

- Ruse, M. (2003). *Darwin and design: Does evolution have a purpose?* Cambridge, MA: Harvard University Press.
- Simpson, G. G. (1944). *Tempo and mode in evolution*. New York: Columbia University Press.
- Simpson, G. G. (1949). *The meaning of evolution: A study of the history of life and of its significance for man*. New Haven, CT: Yale University Press.

## **Lucifer's Last Laugh: The Devil Is in the Details**

### **The Lucifer Effect: Understanding How Good People Turn Evil**

By Philip Zimbardo. New York: Random House, 2007. 551 pp. Cloth, \$27.95.

Philip Zimbardo has studied human behavior in social situations for decades, and his efforts have given us some of the most memorable glimpses into the surprising possibilities of social behavior. Among these, the Stanford Prison Experiment (SPE) has pride of place. Conducted in the basement of the psychology department over 6 days in 1971, the SPE has become one of the icons of situationism, or the idea that ordinary people can all too easily be induced to visit horrors on one another.

Although the findings of the SPE originally were published outside the peer-reviewed mainstream, they quickly became well known and widely disseminated in lecture halls. Then, in 2004, the Abu Ghraib scandal revealed human behavior that was eerily similar to behavior seen at Stanford, although some of it was much more vicious. Zimbardo testified on behalf of one of the accused prison guards, unsuccessfully pleading for a reduced sentence. In this book he offers his cumulative and evolving insights into the power of bad situations to corrupt ordinary people.

### **Synopsis**

In his opening chapter, Zimbardo introduces his new claim that it is the triadic interplay of people, situations, and systems that accounts for social behavior. The concept of system is meant to address the question of how situations arise in the first place. There are "agents and agencies" (p. 438) that select people, define roles, provide feedback, punish, and reward. In the case of the SPE, the system is Zimbardo himself and the institution of Stanford University that backs him up. In the case of Abu Ghraib, the system is the military hierarchy and the civilian leadership in the executive branch of the U.S. government.

In chapters 2 to 9, Zimbardo recounts the SPE in fine detail, offering some material from his notes that has never been published. The broad sequence of events is familiar. Healthy young men, who agreed to participate in a simulated prison study, are randomly assigned to the role of prisoner or guard. On Sunday, the prisoners are rounded up in their homes and transported to the basement of Jordan Hall, which has been outfitted to resemble a prison, complete with a "Hole" for solitary confinement. Guards report to duty in shifts. Both groups are deindividuated by a variety of means (e.g., hoods, smocks, uniforms, shades). The guards quickly assert their authority, and the prisoners are quick to rebel. Riots continue on Tuesday, prisoners break down under unbearable stress, and Zimbardo tries

(unsuccessfully) to plant an informant among the prisoners. Wednesday's main event is a simulated parole hearing, which depends heavily on actors who have not been randomly assigned to their roles and who succeed in keeping the study from collapsing—for the time being. On Thursday, Christina Maslach stares down Zimbardo and gets him to close the study, although not right away. Whether she succeeds by ultimatum or by stirring his conscience is not revealed. On Friday, Zimbardo debriefs and dismisses the remaining participants.

In chapters 10 and 11, Zimbardo considers the fallout of the SPE, acknowledges flaws in the experimental design, and raises the crucial question of how responsible perpetrators are if they have fallen under the spell of a corrosive situation. In chapters 12 and 13, this work of reflection continues with a review of classic studies on conformity, obedience, and bystander apathy. In chapter 14, he turns his lens on Abu Ghraib and the horrors that will be forever associated with its name. The presentation is detailed, careful, and even scientifically detached. Throughout, Zimbardo is mindful of his emerging triadic theory of person, situation, and system. In chapter 14, he places the blame squarely on the system and the individuals who represent it. They, he argues, should be tried; not all guilt should be placed on the people who turned abusive on location. Chapter 16 sounds a note of hope by identifying heroic acts and heroic people and by presenting a 10-step program designed for the rest of us to realize our hero potential.

### Analysis

Although this book is magnificent and timely, it cannot be exempt from critical analysis. Consider the question, first asked by Gordon Bower on the third day of the study, of whether the SPE actually was an experiment. Zimbardo still believes that he could have settled the issue by saying that the independent variable “was the allocation of pretested volunteer subjects to the roles of prisoner and guard, which of course had been randomly assigned” (p. 98). Yet doubts surface a little later when he admits that “we were not testing specific hypotheses” (p. 195), that “the causal directions are uncertain” (p. 196), and that “the data analysis is confounded” (p. 196).

When judged against conventional standards, the SPE does not qualify as an experiment. Consider the issues of contamination, interaction, and replicability. The main source of contamination is Zimbardo himself because he accepts the dual roles of the presumably detached experimenter and the highly involved prison superintendent. Increasingly, the superintendent gets the better of the scientist, which Zimbardo now sees as a point *in favor* of the study's validity. For example, when he is dissatisfied with the performance of two of the guards, he asks the warden “to make them more assertive” (p. 81). The warden, incidentally, is played by one of Zimbardo's own undergraduate students. The most fascinating contamination is indirect. It features Carlo Prescott, an ex-con, who has taught a summer course on prison life with Zimbardo. Now Carlo returns as a consultant and takes to the job with brio. There is no pretense about “random selection.” Carlo upbraids Prisoner 8612, who wants to leave (“Stop it right there, punk!” p. 68), and Zimbardo thanks Carlo “for providing this reality check” (p. 69).

The issue of interaction refers to the dynamic exchanges between the guards, the prisoners, and all participants regardless of role. Any behavior becomes inextricably confounded as part of a personal force, a situational force, and a systemic force. By way of comparison, both the Asch (1956) conformity and the Milgram (1963) obedience paradigm avoided this problem by exposing individual participants, one at a time, to a standardized situation.

Since Wundt (1874) noted its significance, the issue of replicability runs deep in the experimental ethos. Zimbardo cites one successful and one unsuccessful replication of the SPE. Although the successful replication is of somewhat obscure origin, it is hailed with praise. The unsuccessful replication is a serious effort performed under the auspices of the British Broadcasting Company and published in the *British Journal of Social Psychology* (Reicher & Haslam, 2006). Yet this study is summarily dismissed as a “pseudoexperiment” (p. 252). The issue of replicability deserves better than this. Repeated emphasis on the participants being randomly assigned to their roles does not answer the question. Because of the situation’s interactive nature, it is proper to regard the entire study as the unit of analysis. With  $N = 1$ , the SPE (and Abu Ghraib) are single-case samples drawn from large populations of hypothetical studies and real prisons, respectively. How well these cases generalize is unknown. At minimum, it is clear that neither case is random. The extremity of the SPE results are due in part to Zimbardo’s contaminating influence; Abu Ghraib is not the average prison for all the reasons so thoroughly documented in this book.

Now consider the fact that no prisoner leaves the SPE by declaring to quit. Several are released because of their severe stress reactions, but no one simply walks away. To Zimbardo, this is proof that “the prisoners themselves became their own guards” (p. 222). Yet it is clear within a few days that “the majority of the prisoners are willing to leave without pay to get out of this place” (p. 141). Some of them are adamant, screaming and cursing. The prisoners remain confined because they do not find the right words that would remove them from the situation. “Why does he not simply say, ‘I quit this experiment. You must give me my clothes and belongings, and I am out of here!’” Zimbardo wonders (p. 147).

The failure to utter the magic words calls for an explanation. Here is one possibility. Time and again, Zimbardo and his proxy Carlo obscure the prisoners’ options. During the parole hearing, Carlo mocks a prisoner, “So wise guy, you think this is just an experiment?” (p. 139), and Zimbardo coaches him on how to change the deal by asking prisoners, “Would you be willing to forfeit all the money you have earned as a prisoner if we were to parole you?” (p. 140). Even when release finally comes, Zimbardo himself is unable to drop the charade. About a distressed participant he says that “it would be best if he were paroled at this time” (p. 161). The creativity with which Zimbardo and his team hide the option to quit from their participants evolves over time and as prisoners become more desperate. In the end, Zimbardo blames the victims, concluding that they “added to their own dehumanization” (p. 224). In contrast, Milgram determined *before* his experiment which prompts were to be used to induce obedience and at which points the participants’ demands would have to be honored.

## Narratives

The story of the SPE unfolds in two narratives. The original narrative focuses on the competition between personal and situational forces to control behavior, and the revised narrative includes the concept of personal "transformation" from good to evil and the concept of the "system" as a kind of umbrella context that provides structure to both personal and situational forces. Let us consider the original narrative first. Zimbardo announces that "a large body of evidence in social psychology supports the concept that situational power triumphs over individual power" (p. x). Not really. The meta-analysis Zimbardo cites shows that on average the effect size of social influence is  $r = .13$ , which is less than the effect on personality variables,  $r = .21$  (Richard, Bond, & Stokes-Zoota, 2003).

Where does the SPE fit in? Clearly, the guards and the prisoners behave differently, although it is not clear on what dimensions their behavior should best be scored. Suppose an aggression index is derived, and the guards score higher than the prisoners. A simple measure of effect size is the difference between the two means. What might be evidence for the idea that the situation does *not* triumph over the person? Would the difference have to be zero? This possibility of the null hypothesis being zero seems remote because guards might be expected to be somewhat more aggressive if they take their roles seriously.

But Zimbardo wants more. "In the face of strong, common situational forces," he claims, "individual differences shrink and are compressed" (p. 447). The power of the situation should be revealed not only in large differences between groups but also in small differences between group members. These are the ingredients of a standardized effect size, that is, the ratio of group differences and individual differences. The numerator does indeed seem large in the SPE, but Zimbardo's narrative contains interesting news: The denominator is large as well. Over and over, Zimbardo observes dramatic distinctions between mean and good guards and between rebellious and submissive prisoners. To illustrate, consider his almost psychodynamic assessment of Prisoner 3401: He "is more complex than he appears initially. He reveals an interesting mix of traits" (p. 136).

A different picture emerges. Situations do not "overpower" people but rather reveal latent possibilities. Even if everyone becomes more aggressive in an aggression-inducing environment, individuals differ in the degree to which they respond. Hence, person and situation do not compete for a fixed amount of behavioral variance. Behavior is not a zero-sum game. Instead, situations can bring out individual differences that would otherwise never be seen. Zimbardo knows—but often ignores—the point that "people with certain traits that are latent and may never be expressed except when particular situations activate these violent tendencies" (p. 287). His rich description of how people respond to new challenges does not support strict situationism, and it does not support strict trait theories. Instead, it is consistent with contextual theories of social behavior that focus on individual differences in the conditional probabilities with which people respond to certain kinds of social stimulation (Wright & Mischel, 1987).

The revised narrative seeks to go beyond simple notions of person and situation. The power of the situation is such that it can transform a good person to

become bad—at least for a limited time. Guard Hellmann (“John Wayne”), for example, turns from gentle to brutish and back to being a swell guy (although Zimbardo still resents Hellmann’s unwillingness to fully accept his interpretation of the SPE). Zimbardo himself changes from socially conscious scientist to complicit abuser and system agent and back to conscientious scientist. There is no theoretical development of the notion of transformation and no reference to relevant prior work (e.g., Asch, 1952; Darley, 1992). This is unfortunate because the concept is left unprotected from two unappealing possibilities. One is that the term *transformation* is more descriptive than substantive; the other is that the concept eliminates the presumably critical distinction between person and situation. If a person genuinely is one with his or her behavior, no matter how it is caused, both situational and personal attributions are always correct (and it is impossible to be “guilty of the fundamental attribution error,” p. 271).

Throughout the narrative, Zimbardo struggles with the moral implications of his findings. How much blame is to be assigned? If the situation were fully responsible for all behavior, each individual should be beyond freedom and dignity. Zimbardo rejects this option as morally corrupt “excusiology.” “Understanding the psychological foundations of abusive behavior does not excuse the perpetrators” (p. 401). Yet he believes that situational forces should be recognized to mitigate blame, and he testifies on behalf of one of the Abu Ghraib guards. The judge ignores his testimony, perhaps because there is no formula that says how much blame should be subtracted. Perhaps an even bigger question is what to do with a theory that demands to be applied to everyone equally. In Abu Ghraib, Corporal Graner becomes especially notorious. He “far exceeded his role in abusing prisoners both physically and psychologically” (p. 359). He is as viciously creative as guard Hellmann is—albeit on a smaller scale—in the SPE. Perhaps Zimbardo does not testify on Graner’s behalf for fear of damaging his reputation. Alternatively, there is the possibility that Graner does not fit the narrative. “All external constraints on Graner’s antisocial behavior were gone with the wind” (p. 360). Graner already was a high-risk personality. The situation does not make him vicious; it releases his viciousness.

The introduction of the concept of the system deepens these questions. If the system is a conglomeration of situations, people, hierarchies, and roles, it lies beyond experimental study and the powers of analysis of variance. Questions of responsibility become less tractable when they matter the most. My sympathies lie with Zimbardo when he puts the system on trial and, with it, its prominent representatives. It would be gratifying to see ex-secretary Rumsfeld atone for his role in making Abu Ghraib what it was. Who wants to hear how the political situation in the government steered his actions? Nevertheless, the theory that situations matter most must assume that Rumsfeld was caught up in a web of forces outside himself, just like the rest of us.

### **The last temptation of Zimbardo**

As the book draws toward its conclusion, people emerge ever more strongly from situations. Zimbardo concludes with a riff on heroism, saying that it exists, giving poignant examples, proposing a 12-type taxonomy, and ultimately decid-

ing that heroism is banal, just like evil. Perhaps he is right. Consider his favorite example of heroism on the stage of science. On the fifth day of the SPE, Christina Maslach tells her mentor and lover, "What you are doing to these boys is a terrible thing" (p. 168). The next day, Zimbardo ends the experiment. He intimates that there has been an intense argument, that both he and Maslach have much at stake, that Maslach is right, and that thanks to her intervention a grievous situation is rectified. Heroism? Perhaps. But there is another explanation, which requires only information about people's preferences and the ability to think strategically (Brahms, 1994).

Maslach has a choice between confrontation (c) and doing nothing (d). Zimbardo had a choice between relenting (r) and staying the course (s). For Zimbardo, a plausible preference ranking is ds, cr, cs, and dr, in descending order. In other words, he would most like to continue the experiment with Maslach's consent. If she complains, however, he would rather relent than persist because his most significant personal relationship takes priority over a single experiment. He would least like to relent while there is not even a complaint. Maslach, on the other hand, would most prefer to remain quiet and to see Zimbardo quit on his own (dr). She would least like to complain and be turned down (cs). The key assumption here is that she prefers cr (complain and succeed) over ds (be quiet and become complicit). Maslach solves the problem because she can act first and thus predict how Zimbardo will react.

Assuming—correctly, as it turns out—that Zimbardo would rather relent than resist if she complains, she complains. The result is that both obtain their second highest preference. Not a bad outcome for a little mind reading. Perhaps Maslach is cleverer still. She makes her move not at Wednesday's parole meeting but in private on Thursday. She offers Phil an opportunity to save face, and she waits long enough for the experiment to yield enough data to make it an enduring part of psychology's heritage. History has proven Professor Maslach right.

Joachim I. Krueger

*Department of Psychology*

*Brown University*

*Box 1853*

*Providence, RI 02912*

E-mail: joachim@brown.edu

## References

- Asch, S. E. (1952). The transformation of man in society. In S. E. Asch (Ed.), *Social psychology* (pp. 117–138). Upper Saddle River, NJ: Prentice Hall.
- Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, *70* (whole number 416).
- Brahms, S. J. (1994). *Theory of moves*. New York: Cambridge University Press.
- Darley, J. M. (1992). Social organization and the production of evil. *Psychological Inquiry*, *3*, 199–218.
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*, 371–378.
- Reicher, S., & Haslam, S. A. (2006). Rethinking the psychology of tyranny: The BBC prison study. *British Journal of Social Psychology*, *45*, 1–40.

- Richard, F. D., Bond, D. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7, 331–363.
- Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology*, 53, 1159–1177.
- Wundt, W. M. (1874). *Grundzüge der physiologischen Psychologie* [Elements of physiological psychology]. Leipzig, Germany: Engelmann.

## Mind Dynamics

### **A Mind for Structure: Exploring the Roots of Intelligent Systems**

By Ken Richardson. Boca Raton, FL: BrownWalker Press, 2006. 344 pp. Paper, \$25.95.

However much we may despair of it sometimes, human intelligence is quite striking. Yet despite many years of research in cognitive sciences, precious little is known about its operations: How do we end up perceiving a coherent world? How are memories structured by concepts? How exactly does thinking affect behavior (recall the notorious frame problem)? Likewise, the questions of how the brain realizes cognition and what prompted the evolution and the increasing complexity of cognitive systems are still awaiting definitive answers. On the practical side, artificial intelligence is a long way away from designing a machine that could take over all my household chores, leaving me more time to write this book review (better still, write it for me).

In his very stimulating book Ken Richardson diagnoses the mainstream approach to cognition with a lack of unifying perspective—which, in his view, precludes such questions receiving satisfactory answers—and offers a remedy for this deficiency. Wonderfully broad in scope, concerning the very foundations of our thinking about intelligence in the evolutionary context, Richardson's book is an admirable attempt to supply the missing big picture and an important contribution to the growing body of literature developing the dynamic systems approach to cognition. Yet I have reservations, for although the diagnosis is largely correct, the cure may well kill the patient.

Richardson fears that psychology has focussed too much on “closed” behaviors—“well-adapted responses to well-defined stimuli” (p. 147)—thereby ignoring the dynamic quality of everyday thought. The world inhabited by living systems is forever changing, and the sensory inputs they have to rely on are fleeting and fragmentary. However, many theories of organism–environment interactions are content to model them as simple, linear, cue–response or if–then functions. Today this approach is especially pronounced in evolutionary psychology, which tends to view organisms' abilities as adaptations to durable steady-state environments. But as Richardson correctly points out, there are environmental changes occurring within generations, days, and minutes. If we are to understand how living systems can be adaptable—rather than well adapted—in such conditions, the importance of smaller time scales should not be neglected (which holds true for the whole spectrum of phenomena from metabolic regulation to belief revi-