

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 methodologically 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 con-

sequences 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 sec-

ond 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

410 Craig Haney

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 carcéral." *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 Singleton. 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/mental-health/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-*

414 Craig Haney

- 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.

416 Craig Haney

- 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.