

Revisiting the Stanford Prison Experiment: A Case Study in Organized Skepticism

Teresa C. Kulig, Travis C. Pratt & Francis T. Cullen

To cite this article: Teresa C. Kulig, Travis C. Pratt & Francis T. Cullen (2017) Revisiting the Stanford Prison Experiment: A Case Study in Organized Skepticism, Journal of Criminal Justice Education, 28:1, 74-111, DOI: [10.1080/10511253.2016.1165855](https://doi.org/10.1080/10511253.2016.1165855)

To link to this article: <https://doi.org/10.1080/10511253.2016.1165855>



Published online: 12 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 1278



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Revisiting the Stanford Prison Experiment: A Case Study in Organized Skepticism

Teresa C. Kulig, Travis C. Pratt and Francis T. Cullen

Growing evidence exists that the findings of individual studies—including classic experiments—often fail to replicate. Such published results, however, are considered by scholars, and taught to students, as established scientific truth. In this context, citations to Zimbardo and colleagues' classic Stanford Prison Experiment (SPE) in criminology/criminal justice journals (1975–2014) were content analyzed to assess whether the study's conclusions have been embraced or treated with skepticism. The data revealed that scholars were widely accepting of the SPE and, even when voicing concerns, supportive of its message. These results suggest the need to give replications higher priority and for scholars to adhere more closely to the scientific norm of organized skepticism. In the classroom, the continued, uncritical acceptance of the SPE—now more than 40 years old—can serve as an opportunity to teach students about the production and assessment of knowledge within criminology.

In recent years, increasing concern has emerged regarding the validity of findings reported in science, including social science, publications (Freedman, 2010; Open Science Collaboration, 2015). The preference of editors to publish statistically significant results rather than null findings is an ongoing worry (Egger & Smith, 1998; Lipsey & Wilson, 2001), and yet concern is also being expressed that many statistically significant results are, in fact, false positives (Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011). A number of studies reporting seemingly robust effects have later been found not to replicate, including many classic, often-cited experiments (Lehrer, 2010). The very belief that a single study can create established truth is therefore being questioned by those advocating for the meta-analysis of full bodies of empirical work (Schmidt & Oh, 2015). Furthermore, the sociology of knowledge has long cautioned that science is a human enterprise and, as such, is vulnerable to ideological influence (see, e.g. Berger & Luckmann, 1966; Gouldner, 1970). Prominent examples include the role of social Darwinist beliefs in the portrayal

of offenders as racially inferior (Gould, 1981; cf. Lewis et al., 2011) and the widespread embrace of Martinson's "nothing works" doctrine in corrections despite evidence to the contrary (Cullen & Gendreau, 2000, 2001; Gottfredson, 1979).

Taken together, these developments raise the issue of the quality of the research information that criminologists are transmitting in their classrooms. The danger exists that accepted criminological truths may rest upon a shaky empirical foundation. This observation is not intended to suggest that all knowledge is a mere social construction or that every piece of evidence should now be portrayed as a matter of conjecture. For example, a number of meta-analyses and rigorous narrative reviews have been conducted that compile substantial evidence on a number of issues in criminology and criminal justice (Andrews et al., 1990; Braga, Papachristos, & Hureau, 2014; Piquero, Jennings, & Farrington, 2010; Pratt & Cullen, 2000; Pratt, Turanovic, Fox, & Wright, 2014). Still, in light of questions raised about the difficulty of producing solid scientific findings, criminological hubris is an academic sin that should be avoided.

Instead, a more reasonable alternative is the appropriate use of what Merton (1973) called "organized skepticism." A core scientific norm, organized skepticism involves a "methodological and institutional mandate" in which scholars in a discipline collectively engage in the "detached scrutiny of beliefs in terms of empirical and logical criteria" (1973, p. 277). Most often, scientists take this approach when evaluating claims made in other institutional realms, like that of public policy. Organized skepticism thus often generates "public hostility," because it "involves a latent questioning of certain bases of established routine, authority, vested procedures, and the realm of the 'sacred' generally" (p. 264). Those who rely on ideology or faith—for example, in claiming that global warming does not exist or that "scared straight" programs are effective—will attack the organized skepticism of scholars as representing "liberal bias" and "junk science" (Maahs & Pratt, *in press*).

The key point, however, is that the limits of criminologists' own science—especially when rooted in comforting ideological preferences—may result in our embrace of "facts" that may be based more on a matter of faith than of empirical reality. Organized skepticism thus is a scientific norm and practice that should not be reserved only for those outside academia. It also should be applied to us. In particular, we should be wary of celebrated studies that not only have not been widely replicated but are also taken as revealing a fundamental criminological truth.

In psychology, this reconsideration of classic studies is already under way (Lehrer, 2010). Griggs (2014, 2015a, 2015b; see also Griggs & Whitehead, 2014) has recently documented that key studies and interpretations of events have received distorted coverage and/or have largely ignored relevant criticisms. Importantly, he has argued that these problems mark Philip Zimbardo's famous Stanford Prison Experiment (SPE) (Griggs, 2014; Griggs & Whitehead, 2014; see also Haney, Banks, & Zimbardo, 1973; Zimbardo, 1972, 2007; Zimbardo, Banks,

Haney, & Jaffe, 1973). This claim is significant because the Stanford Prison Experiment (herein referred to as "SPE") is a classic study in corrections. The SPE is purported to show that prisons are inherently coercive and inhumane, which lends credence to the deprivation model of imprisonment that links argot roles and an oppositional inmate culture to the pains inherent in institutional life (Sykes, 1958). Phrased more broadly, the SPE has been used, in the past and in the present, as a key piece of evidence in arguments against the use of total institutions specifically and, writ large, against the policy of mass imprisonment (Cullen & Jonson, 2017).

In this context, we begin by reviewing the SPE and by conveying Griggs's (2014; Griggs & Whitehead, 2014) summary of the critique of the SPE that has emerged in the psychological literature. This work is not meant to discount fully the findings of Zimbardo's experiment but only to show that the SPE should be presented as providing a compelling but provisional hypothesis and not incontrovertible truth. We follow this discussion with a case study in organized skepticism. We examine citations to Zimbardo's SPE across 26 criminology and criminal justice journals from 1975 to 2014. The goal is to assess the extent to which citations to the SPE provide a nuanced interpretation of the study's findings or cite the SPE either in a pro forma way or as providing definitive proof that prisons are inherently coercive and inhumane. The broader implications of this analysis are twofold. First, the field might profit from other, systematic efforts to revisit convenient criminological "truths" that have escaped organized skepticism in the past. Second, in our pedagogy, efforts should be made to place classic studies in an appropriate context and to teach students to appreciate the challenges in capturing empirical reality.

Revisiting the SPE

In 1971, Philip Zimbardo set up an innovative experiment in the basement of Jordan Hall (the Psychology building) on the Stanford University campus. He recruited 24 male undergraduate students to pretend to be either guards or inmates in a simulated prison to see what would happen when "normal" people were placed in a situational context characterized by extreme power differentials and proscribed role expectations. The outcome has been well documented and it goes a little something like this: the students who were assigned as guards quickly developed a thirst for abusing their fellow student-inmates and became nasty and coercive, and the students who were assigned as inmates became submissive and damaged, traumatized by their experience of victimization at the hands of their keepers, and it got so bad that it had to be called off after just six days (Haney et al., 1973; Zimbardo et al., 1973). As a result, the message of the SPE was that the prison is an inherently destructive institution, one that is capable of turning otherwise good people into evil-doers, and of turning otherwise well-adjusted people into cowering balls of anxiety.

Of course, Zimbardo and his colleagues were not the only ones to suggest that prisons were harmful institutions—others made similar claims both before and after the SPE (see, e.g. Carroll, 1974; Fleisher, 1989; Goffman, 1961; Hazelrigg, 1968; Rothman, 1980; Sykes, 1958; Toch, 1977). In fact, the view that prisons are harmful, perhaps intractably so, is common among scholars (see Dilulio, 1987). Even if the SPE had never been conducted, this view likely would have prevailed. Still, the SPE has a special place in the prison effects literature because it is the only study to offer experimental evidence using psychologically normal college students—not offenders convicted of an array of crimes who have a history of misconduct—to show the inevitability of the emergence of a coercive and conflictual prison environment. The study’s dramatic events—including “inmates” screaming and begging to be released—were recorded and documentaries made (e.g. by the BBC and one produced by Zimbardo, *Quiet Rage: The Documentary—The Stanford Prison Experiment*). Indeed, it was arguably the vivid and provocative nature of Zimbardo’s findings, combined with how he popularized them, that established the SPE as a classic study in psychology—right up there with Stanley Milgram’s fake electroshock experiment (Gregoire, 2013).

The SPE’s influence can be demonstrated in several ways. Its characterization as a “classic” study is, for example, common in the literature (Aviram, 2009; Bordt, 1999; Brown, 1982; Cheek & Miller, 1983; Maier-Katkin, Mears, & Bernard, 2009; Philliber, 1987), and the citation count for the SPE publications on Google Scholar is nearly 2,000—a total that few criminological studies will ever reach (Cohn & Farrington, 2012). Discussions of the SPE are now even a core part of university curricula for institutional review boards (Fitch, 2005; Tolich, 2014). It is also a staple in psychology and criminology/criminal justice classrooms in that the SPE has long made appearances in undergraduate textbooks (for psychology, see Griggs & Whitehead, 2014; for examples in corrections books, see Cullen & Jonson, 2017; Johnson, 1996; Krisberg, Marchionna, & Hartney, 2015; Latessa & Holsinger, 2016; Schwartz, Clear, & Travis, 1980; Stohr & Walsh, 2016; Welch, 1996; Whitehead, Dodson, & Edwards, 2013). Finally, although not a terribly objective metric of impact, Zimbardo is famous—handsome actor Billy Crudup plays him in the movie *The Stanford Prison Experiment* and Zimbardo himself got to walk the red carpet at the movie’s premiere. It is therefore pretty safe to say that the SPE is kind of a big deal.

But despite its early and continued acclaim, the SPE is marked by key problems. First, the results of the SPE have failed to replicate in other contexts. Specifically, in an effort to redo the SPE (or to come as close to it as they could) by the British Broadcasting Company (BBC), 15 men were randomly assigned as either guards or prisoners in a controlled environment and were observed over the next eight days. But unlike what happened in the basement at Stanford, the guards in the British study failed to act either aggressively or coercively against the prisoners (Reicher & Haslam, 2006). Instead, they failed to even identify much at all—either individually or collectively—with their role as a guard, and as the study proceeded the rowdy prisoners eventually caused

the prisoner–guard system to collapse. It is therefore questionable whether the social and psychological processes that the SPE seemed to be tapping into are general or were instead unique to the Stanford study.

And second, serious methodological criticisms have been raised regarding the behavioral priming that Zimbardo engaged in with the student–guards both before and during the experiment. In particular, the students who were assigned as guards stepped into their roles with a clear set of instructions from Zimbardo himself (Banyard, 2007). They were given an orientation by Zimbardo as to how they were expected to behave; that is, to be a “tough guard” (Zimbardo, 2007). And those guards who failed to live up to that label were taken aside by Zimbardo’s graduate student “warden” and were chastised for being too soft on the prisoners (Griggs, 2014). It was these “demand characteristics” of the experiment that has led some to question the study’s overall conclusion that situational characteristics—not personal ones—drive human behavior in a prison setting (see, e.g. Banuazizi & Movahedi, 1975). After all, would the student–guards really have been so mean if Zimbardo had not told them that they were supposed to be? Maybe, maybe not. But either way, the strong demand characteristics attached to the experiment likely preclude any strong conclusions about what prisons do to people. As Johnson (1996, p. 200) revealed two decades ago:

the students...were implicitly encouraged to play out stereotypical conceptions of guard and inmate roles. The student/guards....were issued uniforms reminiscent of a southern state trooper, complete with silver “Cool Hand Luke” reflective glasses and long nightsticks, and then given blanket authorization to run the prison as they saw fit....By contrast, the student/inmates wore sacklike dresses....they had bracelets for their ankles and stocking caps for their heads....Can anyone doubt that the student/guards felt it was appropriate to play the tough-guy role with the emasculated student/inmates over whom they held complete power?

Current Focus

It is therefore clear that some organized skepticism concerning the SPE is warranted here. So how should we go about it? On the one hand, we do not want to give the impression that Zimbardo’s work was anything less than groundbreaking. It certainly was. It did a lot to humanize a correctional system that at the time was in dire need of it (Haney & Zimbardo, 1998). But on the other hand, it would be useful to see how scholars are thinking and writing about the SPE. Are its findings accepted blindly? Or is there recognition that the SPE cannot be interpreted as simplistically as Zimbardo et al. would suggest and that it is being viewed as a study—like all studies—with limitations and whose results should be treated as tentative and as warranting replication? Put simply, if the field is writing about the SPE as if its findings are beyond question,

it is likely that the same message—one of unquestioned deference to a “classic study”—is being communicated in our classrooms. And if that is the case, we are probably teaching it wrong.

Methods

Sample of Studies

The current study sought to include all instances in which works written by Zimbardo, Haney, and colleagues on the SPE were cited in criminal justice and criminology journals. Toward this end, a Google Scholar reference search for SPE citations was reviewed to find articles published in criminal justice and criminology journals. Based on the Google Scholar search, 26 different criminal justice and criminology journals were selected that included one or more of the SPE citations. Articles were included in the sample if they were published in these journals between 1975 and 2014 and if they contained an accurate citation to the SPE. An example of an “inaccurate” cite was if the article included a reference to the SPE citation but the citation was not in the text of the article. In most cases, the results of the SPE were either discussed or the study was cited in support of a substantive point. Only publications that were printed in English were included. Using these criteria, 59 articles were noted as appropriate for further analysis. A list of the journals that were included are presented in Appendix A.

We recognize that these 59 studies represent a small portion of the broader literature that has cited the SPE studies. We elected to restrict our analyses to this sample of studies for two primary reasons. First, we are interested in how scholars in criminology and criminal justice think about and discuss the SPE in their own work. This is not to say that the opinions of scholars in psychology, sociology, social work, or ethics who have cited the SPE are either unimportant or otherwise irrelevant. We simply concluded that uncovering how scholars in our own discipline are treating the SPE would be of most substantive use to the field and to the readership of *JCJE*. And second, unlike other sources of literature—books, book chapters, technical reports to agencies, conference presentations—the peer review process at journals is most likely to place pressure on authors to be at their critical best; to provide as much depth and nuance as possible so that if scholars are, in fact, being critical of the SPE, articles in peer-reviewed journals may be most likely to contain such criticism.

Coding of Studies

The purpose of the study was to determine the extent to which scholars making reference to the SPE in criminal justice and criminology publications offered a critical analysis or were largely accepting of this classic study’s

findings. In coding studies, we followed the general approach used in a similar study of citations to the SPE in psychology textbooks (Griggs, 2014; Griggs & Whitehead, 2014). Thus, each study was coded according to key dimensions regarding how the SPE was cited and discussed. Three broad categories were employed: (1) accepting; (2) critical but accepting; and (3) not accepting and critical. These are explained in more detail below.

Accepting. In this category, the authors of a study expressed no reservations about the results of the SPE. Rather than simply code studies as “accepting,” an effort was made to determine how the SPE was cited in each study. All content relevant to a SPE-related citation was extracted from the article, recorded, and then assessed. This information is presented in Appendix B. Because the discussion of the SPE was often of limited length, the coding of articles involved a measure of discretion. The materials from the articles are presented in Appendix B, which allows these decisions to be reviewed. Although some variation in coding is possible, the general pattern of results presented here is unlikely to change. Finally, to be as parsimonious as possible Appendix B contains only the most central cite and quoted material from each article. Where multiple cites in an article are made to the SPE, we list the page numbers on which these other cites can be found. For lengthy quoted material, we omitted less essential material, noting omitted text with the conventional use of ellipsis points.

In this regard, one distinction related to the core theme of the SPE. Thus, Haney et al. (1973) conceive of prisons as creating a structural environment that ineluctably leads to guard and inmate social roles that are coercive and conflictual. For example, they state: “The environment of arbitrary custody had great impact upon the affective states of both guards and prisoners as well as upon the interpersonal processes taking place between and within those role-groups” (1973, p. 80). In articles, however, authors differed in terms of whether they focused more on the structural component of prisons or on the resulting social roles. Accordingly, studies that accepted the SPE as showing that prisons are inhuman, coercive, or painful were coded as falling in one of three subcategories.

The first subcategory is that prisons are inhumane/coercive/painful due to the structure or environment of prisons themselves. An example comes from Clements et al. (2007): “The potentially adverse impact of prison environments has been studied for several decades and was famously symbolized in the ‘Stanford Prison Experiment’ (Haney, Banks, & Zimbardo, 1973)” (p. 922). This passage shows how the actual prison environment is what is causing harm while social roles are not explicitly discussed. A second subcategory is that prisons are harmful due to the social roles individuals are induced to play. Hussemann and Page (2011) fall under this second category by including the following passage: “The prison officer role promotes antipathy toward inmates, as depicted so clearly in the well-known ‘Stanford prison experiment’ (Haney, Banks, and Zimbardo, 1973)” (p. 273). A third subcategory is a study sees prisons as problematic due to a combination of the environment and social roles that are induced.

Sundt and Cullen (2002) exemplify this usage: "These dramatic results led to the conclusion that the behavior of the research participants was the result of the roles that they had been assigned and the 'inherent pathological characteristics of the prison situation itself' (Haney et al., 1973, p. 30)" (p. 372). This third subcategory was also used to capture authors who described the effects that resulted from the SPE (e.g. prisoner personality, depressed-helpless prisoner).

Two other kinds of acceptance were detected and thus coded. First, the SPE was at times referred to as part of a "string cite"—as part of two or more works cited in support of a specific point. For example, sometimes the study was cited with other works to justify some type of prison issue, the power of authority, or the influence of social-structural factors that affect personalities. An example of a string cite is illustrated by Blevins, Listwan, Cullen, and Jonson (2010):

As a result, prisoners forfeit their autonomy, a sense of total safety and security, personal identities, access to many material goods and services, privacy, heterosexual relationships, unrestricted interaction with family and friends, and many other general comforts of life (Sykes, 1958; see also Haney, Banks, & Zimbardo, 1973; Toch, 1977). Many of these 'pains' are particularly difficult to adjust to while incarcerated. (p. 151)

Second, one study was coded as accepting based on the use of the SPE to describe a methodological design. Brown (1982) discusses the SPE in terms of involving students in research: "Zimbardo's (1976) classic prison simulation study at Stanford certainly demonstrates the potential of simulation studies for absorbing students in the research process" (p. 14).

Critical But Accepting. We coded whether a given study was in any way critical of the SPE but still accepting of its findings. If a cite fell into this category, we then coded whether (1) the nature of that criticism was "ethical" (e.g. the harm done to human subjects), (2) the nature of the criticism was "artificial" (e.g. the circumstances of the experiment did not reflect real life), or (3) the author stated there were limitations but did not elaborate on any details. As with the analysis of the studies falling into the accepting category just discussed, the citations of the critical but accepting studies were coded into one of three subcategories with regard to how the SPE was discussed as showing the prisons are inhumane, coercive, or painful.

One study could not be coded in this way. Here, Farrington (1983) only mentioned the SPE in passing and that the SPE was not included in their analysis of experiments. However, the article did offer an indirect criticism of the SPE in the form of a citation by Banuazizi and Movahedi (1975), who argued that participants were not in an environment that was equivalent to a real prison and the results from the SPE are most likely based on demand characteristics. Since Farrington (1983) did not actively reference the findings of the SPE in a substantive way, but still offered criticism in the form of a citation, this article was coded under "Critical but Accepting" with its own category explaining

it discussed the SPE in no substantive detail. Additional studies that included a citation of authors who criticized the SPE (e.g. Banuazizi & Movahedi, 1975), even if not explicitly discussing the criticism, were coded as criticizing the study with no specific details.

Not Accepting and Critical. The final category, which would show true skepticism about the SPE, is if an article criticized the experiment and was not accepting of its findings.

Analytic Strategy

Given that we have coded both quantitative and qualitative data from each of the studies in the sample, we apply a mixed methods analytic strategy (Johnson, Onwuegbuzie, & Turner, 2007). This approach is intended to maximize the benefits and to minimize the potential limitations of both quantitative and qualitative methods (Creswell, 2013). Mixed methods models take a number of different forms (Small, 2011). These include collecting and analyzing quantitative and qualitative data simultaneously, as well as more “sequential” models where either a quantitative or qualitative approach is taken first, followed by the other (Fielding, 2012; Johnson & Onwuegbuzie, 2004). For our study, we address two interrelated questions, each of which is approached using a sequential method that begins with a quantitative component that will then be followed by a qualitative component. First, are citations to the SPE in the criminological literature largely accepting or critical of its findings? And second, in those studies that do convey concerns about the SPE, what is the nature of those criticisms? We conclude our results with an analysis of the degree to which the SPE was central to the discussion the authors were making or whether it was instead part of a larger body of literature that scholars were drawing upon to make their point.

Results

Accepting

In all, citations to the SPE were examined for 59 articles. The categorization of each article is presented in the appendices. As seen in Table 1, overall, authors were accepting of the SPE findings. Of the 59 articles, 47 or 79.7% did not state any reservations about the study or its findings.

The category of accepting the study’s findings was broken down into the three subcategories to better illustrate how authors presented the SPE. First, a majority of articles (55.9%) presented the results as inhumane, coercive, or painful. Under this category, 13 articles (22.0%) focused on the structure of prisons, 10 articles (16.9%) discussed the social roles, and 10 articles (16.9%) presented the findings as combination of the two. Second, in other articles,

Table 1 Summary of studies and documentation of the SPE ($N = 59$)

Acceptance of the SPE in publication	Number of publications	Percentage of total publications ^a (%)
<i>Accepting</i>	47	79.7
Prisons are inhumane/coercive/painful	33	55.9
Due to structure/environment	13	22.0
Due to social roles	10	16.9
Due to a combination of both	10	16.9
String cite with other works	13	22.0
SPE as a research method	1	1.7
<i>Critical but accepting</i>	10	16.9
Prisons are inhumane/coercive/painful	9	15.3
Due to structure/environment	2	3.4
Due to social roles	2	3.4
Due to a combination of both	5	8.5
No substantive discussion of the SPE	1	1.7
<i>Not accepting and critical</i>	2	3.4
Critique type of the SPE in publication	12	20.3
Ethical	2	3.4
Artificial	5	8.5
Not specific in details	5	8.5

Note. Publications were defined as “Accepting” if the authors did not state any reservations about the findings.

^aTotal percentages might not equal exactly due to rounding.

authors used a citation that referenced the SPE as part of a “string citation”—that is, an SPE work was cited along with other publications make some point about a prison issue, the power of authority, or the influence of social-structural factors that affect personalities. In sum, 22.0% of articles used a string citation with other works. Third, one study (1.7%) did not discuss the results of the SPE, but focused on the research method that was used as a way to include students in the research process.

Critical But Accepting

As seen in Table 1, there were 10 articles (16.9%) that offered some form of criticism of the SPE, but were ultimately still accepting of the findings. Nine of these articles (15.3%) presented the findings in a manner that illustrated that prisons were inhumane, coercive, or painful in nature. The subcategories of this

group demonstrate that 2 articles (3.4%) discussed this through the structure of the prison, 2 articles (3.4%) reported social roles, and 5 articles (8.5%) presented the findings as a combination of both structure and social roles. There was an exception to this grouping where one article did not discuss the SPE in any substantive detail (see Farrington, 1983). However, since the article offered some criticism and did not dispute the findings of the SPE, the article was still coded as "Critical but Accepting."

Not Accepting and Critical

There were 2 articles (3.4%) that did not accept the findings of the SPE and were also critical. The first article by Jones (1995) stated the following:

One method used is the "controlled experiment," in which subjects are assigned to "inmate" and "guard" roles in a mock prison setting, and attitudinal and behavioral changes are recorded throughout the experiment (See Haney et al., 1973). Although experiments do demonstrate that behavioral changes occur for subjects playing the roles of inmate and guard, these prison simulations are unsatisfactory for understanding the effects of a prison sentence because of their artificiality. The problem of artificiality arises for a number of reasons: the experimental subjects are aware that they are participating in a scientific experiment; they know that their involvement in the "prison" is only for a brief period of time; and they are certain that the researchers would never place them in any physical danger. For these reasons, experiments can provide only the weakest prison simulations. (p. 107)

This passage illustrates that the author does not believe prison simulations, such as the SPE, are sufficient to establish significant effects because they are artificial.

The second article by Suedfeld (1978) discussed the effects of prisons on prisoners with the SPE citation and another article:

At some time, sentimentalists perpetuate and disseminate myths about the current system. For example, it has frequently been stated as a self-evident fact that prisons so degrade their residents as to destroy their feelings of personal pride (Cooper, 1975; Zimbardo, 1973). In fact, there is at least some indication that this is not the case. While findings are mixed, several British, Canadian and U.S. studies have actually found the self-esteem scores of prisoners on well-validated tests are comfortably within the normal range (Bennett et al, 1971; Gendreau, Gibson, Surrige & Hug, 1973; Zimmermann et al, 1974). (p. 50)

Although the author does not specifically address criticisms of the SPE, it does offer evidence contrary to findings by citing researchers that found prisons do not harm prisoner's feelings or self-esteem. Therefore, this article was also coded as being critical of the SPE without specific details.

Type of Critiques

In addition to whether articles were accepting or critical of the SPE, details about the type of criticism were documented. As Table 1 reports, there were three main types of critiques that were recorded. Of the 12 publications (20.3%) that offered a critique of the SPE or its findings, 2 articles (3.4%) mentioned ethical concerns, 5 articles (8.5%) discussed issues of artificiality, and 5 articles (8.5%) offered a criticism (e.g. a citation, suggesting the study had limitations) but did not elaborate in any substantive details. Coding for critiques are included at the end of each article passage in Appendix C, with the exception of the two articles that were not accepting and critical.

The Context of SPE Citations

We consider it important to also explore the extent to which authors are using the SPE as either the sole or primary piece of work to make their points (whatever those points may be). So to that end, we also analyzed the context of SPE citations in terms of whether the SPE was cited alone or whether it was included with other works when discussing its results. Out of the 59 total articles, 13 articles cited an SPE study in conjunction with other research and all of these articles were coded as “accepting” the SPE’s findings (see Appendix B). The remaining articles ($n = 46$) cited the SPE in support of a specific statement related to the SPE or its findings. Scholars that included other references with an SPE citation generally used these studies as additional support for their arguments regarding the potentially inhumane nature of prison environments, the harmful effects of these environments on prisoners, and/or the importance of social roles in institutions. Our results therefore reveal a rather clear broad trend: When the SPE is cited by criminal justice and criminology scholars in their peer-reviewed work, it is the SPE—even if in conjunction with other works—that serves as the primary intellectual source as authors attempt to make their case regarding the nature and consequences of prisons.

Summary of Findings

Based on our analysis of the 59 articles, a high majority (nearly 4 in 5) of the articles were accepting without any criticisms of the SPE or the findings. This finding is not surprising considering that seven articles described the SPE as “classic.” An additional six articles described the SPE as “infamous/famous” ($n = 4$), “venerable” ($n = 1$), or “notorious” ($n = 1$), which further validates the SPE as something that is revered by some scholars. There were 12 articles (20.3%), however, that offered some form of critique of the study itself or of the findings. These criticisms could have been given in specific detail (e.g. the study was not ethical), by including a citation that criticizes the SPE, or by

stating the study had some limitations and not giving further details. Even though some criticisms were offered, authors more often cited the SPE in a favorable way where the findings were still accepted. For example, of those 12 articles that cited some critique, only 2 were not accepting and critical of the SPE. These findings reinforce the idea that the results of the SPE are generally accepted without any substantive overview of the limitations.

Discussion

“Classic” studies do not acquire their status as classics by accident. They are usually highly creative, they are bold in their approach and often ask questions in a fundamentally different way, they tend to inspire new lines of work and new data collection efforts, and they generally cause us to think differently about the things we study. The risk, however, is that status as a “classic” can inoculate a study against critique. Either wittingly or unwittingly, its findings can be treated as sacrosanct—as accepted wisdom that nobody thinks to question. In this regard, mounting evidence exists that the SPE falls into this category, where its findings have gone unquestioned despite both problems with its methodology and the failure of subsequent research to find similar results (Griggs, 2014; Griggs & Whitehead, 2014). The purpose of the present study was to call attention to the lack of organized skepticism from the field in this case by systematically examining how the SPE is cited and discussed in the criminological literature.

On a broad level, the status of the SPE as a classic and widely accepted study suggests that greater attention should be devoted to the replication of influential studies in the discipline. Although replications of classic studies certainly have been undertaken (see, e.g. Boman, Rebellon, & Meldrum, 2016; Lowenkamp, Cullen, & Pratt, 2003; Singer & Levine, 1988), they remain rare (McNeeley & Warner, 2015). A strong disincentive to undertaking replications is scholars’ belief, likely rooted in reality, that peer-reviewed journal editors and reviewers—particularly in top-tier outlets—prefer to publish new and cutting-edge original work (Ferguson & Heene, 2012; Madden, Easley, & Dunn, 1995). The challenge of submitting a replication study is that it will be rejected because it “says nothing new.” As we have noted, this preference for freshly minted articles may undermine scientific advance because such articles may not capture empirical reality; they may not be reproducible. As such, innovative findings—those that seem to reshape our thinking—should inspire not blind faith that new criminological truth has been established but the immediate call for confirmatory research. Other disciplines (e.g. medical and biological sciences, psychology) are now embracing the importance of systematic replication (Ioannidis, Ntzani, Trikalinos, & Contopoulos-Ioannidis, 2001; Open Science Collaboration, 2015). Criminology should as well.

A collateral risk is that a single replication of a prominent work will itself be considered definitive—especially if it seemingly falsifies a previously accepted

view. Accordingly, it is important to consider what it really means when a finding “fails to replicate.” This purported failure could come in many forms: It could simply be a difference in magnitude of effects, where large effects may have been found in an original study and statistically significant but substantially attenuated effects might be found in a replication. Or it could be that the findings of an original study are simply contingent, where results replicate under certain methodological conditions but not others. There could, of course, just be a complete failure to produce any of the original results under any methodological conditions. Still, the inability of a replication to confirm an original study’s findings does not mean that the this study reported misleading data; it could be that the replication study did. Thus, the failure of the BBC study to replicate the SPE may have occurred because the methodological conditions were too different (Zimbardo, 2006) or that other problems plagued it. If so, it is possible that it was the BBC study’s results that were problematic. Without a large body of replications, we really cannot say with confidence either way.

These considerations lead to the conclusion that, substantively, it is not clear whether the SPE’s depiction of prisons as inherently coercive is accurate, accurate to some degree, accurate under some circumstances, or inaccurate and thus misleading. Although the disquieting conditions and effects of prison are well studied, some works within the literature on institutional corrections also paint a more nuanced picture of prison life than the SPE would have us believe. These works show that whereas some institutions are in fact nasty places, others are safe, more attentive to inmate needs, and well managed. They also suggest that while many prisons are deplorable places, concrete steps can be taken to create environments that are less violent, are more therapeutic, and facilitate mature, healthy coping among inmates (see, e.g. Cullen, Jonson, & Stohr, 2014; Dilulio, 1987; Fleisher, 1989; Gendreau & Labrecque, *in press*; Johnson, 1996; Pratt & Maahs, 1999; Toch, 1977). In this context, additional replications of the SPE across time, space, and researchers might have provided invaluable insights into the nature of imprisonment. Of course, the limited replication of the SPE is understandable given the difficulty of conducting such a study and, more so, the potential harm to subjects that the original study seemingly produced. What is less clear, however, is the continued, near-universal acceptance of the original SPE’s findings among criminologists. As we have documented, rarely are the results of the SPE questioned in subsequent studies, and even when the findings are questioned the SPE’s broad conclusions are still generally accepted.

Although speculative, we can suggest a key reason for the SPE’s treatment as a classic study beyond critique: it confirmed what criminologists already thought about the nature and effects of imprisonment. First, as advanced by Sykes (1958), the deprivation model depicted the prison as a painful social world that inevitably produced an oppositional inmate culture and conflict with correctional officers. The SPE seemed to suggest that, as Sykes hypothesized, adaptations to incarceration were not due to the nature of the inmates but to

the nature of the prison environment. Second, the SPE was part of a broad critique of “total institutions”—whether prisons or mental hospitals (see Goffman, 1961). It was fashionable to deconstruct the rehabilitative ideal—to claim that it served as a mask of benevolence that covered up the unavoidable inhumanity of institutional life (see, e.g. Foucault, 1977; Rothman, 1980). Third, exposés of prisons seemingly confirmed the coercive nature of prisons and the willingness of the state to exert brutal power over inmates’ lives. The SPE was conducted in August 1971. September then saw the occurrence of the Attica prison riot, which included the slaughter of inmates (and guards) when those retaking facility used indiscriminate lethal force. This tragic event was struck many as a real-life “replication” of Zimbardo’s laboratory experiment. Fourth, in the decades that followed, the mass imprisonment movement gained steam and introduced a prolonged mean season into corrections (Clear & Frost, 2014). As Simon (2014) notes, a paradigm of “total incapacitation” prevailed in which offenders were seen as uniformly at risk, beyond redemption, and in need of incapacitation. Institutions became crowded, unhealthy, and arguably unconstitutional (Simon, 2014). Little evidence was forthcoming to falsify the central message of the SPE. In fact, as noted, much literature was produced documenting the disquieting conditions of American prisons (see, e.g. Irwin, 1980, 2005). Citing this classic study as incontrovertible truth thus provoked nods of the head and not raised eyebrows.

At this point, we do not know the extent to which professors cover the SPE in their courses—including using either a video documentary or movie—and what conclusions they draw in their lectures. A survey probing this issue would be an ideal candidate for future research to pursue. Regardless, we believe that the SPE offers an interesting and potentially invaluable conduit to use in classrooms to discuss important issues about the criminological enterprise—and especially about the value of organized skepticism. Thus, professors could use the study to illustrate to students how methodological choices can influence a study’s outcome, and how necessary it is to take a critical look at those choices. It could motivate a conversation with them about the importance of replication and what it means when a finding fails to replicate. It could serve as a cautionary tale to our students about the dangers of accepting without question the findings of a single study—even if that study was rigorously conducted and carefully executed. It could be used to remind students to be particularly wary of studies that confirm their preexisting biases—of the need to be just as skeptical of findings they like as those that strike them as viscerally untrue. And it could open up a discussion about whether reforming prisons is more possible than the SPE’s core findings seem to suggest. This is not to say we should engage in knowledge destruction and start claiming that no empirical findings should ever be trusted because they are all mere social constructions. Indeed, we want to stress that we are not saying that the key findings from the SPE are wrong. But they might be. We just do not really know and the SPE could be a very useful teaching tool for getting students to learn how to be skeptical of an empirical finding without being dismissive of it.

In the end, when publishing their own studies, the way that scholars present and discuss the classic works in the field is likely similar—if not identical—to how they present and discuss them to their students. In the case of the SPE, given how it is portrayed in the criminological literature, it is likely that we are explicitly presenting to our students an overly favorable view of the experiment, while at the same time implicitly communicating that healthy skepticism of scientific findings is unnecessary. Our results show that doing so might not be the best thing to do. But we can still do things differently, and the SPE can become topical again for reasons related to how scientific knowledge is created and communicated—reasons that have nothing to do with the fact that there was a movie about it. Admittedly, having a movie made about a criminological experiment is, well, pretty cool. Still, its very existence might well be one added reason to be skeptical. Hollywood takes license with the truth; we should not.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Teresa C. Kulig is a doctoral student in the School of Criminal Justice at the University of Cincinnati. Her recent writings include an evaluation of theories of environmental criminology, an analysis of sexual tourism, the longitudinal study of adolescent victimization, and racial bias in laws named after victims. Her research interests include testing theories of victimization, human trafficking and child exploitation, correctional treatment, and the organization of criminological knowledge through meta-analysis.

Travis C. Pratt is a Fellow in the University of Cincinnati Corrections Institute. He received his PhD in Criminal Justice from the University of Cincinnati. His research focuses primarily on criminological theory and correctional policy, with his most recent work addressing the nature and consequences of violent victimization. He is the author of *Addicted to Incarceration* (2009, Sage), and he has published more than 60 peer-reviewed articles that have appeared in journals such as *Criminology*, *Journal of Research in Crime and Delinquency*, *Journal of Quantitative Criminology*, *Journal of Pediatrics*, and *Crime and Justice: A Review of Research*.

Francis T. Cullen is Distinguished Research Professor Emeritus and a Senior Research Associate in the School of Criminal Justice at the University of Cincinnati. His recent works include *Challenging Criminological Theory: The Legacy of Ruth Rosner Kornhauser*, *Correctional Theory: Context and Consequences* (2nd ed.), and *Environmental Corrections: A New Paradigm for Supervising Offenders in the Community*. His current research focuses on the organization of criminological knowledge and on rehabilitation as a correctional policy. He is a Past President of both the American Society of Criminology and the Academy of Criminal Justice Sciences.

References

- Abrams, A. I., & Siegel, L. M. (1978). The transcendental meditation® program and rehabilitation at Folsom state prison: A cross-validation study. *Criminal Justice and Behavior, 5*, 3–20.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*, 369–404.
- Aviram, H. (2009). Mass atrocity and criminology. *Theoretical Criminology, 13*, 487–493.
- Banuazizi, A., & Movahedi, S. (1975). Interpersonal dynamics in a simulated prison: A methodological analysis. *American Psychologist, 30*, 152–160.
- Banyard, P. (2007). Tyranny and the tyrant: From Stanford to Abu Ghraib. Review of *The Lucifer effect: Understanding how good people turn evil* by P. Zimbardo. *The Psychologist, 20*, 494–495.
- Berger, P. L., & Luckmann, T. (1966). *The social construction of reality*. Garden City, NY: Anchor.
- Blevins, K. R., Listwan, S. J., Cullen, F. T., & Jonson, C. L. (2010). A general strain theory of prison violence and misconduct: An integrated model of inmate behavior. *Journal of Contemporary Criminal Justice, 26*, 148–166.
- Boman, J. H., Rebellon, C. J., & Meldrum, R. C. (2016). Can item-level error correlations correct for projection bias in perceived peer deviance measures? A research note. *Journal of Quantitative Criminology, 32*, 89–102.
- Bordt, R. L. (1999). Simulation as a tool for teaching research methods in a criminology course. *Journal of Criminal Justice Education, 10*, 373–382.
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*, 633–663.
- Brown, S. E. (1982). Research methods and criminal justice curricula: Surmounting the obstacles. *Criminal Justice Review, 7*, 11–16.
- Carlan, P. E., & Byxbe, F. R. (2000). The promise of humanistic policing: Is higher education living up to societal expectation? *American Journal of Criminal Justice, 24*, 235–246.
- Carroll, L. (1974). *Hacks, blacks, and cons: Race relations in a maximum security prison*. Prospect Heights, IL: Waveland.
- Cheek, F. E., & Miller, M. D. S. (1983). The experience of stress for correction officers: A double-bind theory of correctional stress. *Journal of Criminal Justice, 11*, 105–120.
- Clarke, R. V. (2012). Opportunity makes the thief. Really? And so what? *Crime Science, 1*, 1–9.
- Clear, T. R., & Frost, N. A. (2014). *The punishment imperative: The rise and failure of mass incarceration in America*. New York: New York University Press.
- Clements, C. B., Althouse, R., Ax, R. K., Magaletta, P. R., Fagan, T. J., & Wormith, J. S. (2007). Systemic issues and correctional outcomes: Expanding the scope of correctional psychology. *Criminal Justice and Behavior, 34*, 919–932.
- Cohn, E. G., & Farrington, D. P. (2012). Scholarly influence in criminology and criminal justice journals in 1990–2005. *Criminal Justice Review, 37*, 360–383.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Crewe, B. (2011). Soft power in prison: Implications for staff–prisoner relationships, liberty and legitimacy. *European Journal of Criminology, 8*, 455–468.
- Cullen, F. T. (2013). Rehabilitation: Beyond nothing works. In M. Tonry (Ed.), *Crime and justice in America 1975–2025* (Crime and justice: A Review of Research (Vol. 42, pp. 299–376). Chicago, IL: University of Chicago Press.

- Cullen, F. T., & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *Policies, processes, and decisions of the criminal justice system: Criminal justice 2000* (Vol. 3, pp. 109–175). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Cullen, F. T., & Gendreau, P. (2001). From nothing works to what works: Changing professional ideology in the 21st century. *The Prison Journal*, 81, 313–338.
- Cullen, F. T., & Jonson, C. L. (2017). *Correctional theory: Context and consequences* (2nd ed.). Thousand Oaks, CA: Sage.
- Cullen, F. T., Jonson, C. L., & Stohr, M. K. (2014). *The American prison: Imagining a different future*. Thousand Oaks, CA: Sage.
- Cullen, F. T., Smith, P., Lowenkamp, C. T., & Latessa, E. J. (2009). Nothing works revisited: Deconstructing Farabee's rethinking rehabilitation. *Victims and Offenders*, 4, 101–123.
- Dilulio, J. J. (1987). *Governing prisons: A comparative study of correctional management*. New York, NY: The Free Press.
- Egger, M., & Smith, G. D. (1998). Meta-analysis bias in location and selection of studies. *British Medical Journal*, 316, 61–66.
- Esbensen, F. A. (1986). Participant-observation in a county jail. *International Journal of Offender Therapy and Comparative Criminology*, 30, 163–175.
- Farrington, D. P. (1983). Randomized experiments on crime and justice. In M. Tonry & N. Morris (Eds.), *Crime and justice: A review of research* (Vol. 4, pp. 257–308). Chicago, IL: University of Chicago Press.
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7, 555–561.
- Fielding, N. G. (2012). Triangulation and mixed methods designs: Data integration with new research technologies. *Journal of Mixed Methods Research*, 6, 124–136.
- Fitch, K. L. (2005). Difficult interactions between IRBs and investigators: Applications and solutions. *Journal of Applied Communication Research*, 33, 269–276.
- Fleisher, M. S. (1989). *Warehousing violence*. Newbury Park, CA: Sage.
- Foucault, M. (1977). *Discipline and punish: The birth of the prison*. New York, NY: Pantheon.
- Freedman, D. H. (2010, November). Lies, damned lies, and medical science. *The Atlantic*, 76–86.
- Gendreau, P., & Labrecque, R. M. (in press). The effects of administrative segregation: A lesson in knowledge cumulation. In J. Wooldredge & P. Smith (Eds.), *The Oxford handbook of prisons and imprisonment*. New York, NY: Oxford University Press.
- Glaser, D. (1980). Alternatives to incarceration. *International Review of Criminal Policy*, 36, 55–58.
- Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. Garden City, NY: Doubleday.
- Gonzalez, J. L., Dawkins, M. P., & Hokanson, J. E. (1979). Autonomic reactions to stress: The effects of several Therapy analogues on prisoners. *Criminal Justice and Behavior*, 6, 107–117.
- Gottfredson, M. R. (1979). Treatment destruction techniques. *Journal of Research in Crime and Delinquency*, 16, 39–54.
- Gould, S. J. (1981). *The mismeasure of man*. New York, NY: W. W. Norton.
- Gouldner, A. W. (1970). *The coming crisis of Western sociology*. New York, NY: Avon.
- Gregoire, C. (2013, October 18). 10 psychological studies that will change what you think you know about yourself. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/2013/10/18/20-psychological-studies-_n_4098779.html
- Griffith, J. E., Pennington-Averett, A., & Bryan, I. (1981). Women prisoners' multidimensional locus of control. *Criminal Justice and Behavior*, 8, 375–389.

- Griggs, R. A. (2014). Coverage of the Stanford prison experiment in introductory psychology textbooks. *Teaching of Psychology, 41*, 195–203.
- Griggs, R. A. (2015a). The disappearance of independence in textbook coverage of Asch's social pressure experiments. *Teaching of Psychology, 42*, 137–142.
- Griggs, R. A. (2015b). The Kitty Genovese story in introductory social psychology textbooks: Fifty years later. *Teaching of Psychology, 41*, 142–159.
- Griggs, R. A., Whitehead, G. I., III. (2014). Coverage of the Stanford prison experiment in introductory psychology textbooks. *Teaching of Psychology, 41*, 318–324.
- Haney, C. (2008). A culture of harm: Taming the dynamics of cruelty in supermax prisons. *Criminal Justice and Behavior, 35*, 956–984.
- Haney, C., Banks, W. C., & Zimbardo, P. G. (1973). Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology, 1*, 69–97.
- Haney, C., & Zimbardo, P. (1998). The past and future of U.S. prison policy: Twenty-five years after the Stanford Prison Experiment. *American Psychologist, 53*, 709–727.
- Hazelrigg, L. (Ed.). (1968). *Prison within society: A reader in penology*. Garden City, NY: Doubleday.
- Hogan, N. L., Lambert, E. G., Hepburn, J. R., Burton, V. S., & Cullen, F. T. (2005). Is there a difference? Exploring male and female correctional officers' definition of and response to conflict situations. *Women and Criminal Justice, 15*, 143–165.
- Hokanson, J. E., Megargee, E. I., O'Hagan, S. E., & Perry, A. M. (1976). Behavioral, emotional, and autonomic reactions To stress among incarcerated, youthful offenders. *Criminal Justice and Behavior, 3*, 203–234.
- Hussemann, J. M., & Page, J. (2011). Gender diversity and the prospects for progressive prison reform. *Women and Criminal Justice, 21*, 267–289.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*, 696–701.
- Ioannidis, J. P. A., Ntzani, E. E., Trikalinos, T. A., & Contopoulos-Ioannidis, D. G. (2001). Replication validity of genetic association studies. *Nature Genetics, 29*, 306–309.
- Irwin, J. (1980). *Prisons in turmoil*. Boston, MA: Little, Brown.
- Irwin, J. (2005). *The warehouse prison: Disposal of the new dangerous class*. Los Angeles, CA: Roxbury.
- Jacobs, A. K. (2005). "We're all brothers! All worm meat!" A consideration of Kazantzakis' Zorba the Greek in the quest for the just society. *Contemporary Justice Review, 8*, 229–237.
- Jacobs, J. B., & Kraft, L. J. (1978). Integrating the keepers: A comparison of black and white prison guards in Illinois. *Social Problems, 25*, 304–318.
- Johnson, R. (1996). *Hard time: Understanding and reforming the prison* (2nd ed.). Belmont, CA: Wadsworth.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*, 14–26.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*, 112–133.
- Jones, R. S. (1995). Uncovering the hidden social world: Insider research in prison. *Journal of Contemporary Criminal Justice, 11*, 106–118.
- Kolstad, A. (1996). Imprisonment as rehabilitation: Offenders' assessment of why it does not work. *Journal of Criminal Justice, 24*, 323–335.
- Kraut, R. E. (1976). Deterrent and definitional influences on shoplifting. *Social Problems, 23*, 358–368.
- Krisberg, B., Marchionna, S., & Hartney, C. (2015). *American corrections: Concepts and controversies*. Thousand Oaks, CA: Sage.
- Lankford, A. (2009). Promoting aggression and violence at Abu Ghraib: The US military's transformation of ordinary people into torturers. *Aggression and Violent Behavior, 14*, 388–395.

- Lankford, A., & Gillespie, K. (2011). Rehabilitating terrorists through counter-indoctrination: Lessons learned from the Saudi Arabian program. *International Criminal Justice Review*, 21, 118–133.
- Lankford, A., & Hakim, N. (2011). From Columbine to Palestine: A comparative analysis of rampage shooters in the United States and volunteer suicide bombers in the Middle East. *Aggression and Violent Behavior*, 16, 98–107.
- Latessa, E. J., & Holsinger, A. (Eds.). (2016). *Correctional contexts: Contemporary and classical readings*. New York, NY: Oxford University Press.
- Lehrer, J. (2010, December 13). The truth wears off: Is there something wrong with the scientific method? *The New Yorker*, 86, 52–57.
- Lerman, A. E., & Page, J. (2012). The state of the job: An embedded work role perspective on prison officer attitudes. *Punishment and Society*, 14, 503–529.
- Lewis, J. E., DeGusta, D., Meyer, M. R., Monge, J. M., Mann, A. E., & Holloway, R. L. (2011). The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skills and bias. *PLoS Biology*, 9, 1–6.
- Lindquist, C. A., & Whitehead, J. T. (1986). Guards released from prison: A natural experiment in job enlargement. *Journal of Criminal Justice*, 14, 283–294.
- Lipsey, M., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Listwan, S. J., Daigle, L. E., Hartman, J. L., & Guastaferro, W. P. (2014). Poly-victimization risk in prison: The influence of individual and institutional factors. *Journal of Interpersonal Violence*, 29, 2458–2481.
- Lowenkamp, C. T., Cullen, F. T., & Pratt, T. C. (2003). Replicating Sampson and Groves's test of social disorganization theory: Revisiting a criminological classic. *Journal of Research in Crime and Delinquency*, 40, 351–373.
- Maahs, J., & Pratt, T. C. (in press). "I hate these little turds!" Science, entertainment, and the enduring popularity of scared straight programs. *Deviant Behavior*.
- Madden, C. S., Easley, R. W., & Dunn, M. G. (1995). How journal editors view replication research. *Journal of Advertising*, 24, 77–87.
- Maier-Katkin, D., Mears, D. P., & Bernard, T. J. (2009). Towards a criminology of crimes against humanity. *Theoretical Criminology*, 13, 227–255.
- Marquart, J. W. (2005). Understanding the power of social contexts on criminal justice institutions. *Journal of Criminal Justice Education*, 16, 217–225.
- Marquart, J. W., Barnhill, M. B., & Balshaw-Biddle, K. (2001). Fatal attraction: An analysis of employee boundary violations in a southern prison system, 1995–1998. *Justice Quarterly*, 18, 877–910.
- Maruna, S. (2011). Why do they hate us? Making peace between prisoners and psychology. *International Journal of Offender Therapy and Comparative Criminology*, 55, 671–675.
- McNeeley, S., & Warner, J. J. (2015). Replication in criminology: A necessary practice. *European Journal of Criminology*, 12, 581–597.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. N. K. Storer (Ed.). Chicago, IL: University of Chicago Press.
- Morgan, R. D., Van Haveren, R. A., & Pearson, C. A. (2002). Correctional officer burnout: Further analyses. *Criminal Justice and Behavior*, 29, 144–160.
- Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and reoffending. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 38, pp. 115–200). Chicago, IL: University of Chicago Press.
- O'Donnell, I. (2003). A new paradigm for understanding violence? Testing the limits of Lonnie Athens's Theory. *British Journal of Criminology*, 43, 750–771.
- Oleson, J. C. (2005). Book review: The U.S. federal prison system. *Theoretical Criminology*, 9, 499–503.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, pp. 943; aac4716-1–aac4716-8. doi:10.1126/science.aac4716

- Pepinsky, H. E. (1988). Violence as unresponsiveness: Toward a new conception of crime. *Justice Quarterly*, 5, 539–563.
- Perry, A. M., & Hokanson, J. E. (1977). Race factors in responses to interpersonal stress among young adult offenders. *Criminal Justice and Behavior*, 4, 45–61.
- Petronio, R. J. (1982). Role socialization of juvenile court probation officers. *Criminal Justice and Behavior*, 9, 143–158.
- Philliber, S. (1987). Thy brother's keeper: A review of the literature on correctional officers. *Justice Quarterly*, 4, 9–37.
- Piquero, A. R., Jennings, W. G., & Farrington, D. P. (2010). On the malleability of self-control: Theoretical and policy implications regarding the general theory of crime. *Justice Quarterly*, 27, 803–834.
- Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, 38, 931–964.
- Pratt, T. C., & Maahs, J. (1999). Are private prisons more cost-effective than public prisons? A meta-analysis of evaluation research studies. *Crime and Delinquency*, 45, 358–371.
- Pratt, T. C., Turanovic, J. J., Fox, K. A., & Wright, K. A. (2014). Self-control and victimization: A meta-analysis. *Criminology*, 52, 87–116.
- Rector, M. G., & Gilman, D. (1976). How did we get here and where are we going: The future of the juvenile court system. *Criminal Justice Review*, 1, 77–90.
- Reicher, S., & Haslam, S. A. (2006). Rethinking the psychology of tyranny: The BBC prison study. *British Journal of Social Psychology*, 45, 1–40.
- Roberg, R. R. (1981). Management research in criminal justice: Exploring ethical issues. *Journal of Criminal Justice*, 9, 41–49.
- Rothman, D. J. (1980). *Conscience and convenience: The asylum and its alternatives in Progressive America*. Boston, MA: Little, Brown.
- Santos, M. G., & McPherson, R. B. (1995). Transcending the wall. *Journal of Criminal Justice Education*, 6, 41–60.
- Schmidt, F. L., & Oh, I.-S. (2015). *The crisis of confidence in research findings in behavioral and social sciences: Is lack of replication the real problem? Or is it something else?* Expanded version of invited address presented at the 2014 conference of the Association of Psychological Science.
- Schwartz, M. D., Clear, T. R., & Travis, L. F., III. (Eds.). (1980). *Corrections: An issues approach*. Cincinnati, OH: Anderson.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366.
- Simon, J. (2014). *Mass incarceration on trial: A remarkable court decision and the future of prisons in America*. New York, NY: The New Press.
- Singer, S. I., & Levine, M. (1988). Power-control theory, gender, and delinquency: A partial replication with additional evidence on the effects of peers. *Criminology*, 26, 627–648.
- Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology*, 37, 57–86.
- Stinchcomb, J. B. (2002). Prisons of the mind: Lessons learned from home confinement. *Journal of Criminal Justice Education*, 13, 463–478.
- Stohr, M. K., & Walsh, A. (2016). *Corrections: The essentials* (2nd ed.). Thousand Oaks, CA: Sage.
- Suedfeld, P. (1978). Beyond sentimentality. *International Journal of Offender Therapy and Comparative Criminology*, 22, 49–55.
- Sundt, J. L., & Cullen, F. T. (2002). The correctional ideology of prison chaplains: A national survey. *Journal of Criminal Justice*, 30, 369–385.

- Sykes, G. M. (1958). *The society of captives: A study of maximum security prison*. Princeton, NJ: Princeton University Press.
- Timor, U. (1998). Constructing a rehabilitative reality in special religious wards in Israeli prisons. *International Journal of Offender Therapy and Comparative Criminology*, 42, 340–359.
- Toch, H. (1977). *Living in prison: The ecology of survival*. New York, NY: The Free Press.
- Tolich, M. (2014). What can Milgram and Zimbardo teach ethics committees and qualitative researchers about minimizing harm? *Research Ethics*, 10, 86–96.
- Unnithan, N. P. (1986). Research in a correctional setting: Constraints and biases. *Journal of Criminal Justice*, 14, 401–412.
- Vaisman-Tzachor, R. (2007). Profiling terrorists. *Journal of Police Crisis Negotiations*, 7, 27–61.
- van Gelder, J. L., Otte, M., & Luciano, E. C. (2014). Using virtual reality in criminological research. *Crime Science*, 3, 1–12.
- Weiner, N. L. (1975). The effect of education on police attitudes. *Journal of Criminal Justice*, 2, 317–328.
- Welch, M. (1996). *Corrections: A critical approach*. New York, NY: McGraw Hill.
- Whitehead, J. T., Dodson, K. D., & Edwards, B. D. (2013). *Corrections: Exploring crime, punishment, and justice in America* (3rd ed.). Waltham, MA: Anderson Publishing.
- Williams, M. J., & Lindsey, S. C. (2014). A social psychological critique of the Saudi terrorism risk reduction initiative. *Psychology, Crime & Law*, 20, 135–151.
- Wolfgang, M. E. (1981). Surveying violence across nations: A review of the literature, with research and policy recommendations. *International Review of Criminal Policy*, 37, 62–95.
- Worley, R. M., & Barua Worley, V. (2013). Inmate public autoerotism uncovered: exploring the dynamics of masturbatory behavior within correctional facilities. *Deviant Behavior*, 34, 11–24.
- Worley, R. M., & Worley, V. B. (2013). Games guards play: A self-report study of institutional deviance within the Texas department of criminal justice. *Criminal Justice Studies*, 26, 115–132.
- Zimbardo, P. G. (1972). Pathology of imprisonment. *Society*, 9, 6–8.
- Zimbardo, P. G. (2006). On rethinking the psychology of tyranny: The BBC prison study. *British Journal of Social Psychology*, 45, 47–53.
- Zimbardo, P. G. (2007). *The Lucifer effect: Understanding how good people turn evil*. New York, NY: Random House.
- Zimbardo, P. G., Banks, W. C., Haney, C., & Jaffe, D. (1973). A Pirandellian prison: The mind is a formidable jailer. *New York Times Magazine*, 8, 38–60.

Appendix A Alphabetized journals with number of articles published containing the SPE citation (N = 59)

Title of journal	Number of articles published
1. <i>Aggression and Violent Behavior</i>	2
2. <i>American Journal of Criminal Justice</i>	1
3. <i>British Journal of Criminology</i>	1
4. <i>Contemporary Justice Review</i>	1
5. <i>Crime and Justice: A Review of Research</i>	3
6. <i>Crime Science</i>	2
7. <i>Criminal Justice and Behavior</i>	9
8. <i>Criminal Justice Review</i>	2
9. <i>Criminal Justice Studies</i>	1
10. <i>Deviant Behavior</i>	1
11. <i>European Journal of Criminology</i>	1
12. <i>International Criminal Justice Review</i>	1
13. <i>International Journal of Offender Therapy and Comparative Criminology</i>	4
14. <i>International Review of Criminal Policy</i>	2
15. <i>Journal of Contemporary Criminal Justice</i>	2
16. <i>Journal of Criminal Justice</i>	7
17. <i>Journal of Criminal Justice Education</i>	4
18. <i>Journal of Interpersonal Violence</i>	1
19. <i>Journal of Police Crisis Negotiations</i>	1
20. <i>Justice Quarterly</i>	3
21. <i>Psychology, Crime, & Law</i>	1
22. <i>Punishment and Society</i>	1
23. <i>Social Problems</i>	2
24. <i>Theoretical Criminology</i>	3
25. <i>Victims and Offenders</i>	1
26. <i>Women & Criminal Justice</i>	2

Appendix B Alphabetized publications that were accepting of the SPE (N = 47)

Publication citation	Detailed content of the SPE
<i>Prisons are inhumane/coercive/painful due to <u>structure/environment</u></i>	
Abrams and Siegel (1978)	"As it is generally structured, our prison system may tend to exacerbate the offenders' feelings of helplessness and hostility (Haney et al., 1973)" (p. 4).
Clements et al. (2007)	"The potentially adverse impact of prison environments has been studied for several decades and was famously symbolized in the 'Stanford Prison Experiment' (Haney, Banks, & Zimbardo, 1973)" (p. 922).
Cullen, Smith, Lowenkamp, and Latessa (2009)	"Prisons were now portrayed as inherently inhumane, whether in Zimbardo's Stanford prison experiment (Zimbardo, Banks, Haney, & Jaffe, 1973) or in Rothman's (1971) social history of the penitentiary" (p. 104).
Cullen (2013)	"The classic 1971 Stanford Prison Experiment, administered by Zimbardo and his colleagues (1973), cemented the view that prisons were inherently inhumane and thus could not function as therapeutic milieus (see also Zimbardo 2007). A mock prison was constructed in the bottom floor of a Stanford University building, and psychologically healthy subjects were randomly assigned to be inmates or guards. Within a short time, the guards moved to assert their control. They regularly 'insulted the prisoners, threatened them, were physically aggressive, used instruments (night sticks, fire extinguishers, etc.) to keep the prisoners in line and referred to them in impersonal, anonymous, deprecating ways' (Zimbardo et al. 1973, pp. 48–49). To avoid serious harm, the experiment had to be halted prematurely. Upon reflection, it sent a chilling message about the absurdity of seeing prisons as settings for inmate improvement: The potential social value of this study derives precisely from the fact that normal, healthy, educated young men could be so radically transformed under the institutional pressures of a 'prison environment.' If this could happen in so short a time, without the excesses that are possible in real prisons, and if it could happen to the 'cream-of-the-crop of American youth,' then one can only shudder to imagine what society is doing both to the actual guards and prisoners who are at this very moment participating in this unnatural 'social experiment' (Zimbardo et al. 1973, p. 56)" (pp. 317–318).

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Marquart, Barnhill, and Balshaw-Biddle (2001)	<p>“Why would seemingly normal employees become romantically (and even sexually) involved with prisoners? Haney, Banks, and Zimbardo (1973) posed a similar question when they asked why bright, emotionally stable young male college students (most of whom were in the peace movement) would abuse their fellow students in a prisonlike environment. Their answer to this question emphasized the surprising power of institutional environments over the participants (Haney and Zimbardo 1998:710). As Haney and colleagues (1973:90) suggested, ‘[T]he abnormality here [the mock prison] resided in the psychological nature of the situation and not in those who passed through it.’ Over the past 27 years, these researchers have generated a body of research that clearly illustrates how ‘mind altering social psychological dynamics’ can bend and twist human behavior (Haney and Zimbardo 1998:710). If one accepts the situational explanation, then complex issues such as consent become difficult to define and comprehend. The inmates’ version must be researched if we are to fully ascertain how and why these dual relationships develop” (p. 905; other quote on p. 885).</p>
Maruna (2011)	<p>“Indeed, the most important lesson for prison psychology might be the need to recognize that prisons—far from neutral environments ideal for the measurement of psychological health—are themselves deeply pathological, disruptive to normative human development (see Liebling & Maruna, 2005). This emerges clearly in Rocque, Bierie, and MacKenzie’s (2011) important analysis of the impact of incarceration on social bonds and indeed is surely the most critical insight in the field of prison psychology (see Haney, 2006; Zimbardo, 1973)” (p. 674).</p>
Morgan, Van Haveren, and Pearson (2002)	<p>“Prison environments contribute to aggressive, rigid, and power-motivated behaviors, as evidenced in the Zimbardo Prison Experiment (Haney, Banks, & Zimbardo, 1973)” (p. 145).</p>
Nagin, Cullen, and Jonson (2009)	<p>“At about this same time, observers of corrections learned of the Stanford Prison Experiment (Haney, Banks, and Zimbardo 1973). Psychologically normal</p>

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Oleson (2005)	<p>students were randomly assigned roles as guards or inmates. The newly minted guards quickly began to abuse and otherwise mistreat the unlucky students assigned to be prisoners. This reinforced the view of many that prisons were inherently coercive and inhumane (Zimbardo 2007)” (p. 117).</p> <p>“The chapter on work, for example, adequately surveys the literature on prison labor, citing fourteen academic sources and two emic accounts, but the literature review of the prison-industrial complex is not comprehensive. Likewise, the chapter on staff includes a section on brutality, but makes no mention of the notorious Stanford Prison Experiment, which provocatively suggested that brutality was a function of the prison environment (Haney et al., 1973)” (p. 501).</p>
Philliber (1987)	<p>“A discussion of guards’ attitudes toward inmates would be remiss without mention of the classic study by Haney, Banks, and Zimbardo (1973). Using students in a simulated prison, these researchers watched the emergence of punitive and hostile behavior among those assigned to be guards. The implication of their work is that the prison environment itself promotes such attitudes among otherwise normal people” (p. 23; other quote on pp. 25–26).</p>
Rector and Gilman (1976)	<p>“We conclude that incarceration of children in jails, detention centers, reformatories or training schools is destructive of the human spirit, is antithetical to rehabilitative or treatment goals, and severely retards children’s potential for normal growth and adjustment (Zimbardo, 1971)” (p. 88).</p>
Santos and McPherson (1995)	<p>“On the other hand, the Stanford experiment suggests that the prison experience itself influences adjustment (Haney, Banks, and Zimbardo 1977)” (p. 58).</p>
Timor (1998)	<p>“According to many scholars, the traditional prison is not an institution fit for the rehabilitation of criminals (e.g., Allen, 1981; Marshall, 1981). Its special components, such as its structural features (Zimbardo, 1972) and the cultural features of those imprisoned there (Irwin & Cressey, 1962), inhibit its ability to rehabilitate prisoners” (p. 340).</p>

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
<i>Prisons are inhumane/coercive/painful due to social roles</i>	
Carlan and Byxbe (2000)	"Ever since the unveiling of the Zimbardo experiment (Haney, Banks, & Zimbardo, 1973), worries have abounded about the disturbing aspects of the 'police personality.' While early studies substantiated the portrayal of police officers as authoritarians, there are indications a college education mitigates or softens this factor" (p. 243).
Glaser (1980)	"Indeed, experiments have shown that even persons who are kind and considerate elsewhere can readily become extremely cruel when placed in the role of prison guard ³ " (p. 56). Footnote 3: Craig Haney, Curtis Banks and Philip Zimbardo, "Interpersonal dynamics in a simulated prison", <i>International Journal of Criminology and Penology</i> , vol. 1, No. 1 (February 1973), pp. 69-97; Philip Zimbardo, "Pathology of imprisonment", <i>Society</i> , vol. 9, No. 2 (April 1972), pp. 6-8.
Hussemann and Page (2011)	"The prison officer role promotes antipathy toward inmates, as depicted so clearly in the well-known 'Stanford prison experiment' (Haney, Banks, and Zimbardo, 1973)" (p. 273).
Jacobs and Kraft (1978)	"That the relationship between guard and inmate is one of structured conflict, which cannot be surmounted even by racial and ethnic ties, is the dramatic conclusion of a well-known Stanford experiment (Haney, Banks, and Zimbardo, 1973). The experiments randomly assigned white Anglo-Saxon male college student volunteers to roles of guard and prisoner in a 'mock prison' constructed in the basement of a university building. Both groups were dressed appropriately. The guards were given simple rules to enforce and were told to maintain order. Conflict between the two groups grew so intense that the experiment had to be stopped. The 'prisoners' showed signs of psychopathology, especially depression. Many 'guards' acted in an authoritarian manner, even sadistically. This experiment is often cited to support the proposition that the behavior of guards and inmates is a product of their organizational roles, and is independent of such extra institutional variables as education, age, political orientation, and race" (pp. 305–306).

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Lankford (2009)	<p>“Both Milgram and Zimbardo’s experiments provide empirical evidence that—even without being specifically recruited and trained for violence—ordinary people will regularly and reliably carry out violence when they are authorized to do so...Zimbardo’s Stanford prison experiment shows that even without the direct orders present in Milgram’s study, ordinary people can still be compelled to act violently in certain situations. Zimbardo (1972) selected 24 ‘normal’ individuals from a pool of young male volunteers— screening out those with criminal records or any hint of medical or psychological disorders. He then randomly divided them into ‘prisoners’ and ‘prison guards’ and placed them in a mock prison environment to see what would happen. After a few days, the guards began berating and humiliating the prisoners, dragging them around, locking them in closets, shooting them with fire extinguishers, and even forcing them to simulate sodomy. Zimbardo stopped the experiment after six days because multiple prisoners had suffered psychological breakdowns and the guards were becoming increasingly abusive. The relevance of Zimbardo’s study is still quite apparent today: his Stanford prison experiment was cited in multiple government reports as evidence that the situational factors at Abu Ghraib contributed to the abuses on site (Zimbardo, 2007). In addition, Zimbardo’s (2007) detailed examination of Abu Ghraib Staff Sergeant Chip Frederick’s personal history and psychological profile provides virtually indisputable evidence that the young man was an ordinary soldier before he yielded to the extraordinary pressures of the system. Milgram and Zimbardo’s research shows that, at least in certain contexts, ordinary people will act aggressively and violently when authorized to do so...In the Stanford prison experiment, despite never giving direct orders for violence, Zimbardo authorized the guards’ aggressive behavior in many important ways. He told the guards that their goal should be to ‘create in [the prisoners] a sense of powerlessness,’ he encouraged the guards to act as tough and aggressive as real prison guards, and he</p>

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Maier-Katkin et al. (2009)	<p>had his staff reprimand the less aggressive guards for being 'too soft' (Zimbardo, 2007, p. 55, 65). The result was that all 12 ordinary young men who had been assigned to the role of guard participated in the physical and psychological abuses. It appears that at Abu Ghraib, when carefully recruited, well-trained military recruits received similar authorizations from U.S. leaders, it produced similarly violent consequences" (p. 390; other quotes on pp. 388, 392, 394).</p> <p>"For example, classic social psychological experiments on authority (Milgram, 1974) and prisonization (Zimbardo et al., 1973) have established the prima facie plausibility of the banal evil thesis. Specifically, humans respond to authority and the expectations associated with various social roles, and both in turn can transform ordinary people into 'perpetrators of evil' (Zimbardo, 1999, 2007). These social psychological studies demonstrate the plasticity of human behavior which can be highly variable depending on context, but are not particularly helpful in identifying the range of social conditions that give rise to or encourage compliance with authority or that channel hostilities in certain directions" (pp. 232–233; other quotes on pp. 230, 245).</p>
Petronio (1982)	<p>"The power that a role may have for subsequent behavior is dramatized in an investigation by Haney, Banks, and Zimbardo (1973), where the interpersonal dynamics of guards and prisoners were studied in a simulated prison" (p. 147).</p>
Weiner (1975)	<p>"In a situation highly comparable to that of the policemen, Zimbardo (1973) found that perfectly 'normal' college students cast in the roles of prisoner and prison guard for even short periods of time became prisoner-like and guard-like in their attitudes and actions-solely as a result of their role-playing" (p. 324).</p>
Williams and Lindsey (2014)	<p>"One need look no further than the Stanford prison experiment for an example of the ways in which individuals' behavior and attitudes can be quickly and strongly shaped by the identities they assume (Haney, Banks, & Zimbardo, 1973)" (p. 140).</p>

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Wolfgang (1981)	"Civilian police in the United States have been found to develop a cynical 'working personality' (Skolnick, 1966), and ordinary citizens often develop dominant personalities while temporarily in the role of prison custodian (Zimbardo, 1972)" (p. 72).
<i>Prisons are inhumane/coercive/painful due to a combination of both</i> Cheek and Miller (1983)	"Zimbardo's classic study (Haney, Banks, Zimbardo, 1973) dramatically showed the development of aggressive, rigid, power-intoxicated behavior in a mock prison setting, while Crouch and Marquart (1980) describe how officers learn to act authoritarian using 'bluster tactics' in order to control inmates" (p. 117).
Esbensen (1986)	"Haney et al. (1973) have reported that in the Stanford Simulation Project, that in the privacy of their cells, conversations between 'prisoners' revolved around issues relating to the 'prison.' This observation applies not only to behavior and rules, but pertains likewise to conceptions about power and control" (pp. 168–169; other quote on p. 167).
Griffith, Pennington-Averett, and Bryan (1981)	"Summarizing results of the Stanford prison experiment (Haney, Banks, & Zimbardo, 1973), Zimbardo (1976) suggested the more adaptive role available to prisoners is being 'docile,' 'cooperative,' 'uncomplaining,' 'conforming in thought,' and 'complying in deed'" (p. 376).
Hokanson, Megargee, O'Hagan, and Perry (1976)	"This pattern of results offers support for the 'depressed-helpless' model of prisoner characteristics and is consistent with the findings from Haney, Banks, and Zimbardo, (1973)" (p. 226; other quote on p. 204).
Kolstad (1996)	"When ordinary persons are locked up in prison they react with pathological, deviant behavior (Zimbardo, Haney, and Banks, 1973). The prison experiment at Stanford University demonstrated the detrimental consequences of confinement on normal, healthy 'offenders.' Educated college men, selected because of their maturity and personal stability, were radically transformed under the institutional pressure of the prison environment. Their submissiveness and strange behavior arose out of situational factors in an institutional environment. What would otherwise have been termed 'pathological behavior' became 'appropriate' behavior in the prison" (pp. 331–332).

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Pepinsky (1988)	"Violence results from depersonalization, which can be induced in any number of ways, as by putting people in the bureaucratic position of doing things to others without responding to their experience and their feelings. Zimbardo (1978) achieved this result by introducing some students to others as prisoners" (p. 549).
Perry and Hokanson (1977)	"These patterns, reflecting reduced abilities to cope effectively with interpersonal stress, are consistent with 'prisoner' reactions observed by Haney, Banks, and Zimbardo (1973), conditions of learned helplessness described by Seligman (1975), and the culturally imbued feelings of powerlessness among minorities found by Ransford (1968)" (p. 58; other quote on p. 46).
Stinchcomb (2002)	"Thinking back to the Stanford Prison Experiment some 25 years ago (Haney, Banks, and Zimbardo 1973), my first inclination was to create some type of 'mock prison.' But aside from the logistics of accommodating large classes in such a simulation, I was reminded of the extreme pathological results that forced the Stanford faculty to cut their project short. Realizing that their experiment was conducted before society became as litigious as we are today only provoked further misgivings" (p. 469; other quote on p. 466).
Sundt and Cullen (2002)	"This hypothesis is most strikingly illustrated by the Stanford Prison Experiment (Haney, Banks, & Zimbardo, 1973), although this model is also consistent with the <i>prisonization</i> and <i>deprivation</i> models of inmate behavior (see, e.g., Clemmer, 1940; Sykes, 1958). The Stanford Prison Experiment examined the influence of contextual factors and adherence to social roles on the behavior of a homogenous group of college students who were assigned the roles of 'inmates' and 'guards' within a mock prison. After six days, however, the experiment, which was planned to last two weeks, was halted because the behavior of the participants had become 'pathological.' Among the guards, for example, the experimenters observed high levels of aggression and negativity. Some guards exhibited 'tyrannical' behavior and even the 'good' guards tolerated the dehumanizing, brutal behavior of their fellow officers. These dramatic results led to the

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Worley and Barua Worley (2013)	<p>conclusion that the behavior of the research participants was the result of the roles that they had been assigned and the 'inherent pathological characteristics of the prison situation itself' (Haney et al., 1973, p. 30). According to the <i>work role/ prisonization</i> model, then, the work experiences and attitudes of correctional employees are determined by the structural and contextual influences of the prison and the nature of the work role" (p. 372; other quote on p. 370).</p> <p>"Even the Stanford Prison Experiment illustrates that actors in controlled environments tend to conform with the norms and expectations imposed by their captors (Haney et al. 1973; Zimbardo, 2007)" (p. 20).</p>
<i>String cite with other works</i> Aviram (2009)	<p>"Much could be built on the insights gained by Stanley Milgram (1974) and the Stanford Prison Experiment team (Haney et al., 1973) in their classic experimental studies of human cruelty and obedience to authority" (p. 491).</p>
Blevins et al. (2010)	<p>"As a result, prisoners forfeit their autonomy, a sense of total safety and security, personal identities, access to many material goods and services, privacy, heterosexual relationships, unrestricted interaction with family and friends, and many other general comforts of life (Sykes, 1958; see also Haney, Banks, & Zimbardo, 1973; Toch, 1977). Many of these 'pains' are particularly difficult to adjust to while incarcerated" (p. 151).</p>
Clarke (2012)	<p>Endnote e: "Pease and Laycock (2011) point out that <i>Crime as Opportunity</i> failed to mention some important psychological experiments - in particular Zimbardo's (Haney et al. 1973) simulated prison experiments and Milgram's (1974) studies in obedience - showing that deviant behavior is heavily influenced by situational factors. Again, these studies were rarely cited in criminological literature and, for us, they differed importantly from our work on opportunity: the psychological experiments were investigating the effect of pressure to behave cruelly in a laboratory setting, while our studies were concerned with the extent to which opportunity passively provides the occasion for committing any kind of crime in the real world" (p. 7).</p>

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Crewe (2011)	"It is widely agreed, in both official and academic discourse, that at the heart of any prison is the relationship between staff and prisoners. These relationships are generally considered to be more positive in England and Wales than in many other jurisdictions (Liebling and Price, 2001). This is despite a fairly recent history that confirmed many of the dystopian prognoses made by scholars of the pathologies and psychology of power (most notably Haney et al., 1973, and Milgram, 1963)" (p. 455).
Hogan, Lambert, Hepburn, Burton, and Cullen (2005)	"The correctional literature historically described guards as 'hacks,' promoting a stereotype of callousness and brutality (Haney et al., 1973; Irwin, 1980; Johnson, 2002)" (p. 159; other quote on p. 147).
Jacobs (2005)	"Social-psychological studies have demonstrated that social-structural factors can overwhelm a variety of personalities (Asch, 1995; Milgram, 1969; Zimbardo, 1972), and nationalism, as seen in the breakup of the Soviet Union and Yugoslavia and in the current American trends since 9/11, continues as a pervasive motivator (see Gellner, 1983)" (p. 231).
Kraut (1976)	"Most sociological research examining the effects of apprehension has confounded the environmental and cognitive consequences. For example, when institutionalization locks inmates into deviance, it does so through its far reaching effects on both their social environment and their knowledge and beliefs (e.g. Goffman, 1961; Zimbardo, 1973)" (p. 359).
Lankford and Gillespie (2011)	"A great deal of past research has demonstrated that authorizations for violence can lead relatively ordinary people to engage in various forms of brutality. Milgram's experiments on obedience to authority and Zimbardo's Stanford prison experiment were largely successful due to the strategic manner in which their authorizations for violence were crafted. Both studies employed transcendent and mundane authorizations (Lankford, 2009; Milgram, 1963; Zimbardo, 1972), as did Burger's (2007) successful replication of the Milgram experiment" (p. 122; other quotes on pp. 119–120, 124, 127, 129).
Lankford and Hakim (2011)	"Another key similarity between volunteer suicide bombers in the Middle East and rampage shooters in the U.S. is that both appear to be products of

(Continued)

Appendix B (Continued)

Publication citation	Detailed content of the SPE
Lerman and Page (2012)	<p>oppressive social environments. Previous research across a range of cultures has shown that social and contextual factors often increase the likelihood of violence (Lankford, 2009a; Milgram, 1963; Waller, 2002; Zimbardo, 1972; Zimbardo, 2007)” (p. 102).</p> <p>“Complicating (and extending) earlier conceptions of staff culture (e.g. Crouch and Marquart, 1980; Haney et al., 1973; Jacobs and Retsky, 1980), which focused almost exclusively on negative elements, Liebling (2008: 108) argues that this culture has ‘both positive negative characteristics’” (p. 511; other quote on p. 505).</p>
Listwan, Daigle, Hartman, and Guastaferrro (2014)	<p>“Those losses affect some more than others and can range from a loss of material goods, heterosexual relationships, and access to family (Sykes, 1958; see also Haney, Banks, & Zimbardo, 1973; Toch, 1977)” (p. 2462).</p>
Unnithan (1986)	<p>“The nature of the correctional setting, which is an attempt ‘to create and maintain total or almost total social control’ (Sykes, 1958), and the intrusive quality of situational variables on the activities of participants (Haney et al., 1973; Johnson and Toch, 1982) have been discussed” (p. 402).</p>
Vaisman-Tzachor (2007)	<p>“Furthermore, the effects of life in captivity in Israeli, in European, and in American prisons cannot be underestimated in the overall presentation of the psychological profiles that emerge out of such studies (Zimbardo, 1971; Haney, Banks & Zimbardo, 1973; Zagorin & Duffy, 2005)” (p. 42).</p>
<i>SPE as a research method</i> Brown (1982)	<p>“Zimbardo’s (1976) classic prison simulation study at Stanford certainly demonstrates the potential of simulation studies for absorbing students in the research process” (p. 14).</p>

Appendix C Alphabetized publications that were critical but accepting of the SPE (N = 10)

Publication citation	Detailed content of the SPE
<i>Prisons are inhumane/coercive/painful due to <u>structure/environment</u></i>	
Lindquist and Whitehead (1986)	<p>"The first assertion, that there is something about working in a prison that brings out the worst in a person, may reflect the 'situational hypothesis' popularized by Haney, Banks, and Zimbardo (1973). In their analysis of the now famous Stanford Prison Experiment, these authors attributed 'guard' hostility toward 'inmates' to the inherently pathological characteristics of the 'prison' itself (Haney, Banks, and Zimbardo, 1973: 89–97). To a number of observers, this simulated prison experiment seemed to confirm a popular stereotype—that, for whatever reason, people who worked in prisons exhibited hostile behavior when they interacted with inmates. Although they presented some data questioning this stereotype of the hostile guard, Jacobs and Retsky (1975: 8) acknowledged the existence of such a prevailing belief when they commented, 'If, as Bittner suggests, 'police work is a tainted occupation,' then prison guard work is utterly polluted.' Perhaps Johnson and Price (1981: 349) provided a reasonable summation of how the literature generally has depicted the correctional officer when they concluded, 'Even under the best circumstances, social science and media versions of the prison officer feature him as the bane of right-thinking convicts, who emerge contaminated by contact with him'" (p. 285). (<i>Not specific in details</i>)</p>
Marquart (2005)	<p>"The relevance of the 'SPE' and its emphasis on situational dynamics was in my opinion right on the mark as a window through which to view the Iraqi prison scandal. I believe that the findings from this experiment will help us to better understand such perennial problems as police corruption, white collar crime, and the prosecution of the innocent. Criminal justice research too often focuses on the pathologies of the subject rather than the pathology of the institutions and settings in which the behavior unfolds" (p. 224). (<i>Multiple details of the SPE throughout the publication</i>)</p> <p>Criticisms include:</p> <p>"We talked about the Stanford prison experiment, which he regarded as artificial because, as he said, 'If you have good leadership the employee's won't abuse the inmates.'...At that point, I really believed the 'Stanford' findings were artificial, the stuff of a sterile lab setting with no connection to real life" (p. 221; other criticism quotes on pp. 218–219). (<i>Artificial</i>)</p>

(Continued)

Appendix C (Continued)

Publication citation	Detailed content of the SPE
<i>Prisons are inhumane/coercive/painful due to <u>social roles</u></i> Bordt (1999)	"Probably the most notable and most controversial simulation study of relevance to criminal justice scholars is the simulated prison study by Haney, Banks, and Zimbardo (1973), in which undergraduate students were assigned to the roles of guard or inmate to replicate the interpersonal dynamics of prison life. Both parties were overwhelmed by their roles: The guards became aggressive and violent; the inmates became hostile but passive. The simulation was ended prematurely to prevent injury to the participants. Despite the debate and controversy surrounding this classic study, scholars still use the simulation method to study a wide variety of subjects related to criminal justice" (p. 374). (<i>Not specific in details</i>)
Roberg (1981)	"For instance, the Stanford Prison Experiment conducted by Zimbardo and his colleagues (1972) screened volunteers (emotionally stable, intelligent, middle-class, male college students) to participate in a simulated prison study in which half were designated as prisoners and the other half as guards. The purpose of the study was to understand more fully what it means psychologically to be a prisoner or a prison guard; the duration of the study was to be two weeks. However, the mock prison was closed down after only six days because of psychotic reactions among participants. As Zimbardo relates: It was no longer apparent to most of the subjects (or to us) where reality ended and their roles began. The majority had indeed become prisoners and guards, no longer able to clearly differentiate between role playing and self. There were dramatic changes in virtually every aspect of their behavior, thinking and feeling. In less than a week the experience of imprisonment undid (temporarily) a lifetime of learning: human values were suspended, self-concepts were challenged and the ugliest, most base, pathological side of human nature surfaced. We were horrified because we saw some boys (guards) treat others as if they were despicable animals, taking pleasure in cruelty, while other boys (prisoners) became servile, dehumanized robots who thought only of escape, of their own individual survival and of their mounting hatred for the guards (1972: 4). The type of harm discussed to this point can be considered as harm resulting from researchers doing something to subjects, harm by <i>commission</i> " (pp. 44–45). (<i>Ethical</i>)

(Continued)

Appendix C (Continued)

Publication citation	Detailed content of the SPE
<i>Prisons are inhumane/coercive/painful due to a combination of both</i> Gonzalez, Dawkins, and Hokanson (1979)	"These results seem to be consistent with previous studies on the effects of imprisonment (Haney, Banks & Zimbardo, 1973) which, although open to methodological questions (cf. Banuazzi & Movahedi, 1975) suggest a 'prison syndrome' of anxiety, passivity, lowered self-esteem, and hopelessness" (p. 108). (<i>Not specific in details</i>)
Haney (2008)	"Whatever its limitations as a literal simulation of an actual prison setting, the venerable Stanford Prison Experiment demonstrated the potentially destructive dynamic that is created whenever near-absolute power is wielded over a group of derogated and vilified others (e.g., Haney, Banks, & Zimbardo, 1973; Haney & Zimbardo, 1977). When the forces that produce that basic dynamic are amplified and intensified, and there are no countervailing pressures introduced into the situation to regulate or moderate their effects, mistreatment is nearly inevitable. Under these conditions, there is great risk that even good, normal people can be led to do bad, sadistic things" (pp. 958–959). (<i>Not specific in details</i>)
O'Donnell (2003)	"A second important psychological contribution to our understanding of the importance of conformity was the experiment conducted by Philip Zimbardo and his colleagues (e.g. Haney <i>et al.</i> 1973). These researchers allocated at random a group of carefully screened normal young men to the roles of prisoner and guard. Neither side was allowed to address the other by name. Strict impersonality was the rule and a long list of regulations was provided to govern daily interactions. The experiment was scheduled to last a fortnight but was halted after one week for fear of lasting psychological damage being caused to participants. The prisoners had become submissive and helpless. The guards were tyrannical and brutal, glorying in every opportunity to humiliate those in their custody. It was clear that the conditions for cruelty were to be found in the social arrangements and not the characteristics of those involved. ⁷ Had the participants been assigned to the opposite roles the overall result would not have been different. When people are given authority over others, and placed in a context where the exercise of that authority is legitimated and backed up with force, the potential for abuse rapidly becomes apparent" (p. 756). Footnote 7: Fromm (1973: 87–96), is more cautious about the extent to which general conclusions can be drawn

(Continued)

Appendix C (Continued)

Publication citation	Detailed content of the SPE
van Gelder, Otte, and Luciano (2014)	<p>from the Zimbardo experiment which he criticizes for artificiality. (<i>Artificial</i>)</p> <p>“Two examples of well-known social psychological experiments that have used such type of (analogue) VE that precede the modern-day computer-generated VEs and that are of particular interest from a criminological perspective, are Stanley Milgram’s (1963) obedience experiment and Philip Zimbardo’s Stanford Prison Experiment (Haney et al., 1973)...The other early well-known example of a hard scenery ‘virtual environment’ is the Stanford Prison Experiment by Zimbardo and colleagues in which a mock prison was set up in the basement of Stanford University. The participating students in this study were randomly assigned to either guard or prisoner roles. Within one week into the experiment, those students that had been assigned the role of prison guard had started to psychologically abuse their ‘prisoners’. The latter often passively accepted the abuse and, at the request of the guards, readily harassed other prisoners who attempted to prevent it (Haney et al., 1973). Besides raising serious ethical concerns, both studies show the strong effects artificial environments can have on the behaviour of people” (p. 2). (<i>Ethical</i>)</p>
Worley and Worley (2013)	<p>“The infamous Stanford Prison Experiment illustrates that controlled environments can be psychologically devastating on many levels. In this study, virtually all of the student-guards displayed some degree of callousness or indifference towards the inmates (Haney, Banks, & Zimbardo, 1973; Zimbardo, 2007). While this experiment has been criticized for its lack of generalizability, it illustrates that artificial environments, such as prison facilities, may be sufficient to prompt power holders to behave in aggressive ways towards others” (pp. 116–117). (<i>Artificial</i>)</p>
No substantive discussion of the SPE	
Farrington (1983)	<p>“The requirement that experiments should be realistic led to the exclusion of all simulation experiments...Other simulations - for example, the famous Stanford prison experiment¹⁶ (e.g., Haney, Banks, & Zimbardo, 1973; see also Banuazizi and Movahedi, 1975) - have also been eliminated. Any experiment carried out on a campus or involving college students as subjects has been excluded” (p. 293). (<i>Artificial</i>)</p> <p>Footnote 16: In the Stanford prison experiment, people were randomly assigned to be ‘guards’ and ‘prisoners’ in a simulated prison environment.</p>