The Psychological Effects of Solitary Confinement: A Systematic Critique

Article *in* Crime and Justice · March 2018

DOI: 10.1086/696041

1 author:

Craig Haney
University of California, Santa Cruz

Craig Haney

The Psychological Effects of Solitary Confinement: A Systematic Critique

ABSTRACT

Research findings on the psychological effects of solitary confinement have been strikingly consistent since the early nineteenth century. Studies have identified a wide range of frequently occurring adverse psychological reactions that commonly affect prisoners in isolation units. The prevalence of psychological distress is extremely high. Nonetheless, use of solitary confinement in the United States vastly increased in recent decades. Advocates defend its use, often citing two recent studies to support claims that isolation has no significant adverse psychological effects, including even on mentally ill people. Those studies, however, are fundamentally flawed, their results are not credible, and they should be disregarded. Critically and comprehensively analyzing the numerous flaws that compromise this recent scholarship underscores the distinction between methodological form and substance, the danger of privileging quantitative data irrespective of their quality, and the importance of considering the fraught nature of the prison context in which research results are actually generated. Solitary confinement has well-documented adverse effects. Its use should be eliminated entirely for some groups of prisoners and greatly reduced for others.

Doing prison research, Alison Liebling has long reminded us, is deeply emotional and intellectually challenging, with different methodological approaches "competing for epistemological prominence—often from different sides of the prison wall" (1999, p. 148). It takes place in "an in-

Electronically published March 9, 2018

Craig Haney is Distinguished Professor of Psychology, University of California, Santa Cruz.

0192-3234/2018/0047-0004$10.00

tense, risk-laden, emotionally fraught environment" (p. 163) and within a closed environment in which prison administrators tightly control access to data and most prisoners manifest an entirely legitimate and understandable skepticism toward data gatherers.

This helps explain why, in Liebling's words, "the pains of imprisonment are tragically underestimated by conventional methodological approaches to prison life" (p. 165). The more these conventional approaches encourage us to conceive of prisons as more or less traditional research settings and prisoners as mere specimens to be "objectively assessed," the less likely we are to gain useful insights into prison life or accurately represent the experience of those living inside.

These cautions are doubly applicable to research on solitary confinement.[1] It involves involuntary isolation of prisoners nearly around the clock in sparse cells located in remote or inaccessible units. Solitary confinement denies prisoners any meaningful social contact and access to positive environmental stimulation.

These prisons within prisons are nearly impenetrable to outside researchers (or anyone else). Prison officials tightly control access to solitary confinement units and to the prisoners inside them. They typically rebuff attempts by researchers to observe conditions and practices, let alone to carefully assess their potentially harmful effects. Prisoners in solitary confinement tend to be even more self-protective than other prisoners are (as part of their accommodation to harsh and frequently abusive conditions) and reluctant to have their "measure" taken by persons whom they have no reason to trust. They generally subscribe strongly to prisoner norms against displaying or acknowledging vulnerabilities that could be interpreted as weakness. The inapt pejorative designation of them as collectively "the worst of the worst" does not inspire confidence in or candor toward outsiders, and certainly not toward anyone remotely associated with the prison administration.

These realities pose a host of methodological challenges for anyone interested in understanding the nature and effects of prison isolation. This is in part why studies of the effects of solitary confinement on prisoners

[1] I use "solitary confinement" to refer to forms of prison isolation in which prisoners are housed involuntarily in their cells for upward of 23 hours per day and denied the opportunity to engage in normal and meaningful social interaction and congregate activities, including correctional programming. The term subsumes a range of prison nomenclature including "administrative segregation," "security housing units," "high security," and "close management," among others.

have rarely, if ever, approximated experimental research designs (including quasi- or natural experimental designs).

Solitary confinement units not only are largely impenetrable to outsiders but also, of course, are subject to legal and ethical restrictions that preclude random assignment of prisoners into them. The rigid prison rules and operating procedures that govern these places can easily frustrate the use of the kind of meticulous controls over conditions and participants that are needed to carry out anything remotely resembling an experiment. The distinctiveness of solitary confinement units and the nonnegotiable staff mandates under which they operate make it difficult, if not impossible, to implement rigorous conventional research designs (e.g., representative samples, control groups, repeated measures). Efforts to conduct randomized or truly controlled studies inevitably face significant risks that the data collected will be so confounded by inevitable methodological compromises as to be uninterpretable and, therefore, meaningless.

Nonetheless, scholars and researchers know a great deal about the negative effects of solitary confinement. We have firsthand or autobiographical accounts by former prisoners (e.g., Burney 1961) and staff members (e.g., Rundle 1973; Slater 1986); ethnographic, interview, and observational research (e.g., Benjamin and Lux 1975; Toch 1975; Hilliard 1976; Jackson 1983; Rhodes 2004; Reiter 2016); and cross-sectional studies that assess prisoners' psychological reactions at particular times (e.g., Grassian 1983; Brodsky and Scogin 1988; Haney 2003).

Much of the important research is qualitative, but there is a substantial amount of it and the findings are robust. They can also be "triangulated," that is, studied through a range of methods and in settings sometimes similar but not necessarily identical to solitary confinement (e.g., Turner, Cardinal, and Burton 2017). Numerous literature reviews have noted that scientists from diverse disciplinary backgrounds, working independently and across several continents, and over many decades, have reached almost identical conclusions about the negative effects of isolation in general and solitary confinement in particular (e.g., Haney and Lynch 1997; Haney 2003; Grassian 2006; Smith 2006; Arrigo and Bullock 2008). Those robust findings are also theoretically coherent. That is, they are consistent with and explained by a rapidly growing literature on the importance of meaningful social contact for maintenance of mental and physical health.

Largely because of the robustness and theoretical underpinnings of the data, numerous scientific and professional organizations have reached a broad consensus about the damaging effects of solitary confinement. Several years ago, for example, a National Academies of Science committee reviewed the existing research and concluded that solitary confinement can precipitate such "serious psychological change" in prisoners that the practice "is best minimized" (National Research Council 2014, p. 201). The American Psychological Association (2016, p. 1), the world's largest professional association of psychologists, asserted that "solitary confinement is associated with severe harm to physical and mental health among both youth and adults, including: increased risk of self-mutilation, and suicidal ideation; greater anxiety, depression, sleep disturbance, paranoia, and aggression; exacerbation of the onset of pre-existing mental illness and trauma symptoms; [and] increased risk of cardiovascular problems."

Similarly, the National Commission on Correctional Health Care (2016), a highly respected organization of correctional medical and mental health professionals, promulgated a series of "principles" with respect to solitary confinement. They are intended to guide the ethical conduct of its members, including that placement in solitary confinement for longer than 15 days represents "cruel, inhumane, and degrading treatment" that is "harmful to an individual's health" (p. 260) and that "health care staff must advocate" to remove persons from solitary confinement whenever "their medical or mental health deteriorates" (p. 261).

Summarizing this growing consensus, a joint 2016 statement of the Association of State Correctional Administrators (the largest professional association of American prison administrators) and Yale Law School's Liman Public Interest Program observed that demands for change in use of solitary confinement are being made around the world. More specifically,

> Commitments to reform and efforts to limit or abolish the use of isolating confinement come from stakeholders and actors in and out of government. Documentation of the harms of isolation, coupled with its costs and the dearth of evidence suggesting that it enhances security, has prompted prison directors, legislatures, executive branch officials, and advocacy groups to try to limit reliance on restricted housing. Instead of being cast as the solution to a problem, restricted housing has come to be understood by many as a problem in need of a solution. (Association of State Correctional Administrators and the Arthur Liman Public Interest Program 2016, p. 15)

Even more recently, the director of the Colorado Department of Corrections, Rick Raemisch, announced that Colorado has ended use of long-term solitary confinement, so that even prisoners "who commit serious violations like assault will now spend at most 15 days in solitary" (2017, p. A25). This development in Colorado is especially notable, for reasons that become clear in the pages that follow.

Against this backdrop, in 2009 and 2010 word began to circulate among prison researchers and policy makers that a new, supposedly unassailable scientific study—the "Colorado study"—had produced results that contravened many decades of empirical findings on the harmful effects of prison isolation. Lovell and Toch (2011, p. 3) characterized a number of its findings as "flabbergasting," and indeed they were. Among the most startling were that a year-long stay in solitary confinement resulted in no "significant decline in psychological well-being over time"; that on most measures, including cognitive performance, "there was improved functioning over time"; and most remarkably that many more mentally ill prisoners benefited from isolation than were damaged by it (O'Keefe et al. 2010, pp. 54, 78). The Colorado researchers thus reported data indicating that solitary confinement made prisoners feel and think better, especially if they were mentally ill.

In fact, however, the Colorado study was riddled with serious methodological problems that limited its value and made the meaning of the results impossible to decipher. Notwithstanding its authors' frank, albeit at times opaque and oblique, acknowledgments of some of its fundamental weaknesses, defenders of solitary confinement have seized on it. It has become a last bastion of resistance against a widespread and growing consensus that use of solitary confinement should be eliminated or drastically limited.

The Colorado study's influence has been amplified by an equally flawed meta-analysis that relied very heavily on it and significantly mischaracterized the prior literature on the effects of isolated confinement (Morgan et al. 2016). Of course, the influence of a fundamentally flawed study can grow if it and the data it produced are included in literature reviews that overlook glaring weaknesses. This risk is greater in meta-analytic than in narrative literature reviews that focus on decontextualized "effect sizes" irrespective of methodological shortcomings of individual studies. Unlike narrative reviews, meta-analyses include only quantitative outcomes or effects. This elevates the importance of numerical outcomes and often scants nuanced assessments of data quality. This is particularly a problem for prison research, an enterprise that is fraught with emotional and methodological challenges, in which aspects of the institutional context or setting can fundamentally alter the nature of the research and the meaning of its results. That is precisely what happened in the Morgan et al. (2016) meta-analysis.

In the following pages, I first discuss the scientific basis for the broad consensus that solitary confinement has substantial negative psychological effects on prisoners. I then discuss the Colorado study and the Morgan et al. (2016) meta-analysis based largely on it. Both are textbook examples of how things can go terribly wrong when researchers fail to take account of the unique nature of the prison environment, the special emotional and methodological challenges of prison research in general, and the contingent and unpredictable conditions and practices that affect solitary confinement units in particular.

I. Solitary Confinement Research and Practice

Documentation of the damaging nature and psychological effects of solitary confinement has a very long history, dating at least to the early nineteenth century, when solitary confinement was the modal form of imprisonment. The notion that prisoners could be reformed—made "penitent"—by time spent in isolation dominated American correctional thinking and practice and eventually spread throughout Europe. Yet the practice was recognized as a dangerous failure not long after its inception. Haney and Lynch (1997), Toch (2003), Grassian (2006), and Smith (2006) reviewed much of the early historical literature. Reports on solitary confinement at Pentonville Prison in England described "twenty times more cases of mental disease than in any other prison in the country" (Hibbert 1963, p. 160). Accounts of solitary confinement in the Netherlands documented "again and again, reports of insanity, suicide, and the complete alienation of prisoners from social life" (Franke 1992, p. 128). Newspaper reports from Philadelphia observed that prisoners in solitary confinement at the Walnut Street Jail "beg, with the greatest earnestness, that they may be hanged out of their misery" (Masur 1989, p. 83). Charles Dickens concluded that a prisoner kept in that "melancholy house" was like "a man buried alive . . . dead to everything but torturing anxieties and horrible despair" (Dickens 1842, p. 116). A similar regime in Auburn, New York, was described as "a hopeless failure that led to a marked prevalence of sickness and insanity on the part of convicts in solitary confinement" (Barnes 1921, p. 53). Stuart Grassian (2006, pp. 342–43) reported that "between 1854 and 1909, thirty-seven articles appeared in German scientific journals on the subject of psychotic disturbances among prisoners." The "most consistent factor" accounting for prison psychoses, "reported in over half the total literature, was solitary confinement."

Systematic early studies of solitary confinement in the United States used what is now seen as a somewhat outmoded theoretical framework, focusing narrowly on sensory rather than social deprivation (e.g., Scott and Gendreau 1969; Gendreau et al. 1972). Even so, the authors of one early study concluded that "excessive deprivation of liberty, here defined as near complete confinement to the cell, results in deep emotional disturbances" (Cormier and Williams 1966, p. 484). In a review of the sensory deprivation literature, Haney and Lynch (1997) noted that "the dissimilarities between conditions created in these studies and those in solitary confinement or punitive segregation in correctional institutions are obvious." They also observed that, nonetheless, the early research did "emphasize the importance of sensory stimulation in human experience and the dramatic effects that can be produced when such stimulation is significantly curtailed" (p. 502).

More recent research focuses on the psychological damage that results from social deprivation. Hans Toch's large-scale psychological study of prisoners in crisis in New York State correctional facilities included important observations about the effects of isolation. After conducting numerous in-depth interviews, Toch (1975, p. 54) concluded that "isolation panic" was a serious problem in solitary confinement. The symptoms Toch described included rage, panic, loss of control and breakdowns, psychological regression, and build-ups of physiological and psychic tension that led to incidents of self-mutilation. He noted that isolation panic could occur under other conditions of confinement but that it was "most sharply prevalent in segregation." Moreover, it marked an important dichotomy for prisoners: the "distinction between imprisonment, which is tolerable, and isolation, which is not."

Empirical studies have identified a wide range of frequently occurring adverse psychological reactions to solitary confinement.[2] These include

[2] For reviews of the literature documenting these adverse reactions, see Haney and Lynch (1997), Haney (2003), Cloyes et al. (2006), Grassian (2006), Smith (2006), and Arrigo and Bullock (2008).

stress-related reactions (such as decreased appetite, trembling hands, sweating palms, heart palpitations, and a sense of impending emotional breakdown); sleep disturbances (including nightmares and sleeplessness); heightened levels of anxiety and panic; irritability, aggression, and rage; paranoia, ruminations, and violent fantasies; cognitive dysfunction, hypersensitivity to stimuli, and hallucinations; loss of emotional control, mood swings, lethargy, flattened affect, and depression; increased suicidality and instances of self-harm; and, finally, paradoxical tendencies to further social withdrawal.

The prevalence of psychological distress, at least as suffered in certain solitary confinement settings, appears to be extremely high. A study conducted at the Security Housing Unit (SHU) at Pelican Bay State Prison in California (Haney 1993; Reiter 2016), an especially severe solitary confinement facility, is illustrative. Structured interviews were used to assess a randomly selected, representative sample of 100 prisoners to determine the prevalence of symptoms of psychological stress, trauma, and isolation-related psychopathology (Haney 2003). The interviews included demographic questions, brief social and institutional histories, and systematic assessments of 25 items, based in part on the Omnibus Stress Index (Jones 1976) and on other instruments similar to those used in Brodsky and Scogin (1988). Every symptom of psychological stress and trauma but one (fainting) was experienced by more than half of the assessed prisoners; many were reported by two-thirds or more and some by nearly everyone. Well over half of the prisoners reported distress-related symptoms—headaches, trembling, sweaty palms, and heart palpitations.

High numbers of the Pelican Bay SHU prisoners also reported suffering from isolation-related symptoms of pathology. Nearly all reported ruminations or intrusive thoughts, oversensitivity to external stimuli, irrational anger and irritability, difficulties with attention and often with memory, and a tendency to withdraw socially. Almost as many reported symptoms indicative of mood or emotional disorders: concerns over emotional flatness or losing the ability to feel, swings in emotional response, and feelings of depression or sadness that did not go away. Finally, sizable minorities reported symptoms that are typically associated only with more extreme forms of psychopathology—hallucinations, perceptual distortions, and thoughts of suicide.

Social withdrawal, a common reaction to solitary, is related to a broader set of social pathologies that prisoners often experience as they attempt to adapt to an environment devoid of normal, meaningful social contact. In order to exist and function in solitary confinement, where day-to-day life lacks meaningful interaction and closeness with others, prisoners have little choice but to adapt in ways that are asocial and, ultimately, psychologically harmful.

A large international literature has reached similar conclusions on the adverse psychological effects of solitary confinement. Solitary confinement not only is a common form of mistreatment to which prisoners of war have been subjected and been adversely affected (e.g., Hinkle and Wolff 1956) but also is associated with "higher levels of later life disability" among returnees (Hunt et al. 2008, p. 616). It is frequently used as a component of torture (e.g., Foster, Davis, and Sandler 1987; Nowak 2006; Reyes 2007). Solitary confinement has been studied in more traditional international criminal justice contexts as well. For example, Barte (1989, p. 52) concluded that solitary confinement in French prisons had such "psychopathogenic" effects that prisoners placed there for extended periods could become schizophrenic, making the practice unjustifiable, counterproductive, and "a denial of the bonds that unite humankind."

Koch (1986, pp. 124–25) studied "acute isolation syndrome" among detainees in Denmark that occurred after only a few days in isolation and included "problems of concentration, restlessness, failure of memory, sleeping problems and impaired sense of time and ability to follow the rhythm of day and night." If isolation persisted for a few weeks or more, it could lead to "chronic isolation syndrome," including intensified difficulties with memory and concentration, "inexplicable fatigue," a "distinct emotional liability" that included fits of rage, hallucinations, and the "extremely common" belief among prisoners that "they have gone or are going mad."

Volkart, Dittrich, et al. (1983) studied penal isolation in Switzerland. They concluded that, compared with prisoners in normal confinement, those in solitary displayed considerably more psychopathological symptoms, including heightened feelings of anxiety, emotional hypersensitivity, ideas of persecution, and thought disorders (see also Waligora 1974; Volkart, Rothenfluh, et al. 1983; Bauer et al. 1993).

The major reviews of the literature reach the same conclusions as the seminal studies. Haney and Lynch (1997, pp. 530, 537) noted that "distinctive patterns of negative effects have emerged clearly, consistently, and unequivocally from personal accounts, descriptive studies, and systematic research on solitary and punitive segregation." The "psychologically destructive treatment" to which prisoners are exposed in solitary confinement is so severe that it likely "would not be countenanced for any other group in our society."

Grassian's extensive survey of solitary confinement research concluded that "the restriction of environmental stimulation and social isolation associated with confinement in solitary are strikingly toxic to mental functioning, including, in some prisoners, a stuporous condition associated with perceptual and cognitive impairment and affective disturbances" (2006, p. 354).

That same year, Smith's comprehensive review concluded that "the vast majority" of studies on the effects of solitary confinement "document significant negative health effects" (2006, p. 456). He observed that "research on effects of solitary confinement has produced a massive body of data documenting serious adverse health effects" (p. 475) including "anger, hatred, bitterness, boredom, stress, loss of the sense of reality, suicidal thoughts, trouble sleeping, impaired concentration, confusion, depression, and hallucinations" (p. 488).

Similarly, Arrigo and Bullock (2008) concluded that "nearly all investigators acknowledge that long-term segregation, mistreatment by correctional staff, and preexisting psychological vulnerability are all apt to result in negative mental health consequences for convicts" and that "the extreme isolation and harsh conditions of confinement in [solitary confinement] typically exacerbate the symptoms of mental illness" (p. 632).

There is an important, theoretically coherent framework that helps explain the consistency of these conclusions. A burgeoning literature in social psychology and related disciplines shows that solitary confinement is a potentially harmful form of sensory deprivation but also, and more destructively, exposes prisoners to pathological levels of social deprivation. Numerous studies have established the critical psychological significance of social contact, connectedness, and belonging (e.g., Fiorillo and Sabatini 2011; Hafner et al. 2011; Cacioppo and Cacioppo 2012). Meaningful social interactions and social connectedness can have a positive effect on people's physical and mental health in settings outside of prison and, conversely, social isolation in general can undermine health and psychological well-being. Thus, it makes sound psychological sense that exposure to especially severe forms of material, sensory, and social deprivation harms prisoners' mental health.

Indeed, researchers have concluded that human brains are "wired to connect" to others (Lieberman 2013). Thwarting the need to establish and maintain connections to others undermines psychological well-being and increases physical morbidity and mortality. Because "social connection is crucial to human development, health, and survival," experts have called for it to be recognized as a national public health priority (Holt-Lunstad, Robles, and Sbarra 2017, p. 527). The involuntary, coercive, hostile, and demeaning aspects of solitary confinement are likely to exacerbate the negative effects of social isolation that have repeatedly been documented in more benign contexts.

Given these long-standing and theoretically informed findings, a study purporting to show that psychological effects of solitary confinement range from harmless to beneficial would normally not be taken seriously. Sometimes, however, the appearance of seemingly objective scientific findings provides legitimacy to doubtful conclusions, especially when they support contested policy or political agendas. That is precisely what happened in the case of the Colorado study. Its authors described it as a scientific advance over all previous studies, and some commentators prematurely lauded its methodological rigor. It appeared on the surface to be an ambitious and well-designed longitudinal study, with appropriate comparison groups and a host of dependent variables that were to be examined. Data were collected through the repeated administration of instruments said to be validated, and an unusually large number of prisoners were to be assessed over a 1-year period.

The reality was very different. The project could not be, and was not, carried out as planned, partly because of powerful demands and correctional contingencies inherent in prison settings in general and solitary confinement in particular. The problems proved insurmountable: comparison groups were not comparable, and the integrity of the "treatments" each group received was quickly corrupted. I discuss these and numerous other problems in the next section. The fundamental methodological flaws that plagued the study prevented collection of any meaningful data and ensured that no meaningful conclusions could be drawn.

The Colorado study nonetheless has continued to play an outsized role in contentious policy debates in which proponents of solitary confinement draw on it to support positions that are becoming indefensible. Defenders have characterized the study as "an outstanding example of applied correctional research" that was "planned with great care," employed a "rigorous" design, and produced results that "were about as conclusive as possible" showing that solitary confinement has few or no adverse effects (Gendreau and Labrecque 2016, p. 9).

A year after the study's release, the National Institute of Corrections devoted an entire issue of *Corrections and Mental Health* to discussion of it. One writer (other than the Colorado researchers themselves) who endorsed its results and defended its methodology was Paul Gendreau, a well-known Canadian researcher and long-time prison system employee. Despite not having published primary research data on isolation since the early 1970s, he had defended its use over many decades, for example, in a 1984 article entitled "Solitary Confinement Is Not Cruel and Unusual: People Sometimes Are!" (Gendreau and Bonta 1984). In *Corrections and Mental Health*, Gendreau hailed the Colorado study as a "truly significant contribution to our knowledge base about the effects of prison life for one of the most severe forms of incarceration" and asserted that "in terms of its methodological rigor" no other study "comes close" (Gendreau and Theriault 2011, p. 1). Moreover, despite the deep skepticism voiced by all of the other contributors to the special issue except Gendreau and the study's authors, the journal's editor described the Colorado study as "an important report" because it showed that "administrative segregation is not terribly harmful" (Immarigeon 2011, p. 1).

Similarly, when a brief summary of the study appeared in a scholarly journal (O'Keefe et al. 2013), it was accompanied by commentary written by several prominent clinicians who claimed to have witnessed as much as or more psychological improvement among isolated prisoners than decompensation. They praised the study as "groundbreaking" and described its methodology as "solid" (Berger, Chaplin, and Trestman 2013, pp. 61–63). The authors averred that "the extremes of solitary confinement have been misunderstood" and that "people are resilient and are able to thrive under even difficult environmental conditions."

The respected Irish prison researcher Ian O'Donnell, though more circumspect, offered similar observations. Although O'Donnell acknowledged some limitations, he praised the study's methodology and invoked its results to support some of his own views. "However unpalatable they might appear to some parties," he asserted, the study's findings "must be taken seriously" (2014, p. 120). O'Donnell characterized the study as "valuable" because, he said, it "highlights the individual's capacity to adapt" (p. 122). He defended the Colorado researchers against criticism, noting that it is ethically impossible to study solitary confinement with "sufficient scientific rigour to satisfy everyone" (p. 122). The study's results suggest, he wrote, "that segregation was not highly detrimental to those forced to endure it" (p. 120) and that the harmfulness of this form of penal confinement "may have been over-emphasized" (p. 123).[3]

The Colorado study also figures prominently in correctional policy reviews by recalcitrant prison officials who do not want to modify segregation practices and in litigation over the harmful effects of solitary confinement, where those defending it are eager to find support.[4] For example, the US Government Accountability Office conducted a review of segregated housing practices in the federal Bureau of Prisons (BOP): "BOP HQ officials cited the 2010 DOJ-funded study of the psychological impact of solitary confinement in the Colorado state prison system. This study showed that segregated housing of up to 1 year may not have greater negative psychological impacts than non-segregated housing on inmates. While the DOJ-funded study did not assess inmates in BOP facilities, BOP management told us this study shows that segregation has

[3] O'Donnell indicated that the study documented the "benefits" of solitary, ones he suggested derived from "the many hours spent in quiet contemplation" in solitary confinement units. He also suggested that the results buttressed his own belief that "severe forms of trauma are sometimes accompanied by an improvement in functioning" (p. 123).

[4] For example, consider the "Expert Report by Robert Morgan, PhD, Ashker, et al. v. Governor, et al., Case No.:C09-05796 CW (N.D. Cal.)" submitted under oath to a federal district court. Morgan opined that being housed in extremely harsh solitary confinement (the SHU in California's Pelican Bay State Prison) for "*ten or more continuous years* does not place inmates at substantial risk of serious mental harm" (p. 1; emphasis added), a position that he supported in part by citing the Colorado study. He described the study as "the most sophisticated study to date on the topic" of the effects of solitary confinement, claimed it showed "an absence of adverse effects for segregated inmates" (p. 1), and cited the results of his own meta-analysis (which was incorporated into Morgan et al. [2016], which I discuss later in this essay) to buttress his defense of long-term solitary confinement. Similarly, see the "Expert Report Provided in the Matter of BCCLA and JHS v. AGC, Court No.:S150415" by Jeremy Mills, PhD, filed in support of the continued use of solitary confinement in Canadian prisons. The Colorado study is described by Mills as "quite likely the most sophisticated longitudinal study to date examining the effects of segregation on mentally ill and non–mentally ill offenders" (p. 13). He also characterized meta-analyses like the Morgan et al. meta-analysis, of which he was a coauthor, as "a hallmark of the scientific process" (p. 12). Mills embraced the Colorado study's conclusions as supportive of his own, which were gleaned from his "clinical experience" working in segregation units on behalf of the Canadian Correctional Service. These included his view that both mentally ill and non–mentally ill prisoners usually need only "a few days" of "a period of adjustment" to get used to solitary confinement. He suggested that prisoners placed in solitary confinement "more frequently" forgo the adjustment period entirely because "they are familiar with the environment" (p. 14). Neither Morgan nor Mills acknowledged the Colorado study's numerous fundamental methodological flaws or indicated that the Morgan et al. meta-analysis on which they relied was based primarily on it.

little or no adverse long-term impact on inmates" (Government Accountability Office 2013, p. 39).

The Colorado study's continuing cachet in prison policy making and important legal circles means that its scientific bona fides bear especially careful analysis. Examining and deconstructing its methodology is a tedious but worthwhile exercise because it illustrates the difficulty of honoring norms of scientific rigor in a setting in which conventional research designs are nearly impossible to implement and necessary trade-offs are especially costly to the quality of the data collected. I turn to that exercise in Section II and to a deconstruction of the Morgan et al. (2016) meta-analysis in Section III.

II. Interrogating the Colorado Study

Results of the Colorado study appeared in two versions: a lengthy final report to the National Institute of Justice (O'Keefe et al. 2010) and a short article in the *Journal of the American Academy of Psychiatry and Law* (O'Keefe et al. 2013). I mostly discuss the more detailed National Institute of Justice report.[5] I also draw on two depositions, under oath, of Maureen O'Keefe, the lead researcher, in connection with prisoner litigation concerning Colorado's "supermax" facility (where much of the study was conducted). In response to detailed questions, O'Keefe discussed numerous issues not raised in the report or fully addressed in published exchanges following its release.[6]

Why the study was undertaken is unclear. Neither of the primary researchers had prior experience with solitary confinement. Maureen O'Keefe had a master's degree in clinical psychology but no prior involvement in research on the effects of isolation. Kelli Klebe was a psychometrician who also had no direct experience with solitary confinement (O'Keefe 2010, pp. 13–14). Yet they designed the study (pp. 77–79).

The study's impetus may have come from Larry Reid, warden of the Colorado supermax prison that housed prisoners assigned to administra-

[5] A number of brief but highly critical commentaries by prison researchers also questioned aspects of the methodology: Grassian and Kupers (2011), Rhodes and Lovell (2011), Shalev and Lloyd (2011), and Smith (2011). See also the response to at least some of these criticisms by Metzner and O'Keefe (2011).

[6] The two depositions are Deposition of Maureen O'Keefe, *Dunlap v. Zavaras*, Civil Action no. 09-CV-01196-CMA-MEH, October 5, 2010; and Deposition of Maureen O'Keefe at 96, 101 *Sardakowski v. Clements*, Civil Action no. 12-CV-01326-RBJ-KLM, October 25, 2013.

tive segregation. O'Keefe indicated that Reid "kept pushing for the study to be done" and served as a member of the study's advisory board (2010, p. 51). A few years before the Colorado study was planned, administrators at a Wisconsin supermax had lost a lawsuit over their use of solitary confinement (*Jones 'El v. Berge*, 164 F.Supp. 2d 1097 [W.D. Wis. 2001]), and Reid apparently wanted to avoid a similar decision. As O'Keefe (2013, p. 44) observed, "I believe [Reid's] concern was that Wisconsin had lost the case and it had severely restricted their ability to use administrative segregation."

The Colorado researchers said that they expected to find that administrative segregation had negative psychological effects: "We hypothesized that inmates in segregation would experience greater psychological deterioration over time than comparison inmates, who were comprised of similar offenders confined in non-segregation prisons" (O'Keefe et al. 2010, p. viii). If so, Warden Reid did not appear to share that view. The Colorado Department of Corrections then housed "three times as many people in solitary confinement as the average state prison system" (*Correctional News* 2012, p. 1). Moreover, O'Keefe (2013, p. 46) acknowledged that Reid "was very pro administrative segregation and all of us on the project felt that way."

Psychologist John Stoner, the mental health coordinator at the Colorado supermax prison, also strongly supported administrative segregation and served as a member of the study's advisory board. He had testified in the Wisconsin case that administrative segregation was not "as detrimental to mental health as others have found it to be" (*Jones 'El v. Berge*, p. 1104). Among other things, Stoner said that he was not troubled by Wisconsin's use of "boxcar" cells with solid metal doors that closed off visual contact and muffled sound because he thought they were "necessary for the protection of staff and other inmates" (p. 1104). He also observed in written testimony that prisoners in isolation who appeared to be seriously mentally ill were likely not as sick as other experts indicated; he speculated that they might be malingering. Although Stoner told the court in *Jones 'El v. Berge* that the isolated housing conditions at the prison were entirely appropriate, the judge disagreed. She held that the Wisconsin facility was unconstitutionally harsh for mentally ill prisoners and ordered them removed.

In any event, the Colorado researchers started out with a seemingly good idea and what appeared to be a reasonable research design. They would identify groups of prisoners housed in administrative segregation

(AS) and in the general population (GP), subdivided into those suffering from serious mental illness (MI) and not (NMI). Their psychological status would be tracked for 1 year to determine whether and how the different groups were affected by different conditions of confinement.[7] The characteristics of the AS and GP prisoners were not matched at the outset but were expected to be more or less comparable because all had committed rules violations for which they might have received an AS placement.

Assignments to AS were thus not random. The researchers reported that "placement into AS or GP conditions occurred as a function of routine prison operations, pending the outcome of their AS hearing, without involvement of the researchers. . . . Inmates who returned to GP following an AS hearing were assumed to be as similar as possible to AS inmates and, therefore, comprised the comparison groups" (O'Keefe et al. 2010, p. 17). The prisoners whom prison authorities chose to send to administrative segregation became the treatment group and those returned to the general population became the comparison group (again, with each group subdivided into those identified by the prison system as mentally ill and those not).

Unfortunately, the plan fell apart almost immediately. The prison context and "routine prison operations" fundamentally undermined the research design.

A. Contamination of Treatment and Comparison Groups

The study's implementation was compromised in two fundamental ways. It is important at this juncture to acknowledge the distinction between mere methodological "limitations"—respects in which a study is not perfect—and problems that are so fundamental that they make the resulting data uninterpretable. The two flaws from which the Colorado study suffered were fatal—separately and in combination.

1. *All Participants Were Exposed to the Treatment.* All participants in the study, including those in the comparison group, were initially placed

[7] Data for one group of participants—prisoners "with the most acute psychiatric symptoms" housed at a psychiatric treatment facility where they lived and interacted with one another "on their living unit" (O'Keefe et al. 2010, pp. 14–15)—did not bear directly on the issue of whether and how much prisoners were affected by AS. The researchers included them separately "to study inmates with serious mental illness and behavioral problems who were managed in a psychiatric prison setting" (p. 17). The prisoners in this group were not living in conditions remotely comparable to prisoners housed in conventional GP or AS units.

in "punitive segregation," a severe form of solitary confinement, for unspecified but not insignificant periods, before being assigned to administrative segregation or the general population. "At the time leading up to and during their AS hearing," the researchers acknowledged, "inmates have typically been in segregation" (O'Keefe et al. 2010, p. 8).[8] The reason was that Colorado prison officials were required to hold hearings to determine whether prisoners were guilty of infractions and if so whether AS punishment was warranted. Prisoners in Colorado as elsewhere are placed in special housing while they await the outcomes of their disciplinary hearings, often for days or weeks before the process is complete. Thus, the researchers also noted that "offenders reclassified to AS *remain* in a punitive segregation bed until an AS bed becomes available" (O'Keefe et al. 2013, p. 50; emphasis added).

Although this is routine correctional practice, its methodological implications were disastrous. It meant that all members of the comparison group were exposed to a severe dose of the isolation "treatment" before the study began. O'Keefe et al. (2010, p. 9) indicated that the punitive segregation conditions where prisoners were kept while disciplinary proceedings unfolded were so harsh that they were "only intended to be used for a short period of time." This severity distinguished it from AS, which was intended to be used for much longer periods. Here is how they described punitive segregation:

> Punitive segregation offenders remain in their cell for 23 to 24 hours a day, only coming out for recreation and showers, both of which are located in the living unit. Therefore, most do not leave the unit during their segregation time. Services including meals, library, laundry, and even medical and mental health appointments occur at the cell door. If a situation warrants an offender to be out of cell, the offender is placed in full restraints and escorted to a room within the unit

[8] Why "typically" is unclear. The report indicates that all prisoners (including the GP comparison groups) were placed in some form of isolation before, during, and shortly after their AS hearings. It is hard to imagine a procedure in which a prisoner would be taken directly out of GP, immediately given an AS hearing, and immediately returned to GP, without having spent time in some form of isolated housing. In fact, the authors reported that AS participants "on average completed their initial test 7 days (SD = 7.3) after their AS hearing," that GP participants on average "were tested 16 days (SD = 18.9) after their hearing," and that "on average, 43 percent of inmates . . . [had] been confined in segregation (40 percent in AS groups and 3 percent in GP groups) for an average of 18.2 days (SD = 18.1)" (p. 30). These figures are mathematically impossible. Moreover, they are at odds with O'Keefe's deposition testimony and with a statement in a more recent published "reflection" on the study (O'Keefe 2017).

> where he or she can meet privately. Many offenders do not like being taken out of their cells because of the use of full restraints. Additionally, they may not like leaving their cell because officers may take the opportunity to search the cell for contraband.
>
> Due to the disciplinary nature of punitive segregation, offenders are stripped of most privileges during their stay. Punitive segregation inmates are neither allowed to work nor permitted to participate in programs or education. Furthermore, their televisions are removed, and they cannot order canteen beyond essential hygiene items. (O'Keefe et al. 2010, p. 8)

Punitive segregation prisoners were denied visits, which were considered too labor intensive for prison staff to administer.

In contrast to AS, prisoners in punitive segregation also were denied the opportunity to engage in programming or education and were "unable to begin working their way toward leaving segregation" (O'Keefe et al. 2010, p. 9). Thus, even study participants who wound up in AS likely experienced punitive segregation as a much worse form of treatment.

This initial exposure of all participants to an especially harsh form of solitary confinement in punitive segregation made it impossible to draw meaningful inferences about any separate, subsequent effects of GP versus AS. There can be no comparison group in a study in which all of its participants are subjected to a harsh form of the treatment whose effects are being measured.

It is impossible to know whether or how control group prisoners were damaged by the time spent in punitive segregation and whether those effects continued throughout the study. Nor could anyone know whether the AS prisoners were actually relieved to enter the "treatment" because it was less harsh than punitive segregation. These imponderables could account for participants' psychological reactions, including the reported lack of differences between the AS and GP groups and the reported "improvement" or lack of deterioration of many members of the AS group. This was thus no longer a study of administrative segregation compared with no administrative segregation, but of varying and unspecified amounts of segregation experienced by everyone.

A different kind of analysis might have salvaged something by using the exact periods of overall exposure to administrative segregation–like conditions (including time in punitive segregation) as a continuous variable to estimate whether duration had an effect. However, the amount of time in segregation each prisoner experienced is not reported, so this kind of analysis was apparently not conducted. O'Keefe et al. (2010) treated their data as if they had done a classic treatment versus no treatment study, even though they had not.

The likelihood that initial exposure to punitive segregation conditions had significant negative psychological effects on most participants is more than just speculation. The National Institute of Justice report acknowledged that three of the four groups "showed symptoms that were associated with the SHU syndrome" from the outset (O'Keefe et al. 2010, p. viii), which seems a clear indication that the initial period of segregation adversely affected participants before their AS terms began. High levels of psychological distress measured during or after the prisoners' initial exposure to punitive segregation continued throughout the study. O'Keefe emphasized in a deposition that prisoners in all groups reported "pretty high elevations" of psychological distress (2010, p. 171) and that "clearly, very clearly, the offenders responded with very high elevations. They reported high levels of psychological distress" (p. 201).

Symptoms of distress were so elevated that the researchers wondered, and tried to test, whether the prisoners were malingering: "We had this huge rate of offenders who looked like they could be malingering" (O'Keefe 2013, p. 89). O'Keefe recognized, however, that high scores on a malingering scale "could indicate a lot of psychological problems." In the end, the researchers "didn't really believe that [the prisoners] were malingering" and discarded the results of the malingering scale without analyzing them (p. 89).

Thus, although the researchers acknowledged that most of the participants began the study very much affected by emotional and behavioral trauma, they seem not to have considered that much of that trauma resulted from time spent in the punitive segregation units. Nor did they consider that, when participants "naturally got better as time went on" (O'Keefe 2013, p. 91), it was likely because the conditions of punitive segregation that all of them had experienced were now alleviated, even for those who ended up in AS.

The amount of time that the study participants spent in punitive segregation was problematic, especially because even very brief periods of isolation can have damaging psychological effects. The United Nations Special Rapporteur on Torture, Juan Mendez, has noted that "it is clear short-term solitary confinement can amount to torture or cruel, inhuman, or degrading treatment" and recommended that solitary confinement "in excess of 15 days should be subject to an absolute prohibition"

(2011, p. 23). The United Nations adopted that recommendation in the "Mandela Rules," which defined "prolonged solitary confinement" as lasting "for a time period in excess of 15 consecutive days," and mandated prohibition of such prolonged confinement (Commission on Crime Prevention and Criminal Justice 2015, rules 43.1, 44). The National Commission on Correctional Health Care (2016) also characterized "prolonged solitary confinement" lasting for more than 15 days as "cruel, inhumane, and degrading treatment" because it is "harmful to an individual's health" (p. 260). Yet all of the prisoners in GP and AS experienced a nontrivial duration or dose of isolation that lasted well beyond this potentially damaging threshold. A key table in the National Institute of Justice report indicated that, at the time of their first test interval, participants had spent considerable average times in "Other seg": GP MI prisoners 12.4 days, GP NMI 39.8 days, AS MI 88.9 days, and AS NMI 90.3 days (O'Keefe et al. 2010, table 5).

In her deposition testimony, O'Keefe could not remember exactly how long study participants remained in punitive segregation before their charged disciplinary infractions were resolved. At one point, she said, "When an offender acted out, they were put in punitive seg and generally given notice of a hearing pretty quickly, and then the hearing happened, again pretty quickly after that" (2013, p. 93). Later she "guessed" the time was around "the two week mark" (p. 94). That was not remotely accurate, according to table 5 in the report, except for the GP MI group. O'Keefe later offered another estimate, this time that prisoners were kept in various punitive segregation units "an average of 30 days" before their initial testing session (2017, p. 2). This, too, is much less time than the National Institute of Justice report showed. In any event, it appears that all study participants were subjected at the outset to harsh conditions of punitive segregation for at least twice as long as the Mandela Rules would prohibit, even before the study officially began.

2. *Uncontrolled Cross Contamination.* The second fundamental flaw was as important as the first. It, too, occurred because placement and retention in AS were correctional rather than methodological decisions. The researchers admitted that they "lack[ed] control over the independent variable, which in this case is the conditions of confinement" (O'Keefe et al. 2010, p. 35). There was, in their words, "contamination across groups," because some AS participants "were not confined in segregation for their entire period of participation in the study" and because some GP participants "may have at some time during their study partic-

ipation been placed in punitive segregation or even AS" (p. 35). The researchers also acknowledged that prisoners in the various subgroups "may have [been in] multiple locations within a study period" (p. 35).[9] In fact, not only did participants move between AS and GP, but a number of them were housed in other conditions during the study, including the hospital and "community placement" (p. 36).

Transferring prisoners back and forth between locations and custody statuses is routine correctional practice, but it had disastrous methodological consequences. It meant that some AS prisoners in the study were released into GP for good behavior, some GP prisoners were placed in AS (or punitive segregation) for rule violations, and some members of both groups were transferred to other settings. Having both control and experimental group members move back and forth between treatment and control conditions (and other unspecified places) destroyed the integrity of the two groups and made it impossible to compare their experiences meaningfully.

The contamination occurred differently between groups. By the end of the study, only small and very different numbers of "uncontaminated" participants were left in each group.[10] Methodologically speaking, a true, a natural, or even a quasi experiment cannot be completed if researchers lose control of the integrity of their treatment and comparison groups. The researchers, however, simply aggregated the contaminated prisoners' data into the groups in which they were originally placed.

O'Keefe et al. (2010, p. 35) acknowledged that "one of the challenges of applied research is the researchers' lack of control over the independent variables," but that admission does not ameliorate the problem. They

[9] They wrote that "participants remained in their assigned group regardless of their placements throughout the prison system" (O'Keefe et al. 2010, p. 35), but mean by this that individual prisoners were considered to be in those groups for purposes of data analyses even though they did not actually remain housed there.

[10] There were only 26 "pure" cases in the AS MI group (of the original 64), 39 in AS NMI (of 63), 13 in GP MI (of 33), and only 11 in GP MI (of 43) (O'Keefe et al. 2010, p. 35). All the others moved back and forth between treatment, control, and miscellaneous other conditions on an unspecified number of occasions. Thus two-thirds (52 of 76) of the GP control participants spent time in segregation or other non-GP settings during the study period, and their self-reports were used to contrast their prison experiences and reactions with those of the AS prisoners, half of whom (62 of 127) spent unspecified amounts of time in GP or elsewhere. The "pure" cases were pure only in the sense that they were not contaminated by moving back and forth between treatment, control, and other conditions during the study. They were still "contaminated" by being exposed to punitive segregation before the study officially began.

nonetheless asserted that "a significant advantage of this study is the use of comparison groups to determine if [persons in AS] change over time differentially compared to similar groups who are not placed in AS" (p. 59). However, they did not compare similar groups and thus can reach no conclusions about differences in the groups' experiences.

In fact, it is impossible to conclude anything meaningful from the Colorado results. Lovell and Toch (2011, p. 4) in their initial commentary on it correctly concluded that "despite the volume of the data, no systematic interpretation of the findings is possible."

B. Additional Serious Flaws

The researchers' inability to maintain control of key aspects of their research created numerous additional methodological problems. These problems further negated the possibility that any credible or meaningful findings would emerge from the study.

The additional problems pertained to how the participants were selected and how the various groups were composed, what the researchers recorded (or failed to record) about the experiences of members of the different groups, and questionable data collection procedures. Most stemmed from unyielding correctional realities and some from unwise methodological choices.

1. *Sampling and Group Composition.* The initial sample was drawn from among prisoners deemed eligible for the study by virtue of having received a disciplinary write-up and scheduled hearing to determine whether they would be placed in AS or returned to GP. The initial group of eligible prisoners was much larger than the number selected to participate. The decision about whom to approach was made single-handedly and, as she would characterize it, "haphazardly" by O'Keefe: "I would determine who we used, who we included in our study" (2010, p. 116).

The major consideration for inclusion was proximity to the field researcher: "We had one researcher, so we had to be able to manage her workload" (O'Keefe 2010, p. 116). She described the process as "haphazard selection. . . . We didn't do it in a random fashion, but we didn't necessarily do it in a very targeted fashion either" (p. 116). Participants were drawn from only 10 of Colorado's 26 men's GP prisons (O'Keefe et al. 2013, p. 51). A disproportionate number came from Limon Correctional Facility "[because] it's fairly close" (O'Keefe 2013, p. 66). This was not mentioned in either the National Institute of Justice report or the briefer published version of the study. If there was anything significantly different about that prison, for example, if its punitive segregation unit (where participants were housed before the study began) was especially harsh or its GP units (to which many participants were returned) were particularly dangerous, troubled, or inhumane, then a disproportionate number of prisoners would have been affected by being held there.[11] There is no way to tell.

There was also unexplained and unnecessary imprecision in the composition of the groups. In addition to being composed of persons subjected to punitive segregation immediately before they entered GP, the GP group began as an amalgam of prisoners who subsequently lived under different conditions of confinement. Thus, "thirteen participants in the GP groups were selected from the diversion program (for being at risk of AS placement)" (O'Keefe 2010, p. 30). The report elsewhere implied that all of the prisoners were at risk of AS placement because all had AS hearings; apparently that was not true, and some were "diverted" out of the process entirely.

A potentially more serious problem concerned the composition of the AS group. O'Keefe et al. (2010, p. 8) asserted that "Colorado does not house protective custody; therefore, no AS placements occur at the request of inmates." This is a correctional non sequitur. Colorado may not officially house protective custody inmates, but they exist in every American prison system. Protective custody inmates often end up housed in AS, whether or not they formally request it. In the Colorado study, an unusually large group of AS participants were identified as having sex offender needs: 30 percent of the AS NMI prisoners and 44 percent in the full AS group (p. 45). In other prison systems, many, possibly all, such prisoners would be protective custody cases. To be sure, protective custody prisoners are subject to the painful and potentially harmful effects of social and sensory deprivation. However, they are in a very different situation psychologically than prisoners placed in AS for punishment. Protective custody prisoners typically prefer to be housed in AS-type conditions instead of what they regard as more dangerous GP environments. As a result, they are likely to be reluctant to voice complaints about living

[11] O'Keefe understood the implications of the sampling methods. Concerning work by others on the effects of administrative segregation, she wrote, "Of particular concern is that sampling procedures are often not discussed, and thus it is impossible to know if the findings were based on a representative sample" (2008, p. 127).

conditions or adverse emotional reactions, lest they be moved. That a third of the AS NMI prisoners and nearly half of the AS group overall in the Colorado study were probably protective custody cases undermined any straightforward interpretation of the data.

Gang members presented a similar problem. Thirty percent of AS MI prisoners and 43 percent of those in the AS NMI group were identified as gang members (O'Keefe et al. 2010, table 9). Being a gang member would ordinarily reduce a prisoner's willingness to report psychological distress because that would be a sign of vulnerability that might be interpreted as weakness.

Thus, nearly three-quarters of both the mentally ill and non–mentally ill AS prisoners were likely protective custody cases or gang members. Yet the researchers ignored the implications of this entirely.

2. *Uncontrolled Differences in GP Conditions.* The control condition—GP—referred to placement in one of 10 different prisons. However, none of the specific conditions of confinement at any of those prisons is described.[12] Variations in GP environments matter because, obviously, unless all GP prisoners experienced the same environment, they were not really in the same condition. If some of the GP environments were so troubled, dangerous, and harsh that they approximated or were worse than conditions in AS, it would be impossible to make meaningful comparisons.

A disproportionate number of study participants were housed in the Limon Correctional Facility (O'Keefe 2013, p. 66). This appears to have been an especially troubled prison when the study was conducted. In 2010, a journalist wrote about "Limon's long history of inmate violence, including two fatal stabbings in five years and the beating death of a correctional officer" (Mitchell 2010).[13] The prison's 5-year violent history encompassed the entire period of the Colorado study from July 2007 through March 2010 (O'Keefe et al. 2010, p. vii). This meant that many study participants came from (and GP comparison group prisoners remained in) an especially harsh and dangerous GP environment, perhaps one as psychologically stressful as an AS unit. In fact, Limon's vi-

[12] The published article indicated only that "GP inmates have access to significant out-of-cell time (e.g., >10 hours/day), jobs, and programming" (O'Keefe et al. 2013, p. 51). No additional information about the GP environments was provided.

[13] There were also allegations that in 2008 sex offenders at the prison were targeted by gang members who extorted them to pay "rent" and repeatedly threatened and assaulted them (*Davis v. Zavaras*, 2010 WL 625043 [D. Colorado 2010]).

olent history may have been serious enough to have precipitated recurring violence-related lockdowns (e.g., Associated Press 2007), including in the GP units where some of the control inmates were housed. None of this was commented on or taken into account.

3. *Uncontrolled Differences in AS Conditions.* Colorado study AS participants were ostensibly in the same study condition but were nonetheless exposed to very different conditions of confinement. These differences were not recorded or quantified and thus could not be taken into account. First, as I noted, all study participants experienced varying amounts of a harsh form of prison isolation, punitive segregation, before the study began. For a significant number (apparently, the majority) of the AS prisoners, that continued for a quarter or more of the length of the study. Thus, "When the study began, there was a 3-month average wait for inmates to be transferred to [AS]," which was "due to a shortage of beds. While on the waitlist, AS inmates were held in a punitive segregation bed at their originating facility" (O'Keefe et al. 2010, p. 19).

The median stay in punitive segregation for AS participants was reported as 99 days (which means that half were longer), although a very small group of prisoners were moved "quickly" into AS. Despite these very different periods in prestudy punitive isolation, all AS participants were lumped together for purposes of analysis.[14]

There was additional imprecision about how much and what kind of isolation any one AS participant experienced. Some "were not confined in segregation for their entire period of participation in the study" but were released into GP or other less onerous settings (O'Keefe et al. 2010, p. 19).

However, even beyond this, it is impossible to know exactly what conditions of confinement were experienced by participants who remained in AS throughout the study. The reason is that Colorado's AS program operated a "level" system in which a prisoner's "quality of life" (QOL) varied as a function of behavioral compliance and programming. Changes in QOL were meant to be incentives for compliance with unit rules and eventual reassignment to GP. The average length of AS stay was said to be 2 years, with the expectation that prisoners would spend at least 1 year in AS. However, the minimum stays specified for the QOL program

[14] The "distance between when they were ad-seged and when they went to CSP became longer and longer because of the wait list in DOC" (O'Keefe 2010, p. 108). An unspecified but not insignificant number of administrative segregation prisoners "were held in the punitive segregation bed but classified as ad-seg. And that's the—for the study average to be about 90 days, but people could be there pretty short, pretty long" (p. 109).

envisioned much shorter stays: 7 days at level I, 90 at level II, and 90 at level III—187 days altogether—after which prisoners were eligible for consideration for reassignment back to GP (O'Keefe et al. 2010, p. 11).

Providing achievable incentives for good behavior and early release from AS are sensible correctional practices. However, they, too, further compromised any meaningful interpretation of the study results.

This methodological problem was significant because the differences in QOL at different levels of AS were substantial. The researchers acknowledged that "it was expected that [prisoners in AS] might experience varying amounts of isolation based on the amount of time spent at different [QOL] levels" (O'Keefe et al. 2010, p. 40). But these varying amounts of isolation were not documented or taken into account.

O'Keefe acknowledged that the researchers initially wanted information from prison staff on participants' out-of-cell time, "to track every time they left their cell," but could not obtain it because the data "just were not coded consistently or every time" by correctional officers (2013, p. 55). That meant that the researchers were unable to track the basic facts of whether, when, and for how long any one prisoner was at one or another AS level or incorporate these data into their analysis (p. 60). O'Keefe et al. (2010, pp. 40–41) reported that staff records yielded "conflicting information," and "it was often difficult to decipher and/or interpret the records." Thus, "it was not possible to code or use [them] in the study."

4. *Failure to Control or Record Treatment Dose.* There was more to these uncontrolled and unrecorded variations than just minor differences in the amount or duration of isolation. The variations in isolation in the AS condition—including for the relatively few prisoners who stayed in AS continuously—were very significant. The QOL level III AS prisoners were given additional privileges and allowed to have jobs as orderlies or in the barbershop. This permitted significant out-of-cell time, during which the prisoners were presumably unrestrained and in contact with others.[15] These opportunities are rare in prison AS units anywhere and

[15] As O'Keefe et al. (2010, p. 12) noted, "Arguably one of the most important benefits of QOL level three is an offender's ability to have more contact with friends and family. While offenders' visits remain noncontact, they are increased to four 3-hour visits per month and four 20-minute phone sessions. . . . One additional benefit is that offenders may now be eligible to work as a porter or barber. . . . Benefits to being offered a job position include the ability to earn money, increased time out of cell, and two additional phone sessions per month."

constitute a significant modification in the nature of the isolation experienced by an unspecified number of AS prisoners. They introduced even more heterogeneity into the "same" condition in the study than already existed.

The researchers also noted that an AS prisoner who acted out could be even more significantly locked down by being placed "on special controls in the intake unit where he can be carefully monitored" and "additional sanctions may be imposed through the disciplinary process" (O'Keefe et al. 2010, p. 13).

None of these and other variations in actual day-to-day conditions of confinement were taken into account. The researchers also did not record and were unable to estimate other basic, important variations in the experiences and treatment of the study participants. These included the number of social or family visits prisoners had, visits from attorneys (O'Keefe 2010, p. 164), and the nature or amount of mental health services the prisoners (including those who were mentally ill) received. As O'Keefe summarized, "We did not look at any facet of segregation or correctional conditions that might affect the outcome of the study. We merely looked at, based on their conditions of confinement—that is, whether they had originally been coded 'AS' or 'GP'—and then noted 'if they reported worse change over time' " (p. 207). But whether a prisoner had originally been coded AS or GP did not indicate what "conditions of confinement" he had experienced in the course of the study.

C. Miscellaneous Data Collection Problems and Issues

In addition, there were very serious problems with how the Colorado researchers initially structured and eventually implemented the data collection process as well as with the dependent measures they used. Some of these problems were the product of the challenging nature of the prison environment. Others were not.

1. *A Single, Inexperienced Field Researcher.* Almost all the data collection was done by one inexperienced research assistant who had only a bachelor's degree, no graduate training, and no prior experience working with prisoners or in a prison setting. She was single-handedly responsible for conducting five to six separate testing sessions in which she administered between 10 and 12 separate tests with each of 247 participants in 10 different prisons.

The data collection was unusually challenging. O'Keefe noted, "Say when she was at CSP [the AS facility], she might have a whole bunch of [participants] and she would go back and forth checking to make sure that they were all right, and administering the questionnaires when she needed to" (2010, p. 118). Yet no one oversaw her day-to-day work (p. 130). O'Keefe had no recollection of ever observing her administering the tests and indicated Klebe did not (2013, p. 85).

2. *Solicitation and Consent.* When prisoners' participation and consent were solicited, they were told, somewhat misleadingly, that "we're looking at how inmates across the entire DOC are adjusting to prison life" (O'Keefe 2010, p. 199). O'Keefe characterized this as "being cautious without being dishonest" (p. 200). The consent form told prisoners that the "risks of this study to you are very small in contrast with the benefits that are high. This study will help us to figure out what types of men adjust better to prison and how to help those who are struggling with prison life" (O'Keefe 2013, pp. 81–82). This, too, was misleading. The study was not about the types of men who adjust better to prison and how to help them. Moreover, no consideration was apparently given to the possibility that prisoners might want to appear to be "adjusting" rather than "struggling." This would apply with special force to AS prisoners, hoping to advance their QOL level and with that gain additional privileges and earlier release from the unit.

3. *Prison Employee?* The field researcher had to complete "the full CDOC [Colorado Department of Corrections] training academy" and at all times was required "to wear a visible CDOC badge that permitted her unescorted access to the facilities" (O'Keefe et al. 2010, p. 28). Although O'Keefe was "not sure" how the field researcher introduced herself to prisoners, she conceded that "it could be" that prisoners thought the field researcher was a DOC employee (2010, p. 125).

Prisoners in general, and especially in AS units, are typically reluctant to confide in prison staff (including even mental health staff) because of potential adverse consequences. Those consequences can include increased surveillance, placement in degrading "suicide watch" cells, or transfer to or retention in some other form of AS. For these reasons, prisoners frequently avoid admitting that they feel suicidal, depressed, frightened, angry, panicky, out of control, or violent.

That prisoners could reasonably infer that the field researcher/prison employee was checking on their "adjustment" is likely to have dampened their willingness to disclose sensitive feelings. This possibility is nowhere discussed. Despite the fact that while the study was under way, O'Keefe acknowledged awareness of the fraught nature of prisoner-staff relations, especially in AS units: "Administrative segregation facilities are characterized by the complete control exerted over inmates by correctional staff. The typical 'we-they' dynamic between inmates and staff is exacerbated in segregated settings where inmates have almost no control over their environment. Prisoner abuses have been discovered and punished in administrative segregation settings, but in other situations Human Rights Watch found that 'management has tacitly condoned the abuse by failing to investigate and hold accountable those who engage in it'" (2008, p. 126; internal citations omitted).

4. *Undermining Trust.* Little was done to overcome what O'Keefe described as the "we-they" dynamic that she believed was likely to be exacerbated in prison AS units. Two related problems with the Colorado study likely exacerbated the effects of this dynamic. The first was an error of omission: no interviews were conducted to establish rapport with prisoners. O'Keefe indicated that "it was not part of the study to probe and ask them [the prisoners] about themselves" (2013, p. 75). Without rapport-building interactions, prisoners in the study were unlikely to have had much confidence that the field researcher was interested in their well-being or that personal revelations would be handled with sensitivity.

The second problem is more troubling. The field researcher was apparently required (or decided on her own) to challenge prisoners if she thought their answers were "questionable" or "untruthful, or if she found the pattern of their responses abnormal" (O'Keefe et al. 2010, p. 36). There was no explicit or systematic protocol by which this judgment was reached (none is described). In any event, the field researcher reviewed the prisoners' responses on the spot, in their presence, every time they completed a questionnaire. If she was skeptical, the prisoner was asked to redo the test. Prisoners could decide to redo the test or not, but "if the participant said he was being honest and the researcher still did not believe him, she marked the test as questionable" (p. 36).

These practices potentially created very significant data quality problems. They not only jeopardized the development of rapport or trust but also increased the chances that prisoners would give situationally desirable answers. In addition, the problems likely extended to more prisoners than only those who were challenged directly, but to other prisoners who learned through word of mouth that they would be asked to redo their questionnaires if the researcher was skeptical of their answers.

5. *"Untruthful" and Other Questionable Data.* Twelve percent of participants "had a questionable response pattern on any measure at any

time period" (O'Keefe et al. 2010, p. 36). It is unclear whether that figure included all participants who were asked about their answers or only those whose answers were marked "questionable." If challenged prisoners admitted being untruthful and redid the questionnaire, the second versions of their answers were incorporated into the study data. However, even if the field researcher was skeptical and prisoners chose not to redo their questionnaires, "we still included that in the study. . . . In order to increase our statistical power . . . we left those cases in" (O'Keefe 2010, p. 166).

In addition, 23 participants withdrew their consent and dropped out before the study was completed. However, their data were retained and used in the overall analyses (O'Keefe et al. 2010, p. 19). The dropouts constituted nearly 10 percent of the 247 participants. This meant that, in total, more than 20 percent of the participants whose data were included in the study results were adjudged to have given untruthful responses or withdrew from the study.

6. *An AS "Heisenberg Effect"?* The repeated testing procedure changed the conditions of confinement, especially for AS prisoners otherwise subject to extreme social deprivation. The six interactions of approximately an hour each between the field researcher and the prisoners, no matter how strained or superficial they might have been, increased the otherwise minimal social contact that AS prisoners had with people outside the segregated housing unit.[16] In many prison systems, there are many AS prisoners who get no visits at all. The mere act of repeatedly attempting to measure the effects of severe conditions of isolated confinement can change them, if only slightly, for the better.

7. *Miscellaneous Issues.* There were other irregular, questionable, and unexplained research decisions and data anomalies. Exactly why prisoners were assigned to AS or GP was not indicated, even though this was how the treatment and control groups were created. Assignment to AS was apparently nearly automatic: no more than "approximately 10 percent of hearings do not result in AS placement" (O'Keefe et al. 2010, p. 17). This raised questions, never addressed, about what accounted for the unusual outcome in the case of the group that was returned to GP.

[16] It apparently exceeded the contact AS MI prisoners had with mental health staff: "Offenders with mental illness who are stable are offered a one-on-one session at least once every 90 days," which takes place "in a noncontact booth in the visiting room" (O'Keefe et al. 2010, p. 11).

Nor were reasons discussed for why the NMI prisoners who returned to GP had more disciplinary infractions (average 16 each) than those sent to AS (13.2 average). Nor were reasons discussed for why AS MI prisoners had 70 percent more disciplinary infractions on average than the AS NMI inmates (22 infractions compared with 13.2; O'Keefe et al. 2010, table 9). Nor was there discussion of the effects of exclusion of prisoners from the study who did not read English at an eighth-grade level on the representativeness of the final group of participants, especially with respect to ethnicity and the prevalence of cognitive impairments.

D. Troubling Dependent Measures

There were also serious problems in the handling of dependent variables in the study. Dependent measures were said to have been selected on the basis of several important criteria. However, the first two criteria the researchers identified—"(1) use of assessments with demonstrated reliability and validity, (2) use of multiple sources for providing information (e.g., self-report, clinician ratings, files)" (O'Keefe et al. 2010, p. 19)—did not apply to the dependent measures that were actually used in the analyses.

1. *Unvalidated Scales and Instruments.* Some of the study's scientific bona fides were based on its claimed use of validated and objective assessment instruments. The researchers asserted that "the use of a reliable and valid standardized measure in the present study enabled objective assessment of psychological functioning" (O'Keefe et al. 2013, p. 57).

Indeed, O'Keefe acknowledged that "inaccurate judgments" could be made if instruments were not properly validated (2010, p. 22). However, she later conceded that only "a very low number" of the numerous scales and measures used, perhaps no more than one or two, had been normed or validated with a prisoner population (pp. 144–45).[17]

[17] There was no evidence that even the Brief Symptom Index (BSI), on which the researchers relied exclusively in the published version of the study, O'Keefe et al. (2013), had ever been validated with a prisoner as opposed to a "forensic" population. One study that the authors cited to support its psychometric properties (Kellett et al. 2003) concerned the BSI's reliability with persons suffering from intellectual disabilities and did not include a representative sample of prisoners (the "forensic" portion of the sample consisted of 45 "intellectually disabled" convicted persons who were "detained in a maximum security hospital" [p. 129]). The second, Boulet and Boss (1991), was a study of "psychiatric inpatients and outpatients who presented for evaluation at the forensic service of a psychiatric hospital" (p. 434). The third, Zinger, Wichmann, and Andrews (2001), focused on prisoners but did not report reliability or validity data for the BSI.

2. *"Constructs" That Could Not Be Interpreted or Compared.* The near-exclusive reliance on prisoners' self-report assessments was problematic because the researchers chose to separate the various scales into their component parts and then recombine items into eight separate "constructs." Instead of reporting scores on the instruments or scales themselves, only the constructs built from them were presented as standardized composite rather than numerical scores (O'Keefe et al. 2010, p. 22). This meant that the significance of reported overall trends and comparisons between groups was, as Lovell and Toch (2011, p. 4) put it, "difficult to assess because of the degree to which the data have been cooked."

There are a number of unanswered questions concerning construction of composite scales including their basic validity (whether the instruments measured what they purported to measure), whether the various subscales were reliable for this population, and whether the distributions of scores lent themselves to the statistical manipulations and recombinations that occurred. Transformations to the data, the number of instruments, items, and constructs, and the amount of scale and subscale reconstruction that occurred make the results difficult to put in the context of any larger literature using the same self-reported assessments.

3. *Ignoring Behavioral Data.* Researchers who use many rating scales (especially ones not validated for the particular population) generally use other methods of data collection as a validity check. The most basic is a face-to-face interview to establish rapport and acquire background information. When possible, behavioral data (by records reviews or behavioral rating scales completed by others) are included. These different sources of information should be reconcilable, and the interviews provide the glue that binds them. Prison researchers typically take things prisoners say to them very seriously, in part because they contextualize other things being measured or studied. However, no interviews were conducted in the Colorado study, and little or no special effort appears to have been expended to establish rapport. Instead, the researchers engaged in context-free coding and analysis of answers on prepackaged forms associated with tests not typically used with this population. As Lovell and Toch (2011, p. 3) observed, "Readers find themselves swimming in a flood of psychometric data; every so often a clue drifts by, lacking, however, a tether to the context—to what was going on around the prisoners and staff while they carried out this study—we are left to guess what it might mean."

Other kinds of data collection were contemplated including asking corrections officers and clinicians to complete rating scales: "The Brief Psychiatric Rating Scale was completed by clinical staff and the Prison Behavior Rating Scale was completed by correctional officers and case managers" (O'Keefe et al. 2010, p. 26). However, key details about this process were omitted (i.e., exactly who was supposed to complete scales, when, and with what kind of training). In the end, it did not matter. The rating scales were infrequently completed and the responses were too unreliable to be useful. The data were discarded. The researchers ultimately relied only on data from prepackaged, field researcher–administered rating scales.

There was one potential exception. Prison mental health staff kept official accounts of genuine psychiatric emergencies or "crisis events." Any situation that required "immediate psychological intervention is considered a crisis event; crisis events are documented by clinicians" (O'Keefe et al. 2010, p. 42). Because these are typically extreme, clinically significant events, they tend to be reliably recorded. If the prisoners' self-reporting was valid, the results should be more or less consistent with behavioral measures of psychological distress or crisis. In the Colorado study, they were not. Among the 33 GP MI prisoners for whom data were reported, there were only three "crisis events" (on average, one for every 11 inmates). Among the 64 AS MI prisoners, there were 37 "crisis events" (one for every two; O'Keefe et al. 2010, figs. 29, 30). This suggests that at least some mentally ill prisoners were doing much worse in AS than their counterparts were doing in GP.

The researchers dismissed the implications of this incongruity: "Because the number of participants who experienced a crisis event was so small, it was not possible to include this variable as an outcome measure in the change over time analyses" (O'Keefe et al. 2010, p. 42). Thus the significant disparity between self-reports and the behavioral measures was ignored, even though it directly contradicted the study's main finding that AS did not adversely affect the mental health of mentally ill participants. Instead, as they put it, because the mental health crisis data "raise more questions than they provide answers," they were deemed "outside the scope of the current research" (p. 42).

In sum, for all of the above stated reasons, the Colorado study is so methologically flawed that literally no meaningful conclusions can be drawn from it. Drastic compromises necessitated by the complex realities of the prison setting and a series of questionable methodological decisions made by the researchers rendered its results uninterpretable. The Colorado study was not the "most sophisticated" study done to date on the psychological effects of solitary confinement. Its results do not "need to be taken seriously," but cannot be taken for anything at all. Commentators who have praised the study either did not read it very carefully, were unaware of available sources of information on how it was actually conducted, or did not seriously consider the implications of its fundamental flaws.

Ordinarily, a study of this sort would die a quiet death, notwithstanding an occasional prison system's attempt to resuscitate it to defend questionable segregation practices or a scholar overlooking its flaws because its findings comport with his or her own views. However, it has recently been given a second life, figuring prominently in a recently published meta-analysis (Morgan et al. 2016). Its results threaten to live on in another form and to misrepresent the findings of the large, long-established, and frequently reconfirmed literature on the harmful effects of solitary confinement.

III. The Limits and Dangers of Meta-Analysis

Meta-analysis—"a quantitative method of synthesizing empirical research results in the form of effect sizes" (Card 2012, p. 7)—is an important methodological advance that allows researchers to estimate the overall magnitude of relationships between variables. However, it cannot substitute for careful narrative reviews of scientific literature. Meta-analysis comes with substantial limitations, especially for prison research. The prison setting rarely lends itself to collection of meaningful quantitative data capable of generating the kinds of effect sizes on which meta-analyses depend. Most classic book-length treatments of prison life have been primarily ethnographic—not quantitative at all. They contain few if any numerical data, including in the seminal American works by Cressey (1940), Sykes (1958), Toch (1975, 1977), Jacobs (1977), and Irwin (1980) and major comparable British works including Cohen and Taylor (1972) and Crewe (2009).

Similarly, few quantitative effect sizes appear in studies of solitary confinement. This is true of the studies that tell us much of what we know about these institutions, how they operate, and the lengths to which prisoners must go in order to survive inside them, including those from Rhodes (2004), Shalev (2009), Reiter (2016), and Kupers (2017). It is also true of most of the numerous studies of the negative psychological consequences of prison isolation that are discussed in the most-often-cited literature reviews. The nature of the settings and the routine prison operations that govern them make many kinds of conventional research designs impossible to implement.

Because the best prison research is qualitative, or does not lend itself to generating effect sizes, meta-analyses conducted on many important prison topics will be compromised by serious sample bias, resulting in "the drawing of inferences that do not generalize to the population of interest (typically all research conducted on the topic)" (Strube, Gardner, and Hartmann 1985, p. 66).

The concern is not only that meta-analyses on important prison topics almost invariably ignore or underrepresent the larger literature, but also that they privilege certain kinds of studies far beyond their actual scientific merit, and do so in a way that many readers are unlikely to appreciate. One critique rightly observed that readers "might not be motivated to look beyond the meta-analyses themselves due to confidence in the objective, straightforward nature of the tasks of conducting a meta-analysis, reporting findings, and making recommendations" (Coyne, Thombs, and Hagedorn 2010, p. 108). Reducing entire studies to single or multiple effect sizes almost invariably creates a false equivalency between them. Readers can easily be mesmerized by arrays of numbers that appear simply and accurately to represent highly complex and substantially different underlying realities.

The two meta-analyses contained in the Morgan et al. (2016) article suffer from all of these problems and more. They need to be scrutinized carefully because of the stakes involved and the possibility that they will mislead correctional decision makers and policy makers by their "surprising results," ones that, as the authors say, "do not fit with people's intuitive analysis of what happens when you isolate offenders" in solitary confinement. The resulting conclusions are indeed "in marked contrast to the 'fiery opinions' . . . commonly presented in the scientific and advocacy literature" in which solitary confinement "has been likened to torture, with debilitating consequences" (p. 455). They warrant conscientious examination.

A. Truncating the Scope of Literature Reviewed

The first problem with Morgan et al. (2016) is the tiny number and unrepresentative nature of studies included in its two separate meta-analyses. Literature reviews, whether narrative or meta-analytic, are useful only if they faithfully represent the literature being examined. As Card (2012, p. 10) put it, "If the literature reviewed is not representative of the extant research, then the conclusions drawn will be a biased representation of reality." Morgan et al. (2016) excluded a vast number of published studies, including most of the key works.

The first meta-analysis, "Research Synthesis I," reported that over 90 percent of the published material that they found on the topic was eliminated: "Of the 150 studies located, only 14 (or 9.3 percent) were suitable for analysis according to our inclusion criteria" (Morgan et al. 2016, p. 442). The second meta-analysis, "Research Synthesis II," began with an astonishing 40,589 articles, which were reduced by "trained research assistants" using unspecified methods to 61. A "trained research assistant" then used unspecified methods to reduce that number to 19 (0.05 percent of the initial literature; pp. 442–43).

A meta-analysis that includes so little of the available relevant literature is not a synthesis of much of anything. In addition to the drastic reduction in the sheer number of articles included, the selection criteria used by Morgan et al. (2016) excluded key studies but included questionable other ones. Among the articles excluded is Grassian (1983), regarded as one of the seminal studies on the adverse effects of solitary confinement. Morgan et al. also ignored most of the work discussed in widely cited literature reviews by Haney and Lynch (1997), Haney (2003), Grassian (2006), Smith (2006), and Arrigo and Bullock (2008).

Despite the small numbers of studies included, tables reporting effect sizes seem to suggest that a vast number of studies were taken into account. A closer look reveals something different. Many of the studies have little or nothing to do with the key question of whether and when solitary confinement is psychologically harmful. Morgan et al. (2016) included studies that addressed medical outcomes, and behavioral outcomes such as recidivism and institutional misconduct, that have not been widely studied and are not central to understanding solitary confinement's psychological effects. Thus, despite the drastic reduction in overall number of studies, many of the studies actually included were simply beside the main point.

When the largely irrelevant studies are set aside, only six studies on the psychological effects of solitary confinement remain in the first meta-analysis and 10 in the second. Two in the first were excluded from the second and six others were added.[18] No explanation is given for why different sets of articles appeared in the two meta-analyses. In any event, the truncated set of 12 studies was not remotely representative of the larger scientific literature on the psychological effects of solitary confinement.

B. Overreliance on the Colorado Study

Even "the most thorough sampling and complete data recovery cannot make up for basic limitations in the data base" (Strube, Gardner, and Hartmann 1985, p. 68). Indeed, "An experiment that is deficient in either statistical conclusion validity, internal validity, or construct validity is meaningless and, therefore, worthless. Consequently, it should not be used" (Chow 1987, p. 266). Notwithstanding these basic methodological truisms, tables 2 and 4 in Morgan et al. (2016) reveal that both meta-analyses relied primarily on the fatally flawed Colorado study. It provided the bulk of the effect sizes on which their overall conclusions were based.

Thus, in the first meta-analysis, I counted 24 of 50 relevant effect sizes on "psychological outcomes" that came from the Colorado study. In the second meta-analysis, 140 of 210 effect sizes came from the Colorado study.[19] Because of its sample size, the weights given to the multiple effect sizes from the Colorado study dwarf those of most of the other studies included.

As tables 2 and 4 in Morgan et al. (2016) make clear, they repackaged the Colorado results in a way that allowed them to dominate the analyses.[20] Thus, when they claimed that their results "are even more compelling when one considers that primary studies with the strongest designs produced much smaller effects," they were referring primarily to the un-

[18] The first (Morgan et al. 2016, table 2) included six studies that explicitly addressed psychological effects of solitary confinement: Ecclestone, Gendreau, and Knox (1974), Suedfeld et al. (1982), Miller and Young (1997), Zinger, Wichmann, and Andrews (2001), Andersen et al. (2003), and O'Keefe et al. (2010). The second (Morgan et al. 2016, table 4) added six studies: Walters, Callagan, and Newman (1963), Miller (1994), Coid et al. (2003), Cloyes et al. (2006), and Kaba et al. (2014); but it omitted Suedfeld et al. (1982) and Andersen et al. (2003).

[19] "Anti-social indicators" such as "re-admission" and "behavior" like re-arrest and "physical health" outcomes were omitted from this calculation of psychological effects.

[20] Zinger, Wichmann, and Andrews (2001) accounted for another four effect sizes in table 2 and 30 in table 4. It too is fundamentally flawed, as I explain in the next section. By my count, it and the Colorado study account for 28 of 50 relevant effect sizes in the first meta-analysis and 170 of 210 in the second.

interpretable O'Keefe et al. (2010) study. However, few if any of the fundamental defects of the Colorado study were even mentioned and none was seriously engaged. Instead, the authors simply described the Colorado study as "the most sophisticated study" ever done on the topic (Morgan et al. 2016, p. 441) and relied on it for the bulk of their conclusions.[21]

C. Including Other Methodologically Flawed Studies

There are serious problems with a number of the other studies included in the Morgan et al. (2016) analyses. For example, Zinger, Wichmann, and Andrews (2001) accounted for the next-largest number of effect sizes in their meta-analyses. However, there are several problems with how the results of this study were treated and serious issues with how the study itself was conducted, raising questions about whether it should have been included at all. Its sample size is erroneously listed in table 2 as 136. Although 136 was the initial number of participants, only 60 remained at the end of 60 days. The *N* shown in table 4 is, correctly, the 60 who remained, but that also is misleading. That number includes a majority of prisoners in the "administrative segregation" group (13 of 23) who were there voluntarily. Only 10 involuntary prisoners remained in administrative segregation at the end of 60 days. Thus this study was weighted far too heavily in the first meta-analysis and given a misleading weight in the second.

The results of Zinger, Wichmann, and Andrews (2001) are in any case impossible to interpret. They are based on data from a sample that combined "voluntarily" and "involuntarily" segregated prisoners. Voluntarily isolated prisoners (such as protective custody prisoners who "choose" to be in isolation) control their own fates; at least in theory, they can leave. In addition, in most cases they know that by staying they are at least safe from threats to their well-being elsewhere in the prison system, ones they presumably fear and necessarily want to avoid more than the pain and harm they may endure in solitary confinement. They are thus

[21] Morgan et al. (2016) appear to have overweighted the disproportionate number of effect sizes they took from the Colorado study, treating the *N*'s in each group as though their integrity was maintained throughout. However, as I noted, the bulk of the Colorado study participants moved back and forth between groups. Thus the "uncontaminated" cases are far fewer than Morgan et al. cited and used. Because O'Keefe et al. (2010) did not disaggregate their data, Morgan et al. must have relied on the confounded results, treating all participants as if they remained in their original groups for the duration of the study and weighted effect sizes as if this had been the case.

motivated to adapt to their isolation—or to appear to have adapted to it—in ways that involuntarily isolated prisoners are not. They should not be treated as if their experiences represent the effects of solitary confinement on involuntarily segregated prisoners.

A second and more important problem is the significant amount of attrition that occurred. Especially in longitudinal research, participants leave studies for various reasons. This inevitably complicates comparisons over time or between groups because people who remain are likely to be different from those who leave, thereby changing the compositions of the groups in ways that are difficult to specify.[22] This is especially a problem in prison research because prison administrators decide where prisoners are housed, under what conditions, and for how long; they do so on the basis of considerations that have nothing to do with the goals of researchers. In Zinger, Wichmann, and Andrews (2001), the reduction in the number of administrative segregation prisoners after 60 days, from 83 to 23, only 10 of whom were involuntary, means that attrition reduced the number of involuntarily segregated prisoners by 80 percent. The reasons for the attrition were not given.

Attrition is seldom random. That it results largely, if not entirely, from decisions made by prison administrators means that Zinger, Wichmann, and Andrews (2001) wound up with a group that was significantly different, in indeterminate ways, from the group with which they began.[23] They do not report whether and in what ways the prisoners who remained differed from those with whom the study began.[24]

[22] Zinger, Wichmann, and Andrews acknowledge this: "Attrition is a major drawback to psychological research in general. The problem with attrition is especially relevant to the evaluation of the psychological effects of segregation" (2001, p. 56). However, they ignored the extent of this problem in presenting and interpreting their results.

[23] If, for example, disproportionate numbers of transferred prisoners were considered too "vulnerable" to remain in administrative segregation, were reacting especially negatively, or were adjusting poorly and were especially effective at convincing the prison administration to return them to the general prison population, those left behind would be, by definition, those least affected by the experience. Alternatively, if those who remained at the end of 60 days were the most recalcitrant and least compliant, perhaps explaining why the prison administrators were less likely to release them, they may have been especially "difficult" prisoners who were less likely to admit vulnerability or weakness in the assessments they underwent. Or if the voluntary administrative segregation prisoners remaining after 60 days were the least willing or able to return to the general prison population, they may have been unlikely to admit that they were suffering lest this jeopardize their continued safekeeping. Any of these possible scenarios could greatly compromise interpretation of the results, and none of them appear to have been considered.

[24] The assertion that "none of the attrition was attributable to prisoners being incapable of participating in the study because of episodes of delusion or hallucination or suicide at-

An additional methodological problem was acknowledged in passing but not fully discussed, either in the published article or in Zinger's (1998) dissertation, on which it was based. "Practice effects" are a common problem in longitudinal studies because they require repeated administration over time of the same tests or measures. Participants may recall the questions and intentionally or inadvertently try to reproduce the same or similar answers, or lose interest and reply with stock, rote answers, or, if the tests include performance measures, improve (because of practice) each time they take the test. If any of these things occurs, the existence of real changes (especially negative ones) will be masked or minimized.

Zinger (1998) himself recognized that "artifacts of repeated testing" likely played a role in producing apparent improvements in functioning and the lack of signs of deterioration and that practice effects may have accounted for prisoners "report[ing] less problems over time" (p. 93). He also observed that it is well known that "participants lose interest in answering repeatedly to identical questions and tend to report less problems over time" (p. 92).[25] Thus, practice effects may have accounted in large part for the findings of "no change" or "improvement" on the measures used and repeatedly administered.

There are also significant problems with several other studies that were included in the already small group that Morgan et al. (2016) considered. For example, Cloyes et al. (2006) did not compare administrative segregation with nonadministrative segregation at all. Instead, all of the prisoners involved in their study were in solitary confinement. The effect size Morgan et al. reported was the only statistical test of differences between groups that appeared anywhere in Cloyes et al. (2006, p. 772). However, it is a *t*-test of differences in Brief Psychiatric Rating Scale scores between two groups of solitary confinement prisoners—those identified as seriously mentally ill or not, both of which were housed in isolation. Data from this study did not belong in the meta-analysis.

---

tempts" (Zinger, Wichmann, and Andrews 2001, p. 71) sets far too high a threshold and does not adequately address the matter. "Episodes of delusion or hallucination or suicide attempts" are hardly the only measures of whether someone is being so adversely affected that he would seek to be transferred elsewhere or, in the opinion of a correctional administrator or mental health staff member, need to be moved.

[25] Zinger, Wichmann, and Andrews (2001) did acknowledge that reports of "better mental health and psychological functioning over time" are "common in studies which rely on studies with repeated measures designs" (p. 74) but then ignored the implications of this for interpretation of results that showed exactly this.

Walters, Callagan, and Newman (1963) arguably does not belong either. It is over 50 years old and, more importantly, the participants were all volunteers. They were not typical of prisoners involuntarily placed in solitary confinement. In addition, the study lasted only 4 days, not long enough to reach a conclusion that the psychological effects of solitary confinement are minimal. The one effect size Morgan et al. (2016) reported, for "anxiety," is .57 with a weight of .726 (table 4, p. 452). Yet the only mention of numerical data for anxiety in Walters, Callagan, and Newman's study was this: "More isolated than non-isolated prisoners reported an increase in anxiety from the pre-test to post-test period ($p =$ .038, Fisher's Exact Probability Test)." It is impossible to calculate an effect size from this statistic.

Another included study, Andersen et al. (2003, table 2), reported only chi-squares and *p*-values. It is not clear how Morgan et al. (2016) managed to calculate effect sizes from those data.

The decision to include Ecclestone, Gendreau, and Knox (1974) is also questionable. The study is more than 40 years old and, more importantly, included only prisoners who volunteered to spend 10 days in isolation. For previously noted reasons, the experience of volunteers is not comparable to that of involuntary administrative segregation prisoners. In addition, the study used an almost indecipherable measure of psychological functioning—the Repertory Grid Technique—which does not appear to have been used in published prison research before or since.[26] Moreover, half of the initial participants "quit the experiment after two days of solitary confinement" (p. 179), which meant that the assignment of participants was no longer "random," the results suffered from significant attrition bias, and the remaining volunteer participants knew that they could leave whenever they wanted. Notwithstanding these problems, Ecclestone, Gendreau, and Knox concluded that isolated confinement was "not more stressful than normal institutional life" (p. 178). Morgan et al. (2016) included this study in both meta-analyses and singled it out as having one of the stronger research designs (along with Zinger, Wichmann, and Andrews [2001] and O'Keefe et al. [2010]).[27]

[26] Description of the nature and scoring of the Repertory Grid Technique was so complicated that it consumed nearly two full pages of text (Ecclestone, Gendreau, and Knox 1974, pp. 180–81).

[27] The studies deemed to have stronger research designs were identified by name only in Morgan et al.'s (2016) Research Synthesis I, although an estimate of the strength of the designs was also apparently used in Research Synthesis II. Morgan et al. concluded that

In sum, Morgan et al.'s (2016) meta-analyses were based on one fundamentally flawed and uninterpretable study (O'Keefe et al. 2010), another with an attrition rate of 80 percent over a 60-day period (Zinger, Wichmann, and Andrews 2001), two that were four decades old and included only volunteers (Walters, Callagan, and Newman 1963; Ecclestone, Gendreau, and Knox 1974), and one (Cloyes et al. 2006) that could not provide an effect size on the impact of AS.

Few readers are intimately familiar with the solitary confinement literature or willing to invest the effort to read and evaluate each of the studies cited in Morgan et al. (2016). Similarly, few are willing to carefully to examine the hundreds of effect sizes included in the two meta-analyses or are able to make judgments about the propriety of the particular statistical techniques used in the calculations.[28] The presentation of a vast array of numerical data in Morgan et al. gives the impression of an objective representation of equally meaningful effect sizes, but it is not the reality. Their conclusion that solitary confinement has modest or no significant negative psychological effects is not at all what a significant preponderance of the relevant empirical research shows and is at odds with findings

these studies with "stronger designs" were the ones that showed "less impairment" due to isolated confinement (p. 456). My critical discussion of the individual studies in question shows why.

[28] Morgan et al. (2016) appear to have used statistical methods that require very stringent assumptions and will give misleading results if these assumptions are violated (e.g., Aguinis, Gottfredson, and Wright 2011). Furthermore, the meta-analytic method they used requires a large number of studies to assess these assumptions, and there were not enough studies to assess them. Specifically, they used a random-effects meta-analysis model. This model assumes that the included studies are a random sample from some definable universe of studies. For example, are the prisons represented in Morgan et al.'s meta-analysis a random sample of all US prisons? If not, they cannot claim that their results generalize to this universe. Random-effects meta-analyses also assume that weights and sample sizes are uncorrelated with the effect sizes. If they are correlated, the results will be biased. The correlation between the sample sizes and effect sizes reported in their table 1 indicate that the correlation is about $-.5$, which could severely bias the results. In a random-effects meta-analysis, both the mean and the variance of the effect sizes in the universe are key parameters that need to be estimated and both require confidence intervals. Morgan et al. reported only the sample estimate of the variance and not the confidence interval. However, the confidence interval for the variance requires a strong assumption of normally distributed effect sizes, and the confidence interval is very sensitive to minor violations of this assumption. A large number of studies are needed to assess the normality assumption—much larger than the number used. Morgan et al. also appear to have used a new and unproven method for combining multiple effect sizes from a single study. This method requires at least a moderate number of studies (10–20, the more the better), more than the separate meta-analyses that were used. Finally, Morgan et al. also used extremely crude and inaccurate methods to approximate effect sizes in studies that did not provide enough information to correctly compute an effect size.

that are consistent across many decades, theoretically coherent, and buttressed by a very large and growing literature on the harmful effects of social isolation in contexts other than prison.

Misleading repackaging of bad data can ripple through the field and produce an echo chamber in which motivated commentators repeat each others' flawed conclusions. Thus O'Keefe (2017, p. 5) recently asserted that "a recent meta-analysis found small to moderate adverse psychological effects resulting from [solitary confinement] that were no greater in magnitude than the overall effects of incarceration. These findings are consistent with our Colorado results." She was referring to the Morgan et al. (2016) meta-analysis, whose conclusions were not only "consistent" with the Colorado results but based largely on them.

IV. Conclusion

These two studies offer several cautionary tales about the fraught nature of prison research, especially on the methodologically challenging and politically charged topic of solitary confinement. The first of these tales is about the potential influence of bad, uninterpretable data on public discourse and correctional policy. Once the results of research that bear the trappings of science enter into public and policy discourse, it is difficult to correct the record, especially when motivated advocates are willing to overlook fatal flaws in the research. Unfortunately, when this transpires, researchers can lose control of the narrative by which their research is described and the manner in which it is applied. For example, O'Keefe has repeatedly and steadfastly defended her Colorado research but has opposed the uses to which others have put it. She was emphatic that she did "not believe in any way and we do not promote the study as something to argue for the case of segregation. . . . My interpretation is that people believe that this study sanctions administrative segregation for mentally ill and non–mentally ill alike. . . . I do not believe that the conclusions lend to that and that is not the intended use of our study" (2013, p. 96).[29] Yet, that is exactly the use to which a number of interested parties have put it.

[29] Two prominent advisory board members, Jeffrey Metzner and Jamie Fellner (2010), published a "post–Colorado study" article that seemed to contravene the study's findings. They conceded that "isolation can be harmful to any prisoner" and noted that the potentially adverse effects of isolation include "anxiety, depression, anger, cognitive disturbances, perceptual distortions, obsessive thoughts, paranoia, and psychosis" (p. 104)—not at all what

The Colorado study is also a stark reminder that attempts to implement conventional experimental or even quasi-experimental research designs in prison environments face a number of often insurmountable obstacles. The ordinary demands of prison operations nearly always doom even the most carefully planned such studies, and certainly anything resembling a traditional experiment. Savvy prison researchers understand that the desire to treat a prison environment as if it were a research laboratory should be resisted. Real people live (and die) in prison, a setting in which the core dynamics between prisoners and staff are governed by forces beyond the researchers' control.

In separate but related ways, both the Colorado study and the Morgan et al. (2016) meta-analyses underscore the pitfalls of allowing the veneer of scientific rigor to substitute for its reality. They also show the limitations of focusing on quantitative outcomes with little or no concern for precisely how and under what conditions data were acquired. The decontextualized and de-individualized approach to data collection that characterized the Colorado study allowed researchers to treat all participants within each of the study groups as if they were the same, when clearly they—and especially their prison experiences—were not. Ignoring the prison context and individual prisoner trajectories helped render the findings incoherent and uninterpretable.

Similarly, Morgan et al. (2016) illustrate the shortcomings of attempting to apply an otherwise useful approach for summarizing quantitative data to environments as complex and variable as prisons (or especially solitary confinement units). Whatever the benefits of reducing empirical results to effect sizes may be, omitting an entire field's best-known and most in-depth works from consideration because most do not lend themselves to meta-analytic reductions means that nuance and context are inevitably ignored. The compromise in "scientific truth" is far too great.

Some critics of meta-analysis argue that "a literature review should *not* be a formalized or standardized one" (Chow 1987, p. 267; emphasis

---

the Colorado study claimed. Metzner and Fellner's deep concerns led them to recommend that professional organizations "should actively support practitioners who work for changed segregation policies and they should use their institutional authority to press for a nationwide rethinking of the use of isolation" in the name of their "commitment to ethics and human rights" (p. 107). Zinger has become an eloquent critic of the use of solitary confinement in Canada (e.g., Makin 2013) even though defenders of the practice continue to cite his dissertation research to justify its use.

added). As Chow observed, "It is not the case that narrative reviews lack rigor. To the contrary, rigor is maintained by reviewers of the traditional [narrative] approach when they evaluate the validity of individual studies" (p. 268). Meta-analyses, even when done well, risk compromising the richness of the prison data they seek to summarize.

In any event, the magnitude of what can be and often is lost in the course of the compromises made in the kind of research critically discussed in this essay often goes unrecognized. Amid thousands of data entries and hundreds of effect sizes reported in these two studies, there are few references to the core subjectivity, institutional trajectory, or life outcome of a single individual prisoner confined in an isolation unit. Nor is there acknowledgment that the studies focused on human beings rather than on interchangeable data points.

Martha Nussbaum (1995) noted in a different context that regarding people as "fungible" and denying them their subjectivity are powerful ways to ensure their objectification. Objectivity in prison research is a worthy goal, except when it results in objectification of prisoners and others in the prison environment. Feeley and Simon (1992) observed that the era of mass imprisonment occasioned and was facilitated by the emergence of a "new penology" whose key elements—"statistical prediction, concern with groups, strategies of management"—shifted the focus of the prison enterprise "toward mechanisms of appraising and arranging groups rather than intervening in the lives of individuals" (p. 459). This actuarial approach still defines the modern prison. It should not be made worse and reinforced by scholarship that exacerbates rather than alleviates or exposes these depersonalizing tendencies.

Studying only at a distance, as the research criticized in this essay did, requires precisely that kind of objectifying sacrifice. If John Irwin was right, that the close study of people in general and prisoners in particular uncovers their humanity, and I think he was, then the opposite is also true. Studying prisoners at a distance, without trying fully to understand and adequately to convey the conditions in which they live or to gain an "appreciation of their meaning worlds, motivations, and aspirations" (1987, p. 47), leaves us with little insight into basic truths about them. That includes whether and how much they are adversely affected by near-total deprivation of meaningful sensory and social contact.

The insurmountable methodological flaws of the Colorado study and the fundamental inadequacy of the Morgan et al. (2016) meta-analysis should preclude policy makers from using either in debates over the proper use of solitary confinement and the nature of its psychological effects.

REFERENCES

Aguinis, Herman, Ryan Gottfredson, and Thomas Wright. 2011. "Best-Practice Recommendations for Estimating Interaction Effects Using Meta-Analysis." *Journal of Organizational Behavior* 32:1033–43.

American Psychological Association. 2016. *Statement on the Solitary Confinement of Juvenile Offenders*. Washington, DC: American Psychological Association, Public Information Government Relations Office. http://www.apa.org/about/gr/issues/cyf/solitary.pdf.

Andersen, Henrik, D. Setstoft, T. Lillebaek, G. Gabrielsen, and R. Hemmingsen. 2003. "A Longitudinal Study of Prisoners on Remand: Repeated Measures of Psychopathology in the Initial Phase of Solitary versus Non-solitary Confinement." *International Journal of Law and Psychiatry* 26:165–77.

Arrigo, Bruce, and Jennifer Bullock. 2008. "The Psychological Effects of Solitary Confinement on Prisoners in Supermax Units: Reviewing What We Know and What Should Change." *International Journal of Offender Therapy and Comparative Criminology* 52:622–40.

Associated Press. 2007. "Limon Prison under Lockdown after Assault." *Vail (CO) Daily*, September 13. http://www.vaildaily.com/news/limon-prison-under-lockdown-after-assault/.

Association of State Correctional Administrators and the Arthur Liman Public Interest Program. 2016. *Aiming to Reduce Time-in-Cell: Reports from Correctional Systems on the Numbers of Prisoners in Restricted Housing and on the Potential of Policy Changes to Bring About Reforms*. New Haven, CT: Yale Law School.

Barnes, Harry Elmer. 1921. "The Historical Origin of the Prison System in America." *Journal of Criminal Law and Criminology* 12:35–60.

Barte, Henri. 1989. "L'isolement carceral." *Perspectives Psychiatriques* 28:252–55.

Bauer, Michael, Stefan Priebe, Bettina Haring, and Kerstin Adamczak. 1993. "Long-Term Mental Sequelae of Political Imprisonment in East Germany." *Journal of Nervous and Mental Disease* 181:257–62.

Benjamin, Thomas, and Kenneth Lux. 1975. "Constitutional and Psychological Implications of the Use of Solitary Confinement: Experience at the Maine State Prison." *Clearinghouse Review* 9:83–90.

Berger, Ronald, Paul Chaplin, and Robert Trestman. 2013. "Commentary: Toward an Improved Understanding of Administrative Segregation." *Journal of the American Academy of Psychiatry and Law* 41:61–64.

Boulet, Jack, and Marvin Boss. 1991. "Reliability and Validity of the Brief Symptom Index." *Psychological Assessment* 3:433–37.

Brodsky, Stanley, and Forrest Scogin. 1988. "Inmates in Protective Custody: First Data on Emotional Effects." *Forensic Reports* 1:267–80.

Burney, Christopher. 1961. *Solitary Confinement*. New York: Macmillan.

Cacioppo, Stephanie, and John Cacioppo. 2012. "Decoding the Invisible Forces of Social Connections." *Frontiers of Integrative Neuroscience* 6:1–7.

Card, Noel. 2012. *Applied Meta-Analysis for Social Science Research*. New York: Guilford.

Chow, Sui. 1987. "Meta-Analysis of Pragmatic and Theoretical Research: A Critique." *Journal of Psychology* 121(23):259–71.

Cloyes, Kristin, David Lovell, David Allen, and Lorna Rhodes. 2006. "Assessment of Psychosocial Impairment in a Supermaximum Security Unit Sample." *Criminal Justice and Behavior* 33:760–81.

Cohen, Stanley, and Laurie Taylor. 1972. *Psychological Survival: The Experience of Long-Term Imprisonment*. New York: Pantheon.

Coid, Jeremy, Ann Petruckevitch, Paul Bebbington, Rachel Jenkins, Traolach Brugha, Glyn Lewis, Michael Farrell, and Nicola Singelton. 2003. "Psychiatric Morbidity in Prisoners and Solitary Cellular Confinement. I: Disciplinary Segregation." *Journal of Forensic Psychiatry and Psychology* 14:298–319.

Commission on Crime Prevention and Criminal Justice. 2015. *United Nations Standard Minimum Rules for the Treatment of Prisoners*. New York: United Nations Economic and Social Council.

Cormier, Bruno, and Paul Williams. 1966. "Excessive Deprivation of Liberty." *Canadian Psychiatric Association Journal* 11(6):470–84.

Correctional News. 2012. "New Colorado Prison Closed Down." November 14. http://correctionalnews.com/2012/11/14/new-colorado-prison-closed-down/.

Coyne, James, Brett Thombs, and Mariet Hagedorn. 2010. "It Ain't Necessarily So: A Review and Critique of Recent Meta-Analyses of Behavioral Medicine Interventions in *Health Psychology*." *Health Psychology* 29:107–16.

Cressey, Donald. 1940. *The Prison Community*. Boston: Christopher Publishing.

Crewe, Ben. 2009. *The Prisoner Society: Power, Adaptation, and Social Life in an English Prison*. Oxford: Oxford University Press.

Dickens, Charles. 1842. *American Notes*. New York: Harper & Brothers.

Ecclestone, C., Paul Gendreau, and Clifford Knox. 1974. "Solitary Confinement of Prisoners: An Assessment of Its Effects on Inmates' Personal Constructs and Adrenal Cortical Activity." *Canadian Journal of Behavioral Science* 6:178–91.

Feeley, Malcolm, and Jonathan Simon. 1992. "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications." *Criminology* 30:449–74.

Fiorillo, Damiano, and Fabio Sabatini. 2011. "Quality and Quantity: The Role of Social Interactions in Self-Reported Individual Health." *Social Science and Medicine* 73:1644–52.

Foster, Don, Dennis Davis, and Diane Sandler. 1987. *Detention and Torture in South Africa*. New York: St. Martin's.

Franke, Herman. 1992. "The Rise and Decline of Solitary Confinement." *British Journal of Criminology* 32:125–43.

Gendreau, Paul, and James Bonta. 1984. "Solitary Confinement Is Not Cruel and Unusual: People Sometimes Are!" *Canadian Journal of Criminology* 26: 467–78.

Gendreau, Paul, and Ryan Labrecque. 2016. “The Effects of Administrative Segregation: A Lesson in Knowledge Cumulation.” In *The Oxford Handbook of Prisons and Imprisonment*, edited by John Wooldredge and Paula Smith. New York: Oxford University Press.

Gendreau, Paul, and Yvette Theriault. 2011. “Bibliotherapy for Cynics Revisited: Commentary on *One Year Longitudinal Study of the Psychological Effects of Administrative Segregation*.” *Corrections and Mental Health: An Update of the National Institute of Corrections*. https://community.nicic.gov/blogs/mentalhealth/archive/2011/06/21/bibliotherapy-for-cynics-revisited-commentary-on-one-year-longitudinal-study-of-the-psychological-effects-of-administrative-segregation.aspx.

Gendreau, Paul, et al. 1972. “Changes in EEG Alpha Frequency and Evoked Response Latency during Solitary Confinement.” *Journal of Abnormal Psychology* 79:54–59.

Government Accountability Office. 2013. *Improvements Needed in Bureau of Prisons’ Monitoring and Evaluation of Impact of Segregated Housing*. Washington, DC: US Government Accountability Office.

Grassian, Stuart. 1983. “Psychopathological Effects of Solitary Confinement.” *American Journal of Psychiatry* 140(11):1450–54.

———. 2006. “Psychiatric Effects of Solitary Confinement.” *Washington University Journal of Law and Policy* 22:325–83.

Grassian, Stuart, and Terry Kupers. 2011. “The Colorado Study versus the Reality of Supermax Confinement.” *Correctional Mental Health Report* 13(1):1, 9–11.

Hafner, S., R. Emeny, M. Lacruz, J. Baumert, C. Herder, et al. 2011. “Association between Social Isolation and Inflammatory Markers in Depressed and Non-depressed Individuals: Results from the MONICA/KORA Study.” *Brain Behavior and Immunity* 25:1701–7.

Haney, Craig. 1993. “Infamous Punishment: The Psychological Consequences of Isolation.” *National Prison Project Journal* 8(2):3–7, 21.

———. 2003. “Mental Health Issues in Long-Term Solitary and ‘Supermax’ Confinement.” *Crime and Delinquency* 49:124–56.

Haney, Craig, and Mona Lynch. 1997. “Regulating Prisons of the Future: The Psychological Consequences of Solitary and Supermax Confinement.” *New York University Review of Law and Social Change* 23:477–570.

Hibbert, Christopher. 1963. *The Roots of Evil: A Social History of Crime and Punishment*. Boston: Little, Brown.

Hilliard, Thomas. 1976. “The Black Psychologist in Action: A Psychological Evaluation of the Adjustment Center Environment at San Quentin Prison.” *Journal of Black Psychology* 2:75–82.

Hinkle, Lawrence, and Harold Wolff. 1956. “Communist Interrogation and Indoctrination of ‘Enemies of the State.’ ” *Archives of Neurology and Psychiatry* 76:115–74.

Holt-Lunstad, Julianne, Theodore Robles, and David Sbarra. 2017. “Advancing Social Connectedness as a Public Health Priority in the United States.” *American Psychologist* 72:517–30.

Hunt, Stephen, Mack Orsborn, Harvey Checkoway, Mary Biggs, Miles McFall, and Tim Takaro. 2008. "Later Life Disability Status Following Incarceration as a Prisoner of War." *Military Medicine* 173:613–18.

Immarigeon, Russ. 2011. "One Year Longitudinal Study of the Psychological Effects of Administrative Segregation: Introduction." *Corrections and Mental Health: An Update of the National Institute of Corrections*. https://community.nicic.gov/blogs/mentalhealth/archive/2011/06/21/one-year-longitudinal-study-of-the-psychological-effects-of-administrative-segregation-introduction.aspx.

Irwin, John. 1980. *Prisons in Turmoil*. Boston: Little, Brown.

———. 1987. "Reflections on Ethnography." *Journal of Contemporary Ethnography* 16:41–48.

Jackson, Michael. 1983. *Prisoners of Isolation: Solitary Confinement in Canada*. Toronto: University of Toronto Press.

Jacobs, James. 1977. *Stateville*. Chicago: University of Chicago Press.

Jones, David. 1976. *The Health Risks of Imprisonment*. Lexington, MA: Lexington Books.

Kaba, Fatos, Andrea Lewis, Sarah Glowa-Kollisch, James Hadler, David Lee, Howard Alper, Daniel Selling, Ross MacDonald, Angela Solimo, Amanda Parsons, and Homer Venters. 2014. "Solitary Confinement and Risk of Self-Harm among Jail Inmates." *American Journal of Public Health* 104:442–47.

Kellett, Stephen, Nigel Beail, David Newman, and Pat Frankish. 2003. "Utility of the Brief Symptom Inventory in the Assessment of Psychological Distress." *Journal of Applied Research in Intellectual Disabilities* 16:127–34.

Koch, Ida. 1986. "Mental and Social Sequelae of Isolation: The Evidence of Deprivation Experiments and of Pretrial Detention in Denmark." In *The Expansion of European Prison Systems*, edited by Bill Rolston and Mike Tomlinson. Working Papers in European Criminology, no. 7. Stockholm: European Group for the Study of Deviance and Social Control.

Kupers, Terry. 2017. *Solitary: The Inside Story of Supermax Isolation and How We Can Abolish It*. Oakland: University of California Press.

Lieberman, Matthew. 2013. *Social: Why Our Brains Are Wired to Connect*. New York: Random House.

Liebling, Alison. 1999. "Doing Research in Prison: Breaking the Silence?" *Theoretical Criminology* 3:147–73.

Lovell, David, and Hans Toch. 2011. "Some Observations about the Colorado Segregation Study." *Correctional Mental Health Report* 13(1):3–4, 14.

Makin, Kirk. 2013. "Canadian Prisons 'Out of Step' on Solitary Confinement." *Globe and Mail*, March 21. https://www.theglobeandmail.com/news/politics/canadian-prisons-out-of-step-on-solitary-confinement/article10103358/.

Masur, Louis. 1989. *Rites of Execution: Capital Punishment and the Transformation of American Culture, 1776–1865*. New York: Oxford University Press.

Mendez, Juan. 2011. *Interim Report of the Special Rapporteur of the Human Rights Council on Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*. New York: United Nations.

Metzner, Jeffrey, and Jamie Fellner. 2010. "Solitary Confinement and Mental Illness in U.S. Prisons: A Challenge for Medical Ethics." *Journal of the Acad-*

emy of Psychiatry and Law 38:104–8. http://www.hrw.org/sites/default/files/related_material/Solitary%20Confinement%20and%20Mental%20Illness%20in%20US%20Prisons.pdf.

Metzner, Jeffrey, and Maureen O'Keefe. 2011. "The Psychological Effects of Administrative Segregation: The Colorado Study." *Correctional Mental Health Report* 13(1):1–2, 12–14.

Miller, Holly. 1994. "Reexamining Psychological Distress in the Current Conditions of Segregation." *Journal of Correctional Health Care* 1:39–53.

Miller, Holly, and Glenn Young. 1997. "Prison Segregation: Administrative Remedy or Mental Health Problem?" *Criminal Behaviour and Mental Health* 7:85–94.

Mitchell, Kirk. 2010. "Limon Prison Incentives Keep Inmates in Check." *Denver Post*, November 24. http://www.deseretnews.com/article/700085422/Limon-prison-incentives-keep-inmates-in-check.html.

Morgan, Robert, Paul Gendreau, Paula Smith, Andrew Gray, Ryan Labrecque, Nina MacLean, Stephanie Van Horn, Angelea Bolanos, Ashley Batastini, and Jeremy Mills. 2016. "Quantitative Synthesis of the Effects of Administrative Segregation on Inmates' Well-Being." *Psychology, Public Policy, and Law* 22:439–61.

National Commission on Correctional Health Care. 2016. "Position Statement: Solitary Confinement (Isolation)." *Journal of Correctional Health Care* 22:257–63. http://www.ncchc.org/solitary-confinement.

National Research Council. 2014. *The Growth of Incarceration in the United States: Exploring Causes and Consequences*. Washington, DC: National Academies Press.

Nowak, Manfred. 2006. "What Practices Constitute Torture?" *Human Rights Quarterly* 28:809–41.

Nussbaum, Martha. 1995. "Objectification." *Philosophy and Public Affairs* 24:249–91.

O'Donnell, Ian. 2014. *Prisoners, Solitude, and Time*. New York: Oxford University Press.

O'Keefe, Maureen. 2008. "Administrative Segregation from Within: A Corrections Perspective." *Prison Journal* 88:123–43.

———. 2017. "Reflections on Colorado's Administrative Segregation Study." *NIJ Journal* 278:1–8. https://nij.gov/journals/278/Pages/reflections-on-colorado-administrative-segregation-study.aspx.

O'Keefe, Maureen, Kelli Klebe, Jeffrey Metzner, Joel Dvoskin, Jamie Fellner, and Alysha Stucker. 2013. "A Longitudinal Study of Administrative Segregation." *Journal of the American Academy of Psychiatry and Law* 41:49–60.

O'Keefe, Maureen, Kelli Klebe, Alysha Stucker, Kristin Sturm, and William Leggett. 2010. *One Year Longitudinal Study of the Psychological Effects of Administrative Segregation*. Final Report to the National Institute of Justice. Washington, DC: US Department of Justice, National Institute of Justice.

Raemisch, Rick. 2017. "Putting an End to Long-Term Solitary." *New York Times*, October 13. https://www.nytimes.com/2017/10/12/opinion/solitary-confinement-colorado-prison.html.

Reiter, Keramet. 2016. *23/7: Pelican Bay Prison and the Rise of Long-Term Solitary Confinement*. New Haven, CT: Yale University Press.

Reyes, Hernan. 2007. "The Worst Scars Are in the Mind: Psychological Torture." *International Review of the Red Cross* 89:591–616. https://www.icrc.org/en/international-review/article/worst-scars-are-mind-psychological-torture.

Rhodes, Lorna. 2004. *Total Confinement: Madness and Reason in the Maximum Security Prison*. Berkeley: University of California Press.

Rhodes, Lorna, and David Lovell. 2011. "Is Adaptation the Right Question? Addressing the Larger Context of Administrative Segregation: Commentary on One Year Longitudinal Study of the Psychological Effects of Administrative Segregation." *Corrections and Mental Health: An Update of the National Institute of Corrections* (June 21). http://community.nicic.gov/cfs-file.ashx/__key/CommunityServer.Components.PostAttachments/00.00.05.95.19/Supermax-_2D00_-T-_2D00_-Rhodes-and-Lovell.pdf.

Rundle, Frank. 1973. "The Roots of Violence at Soledad." In *The Politics of Punishment: A Critical Analysis of Prisons in America*, edited by Erik Wright. New York: Harper.

Scott, George, and Paul Gendreau. 1969. "Psychiatric Implications of Sensory Deprivation in a Maximum Security Prison." *Canadian Psychiatric Association Journal* 14:337–41.

Shalev, Sharon. 2009. *Supermax: Controlling Risk through Solitary Confinement*. Portland, OR: Willan.

Shalev, Sharon, and Monica Lloyd. 2011. "*If* This Be Method, Yet There Is Madness in It: Commentary on One Year Longitudinal Study of the Psychological Effects of Administrative Segregation." *Corrections and Mental Health: An Update of the National Institute of Corrections* (June 21). http://community.nicic.gov/cfs-file.ashx/__key/CommunityServer.Components.PostAttachments/00.00.05.95.21/Supermax-_2D00_-T-_2D00_-Shalev-and-Lloyd.pdf.

Slater, Robert. 1986. "Psychiatric Intervention in an Atmosphere of Terror." *American Journal of Forensic Psychology* 7(1):5–12.

Smith, Peter Scharff. 2006. "The Effects of Solitary Confinement on Prison Inmates: A Brief History and Review of the Literature." In *Crime and Justice: A Review of Research*, vol. 34, edited by Michael Tonry. Chicago: University of Chicago Press.

———. 2011. "The Effects of Solitary Confinement: Commentary on One Year Longitudinal Study of the Psychological Effects of Administrative Segregation." *Corrections and Mental Health: An Update of the National Institute of Corrections* (June 21). http://community.nicic.gov/cfs-file.ashx/__key/CommunityServer.Components.PostAttachments/00.00.05.95.22/Supermax-_2D00_-T-_2D00_-Smith.pdf.

Strube, Michael, William Gardner, and Donald Hartmann. 1985. "Limitations, Liabilities, and Obstacles in Reviews of the Literature: The Current Status of Meta-Analysis." *Clinical Psychology Review* 5:63–78.

Suedfeld, Peter, Carmenza Ramirez, John Deaton, and Gloria Baker-Brown. 1982. "Reactions and Attributes of Prisoners in Solitary Confinement." *Criminal Justice and Behavior* 9:303–40.

Sykes, Gresham. 1958. *Society of Captives*. Princeton, NJ: Princeton University Press.

Toch, Hans. 1975. *Men in Crisis: Human Breakdowns in Prison*. New York: Aldine.
———. 1977. *Living in Prison: The Ecology of Survival*. New York: Free Press.
———. 2003. "The Contemporary Relevance of Early Experiments with Supermax Reform." *Prison Journal* 83:221–28.
Turner, Scott, Laura Cardinal, and Richard Burton. 2017. "Research Design for Mixed Methods: A Triangulation-Based Framework and Roadmap." *Organizational Research Methods* 20:243–67.
Volkart, Reto, Adolf Dittrich, Thomas Rothenfluh, and Paul Werner. 1983. "Eine Kontrollierte Untersuchung uber Psychopathologische Effekte der Einzelhaft" [A controlled investigation on psychopathological effects of solitary confinement]. *Psychologie—Schweizerische Zeitschrift fur Psychologie und ihre Anwendungen* 42:25–46.
Volkart, Reto, Thomas Rothenfluh, W. Kobelt, Adolf Dittrich, and K. Ernst. 1983. "Einzelhaft als Risikofaktor fur Psychiatrische Hospitalisierung" [Solitary confinement as a risk for psychiatric hospitalization]. *Psychiatria Clinica* 16:365–77.
Waligora, Bogusław. 1974. "Funkcjonowanie Czlowieka W Warunkach Izolacji Wieziennej" [How men function in conditions of penitentiary isolation]. *Seria Psychologia I Pedagogika* 34:1–123.
Walters, Richard, John Callagan, and Albert Newman. 1963. "Effect of Solitary Confinement on Prisoners." *American Journal of Psychiatry* 119:771–73.
Zinger, Ivan. 1998. "The Psychological Effects of 60 Days in Administrative Segregation." PhD diss., Carlton University, Department of Psychology.
Zinger, Ivan, Cherami Wichmann, and D. Andrews. 2001. "The Psychological Effects of 60 Days in Administrative Segregation." *Canadian Journal of Criminology* 43:47–83.