

## SHORT TITLE: IDENTITY LEADERSHIP IN THE STANFORD PRISON

Rethinking the nature of cruelty:

The role of identity leadership in the Stanford Prison Experiment

S. Alexander Haslam<sup>1</sup>, Stephen D. Reicher<sup>2</sup>, & Jay J. Van Bavel<sup>3,4</sup>

<sup>1</sup> School of Psychology, The University of Queensland

<sup>2</sup> Department of Psychology and Neuroscience, University of St Andrews

<sup>3</sup> Department of Psychology, New York University

<sup>4</sup> Center for Neural Science, New York University

in press, *American Psychologist*

*Correspondence:* Alex Haslam, School of Psychology, University of Queensland, St. Lucia  
QLD 4072, Australia; e-mail: [a.haslam@uq.edu.au](mailto:a.haslam@uq.edu.au); Tel.: +61 (0)7 3346 9157.

*Statement of contributions:* SAH conducted initial analysis and interview, and wrote first draft of manuscript. SDR elaborated analysis and extensively redrafted manuscript. JVB provided feedback on analysis and also extensively redrafted manuscript.

*Acknowledgement:* We are grateful to Thibault Le Texier for alerting us to the existence of the meeting between Warden Jaffe and Guard Mark in the Stanford archive. Thanks also to Jolanda Jetten, Mark McDermott, Dominic Packer and Eric Vanman for comments on early drafts of this manuscript.

### **Abstract**

The Stanford Prison Experiment (SPE) is one of the most famous studies in the history of psychology. For nearly a half century it has been understood to show that assigning people to a toxic role will, on its own, unlock the human capacity to treat others with cruelty. In contrast, principles of identity leadership argue that roles are unlikely to elicit cruelty unless leaders encourage potential perpetrators to identify with what is presented as a noble ingroup cause and to believe their actions are necessary for the advancement of that cause. Although identity leadership has been implicated in behaviour ranging from electoral success to obedience to authority, researchers have hitherto had limited capacity to establish whether role conformity or identity leadership provides a better account of the cruelty observed in the SPE. Through examination of material in the SPE archive, we present comprehensive evidence that, rather than Guards conforming to role of their own accord, Experimenters directly encouraged them to adopt roles and act tough in a manner consistent with tenets of identity leadership. Implications for the analysis of conformity and cruelty as well as for interpretation of the SPE are discussed.

192 words (max 200)

Keywords: leadership, followership, social identification, identity entrepreneurship, Stanford Prison Experiment

### Rethinking the nature of cruelty:

#### The role of identity leadership in the Stanford Prison Experiment

In August 1971, Professor Philip Zimbardo and his research team recruited 24 young men to participate in a study on prison life. The volunteer participants were randomly assigned to be either Prisoners or Guards in a mock prison that had been constructed in the basement of the Stanford psychology department. After just a few days Guards began to repress the Prisoners and their cruelty escalated, up to the point where, after six days, the study was prematurely terminated. What happened during that week — and what it means for our understanding of human behavior — has been the focus of scientific and public debate for the past half century.

The Stanford Prison Experiment (SPE; Haney, Banks & Zimbardo, 1973a, 1973b) is one of the most famous psychology studies of all time. It is covered in most introductory psychology and social psychology textbooks and courses (Bartels, Milovich, & Moussier, 2016; Griggs, 2014; Griggs & Whitehead, 2014) and is a standard point of reference for media stories on tyranny and repression. The study has also been presented to government officials and in court cases to help understand events ranging from prison riots to the abuse of detainees during the Iraq War. On top of this, it has provided the material for an extremely influential website (Zimbardo, 1999), a best-selling book on evil (Zimbardo, 2007), and several feature films that have grossed over 12 million dollars at the Box Office (Alvarez, 2015, Scheuring, 2010; see also Hirschbiegel, 2001). As a result, the study has had a major role to play in shaping how millions of people think both about the nature of human cruelty and about the power of the situation to encourage toxic behavior.

The SPE became famous for the striking evidence it appeared to provide that normal, well-adjusted young men could come to harm innocent civilians simply as a consequence of

having been randomly assigned the role of Guard (rather than Prisoner). Zimbardo and his colleagues have expressed this provocative idea in a number of slightly different ways to different audiences. In their first publication, they wrote that Guards made their own decisions about how to run the prison and that their aggression was “emitted simply as a ‘natural’ consequence of being in the uniform of a ‘Guard’ and asserting the power inherent in that role” (Haney et al., 1973a, p.12). Later on, in a summary chapter, Zimbardo argued that “each subject’s prior societal learning of the meaning of prisons and the behavioural scripts associated with the oppositional roles of prisoner and Guard was the sole source of guidance” (2004, p.39). More recently still, in their textbook, Zimbardo, Johnson and McCann claim that “the mere fact of putting on uniforms was sufficient to transform [participants] into passive prisoners and aggressive guards” (2012). This point has now become canonical in the field in seeming to show that toxic situations are sufficient to stimulate toxic behaviour.

Whatever the precise wording, the core of Zimbardo and colleagues’ argument is that what the Guards did they did “without training from [the Experimenters] in how to be Guards” (Zimbardo, Banks Haney, & Jaffe, 1973, p.40; see also Zimbardo, Banks Haney, & Jaffe, 1974, pp.65-66). In the researchers’ words “our results go one step further [than Milgram’s] in removing the immediate presence of the dominant experimenter-authority figure, [and] giving the subjects-as-guards a freer range of behavioural alternatives” (Haney et al., 1973b, p.90). Indeed, it is the fact that the study points to the power of social roles *on their own* to engender cruelty that has helped make the SPE so influential in the scientific literature. The implication is that good people will generally turn bad if they happen to be put in a powerful position in a toxic place.

This is the lesson which Zimbardo himself has explicitly drawn from the SPE, naming it *The Lucifer Effect* — the title of his best-selling 2007 book. It is a lesson that has been

widely propagated, not only in psychology but also to students and researchers in a wide array of other disciplines (e.g., history and criminology; Browning, 1992; Jacoby, Severance, & Bruce, 2004). Indeed, alongside Milgram's (1963, 1974) classic research on Obedience to Authority (OtA), the SPE may have had more impact on the public consciousness than any other piece of psychological research (Banyard, 2007; Blum, 2018; Konnikova, 2015).

### **Problems with the role account**

Over the five decades since it was conducted, the SPE has been subjected to increasing scrutiny and criticism (Griggs 2014; Turner, 2006). Although textbook and popular accounts of the study generally make little or no mention of limitations in either the study's design or Zimbardo and colleagues' interpretation of its findings (Bartels et al., 2016; Carnahan & McFarland, 2007; Griggs 2014; Griggs & Whitehead, 2014), there are two principal lines of criticism in the scientific literature.

First, from the sources that are available — notably Zimbardo's, 1992, film of the study, *Quiet Rage*, and his 2007 book (though, as Griggs, 2014, points out, these fail to include many important details of the study) — it is clear that many participants did not conform to role. Many Prisoners continued to resist authority until the end of the study (Le Texier, 2018; Reicher & Haslam, 2006). Likewise, many Guards refused to assert their authority. While a few were brutal and cruel, others were not, and some even sided with the Prisoners (Zimbardo, 1992). Such variability requires a more nuanced interpretation of the SPE than is typically provided in media and textbook reports (Haslam & Reicher, 2007). For many participants, then, it appears that being given a role was *not* sufficient to elicit the type of behaviour for which the study has become well known.

Second, to the extent that some Guards and some Prisoners did conform, there are suggestions that, rather than doing so of their own accord, the actions of the Experimenters may have played an important part in producing their behaviour. Early on, Banuazizi and

Movahedi (1975; see also Gray, 2013) noted that the behaviour of participants — especially the Guards — could be explained in terms a range of salient demand characteristics that encouraged them to behave in particular ways. The importance of the Experimenters' actions was subsequently demonstrated in a study by Lovibond, Milthiran and Adams (1979) in which Guards were instructed to adopt either authoritarian, democratic, or participatory approaches to their role. Here only the authoritarian instructions produced toxic behaviour similar to that seen in the SPE. Moreover, in a more recent prison study by Reicher and Haslam (2006) Guards were given *no* direct instructions about how to behave (other than prohibiting violence, as in the SPE) and subsequently showed no inclination to treat Prisoners cruelly.

This body of research led Haslam and Reicher (2007, 2012b) to propose that the Experimenters' *leadership* may have been critical to the emergence of Guard cruelty in the SPE. There is some evidence to support this in the materials Zimbardo has previously made available, notably the briefing which he gave to the Guards during the 'orientation day' before the Prisoners arrived. This is included in the 1992 film of the SPE. During the briefing Zimbardo announced: *You can create in the prisoners feelings of boredom, a sense of fear to some degree, you can create a notion of arbitrariness that their life is totally controlled by us, by the system, you, me – and they'll have no privacy. They'll have no freedom of action, they can do nothing, say nothing that we don't permit. We're going to take away their individuality in various ways. In general what all this leads to is a sense of powerlessness.*

Two features of this passage are notable. The first is Zimbardo's use of the terms 'we' 'us' 'they' and 'their'. In effect, he positions himself as part of the Guard's ingroup, standing with them against the Prisoners (for evidence of the power of such rhetoric, see Donnellon, 1996; Steffens & Haslam, 2013). As he said in an interview in 1972 "*I trained the guards*

*and said 'Look, this is a serious situation, we're in this together and it's you and I and us and the people of California versus the prisoners'"* (cited in Le Texier, 2018). In short, he created a sense of multiple overlapping identities which defined the Experimenters and Guards together (as people interested in science, as reformers, and as good Americans) as a united force against the Prisoners. The second feature is the explicit guidance that Guards are given about the general manner in which they should act — being instructed to create a sense of fear and boredom, to remove the Prisoners privacy and freedom, and so on.

In short, then, this briefing can be seen as an act of leadership on the part of Zimbardo. And although this point has never been acknowledged either by Zimbardo and his colleagues (e.g., Zimbardo, 1999, 2004, 2006) or in textbooks that describe the SPE (see Griggs, 2014), it raises the question of whether there were further acts of leadership during the study, something suggested by a number of the people who were directly involved in the study (Eshelman, cited in Ronson, 2015; Mark, 2007) and also by other scholars (Konnikova, 2015; Le Texier, 2018). It also raises the question of how important the leadership of the Experimenters might have been in producing the cruelty of the Guards.

Yet, interesting as these questions are, neither could ever be answered as long as the only clear evidence of leadership was the single passage cited above. Certainly, this provides a thin basis on which to argue that the Experimenters' interventions were crucial to producing Guard cruelty, let alone to develop an alternative theoretical account of both parties' behaviour. We may have suspected that leadership was important in producing toxic behaviour, but for substantive data to support this analysis we have previously had to look elsewhere for evidence.

### **An alternative account: Identity leadership and engaged followership**

Elaborating on the foregoing observations, Haslam and Reicher (2007a, 2012b) suggest that what may have been going on in the SPE is a specific process referred to as

*identity leadership*. This analysis is grounded in hypotheses derived from social identity theorizing (after Tajfel & Turner, 1979; Turner, Hogg, Oakes, Reicher & Wetherell, 1987)

This analysis emerges from a tradition of social identity research (Tajfel & Turner, 1979) which proposes that human beings have the capacity to define themselves not only as individuals (e.g., “Abigail”, “Barbara”, “Colin”) but as group members (e.g., “Americans”, “Buddhists”, “Cubs fans”) and that the way we define ourselves is bound up with social context. In particular, following Bruner (1958), we adopt a given group membership to the extent that it allows us to make sense of our role in the situation that confronts us (i.e., so that it is fitting; Oakes, Haslam & Turner, 1994). For this reason, Reicher & Haslam (1996) argue that people do not automatically take on roles (in ways that Zimbardo et al., 1973b suggest), but do so only to the extent that these roles make sense — or, rather, have been made to make sense — in the context of a salient social identity.

Once we do self-categorize ourselves and others in collective terms (Turner, Hogg, Oakes, Reicher & Wetherell, 1987; Turner, Oakes, Haslam & McGarty, 1994) we perceive and evaluate ourselves, our world, and our goals in terms of the social group. Thus it is the esteem in which the group is held that defines our self-esteem (Tajfel & Turner, 1979) and indeed when the group is defined positively and is a source of pride, we are more likely to identify with it (Tyler & Blader, 2003). Equally, it is the interest of the group which defines our self-interest such that we may sacrifice our personal gain or even our personal existence for the collective cause (Swann et al., 2014).

Perhaps most importantly, it is the understanding of the group (its collective beliefs, values and norms) which guide us in pursuing this interest (Reicher, Haslam & Spears, 2012; Turner & Reynolds, 2011). The significance of this is that it provides the basis for a model of social influence termed *referent informational influence* (Turner, 1982, 1991). According to this, when group identity becomes salient, individuals seek to ascertain and to conform to those understandings which define what it means to be a member of the relevant group. This then provides distinctive answers to the



three key questions for any model of influence. *Who is influential?* Those in a position to know the group beliefs, values and norms by virtue of their being representative of the group (formally, after Rosch, 1978, those who are prototypical of the group). *Who is influenced?* Those who share a common salient social identity. *What achieves influence?* Messages seen to be consonant with group understandings.

This model of influence is also an implicit theory of leadership. For effective leaders can be understood as those who are able (a) to represent themselves as prototypical of the group, (b) to make potential followers represent themselves as members of a common social group, and (c) to represent their proposals as the actualisation of group understandings. Put slightly differently, leaders are able to influence followers largely as a function of their capacity to represent create and advance a sense of social identity that they share with those followers (a sense of ‘us-ness’). As we have put it elsewhere, effective leadership is therefore a process of *social identity management* (Haslam, Reicher & Platow, 2011).

Over the last 20 years this model of leadership has been made explicit, elaborated and tested by a range of authors (Hogg, 2001; Reicher, Haslam & Hopkins, 2005; Steffens et al., 2014; van Dick et al., 2018; Turner & Haslam, 2001). In our own work, we provide an analysis that centres on four core insights (Haslam et al., 2011; Reicher et al., 2005). The first is that *effective leaders (those who influence and harness the energies of followers) need to be seen to be representative of a shared ingroup*. However, it is not enough just to be prototypical of the group. Many leaders who are ingroup members and who understand group norms and values may nonetheless act for their own ends rather than the group’s. Indeed, in contemporary politics this belief is central to the rise of populist ‘anti-politics’ (Eatwell & Goodwin, 2018). Accordingly, second, *effective leaders need to be seen to be advancing ingroup interests*. But what the group and its interests are is never just given. For this reason, third, *leaders, need to be entrepreneurs of identity* (Reicher & Hopkins, 2000) who actively construe followers as part of a common group (Reicher et al., 2005) and actively construe

themselves as both prototypical and working for the group (Haslam et al., 2011). Additionally, they need to present their own messages and projects as an actualization of group beliefs values and norms. This brings us to the fourth and final point: that because leadership is not just about how leaders act but also about their capacity to shape the actions of followers, *effective leaders need to be impresarios of identity* who promote policies and practices that help translate group values into material lived reality.

Support for these four propositions is provided by a large body of research in social, organizational and political psychology (Haslam et al., 2011; Steffens et al., 2014). For example, recent meta-analysis has found that leaders garner more support, exert more influence, and are seen as more charismatic the more prototypical they are of the group they lead (Barreto & Hogg, 2018, 35 studies,  $r=.49$ ; Steffens et al., 2018, 70 studies,  $r=.39$ ). Moreover, a recent study conducted across 21 countries shows that the above four aspects of identity leadership predict leader success better than many other theoretical constructs (notably transformational leadership and authentic leadership; van Dick et al., 2018). Principles of identity leadership thus appear to characterize and underpin effective leadership across a wide range of leaders, groups and social contexts.

But can principles of identity leadership help us better understand the culture of cruelty that emerged in the SPE? Haslam and Reicher (2007, 2012) argued that they might but, as noted above, the lack of relevant detail made this question hard to answer definitively. To test their ideas Haslam and Reicher therefore turned to other classic studies in which ordinary people were led to be cruel to others, notably Milgram's (1963, 1974) classic studies of 'Obedience to Authority' (OtA). On inspection, there are a number of aspects of these studies that were consistent with an identity leadership account. In the first instance, archival material provided clear evidence of the considerable lengths that Milgram went to in order to encourage participants to identify with his research enterprise and to see its goals — ostensibly to improve scientific understanding of the effects of punishment on learning

— as positive, worthy and a source of pride (Gibson, 2013; Haslam, Reicher, Millard & MacDonald, 2015; Russell, 2011). This identity entrepreneurship, in turn, appears to have underpinned participants' willingness to harm the Learner within the OtA paradigm. In particular, correlational and experimental research shows that in both the original Milgram studies and in more recent conceptual replications, participants prove willing to follow destructive instructions only to the extent that they identify with the (male) Experimenter and the scientific community that he represents (Haslam, Reicher & Birney, 2014; Reicher et al. 2012). Indeed, Haslam and Reicher argue — and have shown experimentally (e.g., Haslam et al., 2014, 2015) — that the power of situational cues within the paradigm derives precisely from their capacity to encourage identification with the Experimenter and his scientific agenda (e.g., so that participants identify more with the Experimenter when he is in the same room or from a prestigious institution, or when he states that the Experiment requires that they continue, rather than ordering to them to go on).

As well as validating an identity leadership analysis of toxic behaviour, these studies also point to some distinctive aspects of leadership as it applies to the cultivation of cruelty. The key issue is that, as we now know from the SPE and the BBC Prison Study (Reicher & Haslam, 2006), few people spontaneously identify with roles that require them to be cruel to others. Hence, as Haslam and Reicher (2006, 2012, 2017) have argued, those who advocate harming others need to work hard to do the identity entrepreneurship necessary (a) to get others to identify with them and their group cause, (b) to construe the group's goals as noble, virtuous and a source of pride, and (c) to show how harmful acts are absolutely necessary for the achievement of those goals.

In the case of the OtA studies, archival research makes it clear that Milgram indeed worked hard to do precisely this (e.g., Haslam et al., 2014; Russell, 2001). Yet in the case of the SPE, Zimbardo and colleagues' public statements argued directly against the identity

leadership analysis, and, as we have seen, in the absence of strong evidence to the contrary their role account has continued to serve as the dominant explanation in textbooks, movies, and the media.

### **New evidence from the SPE**

In 2018, the situation described above regarding the lack of evidence from SPE changed. The Zimbardo archive at Stanford University became available online and was subjected to forensic examination by Le Texier (2018). This has allowed many of the previous gaps in knowledge about what happened in the study to be filled in. Of particular relevance to the present argument, the archive presents a much richer picture of how the Experimenters directly engaged with participants in the study. As a consequence, we are now in a position rigorously to examine whether an identity leadership analysis can explain what happened in the SPE.

There are at least four new pieces of evidence provided by Le Texier (2018) which point to the importance of the Experimenters' leadership in the SPE in producing the cruelty of the Guards as well as other related outcomes:

*First*, following Zimbardo's orientation day briefing, the Guards had a much longer session with the 'Chief Warden' David Jaffe. As Le Texier (2018, p.74) notes, this lasted 5-6 hours but only a portion was recorded. In a key passage, Warden Jaffe told them:

*I guess when you come into prison you have to sit and wait and wait and wait and so we'll have that, the waiting with the uniforms, stocking cap and all this business. The idea is to order them around, you don't request them to do anything, you tell them, and I think that all of you could probably put on some sort of a firm policemenlike voice. I don't know but we'll probably try to role play that a little bit this afternoon. ... If we don't do anything, then we're just playing a game, and nothing happens — we don't learn anything. So your inventiveness and whatever else — stick-to-it-ness as far as*

*devising things, working out a schedule, and then rewards and punishments and all of this kind of thing are very important part of making the thing run. To a large extent we will either create this prison environment or not on the basis of what you do during the shift* (cited in Le Texier, 2018, pp. 80-81)

The important thing to note here is that this briefing does not merely reiterate Zimbardo's general guidelines to the Guards, but provides very detailed instructions about to how to act. When one of the Guards expressed concerns that "*some of these sadisms*" might create "*some kind of bad problems*" Jaffe gives further practical advice on what should be done if the damage to the prisoners requires medical attention: "*We have access to the Cowell Health Center Facilities ... and should we have need for that sort of thing, we can get them over [t]here very fast*". This does two things: it implies that the level of abuse might rise to a level that requires medical attention, while simultaneously signalling that this should not be seen as a problem. Moreover, in line with principles of identity leadership, Jaffe stresses that these actions are necessary for the research team to achieve its important goals, and with his repeated references to 'we', he makes it clear that he sees the Guards as part of that team.

*Second*, it is clear that such detailed instructions were not limited to orientation sessions but continued to the end of the study. One of the Guards, Mike Varn, noted in his post-study evaluation that "*the Warden or Prof. Zimbardo specifically directed me (us) to act a certain way (ex. hard attitude Wednesday following Tuesday leniency)*" (cited in Le Texier, in press). In a similar vein, when writing to Zimbardo after the study had ended, another Guard, Andre Cerovina described how Jaffe told them to do such things as stand outside the cells in the middle of the night and blow their whistles. He concludes: "*I thought that the Warden was very creative, not just then but through the experiment, he gave us very good sado-creative ideas.*" (cited in Le Texier, in press). Another Guard, Dave Eshelman, gleaned the Experimenters' expectations not only from what he was told to do, but also from what he

was not discouraged from doing: *“We would ramp up the general harassment, just sort of crank it up a bit. Nobody was telling me I shouldn't be doing this. The professor is the authority here, you know. He's the prison warden. He's not stopping me.”* (cited in Le Texier, 2018, p.81). In this way, multiple guards independently reported direct, concrete instruction from the Experimenters about how to torment the prisoners.

*Third*, the Guards were aware of what they were expected, or indeed required, to do to ensure the goals of the Experimenters were achieved. One, Terry Barnett, wrote: *“I consciously felt that for the experiment to be at all useful ‘guards’ had to act something like guards. [...] I felt that the experiment was important and my being ‘guard-like’ was part of finding out how people react to real oppression”* (cited in Le Texier, in press). Again this speaks to the ways in which the Experimenters' identity leadership ensured that Guards saw their cruelty as necessary for the advancement of a worthy ingroup cause.

Echoing Eshelman's statement that *“it's the Professor who decides here”*, another (unidentified) Guard said to Zimbardo and his colleagues during the debriefing that *“You're the Experimenter and in a sense we're kind of like the employees”* (cited in Le Texier, in press). Yet, by the same token, another Guard who was unconvinced of the study's worth, dropped out because of the requirement to act harshly: *“I object to the way prisoners will be treated. As the orientation meeting went on, the way things were stressed, there will be too much harassment.”* (cited in Le Texier, 2018, p. 105). All Guards thus inferred that the Experimenters were keen to encourage harassment and cruelty. Moreover, the fact that this assumption extended to a Guard who did not behave this way, suggests this is not simply a *post hoc* rationalization or a strategy for shifting blame.

*Fourth*, there is evidence that the Experimenters shaped not only the Guards' understandings but also the way they structured the prison regime. While the traditional account of the SPE suggests that the Guards drew up the prison rules and regulations

themselves, the new information suggests that they were effectively dictated by Warden Jaffe's orientation session. Of the 17 rules in the SPE, 11 were copies of those which Jaffe had drawn up for an earlier prison study of his own and the other 6 were largely adaptations to the specific circumstances of the SPE (e.g., a rule forbidding playing with the light switches because there were no such switches in Jaffe's study; see Le Texier, 2018, p.61; in press). Likewise, many of the punishments were based on those which had been devised in Jaffe's earlier study. It is not plausible that these similarities are mere coincidence.

In combination, these various pieces of evidence show clearly that the level of intervention by the Experimenters in the SPE was greater than has been reported in the literature or media. More importantly, it is apparent that these explicit and implicit instructions had more of an impact on the Guards than has been previously acknowledged. In the light of this, it is very difficult to sustain Zimbardo and colleagues' claim that Guards acted cruelly entirely of their own accord or slipped naturally into cruel roles.

We can therefore now say with confidence that leadership is important to an understanding of cruelty in the SPE. Nevertheless, while there are suggestions that this took the specific form of identity leadership — for instance, in Guards' statements that they felt that they were acting in pursuit of a worthwhile cause — the evidence we have presented thus far is hardly definitive. Not least, this is because the material discussed above does not provide an opportunity for in-depth analysis of the Experimenters' leadership in action.

There is, however, one further piece of evidence in the archive which is particularly pertinent to this issue. This is the recording of a formal meeting early in the SPE between Prison Warden Jaffe, and a Guard, John Mark. Two features of this meeting make it particularly relevant to the key question of how the Experimenters sought to produce Guard toxicity. The first is that Mark had been reluctant to act repressively (as he confirmed to us in a telephone interview on June 13<sup>th</sup> 2018) and so the whole meeting centered on Jaffe trying to

get him to adopt the role of a tough Guard. The second is that a recording of the entire interchange is available (whereas this is true of less than 10% of the entire study; Le Texier, in press) and therefore can be analysed in its entirety. The recording is held in the Department of Special Collections and University Archives in Stanford University Libraries (Source ID: SC0750\_s5\_b2\_21). It can be accessed directly online at <http://purl.stanford.edu/wn708sg0050> and a full transcript is provided in the Supplementary Materials. This allows students and researchers to review the transcript themselves and determine the degree to which it supports our identity leadership analysis rather than Zimbardo's role account.<sup>1</sup>

This interview provides a new and unprecedented opportunity to investigate a series of questions concerning not just *the degree* of leadership in the SPE but also *the form* taken by such leadership. To start with, did Warden Jaffe accept the position taken by Mark, or did he actively attempt to change Mark's stance? That is, to paraphrase Zimbardo's (2004, p.39) outline of the role conformity account, did Jaffe encourage the Guard to "rely on his own prior societal learning of the meaning of prisons and the behavioural scripts associated with the oppositional roles of prisoner and Guard as [his] sole source of guidance"? Or did he instead employ the signature characteristics of the identity leadership account?

More formally, we can break the identity leadership account down into the three elements that we outlined above and then gauge the presence of each of them in the text. The first involves seeking to establish a common cause and a common group membership that

---

<sup>1</sup> Before reading our own analysis of these issues, we recommend that readers listen to the interview and read the transcript for themselves. Given the necessity of applying open science principles to what has previously been a closed debate, we think it important that readers judge independently how the original materials speak to the respective merits of different conceptual accounts of the toxic behavior that unfolded in the SPE and evaluate for themselves the adequacy of our identity leadership account. For reasons of openness and transparency, we also illustrate our points with verbatim extracts from the meeting accompanied with line numbers related to the meeting transcript in the Appendix. This allows readers to check that our use of these extracts is accurate and appropriate.



links the Experimenters and Guards. The second involves presenting ‘tough’ or cruel behaviour as necessary in order to advance this shared cause. The third involves characterizing the group cause as worthy and noble in order to justify the toxic behaviour that advances it. If these elements are present, this would provide strong novel evidence that identity leadership had a role to play in efforts to encourage the Guards to behave cruelly.

### **The Jaffe-Mark meeting**

#### **On leadership**

Warden Jaffe’s meeting with Guard Mark is replete with evidence that the Guard was not conforming (blindly or otherwise) to the role he had been assigned and that, rather than simply accept this, Warden Jaffe sought to get him to change his mind. Faced with the Guard’s admission that “*I’m not too tough*” [line 14] and that “*if it was just entirely up to me, I wouldn’t do anything. I would just let it cool off*” [line 38], the Warden makes it clear at multiple points in the meeting that the Experimenters expect (and require) him both to be more “*involved*” [lines 12, 80] in proceedings and to be more “*tough*” [lines 13, 17, 244]. Indeed, the meeting starts with (and was presumably motivated by) the observation that:

*We noticed this morning that you weren’t really lending a hand ... but we really want to get you active and involved because the Guards have to know that every Guard is going to be what we call a tough Guard.* [lines 1-2, 11-13]

In urging the Guard to be more tough, it is also apparent that the Warden contextualizes this toughness as an aspect of the role that the Guard is being asked to play in order to ensure the experiment is a success. That is, toughness is not valorised as a positive attribute in and of itself. Rather, it is valorised as an important dimension of an identity that the Warden wants the Guard to adopt. Importantly too, the Warden emphasizes that “*every Guard*” is going to play this same role — underscoring the Experimenters’ expectations

about the norms for the entire group and mirroring what both Jaffe himself and Banks had said when briefing the Guards as a whole.

Again, though, Mark is reluctant to take on this identity and the attributes associated with it. Far from passively or naturally adopting the role of tough Guard, he actively and repeatedly resists the pressure from the Warden. This is the exact opposite of what the role account would predict. In the face of this resistance, Warden Jaffe repeats and intensifies his demands for conformity:

*We'd like you to try a little bit more, to get into the action ... instead of sitting in the background, if you can get in and start doing something yourself, get involved. [lines 77-80]... It's your job to make sure these things [episodes of Prisoner revolt] don't happen, to the best of your ability. [lines 171-172]*

Finally, towards the end of the meeting when the Guard is still showing no willingness to fall into line, the Warden urges him to “*forget some of the more sophisticated psychology that you might know*” [lines 245-246], to forgo his “*individual style*” [line 253], and instead to embrace the role of “*the stereotype Guard*” [line 252]. In this way, the Warden emphasizes the importance of the group and encourages the Guard to self-categorize, and behave, as a member of the group. Importantly, rather than allowing the Guard to define the role for himself, the Warden defines it for him, and does so in explicitly stereotypic terms. Here again, though, the Guard resists by refusing to embrace the categorical ‘us–them’ identity that the Warden invokes to structure his understanding of the situation. Indeed, the Guard actively contests the stereotypic self-definition that the Warden proposes. *Well I've met a lot of police that, er, that act a whole lot of different ways. You can't do it just by the movies or something. 'Cause you know I've met plenty of police.* [lines 248-250]. Far from naturally adopting the role or accepting the stereotypic role, the Guard thus makes an explicit case for rejecting them. It is thus with more than a hint of exasperation that the Warden

concludes the meeting with explicit and concrete instructions about the behaviour expected of the Guard: *When there's a situation... [you have] to have to go in there and shout if necessary. To be more into the action* [line 257-258].

In sum, two conclusions can be drawn from this evidence. First, the Guard is certainly not left to his own devices in deciding what to do within the prison. Indeed, he is assertively pressured by the prison leadership to conform to group-based expectations. Second, any theoretical analysis of Guard behaviour must include the active interventions of the Experimenters in seeking to ensure conformity to a brutal role. In short, there clearly is sustained leadership going on. We now focus on the forms that this takes with a view to understanding the strategies through which the Warden sought to shape the Guard's behavior.

### **On identity leadership**

The transcript of the meeting between Jaffe and Mark provides striking evidence that the Warden encouraged the Guard to discard his personal identity, to adopt a collective identity, and to embody stereotypic expectations associated with that collective identity. In this respect, as with Zimbardo's briefing of the Guards that we discussed earlier, one of the more striking features of Warden Jaffe's discourse is his repeated reference to the collective "we". Indeed, the Warden uses the first-person plural pronouns "we", "our" and "us" 57 times in this single meeting, or once every 30 words. Previous research has found that this use of collective pronouns is associated with effective leadership. For example, across the 34 Australian elections that have been held since Federation in 1901 winning candidates use these collective pronouns once every 79 words while losing candidates use them once every 136 words (Steffens & Haslam, 2013). Thus while politicians are often seen to exemplify principles of identity leadership (e.g., Augoustinos & De Garis, 2012; Gleibs, Hendricks, & Kurz, 2018; Reicher & Hopkins, 2000), the Warden appears to enact this aspect of the

process of cultivating a sense of shared identity more vigorously than even the most successful political leaders.

In this regard, it is clear that the Guard is encouraged to see himself, along with the Experimenter, as part of this collective ‘we’ (i.e., ‘us the prison authorities’). In particular, the Warden highlights their shared disdain for the correctional system, for example, by stressing that “*We happen to agree with you that basically it’s rotten*” [lines 200-201]. He also repeatedly portrays the Guard and the Experimenters as having the same interests and goals: *These things [prisons and mental hospitals] are all over the place. And we want to know about them* [lines 83-85]. In this way, the Warden appeals to a shared goal outside the experiment in order to justify cruelty within it.

Throughout the interaction the Warden thus goes to great lengths to represent the Guard and the Experimenters as being ‘on the same side’ while also sharing common enemies — notably those who defend the prevailing criminal justice system and “*the pigs*” [line 246]. He also encourages the Guard to acknowledge and embrace their shared identity and at one point accomplishes this through a four-step rhetorical manoeuvre where, first, he states his own values; second, he states those of the group as a whole; third, he makes a statement about what he sees the Guard’s values to be; and, fourth, he asks the Guard to confirm their alignment: *I’m very deeply committed to that [rehabilitation]. And I think all of us are... I think you feel the same way. Is that true?* [lines 204-209] In line with identity leadership, the construction of shared identity based on a common cause can thus be seen to lie at the core of the Warden’s intervention to shape the Guard’s behaviour.

The Warden also tries to convince Mark that what he is asking him to do — that is, be tough — is essential to the success of their mutual cause. More specifically, in the face of the Guard’s refusal to act tough, the Warden repeatedly reminds him that in doing so he is putting the research enterprise as a whole at risk. Guard toughness, the Warden explains, is

“really important for the workings of the Experiment” [lines 19-20] since “whether or not we can make this thing seem like a prison, which is the aim of the thing, depends largely on the Guards’ behaviour” [lines 22-24]. This is a point that the Warden underlines towards the end of the meeting when he notes:

*I think you could be a better guard. Er, better in the sense of, you know, the Guards have more responsibility at this point for making the thing work. That’s, that’s the real clincher right here. That, er, that if the Guards fall apart, the experiment falls apart. The Prisoners don’t have that power in this study.* [lines 157-158].

In making these observations, Warden Jaffe attunes Guard Mark to two interrelated considerations. First, that while it may be unpleasant and not what the Guard would otherwise want to do, his willingness to be tough is essential for the success of the project. Second, that if he fails to rise to this challenge, the Guard is effectively letting the team down. The dual implication, then, is that the Guard’s reluctance to conform is not only jeopardizing the group’s collective mission but also selfish. At the same time, tough behaviour is framed as something that is not only essential to the advancement of shared group goals but also expected of any committed group member.

The Warden repeatedly points out that the experiment will fail if the Guards are not sufficiently tough and he repeatedly underscores — and asks the Guard to recognize — the profound and progressive significance of the experiment for both science and society. Thus he asks “Do you understand the rationale behind doing something like this? The importance of it?” [lines 186-187]. He then explains in detail how the design of the study will allow the researchers to make important scientific claims, because: “There isn’t any prison in this country is going to let you set up, you know, observational measures 24 hours a day. So that’s why we have to do it here.” [lines 56-58]. He further explains how important the study is in terms of societal impact: “hopefully what will come out of this study is some very

*serious recommendations for reform, at least reform, if not, you know, revolutionary-type reform.*" [lines 49-53]. In this way, rather than being pernicious, cruelty is represented as an essential driver of social progress.

Critically, this impact is also characterised as a social good with reference to the progressive identity that the Warden and Guard ostensibly share. This aligns the Guard with the shared long-term goals of the research. At the same time, the Warden makes it clear that the worthy goal of exposing penal pathology and facilitating penal reform can only be achieved through cruel behaviour on the part of the Guard: *"What we want to do is be able to study the thing that exists, or as nearly as we can make it to what exists and to be able to go to the world with what we've done and say "Now look, this is what happens when you have Guards who behave this way" ... But in order to say that we have to have Guards who behave that way"* [lines 211-215]. In setting out the case for Guard cruelty in this way, the Warden also explicitly rejects the suggestion that such behaviour might reflect the fact that Guards are naturally cruel: *"We're not trying to do this just because we're sadists"* [lines 53-54]. Indeed, the worthiness of the cause is used as an explicit defence against such an accusation:

*If you need an excuse, and I think most of us do really, it is so we can learn what happens in a total institution ... And we want to know about them. So that we can, we can get on the media and, um, and, and into press with it. And, and, and say "Now look at what, what this really about".* [lines 80-86].

Overall, then, the arguments of the Warden serve to validate cruelty in the study in the name of challenging cruelty in society — only if the Guards are toxic can toxicity be exposed and eliminated. What would otherwise be seen as wicked behavior is thereby recast as worthy action performed in the service of a greater good. This helps assign positive and distinct value to the ingroup, and it also facilitates identification with that ingroup. It also makes oppression

of the Prisoners consonant with group norms and serves to offset — and inoculate against — any doubts or hesitations that the Guard might have about acting brutally. In short, this new evidence displays all the hallmarks of identity leadership.

### **The need to rethink the nature of cruelty in the SPE**

The meeting between Warden and Guard provides clear evidence that Guards in the SPE were not left to their own devices when it came to making decisions about how to behave and run the prison. On the contrary, they were subjected to active leadership from the Experimenters. It follows that any account of the SPE which fails to highlight the leadership of the Experimenters and their concerted attempts to make the Guards act in role is both partial and misleading.

As noted above, numerous Guards independently reported that they believed the Experimenters played an active role in guiding their behaviour. Beyond this, though, it is apparent from various pieces of evidence — notably the transcript of the Jaffe-Mark interview and Jaffe's five-hour briefing of all the Guards — that this leadership took a very specific form. In particular, we see that Jaffe's efforts to encourage the Guard to conform to stereotypic role requirements centred on strategies of identity leadership. That is, he sought to influence Mark through appeals to a sense of shared identity which promoted 'toughness' as an ingroup-defining attribute necessary (a) to run the prison properly, (b) to advance science, and thereby (c) to achieve the valued goal of exposing the toxicity of the American penal system. Indeed, by aligning 'toughness' with the goals of multiple identities the Warden can be seen to have increased the likelihood of it being seen as an appropriate way to behave (in ways suggested by identity complexity theory; Roccas & Brewer (2002).

It should be stressed that these identity appeals were not just a subset of themes that are mingled in with others in the meeting. Instead, the meeting can be seen as one long series of identity appeals. Time and again Mark insists on his unsuitability as a Guard and his

unwillingness to be tough. Time and again Jaffe responds by explaining how important it is that he act like a tough Guard to help Jaffe, Zimbardo and others in their noble enterprise. Remove this interplay between Mark's role rejection and Jaffe's identity entrepreneurship and little would be left.

To summarize, then, the new evidence from the SPE archive sustains three theoretical claims. First, the traditional notion that Guards became cruel of their own accord is very hard (if not impossible) to sustain. Second, the Experimenters' leadership was a central feature of the study. Third, and more specifically, we see that the Experimenters engaged in identity leadership in an effort to encourage Guard cruelty. Although support for these conclusions emerges from the Guard's reports and the Jaffe-Mark transcript, it is also supported by a wealth of other new material (see Le Texier, 2018). In particular, full transcripts of the Guard briefings indicate that the Experimenters went to considerable trouble to create a sense of shared identity with *all Guards* (not just one), to persuade them that they were co-producers of important scientific knowledge, and to indicate in some detail exactly how they were expected to behave. What happened to Mark was thus not exceptional, but simply reflects the fact that he was failing to behave in ways expected (and demanded) by the Experimenters.

At the same time that we make these claims, it is important to be clear about what we are not claiming. First, we do not suggest that identity leadership always produces consent or that it did so in the case of Guard Mark. Nevertheless, the efforts of Warden Jaffe did have some effect on him. In particular, this is because Mark recollects being fully aware that his behaviour was at odds with the Experimenters' goals and that toughness was presented, and affirmed, by the Experimenters as a core norm for the Guard group. He therefore recalls being poorly positioned to challenge the cruelty of others — not least because while he himself was not cruel in his dealings with the Prisoners, he was sidelined from the shift in

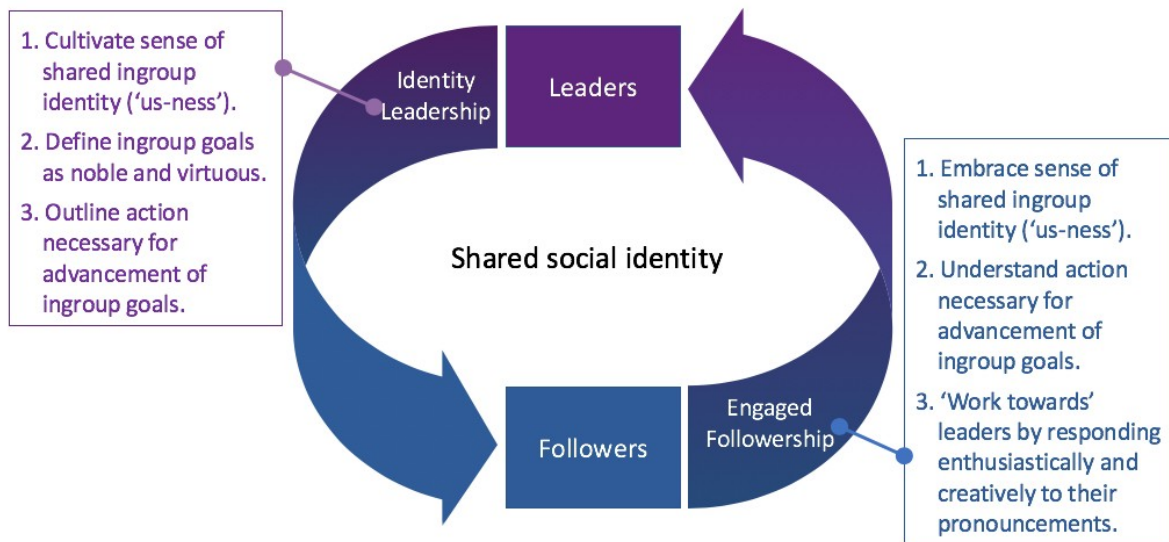


which most of the toxic behaviours reported in the SPE took place. Thus, identity leadership facilitated Guard cruelty even if Guard Mark did not display this behaviour himself.

Second, even though it is clear that the Experimenters were fairly explicit about sanctioning extreme behaviour, and at times provided explicit guidance as to what forms these should take, we are not suggesting that the Guards had no autonomy and were simply following a script. We agree with Zimbardo and colleagues that the Guards clearly improvised and were creative in what they did and how far they went. However, this enriches rather than undermines the identity leadership analysis. For as Haslam and Reicher (2012) have argued, the Guards' creativity can be seen to reflect the fact that effective identity leadership does not produce passive conformity so much as *engaged followership*. That is, when followers identify with a leader and his or her shared cause, they ask themselves what it is that that leader wants them to do and then strive to interpret the instructions they have been given *enthusiastically and creatively*. This is a process that the World War II historian, Ian Kershaw refers to as “working towards the Führer” (Kershaw, 1993; see also Sofsky, 1993). Here, then, we would suggest that those Guards who were inspired by the identity leadership of Zimbardo and his colleagues would have been “working towards the Experimenter”, in ways that translated an appeal to be tough into a willingness to be cruel and oppressive (Haslam & Reicher, 2007).

A critical point here is that the identity leadership analysis we have presented — and which is represented schematically in Figure 1 — does not simply replace one form of automaticity with another (moving from ‘people automatically take on the roles they are thrust into’ to ‘people automatically take on the identities that are thrust upon them’). People are well able to resist the categories and identities proposed to them (as shown above in Mark’s responses to Warden Jaffe), especially if these are at odds with other identities that are important in their lives (as shown in our previous critiques of Zimbardo’s work; see

Reicher & Haslam, 2006). For this and other reasons, identity leadership may elicit very different behaviour from different individuals, and invoking cruelty as a group norm will not always lead people to embrace it. This will be especially true if leaders fail — as Warden Jaffe did with Guard Mark — to persuade would-be followers that cruelty is a normative aspect of an identity that they value.



Can we conclude, then, that even if identity leadership failed to 'turn' Mark, it was responsible for such Guard cruelty as did occur? Again, to make such a claim is beyond the scope of the current evidence. The type of evidence we have from the SPE simply does not allow us to make any definitive causal claims of this form. In particular, the small sample and lack of experimental control mean we cannot be sure that these Guards would not have been cruel in any case (e.g., for reasons suggested by Carnahan & McFarland, 2007), or that their cruelty was not produced by some other feature of the study besides the leadership of the Experimenters. Nevertheless, what we can say with confidence is (a) that the Experimenters undoubtedly did *try* to exercise identity leadership, (b) that several guards independently reported being aware of the Experimenter's efforts to engage in identity leadership, and (c) that identity leadership therefore constitutes a *plausible* framework for explaining Guard cruelty in the SPE. At a very minimum, then, the results of our analysis are plainly more

consistent with an identity leadership account than they are with the standard role account. We would add too that the patterns we have identified above accord with a very large body of research which shows identity leadership to be a critical component of effective leadership (both toxic and benign) in the world at large (e.g., see Steffens et al., 2014; van Dick et al., 2018).

In relation to these various points, the significance of the Stanford archive is how closely it supports theoretical claims that were first made over a decade ago (but which could not be substantiated) — claims which (having now been substantiated) provide an important platform for revising our understanding of the important issues that the SPE addresses. In this regard, it is also clear that there are a great many ways in which the material in the archive might not have substantiated our analysis. Most obviously, this would be the case if it had revealed no evidence of the Experimenters' identity leadership or if the behavior of the participants had mapped closely on to Zimbardo and colleague's role account (e.g., if the Experimenters had provided the Guards with no guidance about how to interpret their role). Indeed, the falsifiability of our analysis is apparent from the fact that, hitherto, Zimbardo has dismissed our evidence-based critiques of the SPE as fundamentally wrong-headed and “scientifically irresponsible” (2006, p.47).

### **Identity leadership and cruelty**

The plausibility of our position is enhanced by two further considerations. First, as well as being applied to the study of leadership in general (e.g., Haslam et al., 2011; van Dick et al., 2018), the identity leadership analysis has successfully been used to explain toxic behaviour in other classic studies (notably, Milgram's OtA research). Indeed, once one acknowledges the role of leadership in both Milgram's paradigm and Zimbardo's, the major distinction between the two disappears (Haslam & Reicher, 2012b). Nevertheless, the value of this corroboration lies the fact that the OtA studies were more carefully controlled than the

SPE, used larger samples, and have been independently replicated (Blass, 2004; Reicher et al., 2017). This advances our claim that the identity leadership framework is useful not only for understanding behaviour in the SPE, but also for understanding toxic (and non-toxic) behaviour in a broad range of experimental contexts.

Second, looking beyond the laboratory to the wider world, it is also clear that identity leadership is a common feature in episodes of human toxicity and brutality, perhaps especially in its most extreme forms (Koonz, 2003; Muller-Hill, 1988; Sofsky, 1993; Vetlesen, 2005). A recurrent observation is that leaders work hard not only to create a common identity with would-be perpetrators, but to convince them that cruelty to outgroups (e.g., Muslims, migrants, dissidents) is necessary for the protection and advancement of the ingroup (e.g., keeping Serbia strong, America great, Turkey safe). However, if we were to select just one example to exemplify the dynamics of human inhumanity, we would choose Himmler's infamous Poznan speech of 6<sup>th</sup> October 1943 in which he rallied SS Officers to persevere with the challenges of exterminating Jews in occupied Poland. The substance of this is exemplified by the following passage:

*It is one of those things that is easily said: 'The Jewish people is being exterminated', every party member will tell you. . . [But] none of them has seen it happen, not one has had to go through with it. Most of you men know what it is like to see 100 corpses lie side by side, or 500 or 1,000. To have stood fast through all this and. . . at the same time to have remained a decent person. . . has made us hard. This is an unwritten and never-to-be-written page of glory in our history. . . All in all, however, we can say that we have carried out this most difficult of tasks in a spirit of love for our people. (Grobmes & Landen, 1983, pp. 454–455; cited by Haslam et al., 2015, p.78)*

By invoking this example we are not suggesting that the toxicity observed in Milgram's studies or in the SPE was in any way comparable with that of the Holocaust

(Miller, 2014). What we believe is striking, though, is the consonance in the processes through which this was encouraged in these different settings. Framing cruelty as essential for the achievement of noble collective goals thus appears to be a critical strategy for mobilizing people to hate and harm others in theatres of conflict both small and large.

It follows from all this that just as cruelty does not inhere simply in the nature of the perpetrators so too it does not inhere only the demands of the situation. An understanding of how it is produced additionally requires an analysis of leadership, of how leaders persuade, and of how they are able to portray toxic behavior as worthy action in defense of a noble group cause. Central to this endeavor are leaders' efforts to construct a sense of shared identity that encompasses both the source and the target of persuasion. Indeed, accounts which seek to naturalize cruelty and harm-doing as an inevitable outcome of human behavior serve to help leaders avoid accountability for the part they have played and so, not surprisingly, are often invoked by repressive leaders in their own defence (Cesarani, 2004). A case in point involves Radovan Karadzic who was indicted before the International Criminal Court for the former Yugoslavia (ICTY) for his war crimes. His defence centred on the claim that the hatred between groups and the violence that was perpetrated "arose of its own accord, from the bottom up and cannot be attributed to Franco Trujman" (cited in Elcheroth & Reicher, 2017, p. 29).

### **Conclusion**

The new evidence that we have examined in this paper makes it clear that the specific narrative of the SPE and our broader understanding of cruelty in the world both need to be rewritten. The totality of evidence indicates that far from slipping naturally into their assigned roles, some of Zimbardo's Guards actively resisted. They were consequently subjected to intense interventions from the Experimenters. These sought to persuade them to conform to group norms for the purpose of achieving a shared and admirable group goal. Where

previously we only had inklings that the behavior of Guards in the SPE could have been produced in response to forces of identity leadership, we now have sufficiently strong and clear evidence of this as to need to place it on the scientific record. This in turn means that we must, of necessity, focus on the role of these forces in spawning cruelty and repression more generally, both in our studies and in the world beyond.

We hope that future presentations of the SPE — whether in the scientific literature, in classrooms, or in boardrooms — will now tell a richer story of what happened in the basement of the Stanford Psychology Department during the summer of 1971. If that happens, then the SPE can continue to help us achieve a richer appreciation of the dark social processes which blight the human condition.

### References

- Alvarez, K. P. (Dir.) (2015). *The Stanford Prison Experiment* [Motion picture]. New York: IFC Films.
- Augoustinos, M., & De Garis, S. (2012). 'Too black or not black enough': Social identity complexity in the political rhetoric of Barack Obama. *European Journal of Social Psychology, 42*(5), 564-577.
- 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. *The Psychologist, 20*, 494-495.
- Barreto, N. B., & Hogg, M. A. (2017). Evaluation of and support for group prototypical leaders: A meta-analysis of twenty years of empirical research. *Social Influence, 12*, 41-55.
- Bartels, J. M., Milovich, M. M., & Moussier, S. (2016). Coverage of the Stanford prison experiment in introductory psychology courses: A survey of introductory psychology instructors. *Teaching of Psychology, 43*, 136-141.
- Blass, T. (2004). *The man who shocked the world: The life and legacy of Stanley Milgram*. New York, NY: Basic Books.
- Blum, B. (2018). The lifespan of a lie. *Medium*. <https://medium.com/s/trustissues/the-lifespan-of-a-lie-d869212b1f62>
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123-152.
- Browning, C. (1992). *Ordinary men: Reserve Police Battalion 101 and the final solution in Poland*. London: Penguin Books.
- Carnahan, T. & McFarland, S. (2007). Revisiting the Stanford Prison Experiment: Could participant self-selection have led to the cruelty? *Personality and Social Psychology Bulletin, 33*, 603-614.
- Cesarani, D. (2004). *Eichmann: His life and crimes*. London: Heinemann.
- Donnellon, A. (1996). *Team-talk: The power of language in team dynamics*. Boston, MA: Harvard University Business School Press.
- Eatwell, R., & Goodwin, M. (2018). *National populism: The revolt against liberal democracy*. London: Penguin.
- Elcheroth, G. & Reicher, S. D. (2017). *Identity, violence and power: Mobilising hatred, demobilising dissent*. London: Palgrave.

- Ellemers, N., De Gilder, D., & Haslam, S. A. (2004). Motivating individuals and groups at work: A social identity perspective on leadership and group performance. *Academy of Management Review*, *29*, 459-478.
- Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. Oxford, UK: Psychology Press.
- Gibson, S. (2013). Milgram's obedience experiments: A rhetorical analysis. *British Journal of Social Psychology*, *52*, 290-309.
- Gleibs, I. H., Hendricks, K., & Kurz, T. (2018). Identity mediators: Leadership and identity construction in campaign speeches of American Presidential candidates' spouses. *Political Psychology*, *39*, 939-956.
- Gray, P. (2013). Why Zimbardo's prison experiment isn't in my textbook. *Psychology Today* (October 18). Retrieved from <https://www.psychologytoday.com/au/blog/freedom-learn/201310/why-zimbardo-s-prison-experiment-isn-t-in-my-textbook>
- Griggs, R. A. (2014). Coverage of the Stanford Prison Experiment in introductory psychology textbooks. *Teaching of Psychology*, *41*, 195-203.
- Griggs, R. A., & Whitehead III, G. I. (2014). Coverage of the Stanford prison experiment in introductory social psychology textbooks. *Teaching of Psychology*, *41*, 318-324.
- Grobmes, A., & Landen, D. (1983). *Genocide: Critical issues of the Holocaust*. West Orange, NJ: Berman House.
- Haney, C., Banks, C. & Zimbardo, P. (1973a). A study of prisoners and guards in a simulated prison. *Naval Research Reviews*. September (pp.1-17). Office of Naval Research: Washington, D.C.
- Haney, C., Banks, C., & Zimbardo, P. G. (1973b). Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology*, *1*, 69-97.
- Haslam, S. A., & Reicher, S. D. (2007a). Identity entrepreneurship and the consequences of identity failure: The dynamics of leadership in the BBC Prison Study. *Social Psychology Quarterly*, *70*, 125-147. doi:10.1177/019027250707000204
- Haslam, S. A., & Reicher, S. D. (2007b). Beyond the banality of evil: Three dynamics of an interactionist social psychology of tyranny. *Personality and Social Psychology Bulletin*, *33*, 615-622.
- Haslam, S. A., & Reicher, S. D. (2012a). When prisoners take over the prison: A social psychology of resistance. *Personality and Social Psychology Review*, *16*, 154-179.



- Haslam, S. A., & Reicher, S. D. (2012b). Contesting the “nature” of conformity: What Milgram and Zimbardo’s studies really show. *PLoS Biology*, *10*(11): e1001426.
- Haslam, S. A., & Reicher, S. D. (2017). 50 years of “obedience to authority”: From blind obedience to engaged followership. *Annual Review of Law and Social Science*, *13*, 59-78. e109015. doi: 10.1146/annurev-lawsocsci-110316-113710
- Haslam, S. A., Reicher, S. D., & Birney, M. (2014). Nothing by mere authority: Evidence that in an experimental analogue of the Milgram paradigm participants are motivated not by orders but by appeals to science. *Journal of Social Issues*, *70*, 473-488. doi:10.1111/josi.12072
- Haslam, S. A., Reicher, S. D., Millard, K., & McDonald, R. (2015). “Happy to have been of service”: The Yale archive as a window into the engaged followership of participants in Milgram’s ‘obedience’ experiments. *British Journal of Social Psychology*, *54*, 55-83.
- Haslam, S. A., Reicher, S. D. & Platow, M. J. (2011). *The new psychology of leadership: Identity, influence and power*. New York and Hove, UK: Psychology Press.
- Hirschbiegel, O. (Dir.) (2001). *Das Experiment* [Motion picture]. Berlin: Senator Film.
- Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, *5*, 184-200.
- Jacoby, J. E., Severance, T. A., & Bruce, A. S. (Eds.). (2004). *Classics of criminology*. Long Grove, IL: Waveland Press.
- Kershaw, I. (1993). Working towards the Führer: Reflections on the nature of the Hitler dictatorship. *Contemporary European History*, *2*, 103–108.
- Konečni, V. J. (2007). Bad apples and bad barrels: Bad metaphors and blind spots regarding evil? *APS Observer*, *20*(5), 9-10.
- Konnikova, M. (2015). The real lesson of the Stanford prison experiment. *The New Yorker*, June 12. <https://www.newyorker.com/science/maria-konnikova/the-real-lesson-of-the-stanford-prison-experiment>
- Koonz, K. (2003). *The Nazi conscience*. New Haven, CT: Harvard University Press.
- Le Texier, T. (2018). *Histoire d’un mensonge: Enquête sur l’expérience de Stanford*. Editions la Découverte: Paris.
- Le Texier, T. (in press). Debunking the Stanford Prison Experiment. *American Psychologist*.
- Lovibond, S. H., Mithiran, X., & Adams, W. G. (1979). The effects of three experimental prison environments on the behaviour of non-convict volunteer subjects. *Australian Psychologist*, *14*, 273–287.

- Mark, J. (2007) Letter to the editor. *Stanford Daily* (June/July) Retrieved from:  
[https://alumni.stanford.edu/get/page/magazine/article/?article\\_id=32561](https://alumni.stanford.edu/get/page/magazine/article/?article_id=32561)
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*, 371-378.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Miller, A. G. (2004). What can the Milgram obedience experiments tell us about the Holocaust? Generalizing from the social psychology laboratory. In A. Miller (Ed.), *The social psychology of good and evil* (pp. 193–239). New York, NY: Guilford.
- Muller-Hill, B. (1988). *Murderous science*. Oxford: Oxford University Press.
- Noelle-Neumann, E. (1974), "The spiral of silence: a theory of public opinion", *Journal of Communication, 24*, 43–51
- Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Oxford: Blackwell.
- Reicher, S. D., & Haslam, S. A. (2006). Rethinking the psychology of tyranny: The BBC Prison Experiment. *British Journal of Social Psychology, 45*, 1-40.
- Reicher, S. D., & Haslam, S. A. (2013). Obedience and tyranny in psychology and history. In A. Golec & A. Cichocka (Eds.), *Social psychology of social problems* (pp.172-195). Houndmills, UK: Palgrave Macmillan.
- Reicher, S. D., Haslam, S. A., & Hopkins, N. (2005). Social identity and the dynamics of leadership: Leaders and followers as collaborative agents in the transformation of social reality. *Leadership Quarterly, 16*, 547-568.
- Reicher, S. D., Haslam, S. A., & Miller, A. G. (2014). What makes a person a perpetrator? The intellectual, moral, and methodological arguments for revisiting Milgram's research on the influence of authority. *Journal of Social Issues, 70*, 393-408.
- Reicher, S. D., Haslam, S.A., & Smith, J. R. (2012). Working toward the experimenter: Reconceptualizing obedience within the Milgram paradigm as identification-based followership. *Perspectives on Psychological Science, 7*, 315–324.
- Reicher, S. D., Haslam, S. A., Spears, R., & Reynolds, K. J. (2012). A social mind: The context of John Turner's work and its influence. *European review of social psychology, 23*, 344-385.
- Reicher, S.D. & Hopkins, N.P. (2001) *Self and Nation*. London: Sage.
- Ronson, J. (2015). *So you've been publicly shamed*. New York: Riverhead Books.
- Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review, 6*, 88-106.

- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- Savin, H. B. (1973). Professors and psychological researchers: Conflicting values in conflicting roles. *Cognition*, 2, 147-149.
- Scheuring, P. (Dir.) (2010). *The Experiment* [Motion picture]. Culver City, CA: Stage 6 Films.
- Sofsky, W. (1993). *The order of terror: The concentration camp*. Princeton, NJ: Princeton University Press.
- Steffens, N., & Haslam, S. A. (2013). Power through 'us': Leaders' use of we-referencing language predicts election victory. *PLoS ONE*, 8(10): e77952.
- Steffens, N. K., Haslam, S. A., & Reicher, S. D. (2014). Up close and personal: Evidence that shared social identity is a basis for the 'special' relationship that binds followers to leaders. *Leadership Quarterly*, 25, 296-313. doi:10.1016/j.leaqua.2013.08.008
- Steffens, N. K., Haslam, S. A., Reicher, S. D., Platow, M. J., Fransen, K., Yang, J., Ryan, M. K., Jetten, J., Peters, K., & Boen, F. (2014). Leadership as social identity management: Introducing the Identity Leadership Inventory (ILI) to assess and validate a four-dimensional model. *Leadership Quarterly*, 25, 1001-1024.
- Steffens, N. K., Munt, K., van Knippenberg, D., Platow, M. J., & Haslam, S. A. (2018). *A meta-analytic review of leader group prototypicality and leadership effectiveness*. University of Queensland: Unpublished manuscript.
- Swann, W. B., Buhrmester, M. D., Gómez, A., Jetten, J., Bastian, B., Vázquez, A., ... & Finchilescu, G. (2014). What makes a group worth dying for? Identity fusion fosters perception of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology*, 106, 912-926.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Thayer, S., & Saarni, C. (1975). Demand characteristics are everywhere (anyway): A comment on the Stanford Prison Experiment. *American Psychologist*, 30, 1015-1016.
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15-40). Cambridge: Cambridge University Press.
- Turner, J. C. (1991). *Social influence*. Milton Keynes: Open University Press.

- Turner, J. C. (2006). Tyranny, freedom and social structure: Escaping our theoretical prisons. *British Journal of Social Psychology*, *45*, 41–46.
- Turner, J. C., & Haslam, S. A. (2001). Social identity, organizations and leadership. In M. E. Turner (Ed.), *Groups at work: Advances in theory and research* (pp. 25-65). Hillsdale, NJ: Erlbaum.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.
- Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. A. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454-463.
- Turner, J. C. and Reynolds, K. J. 2011. "Self-categorisation theory". In P. VanLange, , A. Kruglanski, and T. Higgins, *Handbook of theories in social psychology*, 399–417. New York: Sage.
- Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and social psychology review*, *7*, 349-361.
- van Dick, R., Hirst, G., Grojean, M. W., & Wieseke, J. (2007). Relationships between leader and follower organizational identification and implications for follower attitudes and behaviour. *Journal of Occupational and Organizational Psychology*, *80*, 133–150.
- van Dick, R., Lemoine, J. E., Steffens, N. K., Kerschreiter, R., Akfirat, S. A., Avanzi, L., Dumont, K., Epitropaki, O., Fransen, K., Giessner, S., Gonzales, R., Kark, R., Lipponen, J., Markovits, Y., Monzani, L., Orosz, G., Pandey, D., Roland-Lévy, C., Schuh, S., Vörös, V., Sut, I., Wong, B. I., Xin-an, Z., & Haslam, S. A. (2018). Identity Leadership going global: Validation of the Identity Leadership Inventory (ILI) across 20 countries. *Journal of Occupational and Organizational Psychology*, *91*, 697-728.
- Vetlesen, A. J. (2005). *Evil and human agency: Understanding collective evildoing*. Cambridge, UK: Cambridge University Press.
- Zimbardo, P. (1992). *Quiet rage* (video). Stanford, CA: Stanford University.
- Zimbardo, P. (1999). *The Stanford Prison Experiment website*. ([www.prisonexp.org](http://www.prisonexp.org))
- Zimbardo, P., Banks W. C., Haney, C., & Jaffe, D. (1973). The mind is a formidable jailer: A Pirandellian prison. *The New York Times Magazine*. April 8, Section 6, 38-60.

- Zimbardo, P. G. (2004). A situationist perspective on the psychology of evil: Understanding how good people are transformed into perpetrators. In A. Miller (Ed.), *The social psychology of good and evil* (pp.21–50). New York: Guilford.
- Zimbardo, P. (2006). On rethinking the psychology of tyranny: The BBC Prison Study. *British Journal of Social Psychology*, 45, 47–53.
- Zimbardo, P. (2007). *The Lucifer Effect: How good people turn evil*. London: Random House.
- Zimbardo, P., Johnson, R. L., & McCann, V. (2012) *Psychology: Core Concepts* (7<sup>th</sup> Ed.). London: Pearson.

## Supplementary material

## Transcript of Meeting between Warden Jaffe and Guard Mark

(Side 2 of Tape 136510: <http://purl.stanford.edu/wn708sg0050>.)

<i>line</i>	<i>time</i>	
1	8.38	<i>Warden</i> : OK. We noticed this morning that um you weren't really, you
2		know, lending a hand and I was wondering if there is anything
3		wrong?
4	8.50	<i>Guard</i> : Well this morning it seemed like there was pretty, many guys in
5		there anyhow, in the end. I thought that for security it was better if
6		not everybody went everywhere at one time together. I was just
7		keeping my eye on that person.
8	9.11	<i>Warden</i> : Right, but generally, you've been kind of in the background.
9		Part of that is my fault because I've, you know, I've kind of gone
10		along with when you've wanted to sit outside while they were doing
11		the count and that sort of thing, but we really want to get you active
12		and involved because the Guards have to know that every Guard is
13		going to be what we call a tough Guard and so far, um...
14	9.36	<i>Guard</i> : ... I'm not too tough ...
15	9.38	<i>Warden</i> : Yeah. Well you have to try to kind of get it in you.
16	9.42	<i>Guard</i> : ... I don't know about that ...
17	9.45	<i>Warden</i> : See, the thing is what I mean by tough is, you know, you have
18		to be, um, <i>firm</i> , and you have to be in the action and, er, and that sort
19		of thing. Um, it's really important for the workings of the
20		Experiment ...
21	10.03	<i>Guard</i> : Excuse me, I'm sorry,
22	10.04	<i>Warden</i> : ... because, um, because, the, you know, whether or not we
23		can make this thing seem like a prison, which is the aim of the thing,
24		um, depends largely on the Guards' behaviour.
25	10.20	<i>Guard</i> : Well, I can see that, you know. I can see what it's doing to the
26		experiment. Like all I can do is like think of it in terms of times that
27		I've had dealings with police, like. And it seems the exact same thing
28		to me as any time I've ever seen the police. And it seems that...
29	10.42	<i>Warden</i> : You think it seems pretty realistic?
30	10.44	<i>Guard</i> : I think it's realistic but I think that what you're sort of asking me
31		to do I think is the same reason why I think there is always so much
32		trouble with the police.
33	10.54	<i>Warden</i> : Right...
34	10.55	<i>Guard</i> : Because I think whenever there's something wrong people start
35		yelling and they take away more privileges, just that cycle. I think
36		that makes it worse. And I think there's problems. What I would do
37		if I was a Guard, if it was just entirely up to me, I wouldn't do
38		anything. I would just let it cool off.
39	11.24	<i>Warden</i> : Ah, at some point that may be necessary. But our purpose here
40		is not to devise a <i>better</i> prison. I mean we <i>all</i> know how to do that.
41		Right? Um...
42	11.33	<i>Guard</i> : I don't know. I really don't know.
43	11.35	<i>Warden</i> : Well no, I don't mean a better prison, let's say a better criminal

44 justice system. I mean that's, that's what my whole class that I was  
45 teaching last quarter was dealing with. We were talking about  
46 alternatives, what we could do better. And, the idea is, that this is  
47 supposed to be, as nearly as we can make it, a copy of the existing  
48 one. In other words, we want to see what this does to ordinary  
49 innocent people, and, um, we know it's not nice, but we don't know  
50 how it's not nice. And that's what we're trying to find out. And  
51 hopefully what will come out of this study is some very serious  
52 recommendations for reform, at least reform, if not, you know,  
53 revolutionary-type reform. So this is our goal. We're not trying to do  
54 this just because we're all um sadists. Um. You see we want to be  
55 able to study what a prison is like. And you can't really do that. I  
56 mean no prison, there isn't any prison in this country is going to let  
57 you set up, you know, observational measures 24 hours a day. So  
58 that's why we have to do it here.

59 12.59 *Guard*: Well I don't know about that. Because I just read about a prison  
60 the other day where, er, the guys wanted to stay in until their time  
61 was up because, er, you know [inaudible]...

62 13.16 *Warden*: No, but what I said is that there's no prison that will let us go in  
63 there and observe ...

64 13.19 *Guard*: Oh, oh I see what you mean, Oh I see what you mean.

65 13.18 *Warden*: ... what we want to observe. In other words we can't go in  
66 there. We have no way of studying what, what the typical prison is  
67 like on a thorough basis. So, what we're doing is, on the basis of  
68 what we've read and heard. You know, and what I mean by studying  
69 is, is studying the effects that it has on people. We're trying to make  
70 it, you know, make it here and then study the effects here...

71 13.47 *Guard*: You know, there's also the effects that it has on Guards, you  
72 know. Just, just by wearing this Guard's uniform, you know, that's a  
73 pretty heavy thing for me. I, you know, I don't really get into this  
74 stuff too much. So, you know, I, I, I can do the Guard things, extra  
75 things like getting really tough and yelling and stuff, I don't know  
76 how long I can do that because ...

77 14.14 *Warden*: We'd like you to try a little bit more, to get into the action, you  
78 know, if there's something to be done, instead of sitting in the  
79 background, if you can get in and start doing something yourself, get  
80 involved, um. And, if you, if you need, if you need an excuse, and I  
81 think most of us do really, um, it is, you know, so we can, we can  
82 learn what happens in a total institution, which we've set up, and  
83 which a prison is and a mental hospital is, and these things are all  
84 over the place. And we want to know about them. So that we can, we  
85 can get on the media and, um, and, and into press with it. And, and,  
86 and say "Now look at what, what this really about".

87 15.02 *Guard*: ... Yes ...

88 15.04 *Warden*: This is, this is really I think an important thing, but in order to  
89 make it work we've got to have the kind of staff that, that can ...

90 15.13 *Guard*: Well I think, you know, I'll try. But in a way I was kind of  
91 unfortunate that when this experiment was set up that it was er  
92 random picking like that. Because I think, er, you know, its, I, I don't  
93 think I'm a very good Guard for the experiment, but I could have

- 94           been a really good prisoner.
- 95 15.34 *Warden*: Right, but that part of why, you know, random selection is
- 96           important.
- 97 15.39 *Guard*: Why? Could you explain that to me? Why it's important to do it
- 98           random?
- 99 15.41 *Warden*: Alright. Let's say hypothetically, that... No, well let me start
- 100          over. Let's say that, um, you have, you want to eventually have two
- 101          groups of people and you're going to do different operations on each
- 102          group and you want to find out, or you're going to, you're going to
- 103          do, you know, one group basically is a control group right and one
- 104          group is an experimental group and we do something on them,
- 105 16.61 *Guard*: Yeah
- 106 16.22 *Warden*: Now, let's say that the Prisoners in this group are the
- 107          experimental group and the Guards we can kind of call the control
- 108          group, OK? Um, now, if the, so, and then we, then at the end we
- 109          observe the differences between the two groups. ... Now if you want
- 110          to make sure that the differences are due to the experimental
- 111          manipulation, in other words due to the fact that some people are
- 112          Prisoners in this experiment you have to start out with a group that is
- 113          as similar between groups as possible and the best way to get that is
- 114          random selection. In other words, what we're, what we're basically
- 115          saying when we say we have random selection is that we took two
- 116          groups of people who were similar and did different things to them,
- 117          and we got these effects...
- 118 17.17 *Guard*: OK. But it would seem to me that if you wanted to study an
- 119          experiment that is as closely as possible simulating prison
- 120          experience, like you said, that you wouldn't want two groups of
- 121          people that were that similar to begin with because in a prison you
- 122          have two groups of people that they're about as dissimilar as you can
- 123          find...
- 124 17.37 *Warden*: Well, there are a lot of people who say that the Guards are
- 125          really, not much different...
- 126 17.40 *Guard*: They are different from Prisoners.
- 127 17.42 *Warden*: Well ...
- 128 17.43 *Guard*: They're not like normal people, I'll say that. But they're not like
- 129          Prisoners because they're not people that, that would do, do stuff in
- 130          society that would put them in prison, they're people who would cop
- 131          out to society and take a job, and ...
- 132 17.57 *Warden*: Alright, let's make a different analogy. Let's take a mental
- 133          hospital, which I think is a closer analogy. Most people say that the
- 134          staff in mental hospitals, by and large, are not too dissimilar from the
- 135          patients.
- 136 18.11 *Guard*: I agree, I with that about the doctors too.
- 137 18.14 *Warden*: [Laughs] Right, OK. At any rate, um, I'm not going to hassle
- 138          with you over whether or not they are. All I'm explaining to you is
- 139          why we had random selection. The idea is to, is to get as nearly
- 140          equal, you know the same, er, people in each group to start out with,
- 141          um, so that's why, and you were randomly selected to be a Guard.
- 142          That's the way it was.
- 143 18.43 *Guard*: OK.



- 144 18.47 *Warden*: But we really, you know I'd really like to see you get in there  
145 and try.
- 146 18.53 *Guard*: Well, all I can ... OK I will try. And like, it's just like a situation  
147 where if I had randomly assigned to be a Prisoner I think I would  
148 have done stuff like not eat and I would have been really, really quiet  
149 and just done exercises and things like this. And you know I think  
150 it's pretty much an analogous situation. You know it could have been  
151 that that would have been messing up the results that way too. You  
152 know, if I wasn't in there and being a real prisoner that was trying to  
153 escape and mess it up all the time.
- 154 19.40 *Warden*: No. Alright. All I'm trying to do is to point out, er, how I think  
155 you could be a better guard. Er, better in the sense of, you know, the  
156 Guards have more responsibility at this point for making the thing  
157 work. That's, that's the real clincher right here. That, er, that if the  
158 Guards fall apart, the experiment falls apart. The prisoners don't  
159 have that power in this study. So, because almost anything they do,  
160 you know, they can't ...
- 161 20.22 *Guard*: They rip off the doors. And ripped bed ...
- 162 20.25 *Warden*: No, I don't think they be able to ...
- 163 20.28 *Guard*: One guy just, just broke his door
- 164 20.30 *Warden*: Oh yeah?
- 165 20.32 *Guard*: And I think, I think that, well, I would be surprised if after two  
166 weeks they didn't start ripping up the plastic covers and the beds
- 167 20.40 *Warden*: Well we're going to, yeah ... We may move, you know, over to  
168 the County jail, um,
- 169 20.48 *Guard*: OK
- 170 20.50 *Warden*: ... and that'll solve that problem. But we have to continue until,  
171 until ... You see it's your job to make sure these things don't  
172 happen, to the best of your ability. You have as much responsibility  
173 for that as any of the other Guards on your shift. You're not as big  
174 and you may not be as loud. But you have to, you know, you have to  
175 try and give it what you've got ...
- 176 21.12 *Guard*: OK and you know ...
- 177 21.14 *Warden*: ... and get in there and pitch, and give them a hand because  
178 there aren't that many Guards. We're going to try and get a fourth  
179 Guard for this shift, 'cause we think we need it for the time being  
180 anyway. For today, to help handle. But that doesn't mean you can  
181 fade out into the background. Because this is your shift. OK?
- 182 21.36 *Guard*: OK.
- 183 21.37 *Warden*: ... So let's try and try and get in there and pitch and we'll have  
184 we'll have some of the other Guards on the shift help you, um...
- 185 21.48 *Guard*: Well, you know, you know.
- 186 21.50 *Warden*: And, do you... What I want to add is do you understand, um,  
187 the rationale behind doing something like this? The importance of it.
- 188 22.04 *Guard*: Um. "It" meaning what?
- 189 22.07 *Warden*: The study.
- 190 22.08 *Guard*: The experiment?
- 191 22.10 *Warden*: Yes.
- 192 22.12 *Guard*: Yeah. Yeah. You know I've done, probably, you know, about  
193 20 psych experiments here and in each case I thought they were

- 194 pretty similar. And so I, I pretty much understand this. But er, you  
 195 know, all I can do is try and get into it and do it the best I can as  
 196 myself.
- 197 22.38 *Warden*: No but I'm saying, you know, above and beyond the fact that  
 198 this is just a psych experiment, the importance of it in terms of,  
 199 especially for people who are concerned about the correctional  
 200 system which both Professor Zimbardo and I are. Um, you know, we  
 201 happen to agree with you that basically it's rotten and that it doesn't  
 202 produce the kind of rehabilitation, quote unquote, that you would  
 202 want to produce, and, um, and that perhaps, you know, some of the  
 203 better ways to produce rehabilitation is to remove the things in  
 204 society... You know, all that stuff. I, I am very deeply committed to  
 205 that.
- 206 23.18 *Guard*: OK.
- 207 23.19 *Warden*: ... Um, and I think all of us are. And what we're trying to do,  
 208 and this is what I want to try to get to you because I think you feel  
 209 the same way. Is that true, you know?
- 210 23.20 *Guard*: Yeah.
- 211 23.21 *Warden*: ... is, um, what we want to do is be able to, to study the thing  
 212 that exists, or as nearly as we can make it to what exists and to, and  
 213 to be able to go to the world with what we've done and say "Now  
 214 look, this is what happens when you have Guards who behave this  
 215 way".
- 216 23.56 *Guard*: OK.
- 217 23.58 *Warden*: But in order to say that we have to have Guards who behave  
 218 that way.
- 219 24.00 *Guard*: OK. Yeah. One other thing that I was thinking of is that, er, if it  
 220 came down to like a riotous situation in the prison that would be  
 221 pretty much analogous to, you know, any riot condition and, er, I've  
 222 been in a lot of riots, you know, at Stanford and Chicago, and one at  
 223 Grant Park last summer, and riots in France, in Paris, and of all the  
 224 times I've been in a riot, um, the times, well you know like in Grant  
 225 Park, the police just came right in, and that's not the worst one I've  
 226 ever seen for, you know, all that went on, and, er, the only time that  
 227 anything's ever been a particular potential bad situation nothing's  
 228 ever happened was in Paris where they had, they had something like  
 229 30,000 kids all over the streets of Paris and 15,000 police and, er,  
 230 everybody just sat there the whole day and then when the day ended  
 240 everybody went home. I guess you said it might come to something  
 241 like that later.
- 242 25:23 *Warden*: Well I hope it doesn't. I'm hoping it doesn't. And I don't think  
 243 it will if, we use a bit of sense, you know ... But, um, but for the time  
 244 being, you know, I, we need you to play the part of, you know, tough  
 245 Guard and forget some of the more sophisticated psychology that  
 246 you might know. And, um, and try and react as you picture the pigs  
 247 reacting.
- 248 26:00 *Guard*: Well I've met a lot of police that, er, that act a whole lot of  
 249 different ways. You can't do it just by the movies or something.  
 250 'Cause you know I've met plenty of police.
- 251 26:11 *Warden*: Well, you know, I'm not... We're trying to set up the

252                   stereotype Guard, not, you know ... Alright, and we realize that  
253                   you're going to have your own individual style. But, but so far your  
254                   individual style has been a little bit too soft pedal. And we want you  
255                   to get in there, OK?  
256 26:33 *Guard*: You mean just doing things like in the counts and stuff...  
257 26:36 *Warden*: ... In the counts and in other things. I mean when there's a  
258                   situation like this, um, to have to go in there and shout if necessary.  
259                   And, and to be more into the action. OK? Alright.  
[26.58 Ends]