Medical Journal, 327,* 1376–1380. http://dx.doi.org/10.1136/bmj.327.7428.1376

Kreitman, N. (1976). The coal gas story. United Kingdom suicide rates, 1960–71. *British Journal of Preventive & Social Medicine, 30,* 86–93. http://dx.doi.org/10.1136/jech.30.2.86

Kutcher, S., Wei, Y., & Behzadi, P. (2017). School- and community-based youth suicide prevention interventions: Hot idea, hot air, or sham? *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 62,* 381–387. http://dx.doi.org/10.1177/0706743716659245

Lane, P. W. (2013). Meta-analysis of incidence of rare events. *Statistical Methods in Medical Research, 22,* 117–132. http://dx.doi.org/10.1177/0962280211432218

Large, M. M., Chung, D. T., Davidson, M., Weiser, M., & Ryan, C. J. (2017). In-patient suicide: Selection of people at risk, failure of protection and the possibility of causation. *BJPsych Open, 3,* 102–105. http://dx.doi.org/10.1192/bjpo.bp.116.004309

Leavey, K., & Hawkins, R. (2017). Is cognitive behavioural therapy effective in reducing suicidal ideation and behaviour when delivered face-to-face or via e-health? A systematic review and meta-analysis. *Cognitive Behaviour Therapy, 46,* 353–374. http://dx.doi.org/10.1080/16506073.2017.1332095

Levy, K. N., Yeomans, F. E., & Diamond, D. (2007). Psychodynamic treatments of self-injury. *Journal of Clinical Psychology, 63,* 1105–1120. http://dx.doi.org/10.1002/jclp.20418

Levy, S., & Southcombe, R. H. (1953). Suicide in a state hospital for the mentally ill. *Journal of Nervous and Mental Disease, 117,* 504–514. http://dx.doi.org/10.1097/00005053-195306000-00003

Lilienfeld, S. O., Smith, S. F., & Watts, A. L. (2016). The perils of unitary models of the etiology of mental disorders-The response modulation hypothesis of psychopathy as a case example: Rejoinder to Newman and Baskin-Sommers (2016). *Psychological Bulletin, 142,* 1394–1403. http://dx.doi.org/10.1037/bul0000080

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review, 5,* 356–368. http://dx.doi.org/10.1177/1754073913489750

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion, 6,* 125–138. http://dx.doi.org/10.1037/1528-3542.6.1.125

Lindquist, K. A., Gendron, M., Barrett, L. F., & Dickerson, B. C. (2014). Emotion perception, but not affect perception, is impaired with semantic memory loss. *Emotion, 14,* 375–387. http://dx.doi.org/10.1037/a0035293

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology, 6,* 444. http://dx.doi.org/10.3389/fpsyg.2015.00444

Linehan, M. M. (1987). Dialectical behavioral therapy: A cognitive behavioral approach to parasuicide. *Journal of Personality Disorders, 1,* 328–333. http://dx.doi.org/10.1521/pedi.1987.1.4.328

Linthicum, K. P., Harris, L., & Ribeiro, J. D. (2019). *An experimental investigation of the effects of uncertainty on virtual suicidal behavior*. 53rd Annual Conference for the Association for Behavioral and Cognitive Therapies, Atlanta, GA. Retrieved from https://eventscribe.com/2019/ABCT/fsPopup.asp?Mode=presInfo&PresentationID=603434

Litman, R. E., Farberow, N. L., Shneidman, E. S., Heilig, S. M., & Kramer, J. A. (1965). Suicide-prevention telephone service. *Journal of the American Medical Association, 192,* 21–25. http://dx.doi.org/10.1001/jama.1965.03080140027006

Ludäscher, P., Greffrath, W., Schmahl, C., Kleindienst, N., Kraus, A., Baumgärtner, U., . . . Bohus, M. (2009). A cross-sectional investigation of discontinuation of self-injury and normalizing pain perception in patients with borderline personality disorder. *Acta Psychiatrica Scandinavica, 120,* 62–70. http://dx.doi.org/10.1111/j.1600-0447.2008.01335.x

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly, 2,* 245–264.

Mann, J. J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., . . . Hendin, H. (2005). Suicide prevention strategies: A systematic review. *Journal of the American Medical Association, 294,* 2064–2074. http://dx.doi.org/10.1001/jama.294.16.2064

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22,* 719–748. http://dx.doi.org/10.1093/jnci/22.4.719

Menninger, K. A. (1938). *Man against himself*. New York, NY: Harcourt, Brace, & World.

Miller, D. N., & Gould, K. (2013). Forgotten founder: Harry Marsh Warren and the history and legacy of the Save-A-Life League. *Suicidology Online, 4,* 12–15. Retrieved from http://www.suicidology-online.com/pdf/SOL-2013-4-12-15.pdf

Milner, A. J., Carter, G., Pirkis, J., Robinson, J., & Spittal, M. J. (2015). Letters, green cards, telephone calls and postcards: Systematic and meta-analytic review of brief contact interventions for reducing self-harm, suicide attempts and suicide. *The British Journal of Psychiatry, 206,* 184–190. http://dx.doi.org/10.1192/bjp.bp.114.147819

Ministry of Health. (2019). *Every life matters—He Tapu te Oranga o ia tangata: Suicide Prevention Strategy 2019–2029 and Suicide Prevention Action Plan 2019–2024 for Aotearoa New Zealand*. Retrieved from https://www.health.govt.nz/publication/every-life-matters-he-tapu-te-oranga-o-ia-tangata-suicide-prevention-strategy-2019-2029-and-suicide

Mishara, B. L. (1999). Conceptions of death and suicide in children ages 6–12 and their implications for suicide prevention. *Suicide and Life-Threatening Behavior, 29,* 105–118.

Moniz, E. (1937). Prefrontal leucotomy in the treatment of mental disorders. *The American Journal of Psychiatry, 93,* 1379–1385. http://dx.doi.org/10.1176/ajp.93.6.1379

Motto, J. A. (1976). Suicide prevention for high-risk persons who refuse treatment. *Suicide and Life-Threatening Behavior, 6,* 223–230.

Motto, J. A., & Bostrom, A. G. (2001). A randomized controlled trial of postcrisis suicide prevention. *Psychiatric Services, 52,* 828–833. http://dx.doi.org/10.1176/appi.ps.52.6.828

Murray, S. B., Quintana, D. S., Loeb, K. L., Griffiths, S., & Le Grange, D. (2019). Treatment outcomes for anorexia nervosa: A systematic review and meta-analysis of randomized controlled trials. *Psychological Medicine, 49,* 535–544. http://dx.doi.org/10.1017/S0033291718002088

Naghavi, M., & the Global Burden of Disease Self-Harm Collaborators. (2019). Global, regional, and national burden of suicide mortality 1990 to 2016: Systematic analysis for the Global Burden of Disease Study 2016. *British Medical Journal, 364,* l94. http://dx.doi.org/10.1136/bmj.l94

Nock, M. K., Borges, G., Bromet, E. J., Alonso, J., Angermeyer, M., Beautrais, A., . . . Williams, D. (2008). Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *The British Journal of Psychiatry, 192,* 98–105. http://dx.doi.org/10.1192/bjp.bp.107.040113

Nock, M. K., Green, J. G., Hwang, I., McLaughlin, K. A., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2013). Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: Results from the National Comorbidity Survey Replication Adolescent Supplement. *Journal of the American Medical Association Psychiatry, 70,* 300–310. http://dx.doi.org/10.1001/2013.jamapsychiatry.55

Nock, M. K., & Prinstein, M. J. (2004). A functional approach to the assessment of self-mutilative behavior. *Journal of Consulting and Clinical Psychology, 72,* 885–890. http://dx.doi.org/10.1037/0022-006X.72.5.885

Olfson, M., Wall, M., Wang, S., Crystal, S., Liu, S. M., Gerhard, T., & Blanco, C. (2016). Short-term suicide risk after psychiatric hospital discharge. *Journal of the American Medical Association Psychiatry, 73,* 1119–1126. http://dx.doi.org/10.1001/jamapsychiatry.2016.2035

Oliver, R. G., & Hetzel, B. S. (1972). Rise and fall of suicide rates in Australia: Relation to sedative availability. *The Medical Journal of Australia, 2,* 919–923. http://dx.doi.org/10.5694/j.1326-5377.1972.tb103635.x

Oud, M., de Winter, L., Vermeulen-Smit, E., Bodden, D., Nauta, M., Stone, L., . . . Stikkelbroek, Y. (2019). Effectiveness of CBT for children and adolescents with depression: A systematic review and meta-regression analysis. *European Psychiatry, 57,* 33–45. http://dx.doi.org/10.1016/j.eurpsy.2018.12.008

Pearson, M., Metcalfe, C., Jayamanne, S., Gunnell, D., Weerasinghe, M., Pieris, R., . . . Konradsen, F. (2017). Effectiveness of household lockable pesticide storage to reduce pesticide self-poisoning in rural Asia: A community-based, cluster-randomised controlled trial. *The Lancet, 390,* 1863–1872. http://dx.doi.org/10.1016/S0140-6736(17)31961-X

Prien, R. F., Klett, C. J., & Caffey, E. M., Jr. (1973). Lithium carbonate and imipramine in prevention of affective episodes. A comparison in recurrent affective illness. *Archives of General Psychiatry, 29,* 420–425. http://dx.doi.org/10.1001/archpsyc.1973.04200030104017

Pritchard, J. B. (1950). A Dispute over Suicide. In J. B. Pritchard (Ed.), *Ancient Near Eastern texts relating to the Old Testament* (pp. 405–406). Princeton, NJ: Princeton University Press.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Resnik, H. L. P., Davison, W. T., Schuyler, D., & Christopher, P. (1973). Videotape confrontation after attempted suicide. *The American Journal of Psychiatry, 130,* 460–463.

Riblet, N. B. V., Shiner, B., Young-Xu, Y., & Watts, B. V. (2017). Strategies to prevent death by suicide: Meta-analysis of randomised controlled trials. *The British Journal of Psychiatry, 210,* 396–402. http://dx.doi.org/10.1192/bjp.bp.116.187799

Rosen, L. W., & Thomas, M. A. (1984). Treatment technique for chronic wrist cutters. *Journal of Behavior Therapy and Experimental Psychiatry, 15,* 33–36. http://dx.doi.org/10.1016/0005-7916(84)90119-8

Ross, C. P. (1980). Teaching children the facts of life and death: Suicide prevention in the schools. In M. L. Pech, L. N. Farberow, & R. E. Li (Eds.), *Youth suicide* (pp. 147–169). Dordrecht, the Netherlands: Springer.

Rush, A. J., & Beck, A. T. (1978). Cognitive therapy of depression and suicide. *American Journal of Psychotherapy, 32,* 201–219. http://dx.doi.org/10.1176/appi.psychotherapy.1978.32.2.201

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84,* 328–364. http://dx.doi.org/10.3102/0034654313500826

Schleider, J. L., Abel, M. R., & Weisz, J. R. (2015). Implicit theories and youth mental health problems: A random-effects meta-analysis. *Clinical Psychology Review, 35,* 1–9. http://dx.doi.org/10.1016/j.cpr.2014.11.001

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

Simpson, D. T. (1984). Involuntary civil commitment: The dangerousness standard and its problems. *North Carolina Law Review, 63,* 241–245.

Spencer-Thomas, S., & Jahn, D. R. (2012). Tracking a movement: U.S milestones in suicide prevention. *Suicide and Life-Threatening Behavior, 42,* 78–85. http://dx.doi.org/10.1111/j.1943-278X.2011.00072.x

Stanley, B., & Mann, J. J. (2019). The need for innovation in health Care systems to improve suicide prevention. *Journal of the American Medical Association Psychiatry, 77,* 96–98. http://dx.doi.org/10.1001/jamapsychiatry.2019.2769

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior, 106,* 5–15. http://dx.doi.org/10.1016/j.physbeh.2011.06.004

Swannell, S. V., Martin, G. E., Page, A., Hasking, P., & St John, N. J. (2014). Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-Threatening Behavior, 44,* 273–303. http://dx.doi.org/10.1111/sltb.12070

Testa, M., & West, S. G. (2010). Civil commitment in the United States. *Psychiatry, 7,* 30–40.

The Scottish Government. (2018). *Every life matters: Scotland's suicide prevention action plan*. Retrieved from https://www.gov.scot/Publications/2018/08/8874

Thompson, C. G., & Becker, B. J. (2014). The impact of multiple endpoint dependency on Q and I(2) in meta-analysis. *Research Synthesis Methods, 5,* 235–253. http://dx.doi.org/10.1002/jrsm.1110

Touroutoglou, A., Andreano, J. M., Adebayo, M., Lyons, S., & Barrett, L. F. (2019). Motivation in the service of allostasis: The role of anterior mid-cingulate cortex. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 6, pp. 1–25). Amsterdam, the Netherlands: Elsevier.

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (1990). *Healthy people 2000*. Washington, DC: Author.

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2000). *Healthy people 2010*. Washington, DC: Author.

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2010). *Healthy people 2020*. Washington, DC: Author.

van Bronswijk, S., Moopen, N., Beijers, L., Ruhe, H. G., & Peeters, F. (2019). Effectiveness of psychotherapy for treatment-resistant depression: A meta-analysis and meta-regression. *Psychological Medicine, 49,* 366–379. http://dx.doi.org/10.1017/S003329171800199X

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E., Jr. (2010). The interpersonal theory of suicide. *Psychological Review, 117,* 575–600. http://dx.doi.org/10.1037/a0018697

Viechtbauer, W. (2020). *Package "metafor"* (2.4–0). Retrieved from https://cran.r-project.org/web/packages/metafor/metafor.pdf

Waterhouse, J., & Platt, S. (1990). General hospital admission in the management of parasuicide. A randomised controlled trial. *The British Journal of Psychiatry, 156,* 236–242. http://dx.doi.org/10.1192/bjp.156.2.236

Weintraub, D., Duda, J. E., Carlson, K., Luo, P., Sagher, O., Stern, M., . . . the CSP 468 Study Group. (2013). Suicide ideation and behaviours after STN and GPi DBS surgery for Parkinson's disease: Results from a randomised, controlled trial. *Journal of Neurology, Neurosurgery, and Psychiatry, 84,* 1113–1118. http://dx.doi.org/10.1136/jnnp-2012-304396

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., . . . Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist, 72,* 79–117. http://dx.doi.org/10.1037/a0040360

Wilcox, H. C., Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., & Anthony, J. C. (2008). The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug and Alcohol Dependence, 95,* S60–S73. http://dx.doi.org/10.1016/j.drugalcdep.2008.01.005

World Health Organization. (2013). *Mental health action plan 2013–2020*. Geneva, Switzerland: Author. Retrieved from https://www.who.int/mental_health/publications/action_plan/en/

World Health Organization. (2014). *Preventing Suicide: A global imperative*. Geneva, Switzerland: Author. Retrieved from https://www.who.int/mental_health/suicide-prevention/world_report_2014/en/

World Health Organization. (2015). *Japan turning a corner in suicide prevention*. Geneva, Switzerland: Author. Retrieved from https://www.who.int/mental_health/suicide-prevention/japan_story/en/

Yip, P. S. F., Caine, E., Yousuf, S., Chang, S. S., Wu, K. C. C., & Chen, Y. Y. (2012). Means restriction for suicide prevention. *The Lancet, 379,* 2393–2399. http://dx.doi.org/10.1016/S0140-6736(12)60521-2

Ziskind, E., Somerfeld-Ziskind, E., & Ziskind, L. (1945). Metrazol and electric convulsive therapy of the affective psychoses: A controlled series of observations covering a period of five years. *Archives of Neurology and Psychiatry, 53,* 212–217. http://dx.doi.org/10.1001/archneurpsyc.1945.02300030049007

Received December 16, 2019
Revision received August 3, 2020
Accepted August 11, 2020 ■