



Rethinking the psychology of tyranny: The BBC prison study

Stephen Reicher^{1*} and S. Alexander Haslam^{2**}

¹University of St. Andrews

²University of Exeter

This paper presents findings from the British Broadcasting Corporation (BBC) prison study – an experimental case study that examined the consequences of randomly dividing men into groups of prisoners and guards within a specially constructed institution over a period of 8 days. Unlike the prisoners, the guards failed to identify with their role. This made the guards reluctant to impose their authority and they were eventually overcome by the prisoners. Participants then established an egalitarian social system. When this proved unsustainable, moves to impose a tyrannical regime met with weakening resistance. Empirical and theoretical analysis addresses the conditions under which people identify with the groups to which they are assigned and the social, organizational, and clinical consequences of either doing so or failing to do so. On the basis of these findings, a new framework for understanding tyranny is outlined. This suggests that it is powerlessness and the failure of groups that makes tyranny psychologically acceptable.

In the introduction to his text on *The Roots of Evil*, Staub writes: ‘the widespread hope and belief that human beings had become increasingly ‘civilized’ was shattered by the events of the Second World War, particularly the systematic, deliberate extermination of six million Jews by Hitler’s Third Reich’ (1989, p. 3). The impact of this realization was as marked in academia, and more particularly within academic psychology, as it was in society at large. Indeed, it is arguable that the shadow of the Holocaust lies over the last half century of social psychology and, either indirectly or directly, informs many of the core issues that are of concern to the discipline’s practitioners: questions such as how we come to hate and to discriminate against members of other groups (Tajfel, Flament, Billig, & Bundy, 1971), how people come to see others as less human and less deserving than themselves (Leyens *et al.*, 2003), how the seeds of authoritarianism, social dominance, and power abuse are sown and cultivated (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950), and – the question which concerns us most directly in this

*Correspondence should be addressed to Stephen Reicher, School of Psychology, University of St. Andrews, St. Andrews, Fife, KY16 9JU, UK (e-mail: sdr@st-andrews.ac.uk).

**Correspondence should be addressed to Alex Haslam, School of Psychology, University of Exeter, Exeter, Devon, EX4 4QG, UK (e-mail: A.Haslam@exeter.ac.uk).

paper – how we come to condone the tyranny of others or else act tyrannically ourselves.

While the literature in these various areas is both broad and varied, it is possible to identify at least one major trend. That is, there has been a shift away from explanations that focus on the individual characteristics of those who are prejudiced, discriminatory, or even genocidal, towards those that concentrate on the nature of group processes which can induce the most inoffensive of individuals to commit the most offensive of acts (Billig, 1978; Brown, 1965; Milgram, 1974; Sherif, 1966; Tajfel, 1982). In many cases, and certainly when it comes to the psychology of tyranny, theorists have taken the argument one step further and proposed not only that extreme antisocial behaviour must be analysed at the group level, but also that group psychology necessarily tends in the direction of extreme antisocial behaviour. While we strongly endorse the need for a group-level psychology of tyranny (which we define *as an unequal social system involving the arbitrary or oppressive use of power by one group or its agents over another*), we will take issue with the notion that groups *per se* are the root of the problem. Indeed, we will argue that powerful and effective groups provide an effective psychological bulwark against tyranny and that it is when groups prove ineffective that tyrannical forms of social organization begin to become attractive.

The equation of groups and tyranny has a long history both within social thought and within social psychology. Hannah Arendt (1998) describes the classical view that ‘the rule by many is not good’ and traces this back to Aristotle’s contention that collective rule leads to haphazardness, moral irresponsibility, and is but a disguised form of tyranny. More recently, such ideas were given substance by crowd psychologists such as Gustave LeBon (1895/1947) who argued that, through submergence in the crowd, individuals lose their individual identity and their sense of responsibility and hence become capable of barbaric and atavistic acts. The notion of submergence was directly transposed into the modern social psychological concept of de-individuation, which is seen to arise from anonymity within a group (Postmes & Spears, 1998; Reicher, Spears, & Postmes, 1995). As one influential de-individuation researcher has put it, ‘mythically, deindividuation is the ageless life force, the cycle of nature, the blood ties, the tribe, the female principle, the irrational, the impulsive, the anonymous chorus, the vengeful furies.’ (Zimbardo, 1969, p. 249).

Tyranny as role and power

Though well known as a de-individuation theorist, Zimbardo is better known for his work on the Stanford prison experiment (SPE; Haney, Banks, & Zimbardo, 1973; Zimbardo, 1989; Zimbardo, Maslach, & Haney, 1999). Indeed, this is one of the most famous social psychological investigations ever conducted, representing the culmination of a series of ‘classic’ field studies into the roots of extreme behaviours that were conducted in the aftermath of World War Two (Milgram, 1963; Sherif, 1956). Building on earlier studies, it played a critical part in cementing the shift that we have described from individual to group-level explanations of extreme behaviours (Banuazizi & Movahedi, 1975). Moreover, it was one of the few studies that not only addressed the issue of tyranny but, due to the power of the research paradigm, also produced direct evidence of tyrannical behaviour. While it is not, strictly speaking, a study of de-individuation, Zimbardo certainly used his general understanding of the group as a corrosive force to explain events in the SPE.

The SPE is remembered for showing that, simply as a consequence of assigning college students the role of guard or prisoner, the former became increasingly brutal while the latter became passive and began to show signs of psychological disturbance. Such was the severity of these phenomena that the study, originally scheduled to last 2 weeks, had to be stopped after 6 days. Zimbardo and colleagues explained their findings by commenting that guard 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.*, 1973, p. 12). Thus, immersion in a group is seen to undermine the constraints that normally operate upon people when they act as individuals. In addition, when those groups have power at their disposal, this is believed to encourage extreme antisocial behaviour (Zimbardo, 1969).

Although these findings were significant in their own right, the impact of the SPE was as much ethical as theoretical (e.g. see Smith & Mackie, 2000, p. 49). Indeed, the very extremity of the results led many (including Zimbardo himself) to question the legitimacy of subjecting participants to such situations. The acceptability of conducting any sort of large-scale field interventions thus became a focus for vigorous debate (e.g. Herrera, 1997; Lindsay & Adair, 1990; Sieber, Iannuzzo, & Rodriguez, 1995; Smith & Richardson, 1983). Paradoxically, then, at the same time that the SPE marked the culmination of post-war field studies, it also led to their cessation.

Accordingly, since the 1970s, social psychology has been increasingly dominated by laboratory experiments in which there is minimal or no interaction between participants and scant attention paid to the role of personal and group history or to the development of interactions over time (Bar Tal, 2004; Doosje, Spears, & Ellemers, 2002; Haslam & McGarty, 2001; Levine, 2003; Moreland, Argote, & Krishnan, 1996). Moreover, this unwillingness to undertake studies that create, manipulate and systematically investigate the effects of social environments on human interaction can be seen to have contributed to the increasing dominance of explanations based upon inherent and essentially unavoidable genetic, biological, or psychological propensities. It has also led to an increasing disjunction between the issues that motivate social psychological studies and the nature of those studies themselves. Research reports (and certainly most bids for research funding) typically start by alluding to large-scale topics such as oppression, discrimination, and genocide, but then go on to pursue an empirical strategy that seems very remote from the social realities of such phenomena (e.g. seeking to explain these phenomena in terms of individual-level subconscious processes from a cognitive or, more recently, neuroscientific perspective; Ito, Thompson, & Cacioppo, 2004). As Zimbardo (quoted in Brookes, 2001, p. 2) has argued, partly as a result of these trends, psychology has become increasingly marginal to, and marginalized from, debate surrounding important social issues.

In terms of the specific issue of tyranny, the ethical concerns that have placed the SPE 'off-limits' (with the exception of a partial replication by Lovibond, Mithiran, & Adams, 1979) have led to a situation in which the conclusions of that study have become almost inviolate and social psychological inquiry into tyranny has effectively ground to a halt. Barred from employing the power of the SPE paradigm, it is all but impossible to produce behaviours that are powerful enough to match those found by Zimbardo and his colleagues. Hence, even if researchers harbour doubts about the extreme situational determinism and negative views of the social group, which are used to explain these findings (and many do; e.g. see Ellemers, Spears, & Doosje, 1999; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Haslam, 2001), it has not been possible to produce data that

can give substance to those doubts and hence reopen scientific debate about the psychological bases of tyranny.

However, quite apart from the intrinsic importance of the topic, there are at least two sets of reasons why revisiting the SPE is long overdue. First, any assessment of the conclusions drawn from the SPE is inevitably limited by the fact that only a small proportion of the interactions in the study were recorded (because filming was intrusive and limited) and, of these, only a very small number are in the public domain. Moreover, observational data were never complemented by other data sources that would allow for controlled measurement of key behaviours and the psychological states seen to underlie them. At the very least, there is a need for a fuller and more transparent data set, which might progress empirically grounded and open debate about the psychological bases of tyranny.

However, even the limited amount of data that is available from the SPE casts doubt on the analytic conclusions that have been drawn from it. Where participants did behave in role, it is unclear whether, as Zimbardo and his colleagues claim, this was due to their 'natural' acceptance of role requirements or due to the *leadership* provided by the experimenters (Baron, 1984; Banuazizi & Movahedi, 1975). This is because during the study, the guards were given clear guidance as to how they should behave. Notably, when Zimbardo briefed his guards, he told them:

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 (Zimbardo, 1989).

The importance of such guidance is demonstrated by the research of Lovibond *et al.* (1979) who conducted a study in which the guards were trained to respect the prisoners as individuals and to include them in decision-making processes. Under these conditions, the ensuing behaviour of both guards and prisoners was far less aggressive and extreme (Lovibond *et al.*, 1979).

Yet, even with guidance, many of the participants in the SPE behaved out of role for much of the time (Baron, 1984). The available video material shows that both prisoners and guards challenged their roles not only at the start, but throughout the entire study. In the case of the guards, Zimbardo (1989) notes that, while some exploited their power, others sided with the prisoners and yet others were tough but fair. Such diversity sits uneasily with the notion that role acceptance is simply determined by the situation. It suggests that the emphasis on role acceptance and tyranny is one-sided and that there is a need to focus on (a) the conditions under which people do *or do not* assume their roles and (b) the balance between tyranny and *resistance*.

An alternative analysis: The social identity approach

It is not only that some of the data from the SPE appear to sit uneasily with a role account. Increasingly, the role account – and indeed the generally negative view that group membership leads to a loss of constraints on antisocial behaviour – is at odds with developments in group psychology. One of the most influential of these is the social identity approach (incorporating *social identity theory*; Tajfel, 1978, 1982; Tajfel & Turner, 1979; and *self-categorization theory*; Turner, 1985; Turner *et al.*, 1987; Turner, Oakes, Haslam, & McGarty, 1994). According to this approach, people do not

automatically act in terms of group memberships (or roles) ascribed by others. Rather, whether or not they do so depends upon whether they internalize such memberships as part of the self-concept (Turner, 1982).

Self-categorization theory in particular has argued that this act of self-definition in terms of group membership (social identification) forms the psychological basis of group behaviour and that the character of such behaviour depends upon the norms, values, and understandings that characterize the particular category in question (Turner, 1982, 1999). Thus, while members of certain groups may indeed use their power to act in discriminatory and oppressive ways, members of other groups may act more prosocially and use their power for constructive purposes (Pfeffer, 1981; Postmes & Spears, 1998). Moreover, even if some groups are tyrannical, group action is also the basis on which people gain the strength and confidence to resist, to challenge, and even to overthrow tyranny (Reicher, 1996; Tajfel, 1978).

Consistent with such emphasis, the greater part of early work informed by social identity theory has focused on the conditions under which people act to change inequalities between groups (e.g. Robinson, 1996). In broad terms, it is assumed that people who are positively valued by virtue of their group membership (e.g. members of dominant groups) would identify with and act in terms of the group. For people who are negatively valued by virtue of their group membership (e.g. members of subordinate groups), collective action is contingent upon two sets of factors in particular (Tajfel, 1978; Tajfel & Turner, 1979). The first relates to beliefs about one's ability to advance through the social system despite one's group membership (i.e. the *permeability* of category boundaries). The second concerns the perceived *security* of intergroup relations and comprises two further elements: the perceived fairness of intergroup inequalities (*legitimacy*) and their perceived *stability*. When relations are perceived to be insecure, this is characterized by the fact that individuals are aware of *cognitive alternatives* to the *status quo* and hence can envisage specific ways in which it could be changed.

Permeability affects whether people act individually or collectively, so that a belief that movement across boundaries is possible encourages strategies of individual mobility, but a belief that such movement is impossible encourages people to perceive themselves and act as group members (e.g. Ellemers, 1993; Wright, Taylor, & Moghaddam, 1990). Whether or not people then challenge inequality is also dependent upon intergroup relations being perceived as insecure. That is, people should be most inclined to resist domination when they see inequality as both illegitimate and unstable and can thus envisage cognitive alternatives to it (Turner & Brown, 1978; see also Ellemers, Wilke, & van Knippenberg, 1993; Tajfel, 1978; Tajfel & Turner, 1979).

It is important to stress that social identity theory does not constitute a comprehensive theory of domination and resistance. Most notably, it has little to say about the concomitants of identity processes (e.g. organizational and clinical factors), which may impact upon the ability of group members to act effectively. These are critical issues that we aim to investigate here. Nonetheless, the social identity approach provides a well-articulated and contemporary perspective from which to revisit the issues raised by the SPE: What are the psychological consequences of intergroup inequality? When do people seek to impose such inequality? And when do they resist it?

Before explaining how we addressed these issues, it is important to consider a second set of reasons for revisiting the SPE. These have less to do with the explanation of the findings themselves than with their broader social relevance. For Zimbardo and his colleagues, the results of the SPE were intended to bear directly on the nature of

prison regimes in the United States. Thus, they refer to the setting as a 'simulated prison' and claim that 'this simulated prison environment developed into a psychologically compelling prison environment' (Haney *et al.*, 1973, p. 69). However, as Banuazizi and Movahedi (1975) point out, there are good reasons to doubt these claims, firstly, because there are critical features of the SPE that are very different from a real prison (e.g. participants know they have committed no crime and can ask to leave at any time) and, secondly, because even where there are material similarities, the two are very different phenomenologically (e.g. where the walls of a prison remind the inmates that they must be kept apart from 'decent' people, they remind the participants in the SPE that they are honourable participants in adventurous scientific research).

While these arguments are in themselves controversial (e.g. Thayer & Saarni, 1975), our point is that, even if they are accepted, they do not render the SPE practically irrelevant. For there is another level beyond that of phenomenal equivalence through which the study can (and is) claimed to have real-world implications. That is, as is the case with most psychological research, generalization is theoretically, not empirically, based (Haslam & McGarty, 2004; Turner, 1981). Thus, Zimbardo (e.g. 2001) uses his study to establish the theoretical claim that people 'naturally' assume roles, and then uses this theoretical analysis to explain a wide range of phenomena from prison behaviour to terrorism. Of course, the resultant behaviours will not be the same as those observed in the SPE, but will depend upon the precise nature of the roles and role requirements in the relevant domain of application.

It is precisely this ability to apply role theory to a broad range of domains, despite phenomenal and behavioural differences, that has ensured the impact of the SPE within and beyond psychology amongst those who have little interest in matters relating to prisons. It is the validity of this theoretical account that we wish to address in the present paper, not only for the empirical and theoretical reasons discussed above, but also because we feel that the traditional analysis of the SPE has profound and troubling social implications. If people cannot help but act in terms of assigned role, it implies that they have little choice, and hence little responsibility, for their social actions. This makes it more difficult to hold tyrants to account for what they do. Moreover, in communicating the message that resistance is futile, the analysis discourages the oppressed from attempting to challenge tyranny.

The BBC prison study

In December 2001, we conducted a major social psychological field study (possibly the largest such study in the three decades since the SPE) in conjunction with the documentaries unit of British Broadcasting Corporation (BBC; hence the title 'the BBC prison study'). The study was independently designed, operationalized, run, and analysed by the authors. In this sense, it was an ordinary piece of scientific research, which went through all the normal scientific procedures (including ethical approval), and whose features were designed in relationship to the theoretical issues that concerned us. The contribution of the BBC was to coordinate and manage the logistical task of (a) creating the study environment (in line with our guidelines), (b) filming the study, and (c) preparing some of the resultant material for broadcast. In short, the project can be described as 'original science filmed'. This made it a unique collaboration that was markedly distinct from 'reality television' enterprises in which a television company devises certain scenarios with issues of entertainment in mind and then invites academics to comment on them.

Over a period of 8 days, the study examined the behaviour of 15 men who were placed in a social hierarchy of guards and prisoners within a purpose-built environment. Their behaviour was video- and audio-recorded over the entire period, and this was complemented by daily psychometric and physiological measures. The video data were edited into four 1-hour long documentaries screened in May 2002 (Koppel & Mirsky, 2002).

The aim of the study was not to simulate a prison (which, as in the SPE, would have been impossible on ethical and practical grounds) but rather to create an institution that in many ways resembled a prison (but also other hierarchical institutions such as a school, an office, a barracks; see Morgan, 1979) as a site to investigate the behaviour of groups that were unequal in terms of power, status, and resources. What is critical, then, is not that the study environment replicated a real prison (which no such environment ever could), but that it created inequalities between groups that were real to the participants.

Similarly, our aim was not conduct an exact replication of the SPE (which, would also have been impossible for ethical reasons). Rather, it was to use a different system of intergroup inequality in order to revisit the conceptual issues raised by the SPE. We therefore do not invite comparison with the SPE in terms of the exact details of how people behaved but rather in terms of the ability of different explanatory frameworks to make sense of what happened. Do participants accept their roles uncritically? Do those accorded group power exercise it without constraint, and do those without group power accept their subordination without complaint? After all, if the process of role enactment is indeed 'natural', then it should apply in all cases and any exception is troubling for the overall claim. Do the concepts used by social identity and self-categorization theorists provide a more satisfactory account of when people do (and do not) adopt the social positions ascribed to them?

Our study can thus be seen as an *experimental case study* of the behaviour of members in dominant or subordinate positions and of the developing relations between them. Unlike the SPE, it is not purely exploratory but rather is theoretically informed (by a social identity perspective). Hence, and again in contrast to the SPE, we included manipulations of theoretically relevant variables. Given the practical impossibility of running multiple sets of groups (due to the massive resources required in terms of equipment, personnel, and money), a time-series approach was adopted whereby interventions were introduced at predefined points in the study and their effects on the development of intra- and intergroup relations then analysed.

To summarize, the overall aims of our study were as follows:

- (a) To provide comprehensive and systematic data pertaining to the unfolding interactions between groups of unequal power and privilege.
- (b) To analyse the conditions under which people (i) define themselves in terms of their ascribed group memberships and act in terms of group identities, and (ii) accept or else challenge intergroup inequalities. Specifically, we predict that dominant group members will identify with their group from the start and impose their power. However, subordinate group members will only identify collectively and challenge intergroup inequalities to the extent that relations between groups are seen as impermeable and insecure.
- (c) To examine the relations between social, organizational and clinical factors in group behaviour.
- (d) To develop protocols that provide a practical and ethical framework for examining social psychological issues in large-scale studies.

Method and ethics

The study was designed to create a hierarchical society in which people would live for up to 10 days. It was conducted within an institutional environment that was constructed inside Elstree Film Studios in north London. Prisoners were allocated to lockable 3-person cells that were located, together with showers, off a central atrium. This was separated by a lockable steel mesh fence from the guards' quarters (a dormitory, bathroom, and mess room). A plan of the prison is presented in Figure 1.

Comprehensive details of the procedures are available in Haslam and Reicher (2002) or from the authors. What follows is a description of the key features of the study.

Ethics

The study aimed to create a system of intergroup inequality that was meaningful but was not harmful to participants either physically or mentally. To ensure that no harm eventuated, details of the experimental set-up and planned manipulations were discussed with colleagues and submitted both to the University of Exeter's ethics panel and to the Chair of The British Psychological Society's Ethics Committee prior to the study being conducted. The novelty of the experimental manipulations and theoretical analysis constituted central components of the scientific case that was presented in order to justify the research. Moreover, the following safeguards were built into the study:

- (a) Potential participants went through 3-phase clinical, medical, and background screening to ensure that they were neither psychologically vulnerable nor liable to put others at risk (see below).

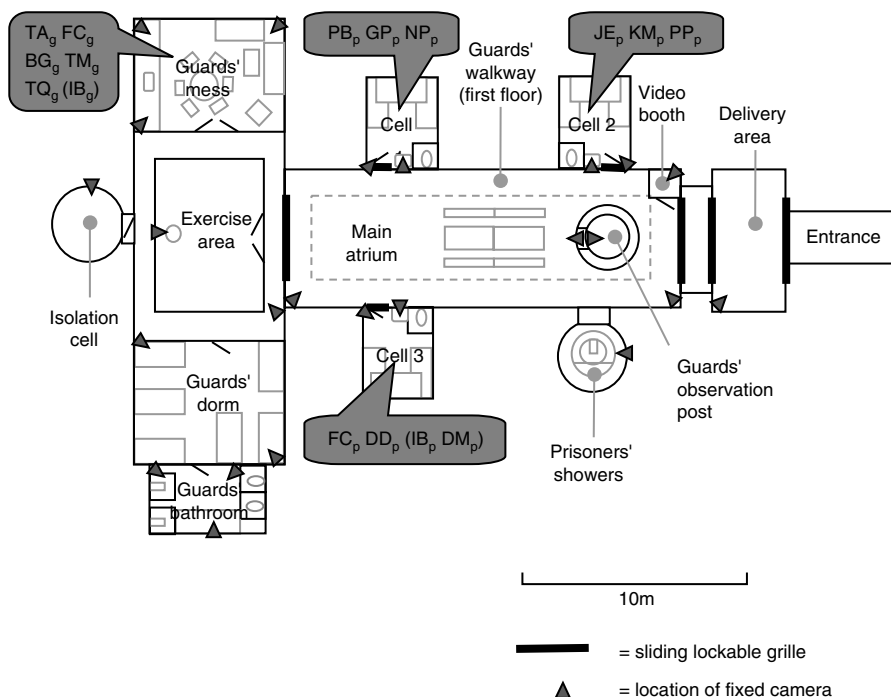


Figure 1. Plan of the prison.

- (b) Participants signed a comprehensive consent form. Amongst other things, this informed them that they may be subject to a series of factors – including physical and psychological discomfort, confinement, constant surveillance and stress – which may involve risk.
- (c) Two independent clinical psychologists monitored the study throughout, and had the right to see any participant at any time or to demand that any participant be removed from the study.
- (d) A paramedic was on constant standby in case of illness or injury.
- (e) On-site security guards were provided with detailed protocols clarifying when and how to intervene in cases of dangerous behaviours by participants.
- (f) An independent 5-person ethics committee – chaired by a British Member of Parliament – monitored the study throughout. This committee had the right to demand changes to the study's set-up or to terminate it at any time.

Apart from minor ailments (blisters, etc.) that were treated by the paramedic, no interventions were necessary to address ethical concerns. After the study, the ethical committee published an independent report (McDermott, Öpik, Smith, Taylor, & Wills, 2002) and characterized the conduct of the study as 'exemplary'.

Selection of participants

Male participants were recruited through advertisements in the national press and through leaflets. Applicants went through three phases of screening. First, they completed a battery of psychometric tests that measured both social variables (authoritarianism, social dominance, modern racism) and clinical variables (depression, anxiety, social isolation, paranoia, aggressiveness, demotivation, self-esteem, self-harm, drug dependence). Second, they underwent a full weekend assessment by independent clinical psychologists. Third, medical and character references were obtained, and police checks were conducted.

For ethical reasons (noted above), we sought to include in the study only people who were well-adjusted and prosocial, scoring at low levels on all social and clinical measures. Additionally, we wished to ensure that the individual dispositions of our participants were such that, if the dynamics of the study produced antisocial actions in this sample (as in the SPE), then it could reasonably be supposed that they would have such an effect upon almost anybody in the population.

The screening reduced an initial pool of 332 applicants to 27 men (we recruited only men to ensure comparability with the SPE and to avoid ethical issues that would arise from placing men and women together in cells). The final sample of 15 was chosen to ensure diversity of age, social class, and ethnic background. They were randomly divided into two groups of 5 guards and 10 prisoners but in such a way as to ensure that the two groups were matched on key dimensions. More specifically, the 15 participants were first divided into five groups of 3 people who were as closely matched as possible on personality variables potentially implicated in tyranny: modern racism, authoritarianism and social dominance.¹ From each group of three, one participant was then randomly selected to be a guard (and the remaining two to be prisoners). This procedure

¹ This procedure – which ensures equivalence on theoretically relevant variables – was considered superior to 'pure' random assignment (i.e. one which does not take account of any individual differences, as in the SPE), as the study's relatively small sample size means that random assignment has the potential to produce large differences between groups (due to the law of large numbers; e.g. Haslam & McGarty, 2003, pp. 180–183).

was conducted blind (i.e. the identities of the participants were not known to the experimenters). Note, however, that due to our experimental manipulations (explained below), the actual number of people in the two groups varied over time. Thus, at the start of the study, there were 5 guards but only 9 prisoners – the 10th being introduced at a later stage.

Data sources

The prison environment was designed in such a way that participants could be both video- and audio-recorded wherever they were. At all times, four channels were video-recorded and all audio channels were recorded. There was also daily psychometric testing. Measures were taken from a battery of scales. For the purposes of the present analyses the critical measures were:

- (a) *social variables*: social identification, awareness of cognitive alternatives, right-wing authoritarianism;
- (b) *organizational variables*: compliance with rules, organizational citizenship; and
- (c) *clinical variables*: self-efficacy, depression.

In order to minimize fatigue, not every measure in the full battery was administered every day. However, each was administered on multiple occasions to allow for an analysis of development over time. Finally, daily swabs of saliva were taken in order to ascertain cortisol levels (as an indication of stress; Laudat *et al.*, 1988). For reasons of space, however, we will not consider the cortisol data here (see Haslam & Reicher, *in press*).

Set-up

Five participants were invited to a hotel the evening before they entered the prison. On arrival, they were told that they would be guards in the study. They were shown the prison timetable – which included such elements as cleaning chores, work duties, prisoner roll calls, exercise time and a recreational hour – and were told that their responsibility was to ensure that the institution ran as smoothly as possible and that the prisoners performed all their tasks. The five guards were then asked to draw up a series of prison rules under headings provided by the experimenters and to draw up a series of punishments for rule violations.

The guards were given no guidance about how they should achieve their goals. The only limits on what they could do were a set of ethically determined ‘basic rights’ for prisoners. In particular (and as in the SPE), all participants were told that physical violence would not be tolerated (for rules, punishments, and rights, see Haslam & Reicher, 2002). Beyond this, however, it was stressed that the guards could act as they pleased.

On the morning of the study itself, the guards were taken in a blacked-out van to the prison (since this was meant to be their entire experiential world for the duration of the study, it was important that they could not imagine the outside). Once inside, they were given a full briefing by the experimenters on the prison layout and the resources available to them.

The guards had a series of means by which to enforce their authority, including keys to all doors inside the prison (including a punishment isolation cell), sole access to an

upper level, a 'guards' station' with a surveillance system from which they could see into the prisoners' cells, resources (including snacks and cigarettes) to use as rewards or withdraw as punishments – and, in addition, the ability to put prisoners on a bread and water diet. They also had far better conditions than the prisoners, including superior meals, extra supplies of drinks and snacks, superior living conditions and well-made uniforms as opposed to the prisoners' uniform of a t-shirt printed with a 3-digit number, loose trousers and flimsy sandals. The prisoners also had their hair shaved on arrival.

After their briefing, the guards changed into their uniforms and practiced the procedure for admitting the prisoners. The nine prisoners then arrived one at a time. They were given no information apart from the prison rules, a list of prisoners' rights, which was posted in their cells, and a very brief loudspeaker announcement from the experimenters. This introduced the permeability intervention (see below) and stressed that violence was not permissible.

Planned interventions

Permeability

At their initial briefing, the guards were told that they had been selected on the basis of their reliability, trustworthiness and initiative as gleaned from pre-selection assessment scales. However, they were also told that while these scales were reasonably reliable, they were not perfect. In particular, the experimenters stated that it was possible that they had misassigned one or more of the prisoners. Hence, the guards were told that they should observe the behaviour of the prisoners to see if anyone showed guard-like qualities. If they did, they were told that there was provision for a promotion to be made on Day 3. This information was also announced to the prisoners over the loudspeaker. In the initial days of the study, participants were thus led to believe that movement between groups was possible (see also Ellemers, van Knippenberg, & Wilke, 1990; Wright *et al.*, 1990). After the promotion of one prisoner to guard actually took place (the selection of the individual being made by the guards on the basis of a procedure suggested by the experimenters), the possibility of movement was removed by announcing that there would be no further promotions (or demotions).

Legitimacy

Three days after the promotion, participants were to be informed by the experimenters that observations had revealed that there were in fact no differences between guards and prisoners on the key group-defining qualities. However, they would be told that it was impractical to reassign them and hence the groups would be kept as they were. Accordingly, whereas previously the group division had been legitimate, this would no longer be the case (Ellemers *et al.*, 1993).

Cognitive alternatives

Within a day of the legitimacy intervention, a new prisoner was to be introduced. Although he was as naïve as the others, he was chosen for this role (from the pool of 10 participants randomly assigned to be prisoners) because of his background as an experienced trade union official. On this basis, we expected that he would introduce a new perspective to the prison based on notions of group-based negotiation and collective- and equal rights (i.e. a perspective that suggested the existing regime was both illegitimate and changeable). It was also thought that he might provide skills

necessary to organize collective action (e.g. see Haslam, 2001). Hence, it was envisaged that his introduction would enable the prisoners (and the participants more generally) to envisage the achievement of a more equal set of social relations.

Results

The findings of the study can be divided into two phases. In the first phase, the guards failed to identify with each other as a group and to cohere collectively. By contrast, after the promotion on Day 3, the prisoners did increasingly identify as a group and work collectively to challenge the guards. This led to a shift of power and ultimately to the collapse of the prisoner-guard system. In the second phase, participants decided to continue as a single self-governing 'commune'. However, they were unable to deal with internal dissent and lost confidence in the communal system. By the end of the study, they were increasingly disposed to tolerate a new and much more draconian system of inequality that some participants now wished to impose.

Results that pertain to these two phases will be presented in turn. For each phase, we combine a description based on the observational data with statistical analyses based on the quantitative measures. The statistical analyses are based on the data of individual participants within groups (excluding data from the participant who was promoted from prisoner to guard on Day 3 and from the prisoner who was introduced on Day 5). Given the interaction between participants, it could be argued that the group, rather than the individual group member, should be the unit of analysis here. For this reason, the present data were also analysed using methods advocated by McGarty and Smithson (2005), which do not assume (or require) independence of observations. These analyses confirmed the reliability of all the patterns reported below. However, for reasons of space and in light of the novelty of these alternative methods, we accord with general usage by presenting statistics that are commonly used even where there is interaction between participants (Hoyle, Georgesen, & Webster, 2001).

Phase 1: Rejecting inequality

Social identification

For the prisoners, the development of social identification was consistent with predictions. From the start, they were clearly dissatisfied by their inferior conditions. Initially though, as predicted in light of the permeability of group boundaries, many sought to improve their lot by displaying the individual qualities necessary for promotion as opposed to mounting a collective challenge to the guards. As a result, there was no shared identity among the prisoners and no consensus about how they should behave (see Photograph 1).

However, after the promotion, when group boundaries were impermeable and participants could only alter their position by changing the general prisoner-guard relationship, the prisoners began to develop a much stronger sense of shared social identity and to develop more consensual norms – particularly in relation to their treatment of the guards. This contrast between the pre- and post-promotion periods is exemplified in exchanges between participants in Cell 2. Before the promotion, two occupants of this cell, JE_p² and KM_p, worked conscientiously and explicitly sought to improve their position by displaying behaviours required to become a guard. As JE_p put

² The subscripts _p and _g after participants' initials denote their status as prisoner or guard, respectively.



Photograph 1. Day 2: The prisoners comply with the guards' regime.

it, 'I'd like to be a guard because they get all the luxuries and we are not'. However, almost immediately after the promotion, all three occupants of this cell (i.e. including PP_p) recognized that the only way to improve their position was to change the system. Accordingly, they began to discuss how they could achieve this together. If anyone expressed doubts about this objective, they were reminded of the collective inequalities in the prison and of the need for collective resistance. This is exemplified by the following interchange between the cellmates:

Extract 1

JE_p: Hopefully we'll get [TQ_g] in. That's the person, he's the target.

KM_p: No. I mean obviously I think it's going to be a lot of fun for us to do this but I don't think [TQ_g] . . . I feel so . . . I just feel . . .

PP_p: Listen, listen mate I, you've got to, you've got to start forgetting about other people's feelings and what they're doing because the days when you're sitting here starving hungry and you've got fuck all and you've got nothing mate and you've got a ratty little bed and a stupid little blanket to sit under and they're under there in their duvets, they've got everything they want and they're not giving two fucks about you. So – think on and *fuck them*.

KM_p: I think they do care about us. But guys I'm going to back you all the way. You should no' doubt me.

For the prisoners, then, the promotion led to a perception of impermeability. This was accompanied by a shift from individual action and identification (i.e. a stress on what 'I' will do) to collective action and identification (i.e. discussion of what 'we' will do) and from compliance to conflict with the guards.

For the guards themselves, the results were very different. Moreover, the patterning of their social identification went against our predictions. Along lines reported by Zimbardo and his colleagues (Haney *et al.*, 1973), we had expected that, from the outset, they would identify with what was a high-status and positively valued group within the prison. There was some evidence of this in the first day of the study, but it was also clear that several guards were wary of assuming and exerting their authority.

Consequently, some guards were always ambivalent about internalizing their assigned identity. Even to the extent that they overcame this ambivalence, the group as a whole could never reach consensus about its norms and priorities. In addition, over

time – particularly once the boundaries between-groups became impermeable and the prisoners had developed a stronger sense of shared identity – these cleavages were exacerbated and the fragility of the guards' collective identity became more apparent.

However, as noted above, in contrast to the SPE, our systematic collection of quantitative data allowed us to complement behavioural observations that speak to these patterns with inferential statistical analysis (and hence to triangulate behavioural and psychometric data). To this end, social identification with prisoner and guard groups was measured every day by means of two 3-item scales containing items used widely in previous research (e.g. Doosje, Ellemers, & Spears, 1995; 'I feel strong ties with the prisoners [guards]'/ 'I identify with the prisoners [guards]/I feel solidarity with the prisoners [guards]'). As with all other measures, responses were made on 7-item scales with appropriately labelled end-points (e.g. *not at all*, *extremely*).

Identification scores were computed by averaging responses on the three measures that comprised each scale and subtracting the out-group identification score from the in-group identification score. Note too that, as with all other measures, data were collected early each morning. Accordingly, scores relate primarily to events of the previous day.

Given the small number of participants, the study inevitably has a low level of statistical power. In light of this, statistical significance was considered in conjunction with effect size (Cohen, 1977; Smithson, 2000). For all analyses, effects were only considered meaningful when, as well as being significant at conventional levels ($\alpha = .05$), effect size was strong by Cohen's (1977, p. 283) criteria (i.e. $\eta^2 > .14$).

Figure 2 presents mean social identification scores as a function of participant group and time. As can be seen from this graph, the data here are consistent with the above observations. In the very first days of the study, the guards identified more strongly with their social position than the prisoners. However, as soon as they had to implement the disciplinary regime (from Day 2 onwards), the guards' identification fell while that of the prisoners rose. Prisoner identification was also particularly high after the promotion (Day 4).

Statistical analysis of in-group identification scores was conducted by means of a 2 (participant group: guards, prisoners) $\times$ 6 (study phase: Day 1 to Day 6) ANOVA with repeated measures on the second factor. This analysis supported observational data in revealing a significant and large interaction between group and time, $F(5, 55) = 3.05$, $p < .05$; $\eta^2 = .22$. Analyses of polynomial trends were used as the most practical way of decomposing this effect (Norusis, 1985). Consistent with observations, these showed

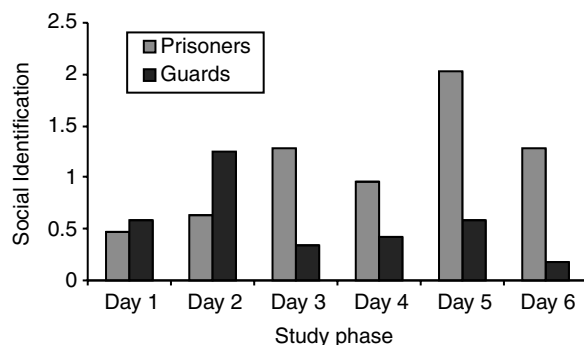


Figure 2. Social identification as a function of assigned group and time.

that social identification among the prisoners increased linearly over time, $t(7) = 2.46$, $p < .05$. On the other hand, identification among the guards declined as the study progressed, but non-significantly, $t(4) = -0.77$, *ns*.³

Security of intergroup relations

After the promotion, the normative consensus among prisoners led to effective organization based on both the expectation and the provision of mutual social support between group members (Haslam, O'Brien, Jetten, Vormedal, & Penna, 2005; House, 1981; Underwood, 2000; see Haslam & Reicher, in press, for a fuller treatment of this point). Conversely, the guards' inability to agree on group norms and priorities meant that they could not trust others to represent them or act appropriately. As a consequence, they felt (and were) weak, inconsistent, and ineffective as a group.

These developments directly undermined the perceived legitimacy of the intergroup inequality as the guards showed no evidence of the qualities by which their selection was supposedly justified. Such developments also contributed directly to the emergence of cognitive alternatives; not only could the prisoners envisage changes to the existing hierarchy, but they also had a growing confidence in their ability to achieve them. Consider, for instance, the outcome of the first confrontation between prisoners and guards. The three prisoners of Cell 2 had orchestrated this mainly to see how the guards would respond when challenged. As part of a pre-arranged plan, JE_p threw his lunch plate to the ground and demanded better food. As the guards tried to intervene, KM_p and PP_p then joined in with further demands for smoking rights and treatment for a blister. The guards were totally disunited in their response, with some wanting to take a disciplinarian line and others wanting to make concessions (see Photographs 2 and 3). This continued until one guard, TQ_g, decided to resolve the situation by acceding to PP_p's requests for a cigarette in order to encourage the other prisoners to return to their cell.

After the incident, the prisoners back in their cell and the guards back in their mess expressed a growing realization that the system was open to change. However, they did so with a very different evaluative tone. The prisoners literally danced with joy. PP_p exclaimed 'that was fucking sweet', to which KM_p responded (admiringly) 'you was fucking quality man'. Then all three cellmates exchanged 'high fives'. The guards bickered despondently. TA_g started by saying 'This is only Day 4. They can see what happened today and now they know they can do whatever they want'. As BG_g and then TQ_g weakly disagreed, TA_g only insisted more, culminating in the claim that 'It's happened. I mean, come on. You know what that has done? That has lit the fuse on [PB_p's] arse'.

Hence, the induction of insecure relations between the groups was not dependent upon our interventions but was an emergent property of the intergroup dynamics. Correspondingly, planned interventions were not necessary for the prisoners to start challenging the guards. The challenge started mounting immediately after the promotion. Consequently, the legitimacy intervention was not implemented. However, because the new prisoner had already been told he would be brought in, this went ahead on Day 5 (and he was later withdrawn on Day 6). Given the situation, he did not so much suggest cognitive alternatives where none existed as suggest additional alternatives to the *status quo*. Most notably, he began to question the prisoner-guard

³ In this instance (and with depression data below), analysis of polynomial contrasts was based on a random subset of five phases because the number of levels of the within-subjects variable (study phase) exceeded the number of subjects.



Photographs 2 and 3. Day 4: The prisoners start to challenge the guards' regime.

division itself and, by questioning the legitimacy of certain aspects of the study (particularly the heat in the prison), he suggested both to his fellow cellmates (DM_p and FC_p) and to one of the guards (TQ_g) that the participants mount a united challenge to the experimenters. As the following extract shows, this provided new ways of envisaging how the system might work:

Extract 2

DM_p: If this was a real-life situation . . .

TQ_g: Yes.

DM_p: . . . and you were working in this kind of heat, then you as an employee could well go to the employer and say 'The condition is unacceptable, I'm not prepared to work in it'. Now let's treat this as a real-life situation. You and I – your group and the group I'm in – both have this problem of the heat. And if I've got to sleep in this, there is no way I will. And I, you know, won't bear it. And I think collectively we should do something about it to the people who are running the experiment. Now you know in a normal, day-to-day, real-life situation, that's what would happen.

TQ_g: Well, I am most impressed with your new-found kind of angle on this, which possibly shouldn't come as a surprise to me. But I think that is a very very valid point you are making and I'm going to go along with it completely.

As a first practical step towards participant unity, DM_p created a new negotiating structure, which brought guards and prisoners together on a basis that was far more equal than had hitherto been the case. The guards were eager and pleased to accept this arrangement, for even if they surrendered much of their hierarchical advantage by doing so, the new system confirmed their position in what promised to be a viable social order.

Quantitative confirmation of these patterns emerges on measures of participants' awareness of cognitive alternatives to the *status quo* (Tajfel & Turner, 1979). These were administered on Days 1, 3, 4, and 6 using a 4-item scale ('I cannot imagine the relationship between guards and prisoners being any different (reverse-scored)'/ 'I think that the guards will always have more privileges than the prisoners (reverse-scored)'/ 'I think that the relationship between guards and prisoners is likely to change'/ 'I think that it would be possible for the prisoners to have more power than the guards'). A single score was computed by averaging responses to these 4 items.

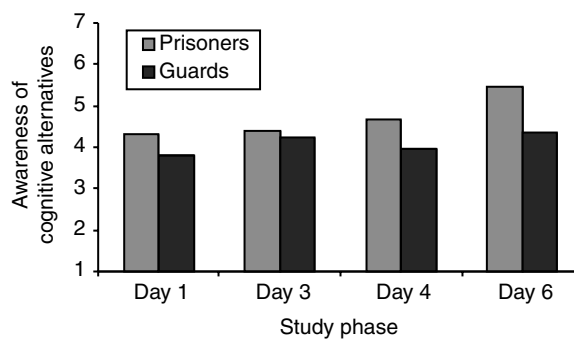


Figure 3. Awareness of cognitive alternatives as a function of assigned group and time.

Mean scores are presented in Figure 3 as a function of participant group and time. As can be seen from this graph, at the start of the study, prisoners and guards had an equal and relatively low sense of cognitive alternatives, but this increased as the study progressed. These patterns were confirmed by statistical analysis of cognitive alternatives data that was conducted by means of a 2 (participant group: guards, prisoners) $\times$ 4 (study phase: Days 1, 3, 4, 6) ANOVA with repeated measures on the second factor. This analysis revealed a significant and large main effect for time, $F(3, 33) = 3.36, p < .05; \eta^2 = .34$. Consistent with the above observations, analysis of polynomial trends to decompose this effect indicated that participants' sense of cognitive alternatives increased linearly over time, $t(12) = 2.83, p < .05$.

Acceptance of the unequal regime: Compliance and organizational citizenship

Associated with the above effects, as the prisoners developed a sense of shared social identity that was defined in opposition to the guards, and as they became aware of the possibility of alternatives to existing status relations, they started to work actively against the regime that the guards were trying to uphold. This was manifest both in minor challenges to the guards' status (e.g. insubordination during roll call) and in more overt challenges – including the incident alluded to above, in which a number of prisoners collectively protested about the quality of their food.

Data consistent with this analysis emerge from measures of (a) participants' willingness to comply with authority and (b) their willingness to engage in acts of organizational citizenship (i.e. to do more than was asked of them in order to make the prison system work; Organ, 1988, 1997).

Compliance was measured on Days 1, 3 and 5 by means of a 2-item scale ('I try to do what the guards want'/'I try to comply with the guards' rules'). Responses on these measures were averaged to provide a single score. The resultant mean scores are presented in Figure 4 as a function of participant group and time. As can be seen from this graph, in the early stages of the study, both guards and prisoners were willing to comply with the rules, but the prisoners became more reluctant to comply after the promotion.

These patterns were confirmed by statistical analysis of compliance scores that was conducted by means of a 2 (participant group: guards, prisoners) $\times$ 3 (study phase: Days 1, 3, 5) ANOVA with repeated measures on the second factor. This revealed a significant

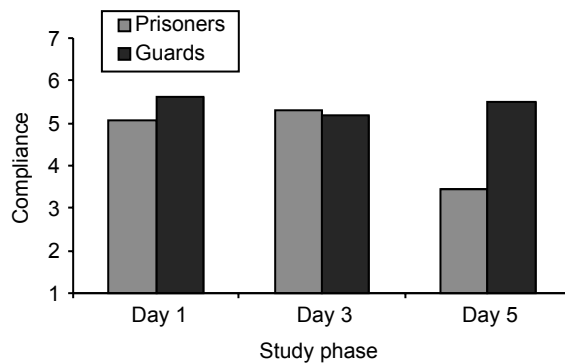


Figure 4. Compliance with prison rules as a function of assigned group and time.

and large interaction between group and time; $F(2, 22) = 3.62$, $p < .05$; $\eta^2 = .32$. Analyses of polynomial trends indicated that compliance on the part of the guards did not vary significantly over time, $t(4) = -0.12$, *ns*, but that of the prisoners declined in a linear fashion, $t(7) = -4.48$, $p < .01$. There was also evidence of a quadratic trend, $t(7) = -2.18$, $p < .07$, suggesting that the prisoners' decline in compliance was particularly marked after group boundaries had become impermeable.

Organizational citizenship (Organ, 1988, 1997) was measured on Days 2, 4, and 5 by means of a 3-item scale containing items used widely in previous research (e.g. Haslam, Powell, & Turner, 2000; Tyler & Blader, 2000; 'I am willing to do more than is asked of me by the guards'/'I will do whatever I can to help the guards'/'Whenever possible I will try to make the guards' work difficult'; reverse scored). A single score was computed by averaging responses to these 3 items.

The resultant mean scores are presented in Figure 5 as a function of participant group and time. As can be seen from this graph, the guards were always more willing than the prisoners to engage in citizenship behaviours that would help them run the regime. However, while the guards maintained their willingness to work for the regime across the three testing phases, over time the prisoners became much more reluctant to support the guards' regime in this way. Again, this was particularly true after the promotion on Day 3.

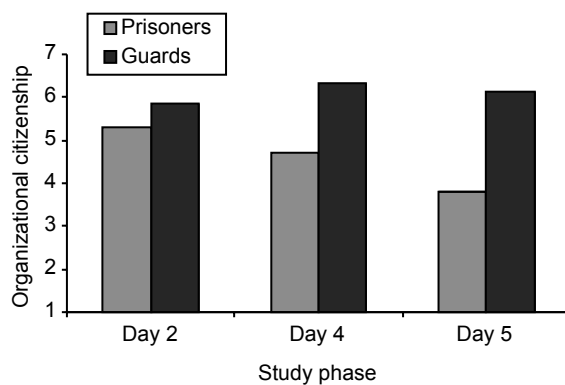


Figure 5. Organizational citizenship as a function of assigned group and time.

These patterns were confirmed by statistical analysis of organizational citizenship scores performed by means of a 2 (participant group: guards, prisoners) $\times$ 3 (study phase: Days 2, 4, 5) ANOVA with repeated measures on the second factor. This analysis revealed a significant and large main effect for group, $F(1, 11) = 7.84, p < .01; \eta^2 = .42$, but this was qualified by a significant and large interaction between group and time; $F(2, 22) = 5.10, p < .05; \eta^2 = .32$. Analyses of polynomial trends revealed that the guards' willingness to display organizational citizenship did not vary across the three phases of testing, $t(4) = 0.69, ns$, but that of the prisoners declined linearly over time, $t(7) = -3.74, p < .01$.

Collective self-efficacy and mental health

The disorganization of the guards and the unity of the prisoners did not just produce conditions where the latter challenged the former; it also led to the prisoners becoming progressively more dominant. Through planning and mutual support, they became increasingly more extreme and more successful in their efforts to undermine the guards' control. By contrast, the guards' attempts to impose their authority became increasingly unsuccessful, and they became more divided and mutually recriminatory. Moreover, as we have already illustrated with reference to observational data, the effectiveness of the prisoners in pursuing their collective goals led to strong positive affect while the inability of the guards to act collectively led to despondency.

These dynamics are reflected in quantitative data obtained on measures of (a) collective self-efficacy and (b) depression. Collective self-efficacy (Bandura, 1995) was measured on Days 2, 4, and 6 by means of a 5-item scale containing items used widely in previous research (e.g. Chen & Bliese, 2002; 'My prison group is confident that we could deal efficiently with unexpected events'/'My prison group can remain calm when facing difficulties because we can rely on our coping abilities'/'My prison group can always manage to solve difficult problems if we try hard enough'/'When my prison group is confronted with a problem, we can usually find several solutions'/'My prison group can usually handle whatever comes our way'). A single score was computed by averaging responses to these 5 items.

Figure 6 presents the resultant mean collective self-efficacy scores as a function of participant group and time. As can be seen from this graph, these data mirror those obtained on measures of social identification, so that, at the start of the study, the guards

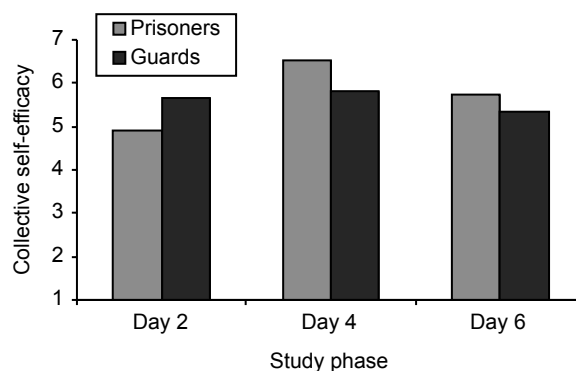


Figure 6. Collective self-efficacy as a function of assigned group and time.

had a greater sense of collective self-efficacy than the prisoners. However, by Day 4, the prisoners' self-efficacy had increased markedly and was now greater than that of the guards. This pattern was maintained at Day 6, although it is interesting to note that after the trade unionist was removed from the prison, both groups showed some decline in self-efficacy (seemingly because the opportunities for group-based negotiation and order that he presented were removed; see Eggins, Haslam, & Reynolds, 2002; Haslam, Eggins, & Reynolds, 2003). These patterns are consistent with our observational account – although the general decline in the guards' sense of self-efficacy is less pronounced in these quantitative data.

Collective self-efficacy scores were statistically analysed by means of a 2 (participant group: guards, prisoners) $\times$ 3 (study phase: Days 2, 4, 6) ANOVA with repeated measures on the second factor. This analysis revealed a significant main effect for time, $F(2, 22) = 6.08$, $p < .01$; $\eta^2 = .36$, that was conditioned by a significant and large interaction between group and time, $F(2, 22) = 4.31$, $p < .05$; $\eta^2 = .28$. In line with the above observations, analyses of polynomial trends revealed that the prisoners' sense of collective self-efficacy increased linearly over time, $t(7) = 2.55$, $p < .05$ – an effect that was also qualified by significant quadratic variation, $t(7) = 4.25$, $p < .01$. On the other hand, the collective self-efficacy of guards declined over time, although this effect was not significant, $t(4) = -0.90$, *ns*.

Depression was measured during the screening process and then every day during the study (partly in order to monitor the ongoing welfare of participants). For this purpose, a 7-item scale was administered ('In general, how has your mood been over the last few days?'/ 'Do you ever feel low or depressed?'/ 'Do you feel hopeless about the future?'/ 'Do you have difficulty dealing with everyday problems?'/ 'Are you self-confident?'/ 'Do you think that you are a worthwhile person?'/ 'Do you think about harming yourself?'). During the screening phase ($N = 332$), these items were found to form a reliable scale ($\alpha = .80$) and so the items were averaged to form a single measure.

Mean depression scores are presented in Figure 7. From this graph, it can be seen that although overall levels of depression were low, they clearly varied as a function of participant group and time. Specifically, while the prisoners were more depressed than the guards at the start of the study, by its end, this situation had reversed. This pattern was confirmed by a 2 (participant group: guards, prisoners) $\times$ 7 (study phase: pre-test, Days 1 to 6) ANOVA with repeated measures on the second factor. This revealed a significant and large interaction between group and time; $F(6, 66) = 3.73$, $p < .01$;

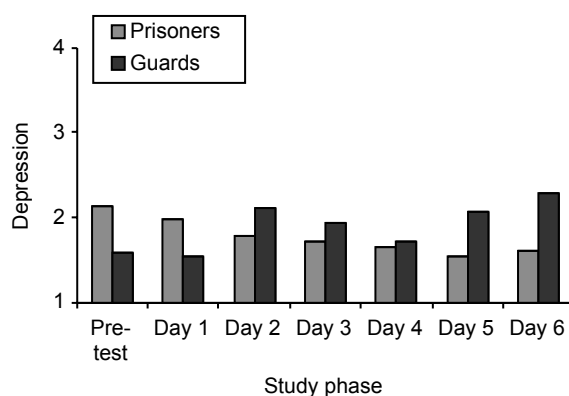


Figure 7. Depression as a function of assigned group and time.

$\eta^2 = .25$. Consistent with observations, polynomial contrasts showed that the prisoners' depression decreased linearly over time, $t(7) = -4.22$, $p < .01$. At the same time the guards' depression increased, though not significantly, $t(4) = 1.25$, *ns*.

Combined impact

The combined outcome of these interrelated dynamics was that as the prisoners became increasingly aware of, and confident in, their collective identity, and as the guards' confidence and collective self-efficacy declined, the prisoners developed and executed plans to take on and destroy the guards' regime. The guards were divided, exhausted and demoralized. They were unable to organize themselves or the prisoners effectively. As a result, late on the evening of Day 6, the prisoners in Cell 2 broke out of their cell and occupied the guards' quarters. At this point, the guards' regime was seen by all to be unworkable and at an end (see Photograph 4).

Phase 2: Embracing inequality

After the collapse of the guard-prisoner hierarchy, all but two of the participants decided that they wanted the study to continue through the institution of a single self-governing commune. With FC_p in the chair, they met with the experimenters and drew up the terms under which the commune would operate. Initially, the new system was highly effective. A number of participants who had been mutually hostile when they were divided into prisoners and guards formed strong and positive affective ties now they were recategorized part of a common group (Gaertner, Mann, Murrell, & Dovidio, 1989; Turner *et al.*, 1987). Collectively, they performed their work tasks and chores with more effort and to a higher standard than at any other point in the study. However, those participants who had been centrally involved in challenging the old regime felt marginalized in this new system. At first, they simply failed to contribute to collective tasks. Later, they began to violate communal rules. Within a day, they were plotting to destroy the commune. The problems this created were exacerbated because the commune's members had never developed procedures for dealing with dissidence, and hence they had no means of responding to threats to their social order.

By the morning of the commune's second day, it was clear to many participants that the new social structure was in crisis. This situation was exacerbated when, by chance,



Photograph 4. Day 6: Cell 2 break out bringing the guards' regime to an end.

the breakfast was of very poor quality. This was (incorrectly) taken as a sign that the experimenters disapproved of the commune system. It led to despondency among the commune's supporters who felt, as one later put it, that they had the burden of responsibility for the system without the ability to make it work. Moreover, the emergent crisis was exploited by opponents of the commune, four of whom (one ex-guard and three ex-prisoners) formulated a plan to create a new and harsher guard-prisoner hierarchy. The nature and tone of this new regime was made abundantly clear in discussions about the form this would take. As PB_p put it, 'We want to be the guards and fucking make them toe the line, I mean *on the fucking line*. No fucking talking while you are eating. Get on with your food and get the fucking hell back to your cell'.

Shortly after breakfast, this group convened a meeting (see Photograph 5). Their leader berated the commune and its supporters and he introduced the idea of the new hierarchy. The supporters of the commune were largely passive in response. They looked despondent and listened in silence until he had nearly finished. During debriefings, a number of them acknowledged that, although they would not have openly endorsed such a hierarchy, they were less opposed to it than they had been previously and that they felt less repulsed by the idea of a strong social order in which someone else assumed responsibility for making the system work.

Again, these trends are supported by analysis of psychometric data. Most notably, this is apparent when one looks at the level of participants' right-wing authoritarianism (Altmeyer, 1981, 1996). This was measured using 8 items abstracted from a 30-item scale administered during pre-testing (Reynolds, Turner, Haslam, & Ryan, 2001; 'Things would go better if people talked less and worked harder'/'It is better to live in a society in which the laws are vigorously enforced than to give people too much freedom'/'People should always comply with the decisions of the majority'/'You have to give up an idea when important people think otherwise'/'There are two kinds of people: strong and weak'/'What we need are strong leaders that the people can trust'/'Our social problems would be solved if, in one way or another, we could get rid of weak and dishonest people'/'People should always keep to the rules'). In pre-testing ($N = 332$), this scale was found to be reliable ($\alpha = .71$) and to correlate very highly with the full 30-item scale ($r = .97$). Accordingly, scores on the above items were averaged to form a single measure. This was administered during the screening process and then on Days 1, 3 and 7.



Photograph 5. Day 8: New guards make the case for an authoritarian regime.

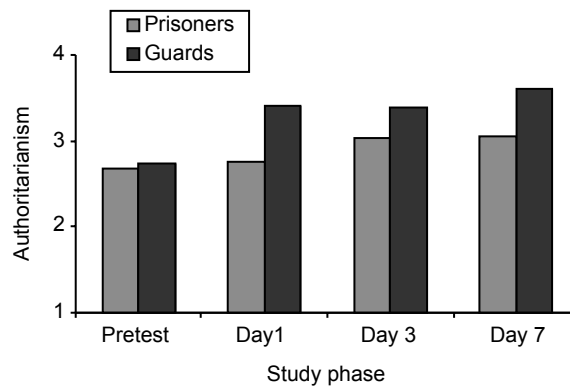


Figure 8. Right-wing authoritarianism as a function of assigned group and time.

In the first instance, it is relevant to examine levels of authoritarianism over time as a function of the groups to which participants were assigned by the experimenters. Data pertaining to this analysis are presented in Figure 8. These were analysed by means of a 2 (participant group: guards, prisoners) $\times$ 4 (study phase: pre-test, Days 1, 3, 7) ANOVA with repeated measures on the second factor. Consistent with qualitative observations, the only effect to emerge from this was a significant and large effect for time, $F(3, 33) = 2.75, p < .05; \eta^2 = .20$. Analysis of polynomial trends revealed that there was a linear increase in all participants' authoritarianism as the study progressed, $t(12) = 2.61, p < .05$.

However, as a fascinating variant on the above analysis, it is also possible to look at authoritarianism as a function of the groups to which the remaining participants *assigned themselves* at the end of the study – that is, as participants who either supported the commune or who proposed setting up a new hierarchy (with themselves as the new guards within it). Data pertaining to this analysis are presented in Figure 9. These were again analysed by means of a 2 (participant group: new guards, new prisoners) $\times$ 4 (study phase: pre-test, Days 1, 3, 7) ANOVA with repeated measures on the second factor. The only effect to emerge from this analysis was a significant and large interaction between group and time; $F(3, 36) = 3.07, p < .05; \eta^2 = .20$. Analysis of

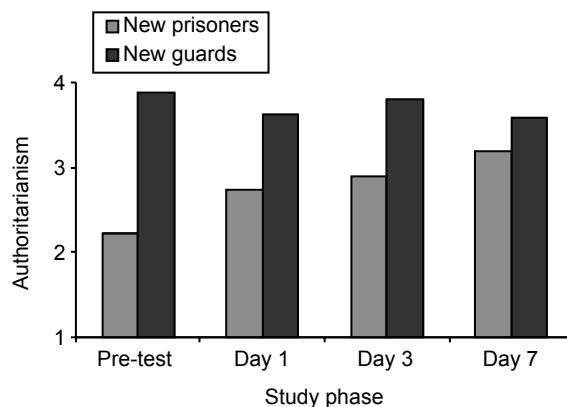


Figure 9. Right-wing authoritarianism as a function of self-selected group and time.

polynomial contrasts indicated that this interaction arose from the fact that the authoritarianism of those who sought to assume the guards' role at the end of the study had declined non-significantly as the study progressed, $t(3) = -1.05$, *ns*. However, the authoritarianism of the remaining participants had increased steadily over time, $t(10) = 3.67$, $p < .01$ (so that the two groups were demonstrating a very similar level of authoritarianism at the time that the new guards were actually seeking to take power).

Despite the general shift towards authoritarianism, and although it is probable that the new guards would have had the force to impose their regime in the face of weakening resistance, in this study such force was prohibited under the pre-established ethical guidelines. Hence, in our judgment, the study was gridlocked and had reached what we saw as a natural point of termination; the existing system was not working and the new system could not be imposed. Accordingly, the study was brought to a conclusion at noon on Day 8. However, the participants remained for a further day in order to undertake a series of structured debriefings designed to obtain and provide feedback on their experience, to explain the rationale for the study and to overcome any hostility between individuals deriving from events in the study.

Discussion

The BBC prison study was designed to examine the factors that determine how people respond when a system of inequality is imposed upon them by others. At the start, almost all the participants rejected this system. However, by the end, they were close to instituting a new and more tyrannical social system. In addition to our original questions concerning the way in which people respond to a system of inequality that has been imposed upon them – do they accept it or do they resist it? – this raises a new and unexpected issue. What are the conditions under which people create a system of inequality *for themselves*?

We will argue that our findings concerning reactions to inequality cannot be explained through a general or 'natural' tendency to assume roles and assert power. Rather, they require an understanding of the conditions under which an externally imposed categorization becomes a subjective self-perception (Turner, 1982). The existing terms of social identity and self-categorization theories provide the basis for such an understanding, although it is necessary to elaborate upon notions of context in order to account for the failure of the guards to cohere as a group. Our findings concerning the creation of inequality lead us to more original theoretical conclusions. On the one hand, we suggest that groups are the basis for collective self-realization – that is, the creation of a social order based on shared values and norms (Drury & Reicher, 2005). However, where groups fail, we argue that people will be more inclined to accept the imposition of a social order by others, even where that violates their values and norms. Therefore, in contrast to those who explain tyranny and other extreme social phenomena in terms of the psychological dysfunctionality of groups, we interpret them in terms of the dysfunctionality of group failure.

Before we are in a position to develop these arguments, however, it is necessary to deal with four potential critiques of the study and its ability to say anything meaningful about the group processes relating to inequality and tyranny. The first is that the behaviour of participants was determined by the fact that they knew they were being observed by television cameras and that this renders the study so artificial as to have little or no general value. The second is that the effects we observed were a product of

participants' personalities and that therefore the study has little to say about group processes. The third is that we failed to create real power differences or meaningful inequalities between groups and hence our study has little to say about the psychology of either power or inequality. The fourth is that the variables upon which our predictions (and hence our planned interventions) focused were not responsible for the effects obtained and, therefore, even if group and power processes were at play, we cannot be sure what they are.

Four critiques

The role of television

From their moment of first contact, volunteers knew that the BBC was involved in this project, and by the time that the study started, all participants were aware that they would be constantly observed by television cameras and that anything they did might be shown on national television. Although it was stressed throughout that this was a specialist scientific project that would not make 'stars' out of those involved, and although the screening process was used to exclude anyone who was motivated by the desire for publicity, this is a highly unusual situation and was bound to impact on behaviour – but what impact, and with what implications?

The most damning argument would be that participants were simply faking their behaviours for the cameras. We believe such an explanation to be implausible in light of the immense effort that would be required to continuously monitor and fake one's own behaviour for nearly 9 days. Moreover, it would be much harder, if not impossible, to fake the psychometric and physiological data. However, even if participants were capable of such play-acting, one would still need to explain the complex pattern of results not only between groups, but also within groups across time. One might, for example, suggest that there is a certain discredit in being a tyrant and a certain glamour in being a rebel, which may explain why our guards were so mild and our prisoners so rebellious. But, why then did the prisoners become more rebellious after the system was made impermeable and insecure? Even more problematically, why did the participants move towards tyranny at the study's end? To argue that our findings can be explained by suggesting that participants were merely play-acting or seeking 'celebrity status' is not only implausible but also unhelpful. In particular, it fails to explain (a) why people acted as they did when they did and (b) why their behaviour and attitudes clearly changed in particular (predicted) ways over the course of the study. By the same logic, surveillance cannot be the whole story, as it remained constant at the same time that behaviour itself was changing.

Having said this, we readily accept (see below) that the televising of the study is *part* of the story – particularly towards the start of the study (participants reported that they were acutely aware of the cameras on entering the 'prison', but that, as time went by, they increasingly forgot about them except during quiet moments; e.g. when seeing a camera move late at night in their cells). However, we suggest that the nature of this impact adds to, rather than detracts from, the richness and wider relevance of our findings. This is because although it may be unusual to be in a position where anything one does might be broadcast into millions of homes, this is an extreme example of something that is an increasingly common feature of our everyday lives – namely, surveillance. For most of our social lives, we are under observation and our behaviour can be examined by audiences who are not present (Reicher & Emler, 1985). Sometimes, the surveillance is a matter of surveillance cameras, scrutiny of computer

transactions, workplace review, or unobtrusive profiling (Lyon, 1994; Lyon & Zuriek, 1996). At other times, it is simply a matter of those who are with us in one place talking about our behaviour to other people in other places (e.g. in gossip, a universal human behaviour; Emler, 1994). However, whichever is the case, we can rarely ignore surveillance when we are alone and almost never when with others.

What the television cameras do, therefore, is to highlight in dramatic form an aspect of human experience, which is all too often overlooked in psychology. Indeed, it is arguable that the typical psychology experiment in which people are isolated and guaranteed absolute anonymity encounters greater problems of artificiality (Cronin & Reicher, 2006). Of course, the size and variety of audiences to which one is made accountable by television is much larger than we commonly experience. However, as we will discuss in more detail below, the fact of accountability, and the fact that our participants had to consider other contexts and other audiences even when acting within the context of the study, enhances rather than diminishes its relevance and meaningfulness. In this respect, we are reminded of Thayer and Saarni's (1975) retort to those who impugn the validity of the SPE on the grounds that participants were guided by the expectations of others. This is summarized in the title of their paper: 'demand characteristics are everywhere (anyway)'. Likewise, to those who criticize the BBC study on the grounds that it was broadcast, we would respond that surveillance is everywhere (anyway).

The role of personality

For those who have watched *The Experiment* (Koppel & Mirsky, 2002), it is hard not to be struck by the force of many of the characters who contributed to the various outcomes described above. In part, this has to do with the visual nature of the medium, where the concrete and visible impact of individuals is inevitably more salient than that of more intangible and 'invisible' elements such as group processes (Asch, 1952). Nonetheless, it could clearly be argued that all the observed effects were a product of strong (and weak) personalities and that the more powerful characters (in particular, JE_p, PP_p, and PB_p) ended up as prisoners. One could argue that their personalities drove these individuals to overwhelm the guards and then, as they grew bored of the commune, to destroy it in turn.

The first response to such an argument is to remember that we matched the prisoners and guards on the obvious individual difference factors relating to tyranny: modern racism, authoritarianism, and social dominance. Nonetheless, one could still argue that there were other characteristics on which we did not match the groups. As a result, the 'stronger' individuals might still, by chance, have been allocated to the prisoner group. However, even if this were the case, as with our discussion of surveillance, it does not prove particularly helpful in explaining how events in the study unfolded. Most fundamentally, this due to the interrelated observations (a) that people's 'character' on relevant dimensions appeared to change over the course of the study and (b) that relevant individual differences were much less apparent at the outset than they were as the study progressed.

We readily accept, however, that individual differences are *part* of the story (see Reicher & Haslam, in press; Turner, Reynolds, Haslam, & Venstra, in press). However, as with the analysis of leadership more generally (e.g. see Haslam, 2001; Haslam & Reicher, 2005; Reicher & Hopkins, 2001b; Turner & Haslam, 2001), the problem is that it is hard to explain changing patterns of behaviour in our study with reference to a blunt and constant factor such as 'personality'. To claim that certain people were 'rebellious' does

not explain why rebellion was muted until after the promotion and fails to explain why those prisoners who rebelled against one form of authority (the guards) nonetheless remained deferential to another (the experimenters). In the case of JE_p, for example, we see that before the promotion, he invested his energies in supporting the system and it was only after promotion was ruled out that he put them into undermining the system. Therefore, it is one thing to call him 'forceful', but to explain how that force was directed, we need to invoke systemic factors. Similarly, in the case of JE's cellmate PP_p, while he was opposed to the guards' authority from the start of the study, his ability actually to undermine their regime was contingent on collective will and support, which only materialized after the promotion (in ways predicted by social identity theory; Tajfel & Turner, 1979; see below).

More generally, there were many points in the study where the prominence and impact of individuals depended upon their relationship to the group (Reicher & Haslam, in press). Thus, when they were seen to reflect shared identity and shared values, even the mildest of men came to the fore. This is apparent in the case of FC_p, an environmentalist who had been quiet throughout the first 6 days of the study, but who became a leading and vocal figure in creating and managing the commune and without whose commitment, knowledge, and experience it may never have come into being.

Furthermore, it is also apparent that when there was no shared identity or shared values, even the most heroic of individuals was bound to fail. This was seen early in the study when FC_g made a strenuous effort to impose the guards' authority and maintain strict discipline. However, as it became clear that he lacked the support of his fellow guards, he gave up, became increasingly passive, and withdrew into the background. Thus, the manifest differences in the forcefulness of prisoners and guards derived not from their inherent personalities but were instead an *emergent* product of the success or failure of their respective groups.

Again, this is not to dismiss the importance of individual variables. As we have repeatedly stressed, the skills, knowledge, values, and pre-existing commitments of individuals played a crucial part in facilitating, shaping (and indeed in blocking) group formation (Postmes, Baray, Haslam, & Morton, in press; Reicher, 2004; Reicher, Drury, Hopkins, & Stott, 2001; Reicher & Hopkins, 1996, 2001b, 2004). Far from being mindless, our participants were creative and thoughtful as group members, continuously striving to shape group and intergroup realities. However, if individuals played a key part in shaping groups, the converse is equally true – the ability of individuals to affect outcomes was dependent upon both the existence of groups and their prototypicality within the group (Turner, 1991). To put it slightly differently, individual agency was not destroyed by the group, but rather achieved through it (Reicher & Haslam, in press). Thus, while personality and other individual difference factors are an important aspect of our study (and we will consider them as part of our analysis below), they cannot substitute for that analysis (Asch, 1952; Brown, 1965; Sherif, 1966; Tajfel, 1978).

It follows that had different individuals participated in the study, we may well have observed quite different outcomes from those reported above. This is because many individual (and, indeed, chance) factors (such as the poor-quality breakfast on Day 8) impacted on the fate of groups and their subsequent trajectory in the study. However, as we explained in the Introduction, this is why we are not seeking to make empirical generalizations on the basis of findings alone but rather are attempting to make generalizations on the basis of the theoretical analysis that those findings support (Turner, 1981). Our claims to generality thus relate to (a) the effect of particular factors

(e.g. permeability, cognitive alternatives, roles) on group formation and (b) the consequences of group success and failure for both individuals and the system as a whole.

The reality of inequality and power

A third potential criticism of our study is that the results and, in particular, the failure of the guards to exert their authority (or the prisoners to accept it), reflects the twin facts that (a) the study context was simply too pallid for our participants to become engaged in it and (b) the guards had no authority to wield even if they wanted to. If, as Zimbardo and his colleagues argue, tyranny is a function of groups and power, then one might assert that our study lacked a meaningful degree of either and that we are therefore in no position to say anything worthwhile about tyranny.

Let us consider these issues of engagement and power in turn. In the case of the former, it is important to distinguish between engaging with the situation and engaging with one's group. The evidence suggests that the prisoners experienced and resented their subordinate status from the start of the study. They disliked their food. They disliked being locked up. They disliked the boredom. The smokers in particular disliked being deprived of cigarettes. Indeed, after a day in his cell, PP_p admitted that he was 'falling for it [the experimental situation] hook, line and sinker'. At the first meal, several prisoners expressed outrage at the poor quality of the food, especially given the superior fare of the guards. The lack of group identification amongst prisoners in the early days cannot therefore be put down to apathy about their plight.

Equally, the guards engaged with the situation from the start. Their initial conversations focused on the power of the situation, the impact of wearing a guard's uniform, and the dangers of becoming tyrannical. At their first meal, they also noted with discomfort the disparity in food quality and quantity. They then tried to ease this discomfort by offering their leftovers to some of the prisoners. Therefore, if the guards failed to act as a group, it was not because they did not care about the disparities in the study but, on the contrary, because they were troubled by them. Their disidentification did not reflect the fact that the study was not compelling, but rather that it was all too compelling (albeit in very different ways from the SPE).

The issues are similar when it comes to the matter of power. It is certainly true that the guards failed to exercise power. However, this was not because they had no power to exercise but precisely because they had so much. In discussion, the guards recognized the various options that were open to them. These included individual punishments, collective punishments, removal of privileges, extra tasks and roll calls and, most particularly, the power to promote a prisoner who would help them run the system in the way they wanted. However, their fear of the guard identity – of being authoritarian and of being seen as authoritarian – made them shun these options, to promote a prisoner who embodied their ambivalence (Herriot, 2002), and even to give away some of the sources of their power. In other words, the participants' psychological state explains their failure to exercise power much more than their lack of power explains their psychological state (Turner, 2005). We will address the reasons for this state shortly. For now, we simply note that our participants did engage with a situation in which there were real inequalities of resources and power. Hence, the study can be viewed as a meaningful exploration of the psychology of unequal groups.

The impact of interventions and key variables

It may be that our study says something of interest, it may be that it says something interesting about group processes, it may even have something important to say about

tyranny and resistance, but what of our claim to have conducted an 'experimental case study' in which we investigated the impact of specific theoretically informed variables? Can our interventions really be described as operationalizations of permeability and cognitive alternatives? Are there not a host of other ways in which these interventions and their consequences can be interpreted, and does this fact not cast doubt on any particular theoretical gloss we might provide?

This is a reasonable concern. In the present study, there is certainly a disjunction between the simplicity of our variables and the complexity of our interventions. In particular, the promotion and the introduction of DM_p could have had many effects, such as making participants uncertain about what might happen next and making them feel helpless and distrustful of the experimenters. However, such problems are not unique to the present study – not least because the theoretical status of independent variables is a potential concern in *all* experimental research (Haslam & McGarty, 2004). Indeed, whenever researchers claim that a concrete instance (a specific operationalization of an independent variable) reflects an abstract generality (the variable as it operates in the world at large), interpretation is involved and alternatives are possible (Billig, 1987). When the operationalization is as complex as ours, however, the space for argument is all the greater.

In experimental practice, this issue is normally addressed in two ways (Haslam & McGarty, 2003). The first is through the use of manipulation checks. As we have seen above (Figure 3), checks of this form provide some evidence consistent with our theoretical interpretation (specifically in relation to the existence of cognitive alternatives). Moreover, one of the advantages of having multiple data sources (i.e. behavioural as well as psychometric) is that we are able to see whether participants spontaneously reacted to our interventions in ways that were expected. We have provided qualitative evidence to suggest that they did. Specifically, after the promotion prisoners saw no point in placating the guards and realized that the only way to improve their position was to challenge the system (Extract 1). Equally, we saw that DM_p transformed the way in which participants conceptualized their situation and their options (Extract 2). As with any manipulation check, such data support the notion that we did manipulate the variables of concern to us.

However, while manipulation checks exclude false positives, they cannot exclude all possible false negatives – that is, additional ways in which the manipulation may have had an impact. This is where the second way of reducing uncertainty about the status of independent variables comes into play. Here, it is not enough merely to argue for the importance of a particular independent variable (or a particular confound; e.g. personality, surveillance, the passing of time). It is also necessary to explain how that variable (or confound) might *plausibly* and *comprehensively* account for the patterns of observed results. That is, it is necessarily to have an integrated theoretical account of how the effects occurred. As we explained when introducing them, in this respect, our claims accord with a well-developed and extremely influential theoretical tradition (after Tajfel & Turner, 1979) and with a well-established and robust pattern of previous experimental findings (e.g. Ellemers, 1993; Wright *et al.*, 1990). This theoretical and empirical tradition lends coherence and plausibility to our analysis.

Having said that, we certainly do not view our analysis as inviolate, and we fully accept that our account of the impact of particular interventions is open to argument. However, for critics who wish to challenge this account, it is incumbent upon them to do as we have and explain how an alternative understanding of our interventions provides a better understanding of the study's findings in their entirety.

With this in mind, it is time to move beyond potential criticisms and to be explicit about what our explanation of events is. This involves three elements. The first relates to the conditions of social identification, the second addresses the consequences of social identification, and the third concerns reactions to group failure.

A social identity account of tyranny

The conditions of social identification

The simplest and clearest finding of our study is that people do not automatically assume roles that are given to them in the manner suggested by the role account that is typically used to explain events in the SPE (Haney *et al.*, 1973; Zimbardo, 1989). Instead, there are a range of factors that determine whether people themselves identify with the social positions to which they are ascribed by others. Some of those factors operated in ways specified by social identity theory (Tajfel & Turner, 1979). Thus, the shift from permeability to impermeability of group boundaries had a strong impact on prisoner identification (along lines previously demonstrated in separate programmes of research by Ellemers and Wright and their colleagues; Ellemers, van Knippenberg, De Vries, & Wilke, 1988; Ellemers *et al.*, 1990, 1993; Wright, 1997; Wright & Taylor, 1998; Wright *et al.*, 1990; see also Lalonde & Silverman, 1994). Equally, the increasing insecurity of intergroup relations affected the willingness of the prisoners to work as a group in order to challenge the guards.

In addition, changing perceptions of security, and of the component of legitimacy in particular, help explain not only why the prisoners challenged the guards, but also why the participants as a whole did not challenge the experimenters, despite the adverse conditions that living in the prison imposed upon them (cf. Spears & Smith, 2001). Most of the time, when participants discussed the matter, they saw these conditions as a legitimate part of the study in which they had agreed to participate. On only two occasions were the conditions seen to violate that contractual legitimation – when the new prisoner, DM_p, began to question the heat in the prison, and when one night, the experimenters failed to replenish basic supplies for the guards. On both occasions, this led to explicit consideration of ‘mutiny’.

However, if the study supported social identity theory in demonstrating the role of contextual factors in moderating the relationship between role and identity, it also shows the importance of extending the way in which we conceptualize the nature of context and its relation to human action. Thus, although the role of guard was positively valued in the immediate context of the prison, those assigned to this position were concerned with the possibility of negative evaluations by future audiences, and hence some of them were reluctant to identify with the role.

It is here that the prospective televising of the study most clearly had an impact on the behaviour of participants and serves to raise important practical and theoretical points. Practically, it demonstrates how extreme behaviours can be restrained by rendering actors visible, and hence accountable, to broader or yet-to-be encountered audiences (Postmes, Spears, Lea, & Reicher, 2000; Reicher & Levine, 1994; Tedeschi, 1981; Tetlock, 1985).⁴ Theoretically, it shows that context, for human beings, cannot

⁴ Some commentators have argued that this analysis is inconsistent with evidence arising from the torture of inmates by US troops at Abu Ghraib prison in 2004. However, as we have argued elsewhere (Reicher & Haslam, 2004), it appears that these acts were performed for the benefit of an in-group audience that (like the experimenters in the SPE) was assumed by the perpetrators to approve of the action.

simply be assumed to be where we are in the here and now. People have imaginations that allow them to think of other times and places, and to evaluate their behaviour in relation to these. Indeed, surveillance has an impact on behaviour precisely through its capacity to make salient such alternative contexts and alternative audiences (Reicher & Haslam, 2003; Reicher & Levine, 1994).

Two points follow from this. First, under conditions of surveillance, individual differences relating to the different social commitments people have outside a given context have the capacity to impact upon behaviour within the context. Thus, in our study, the fact that TQ_g was publicly visible as a successful entrepreneur who owned a liberal enterprise goes some way towards explaining why he was particularly ambivalent about adopting the identity of guard. However, if individual factors were important here, then they operated in a paradoxical way. First, TQ_g's very effectiveness outside the study is what rendered him ineffective within it. Second, once human imaginative capacities are acknowledged, forms of extreme situationism (such as the role account), which suggest that behaviour is always dominated by the present context, become untenable. Instead, it is important to take past and future contexts into account as well (Reicher & Hopkins, 2001a).

The consequences of social identification

One of the distinctive features of the present study was that its intensive nature allowed established disciplinary boundaries to be transcended so that it was possible to investigate the interrelationships between social, organizational and clinical variables. The results presented above point to the existence of such links and to their richness and complexity.

On the organizational side, the achievement of a common social identity was seen to be necessary in order for group members to trust others to act appropriately, to support others in their actions and to expect support from them in return (Haslam *et al.*, 2005). In this way, the rudiments of effective organization (e.g. task differentiation, delegation, leadership, trust) were seen to be contingent upon shared social identification in a manner argued by a range of organizational theorists (e.g. Ellemers, De Gilder, & Haslam, 2004; Haslam, 2001; Haslam, van Knippenberg, Platow, & Ellemers, 2003; Hogg & Terry, 2001; Postmes, 2003; Smith, Tyler, & Huo, 2003; Turner & Haslam, 2001; Tyler & Blader, 2000). For this reason, the effectiveness of the prisoners, who were able to develop a sense of shared social identity, contrasted markedly with the ineffectiveness of the guards, who were not (Drury & Reicher, 2005).

On the clinical side, the increasing support amongst prisoners and their successful challenges to the guards contributed to mental states that were increasingly positive. On the other hand, the atomization of the guards and their failure to fulfil their collective tasks led to increasingly negative states. This is reflected in the depression data (and also on other measures not reported here including paranoia, anxiety and burnout; see Haslam & Reicher, *in press*). Although the nature of our quantitative data does not allow for analysis of causal relations between these variables (Haslam & McGarty, 2004), we can say with confidence that participants' mental states evolved in relation to the social dynamics between groups, and that variables such as social support and collective self-efficacy are implicated in that relationship (e.g. as argued by Branscombe, Schmitt, & Harvey, 1999; Cohen & Wills, 1985; Contrada & Ashmore, 2000; Hall & Cheston, 2002; Haslam *et al.*, 2005; Orford, 1992; Schwarzer, 2001; Terry, Carey, & Callan, 2001). Clearly though, the exact nature of the relationship merits further investigation.

Reactions to group failure

There were two instances of group failure in the study. The first concerned the failure of the guards to fulfil their assigned task of making a hierarchical prison system work, and the second concerned the failure of the commune's supporters to establish a viable egalitarian social order.

It is notable that in each case, group members became more willing to accept a system that promised to be viable even if it meant ceding some of the core principles of the group. On Day 5, after the introduction of the new prisoner, the guards were willing to cede their authority and accept a more equal social system and on Day 8 commune supporters were prepared to cede equality and tolerate a more hierarchical social system. Thus, rather than people 'naturally' preferring any given form of social order, it appears that, when group members fail to impose an order based on their own existing norms and values, they are willing to adapt those values (or to adopt new ones) in order to create a viable order rather than have no order at all.

In this regard, the fact that the participants' authoritarianism increased significantly over time constitutes one of the most important findings of the study as a whole. Traditionally, authoritarianism has been viewed a stable individual difference variable that has the capacity to explain the emergence of hierarchical and tyrannical social structures (Adorno *et al.*, 1950; Altmeyer, 1981, 1996; for reviews see Billig, 1978; Brown, 1965; Duckitt, 1994). However, in contrast to this analysis, the present study illustrates that authoritarianism is a variable *outcome* of social structure (see also Oakes, Haslam, & Turner, 1994; Reynolds *et al.*, 2001). More specifically, we see that authoritarian solutions – and the personalities that would promote them – appeared more attractive after attempts to make democracy work were seen to have failed.

Against this argument, it might be suggested that the system we observed moved towards tyranny simply because the more authoritarian individuals were waiting for an opportunity to impose their preferred system on others (indeed, a similar point could be made in relation to the SPE). However, along lines intimated above, two points speak against this conclusion and against related arguments couched in terms of individual personality-based dynamics. The first is that tyranny arose out of the social (dis)order of the prison, and authoritarianism (like other dimensions of raw personality) was clearly not a straightforward determinant of that (dis)order. The second is that although it may be true that those who proposed the new regime were those who had initially been more disposed towards authoritarianism, their influence and leadership was clearly contingent upon their views having become more representative of the views of the participants *as a whole* as these had evolved over the course of the study (for more in-depth treatment of this issue see Haslam & Reicher, 2005; Reicher & Haslam, in press; see also Haslam, 2001; Haslam & Platow, 2001; Navas, Morales, & Moya, 1992; Reicher & Hopkins, 1996, 2001b; Turner, 1991; Turner & Haslam, 2001; Turner *et al.*, in press). Indeed, the interaction between final group membership and study phase suggests that, if anything, the authoritarianism of the guards in the vaunted new regime had declined over time, whereas that of other participants was increasing (see Figure 9). It also is worth adding that previous research has been unable to observe dynamics of this form at work or to generate empirical support for arguments such as these, precisely because it has been unable to examine the impact of evolving group history on both individual psychology and social structure (Haslam & McGarty, 2001; Levine, 2003).

Conclusion: Rethinking the relationship between groups, power and tyranny

At one level, our study confirms the findings of the SPE. It shows that an understanding of collective conflict and tyranny cannot be achieved simply by looking at individuals but requires an analysis of group processes and intergroup relations. In this sense, we agree with Zimbardo (and many others; e.g. Asch, 1952; Sherif, 1966; Tajfel, 1978) that such phenomena can only be explained through group-level analysis. Our disagreement with prior analysis of the SPE thus relates to the *nature* of group processes and of the conditions under which they lead to social pathologies.

As almost every psychology student (and an unusually large proportion of the general public) knows, the message of the SPE is that the toxic combination of groups and power leads to tyranny. The implications of the BBC prison study are different. In common with recent theoretical developments in social psychology, they contest the premise that group behaviour is necessarily uncontrolled, mindless and antisocial (Ellemers *et al.*, 1999; Oakes *et al.*, 1994; Postmes *et al.*, 2000; Reicher, 1982, 2001; Spears, Oakes, Ellemers, & Haslam, 1997; Turner, 1999). In contrast, the results of the BBC prison study suggest that the way in which members of strong groups behave depends upon the norms and values associated with their specific social identity and may be either anti- or prosocial (Jetten, Spears, & Manstead, 1997).

However, based on the present data, we would argue that *failing* groups almost inevitably create a host of problems for their own members and for others. These problems have a deleterious impact on organization, on individuals' clinical state, and – most relevant here – on society. For it is when people cannot create a social system for themselves that they will more readily accept extreme solutions proposed by others. It is when groups lack the power to exercise choice that an authoritarian ideology that promises to create order for them appears more seductive. In short, it is the *breakdown* of groups and *powerlessness* that creates the conditions under which tyranny can triumph (for related arguments see Kanter, 1979; Pfeffer, 1981; Reynolds & Platow, 2003).

We would argue that as well as being consistent with contemporary thinking in social psychology (e.g. after Tajfel & Turner, 1979), this analysis also articulates more closely than Zimbardo's original role account with the analysis of tyranny put forward by researchers in other academic disciplines. Most notably, it accords with influential analyses proposed by modern historians (e.g. Abel, 1986; Gellately, 2001; Hobsbawm, 1995; Rees, 2002). Consider, for instance, Hobsbawm's account of the conditions that gave rise to the fall of the Weimar republic and the emergence of Nazism in 1930s Germany:

The optimal conditions for the triumph of the ultra-right were an old state and its ruling mechanisms which could no longer function; a mass of disenchanting, disoriented and disorganized citizens who no longer knew where their loyalties lay; strong socialist movements threatening or appearing to threaten social revolution, but not actually in a position to achieve it.... These were the conditions that turned movements of the radical right into powerful, organized and sometimes uniformed and paramilitary force (1995, p. 127).

We would also argue that this analysis can be used to make sense of what happened in the SPE when one looks more closely at the events that unfolded there. As we have noted, that study, like ours, appears to have started off with prisoners threatening to become ascendant over the guards. Things changed when Zimbardo intervened in such a way as to lead the prisoners to believe that they could not leave the study. At this point they became disoriented as to their position – in Zimbardo's own words, they

experienced 'role confusion' (Zimbardo, 1989). They ceased to support each other against the guards. They collapsed as a group and this allowed tyrannical guards to prevail. Thus, whereas in the SPE, the failure of the prisoners allowed an existing tyranny to be consolidated, in the BBC study, the failure of the commune paved the way for the emergence of a new tyranny.

These, of course, are large and controversial claims. However, they have important theoretical implications along with considerable practical implications. They point to new ways of thinking both about the nature of group psychology and about the psychological underpinnings of tyranny. Because the scope of these analyses is so large, it would be both presumptuous and foolish to expect them to be accepted on the basis of a single data set – especially one reflecting the complexities of the present study. Yet what is true of our study is largely true of the SPE. As we have argued, one of the great problems arising from that study was that it made strong claims (that had an enormous impact on public consciousness) but then led to debate being closed off because further research was declared to be ethically unacceptable.

Hence, irrespective of our results and of our analysis, we would claim that one of the significant achievements of the BBC prison study is to show that, if sufficient care is taken, it is possible to run powerful and influential field studies into social processes that are also ethical. In a field increasingly dominated by reductionist accounts of human behaviour, such studies can restore balance by demonstrating the impact of systematic variations in social relations upon human behaviour. The richness, immediacy and relevance of their findings can also help reconnect psychologists with policy makers. For these reasons, we offer our conclusions not so much as a final word on the matter, but more in the hope of stimulating researchers to renew their interest in the important debates that our discipline can, and must, advance.

Acknowledgements

We would like to thank the participants as well as the staff of the BBC for their commitment to this research and for comments on ideas in this paper. The assistance and input of Nyla Branscombe, Lloyd Carson, David Corner, Andrew Eagle, Liberty Eaton, Scott Galloway, Nick Hopkins, Mike Howe, Jolanda Jetten, Andrew Livingstone, Eva Loth, Grant Muir, Lucy O'Sullivan, Michael Schmitt, Heather Smith, Tom Postmes, Fabio Sani, Denis Sindic, Stephanie Sonnenberg, Paul Webley, Stephen Wilks, Huw Williams and Brian Young is also gratefully acknowledged, together with detailed and very helpful comments from several anonymous reviewers.

References

- Abel, T. (1986). *Why Hitler came to power*. Cambridge, MA: Harvard University Press.
- Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper.
- Altmeyer, B. (1981). *Right wing authoritarianism*. Winnipeg: University of Manitoba Press.
- Altmeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.
- Arendt, H. (1998). *The human condition*. Chicago: University of Chicago Press.
- Asch, S. E. (1952). *Social psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). New York: Cambridge University Press.
- Banuazizi, A., & Movahedi, S. (1975). Interpersonal dynamics in a simulated prison: A methodological analysis. *American Psychologist*, 30, 152–160.

- Baron, R. S. (1984). Group dynamics. In A. S. Kahn (Ed.), *Social psychology* (pp. 260–289). Dubuque, IA: W.C. Brown.
- Bar Tal, D. (2004). The necessity of observing real life situations: Palestinian–Israeli violence as a laboratory for learning about social behaviour. *European Journal of Social Psychology*, 34, 677–701.
- Billig, M. (1978). *Fascists: A social psychological view of the National Front*. London: Academic Press.
- Billig, M. (1987). *Arguing and thinking*. Cambridge: Cambridge University Press.
- Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, 77, 135–149.
- Brookes, E. (2001, October 16). The Experiment. *The Guardian*, G2, pp. 1–3.
- Brown, R. (1965). *Social psychology*. New York: Free Press.
- Chen, G., & Bliese, P. D. (2002). The role of different levels of leadership in predicting self- and collective efficacy: Evidence for discontinuity. *Journal of Applied Psychology*, 87, 549–556.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Wiley.
- Cohen, S., & Wills, T. A. (1985). Stress, social support and the buffering hypothesis. *Psychological Bulletin*, 98, 310–357.
- Contrada, R. J. & Ashmore, R. D. (Eds.), (2000). *Self, social identity, and physical health: Interdisciplinary explorations*. New York: Oxford University.
- Cronin, P. & Reicher, S. (2006). A study of the factors that influence how senior officers police crowd events: On SIDE outside the laboratory. *British Journal of Social Psychology*, 45, 175–196.
- Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology*, 31, 410–436.
- Doosje, B., Spears, R., & Ellemers, N. (2002). The dynamic and determining nature of group identification: Responses to anticipated changes in the status hierarchy. *British Journal of Social Psychology*, 41, 57–76.
- Drury, J., & Reicher, S. (2005). Explaining enduring empowerment: A comparative study of collective action and psychological outcomes. *European Journal of Social Psychology*, 35, 35–58.
- Duckitt, J. (1994). *The social psychology of prejudice* (2nd ed). New York: Preager.
- Eggs, R. A., Haslam, S. A., & Reynolds, K. J. (2002). Social identity and negotiation: Subgroup representation and superordinate consensus. *Personality and Social Psychology Bulletin*, 28, 887–899.
- Ellemers, N. (1993). The influence of socio-structural variables on identity enhancement strategies. *European Review of Social Psychology*, 4, 27–57.
- 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.
- Ellemers, N., Spears, R., & Doosje, B. (1999). *Social identity: Context, content and commitment*. Oxford: Blackwell.
- Ellemers, N., van Knippenberg, A., De Vries, N. K., & Wilke, H. (1988). Social identification and permeability of group boundaries. *European Journal of Social Psychology*, 18, 497–513.
- Ellemers, N., van Knippenberg, A., & Wilke, H. (1990). The influence of permeability of group boundaries and stability of group status on strategies of individual mobility and social change. *British Journal of Social Psychology*, 29, 233–246.
- Ellemers, N., Wilke, H., & van Knippenberg, A. (1993). Effects of the legitimacy of low group or individual status on individual and collective identity enhancement strategies. *Journal of Personality and Social Psychology*, 64, 766–778.
- Emler, N. (1994). Gossip, reputation, and social adaptation. In B. F. Goodman & A. Ben-Ze'ev (Eds.), *Good gossip* (pp. 117–138). Lawrence: University of Kansas Press.

- Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology*, 57, 239–249.
- Gellately, R. (2001). *Backing Hitler*. Oxford, UK: Oxford University Press.
- Hall, S., & Cheston, R. (2002). Mental health and identity: The evaluation of a drop-in centre. *Journal of Community and Applied Social Psychology*, 12, 30–43.
- Haney, C., Banks, C., & Zimbardo, P. (1973). A study of prisoners and guards in a simulated prison. *Naval Research Review* 9, 1–17 [Reprinted in E. Aronson (Ed.), *Readings about the social animal* (3rd ed., pp. 52–67). San Francisco, CA: W.H. Freeman].
- Haslam, S. A. (2001). *Psychology in organizations: The social identity approach*. London: Sage.
- Haslam, S. A., Eggins, R. A., & Reynolds, K. J. (2003). The ASPIRe model: Actualizing Social and Personal Identity Resources to improve strategic planning and encourage sustainable productivity in organizations. *Journal of Occupational and Organizational Psychology*, 76, 83–113.
- Haslam, S. A., O'Brien, A., Jetten, J., Vormedal, K. & Penna, S. (2005). Taking the strain: Social identity, social support and the experience of stress. *British Journal of Social Psychology*, 44, 355–370.
- Haslam, S. A., & McGarty, C. (2001). A hundred years of certitude? Social psychology, the experimental method and the management of scientific uncertainty. *British Journal of Social Psychology*, 40, 1–21.
- Haslam, S. A., & McGarty, C. (2003). *Research methods and statistics in psychology*. London: Sage.
- Haslam, S. A., & McGarty, C. (2004). Experimental design and causality in social psychological research. In C. Sanson, C. C. Morf, & A. T. Panter (Eds.), *Handbook of methods in social psychology* (pp. 235–264). Thousand Oaks, CA: Sage.
- Haslam, S. A., & Platow, M. J. (2001). The link between leadership and followership: How affirming a social identity translates vision into action. *Personality and Social Psychology Bulletin*, 27, 1469–1479.
- Haslam, S. A., Powell, C., & Turner, J. C. (2000). Social identity, self-categorization and work motivation: Rethinking the contribution of the group to positive and sustainable organizational outcomes. *Applied Psychology: An International Review*, 49, 319–339.
- Haslam, S. A., & Reicher, S. (2002). A user's guide to The Experiment – *Exploring the psychology of groups and power: Manual to accompany the BBC video*. London: BBC Worldwide.
- Haslam, S. A., & Reicher, S. D. (2005). *Identity entrepreneurship and the consequences of identity failure: The dynamics of leadership in the BBC Prison Study*. Manuscript submitted for publication.
- Haslam, S. A., & Reicher, S. D. (in press). Stressing the group: Social identity and the unfolding dynamics of stress. *Journal of Applied Psychology*.
- Haslam, S. A., van Knippenberg, D., Platow, M. & Ellemers, N. (Eds.), (2003). *Social identity at work: Developing theory for organizational practice*. Philadelphia, PA: Psychology Press.
- Herrera, C. D. (1997). A historical interpretation of deceptive experiments in American psychology. *History of the Human Sciences*, 10, 23–26.
- Herriot, P. (2002). Selection and self: Selection as a social process. *European Journal of Work and Organizational Psychology*, 11, 385–402.
- Hobsbawm, E. (1995). *Age of extremes: The short twentieth century 1914–1991*. London: Abacus.
- Hogg, M. A., & Terry, D. J. (Eds.), (2001). *Social identity processes in organizational contexts*. Philadelphia, PA: Psychology Press.
- House, J. S. (1981). *Work stress and social support*. Reading, MA: Addison-Wesley.
- Hoyle, R. H., Georgesen, J. C., & Webster, J. M. (2001). Analysing data from individuals in groups: The past, the present, the future. *Group Dynamics: Theory, Practice and Research*, 5, 41–47.
- Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the timecourse of social perception: The effects of racial cues on event-related brain potentials. *Personality and Social Psychology Bulletin*, 30, 1267–1280.

- Jetten, J., Spears, R., & Manstead, A. R. (1997). Strength of identification and intergroup differentiation: The influence of group norms. *European Journal of Social Psychology*, 27, 603–609.
- Kanter, R. (1979). Power failure in management circuits. *Harvard Business Review*, July–August, 65–75.
- Koppel, G. (Series producer), & Mirsky, N. (Executive producer). (2002, May 14, 15, 20, 21). *The Experiment*. London: British Broadcasting Corporation.
- Lalonde, R. N., & Silverman, R. A. (1994). Behavioral preferences in response to social injustice: The effects of group permeability and social identity salience. *Journal of Personality and Social Psychology*, 66, 78–85.
- Laudat, M. H., Cerdas, S., Dournier, C., Guiban, D., Guilhaume, B., & Luton, J. P. (1988). Salivary cortisol measurement: A practical approach to assess pituitary-adrenal function. *Journal of Clinical Endocrinology and Metabolism*, 66, 343–348.
- LeBon, G. (1895, translated 1947). *The crowd: A study of the popular mind*. London: Ernest Benn.
- Levine, R. M. (2003). Times, theories and practices in social psychology. *Theory and Psychology*, 13, 53–72.
- Leyens, J. P., Cortes, B., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R., Paladino, M. P., Rodriguez-Perez, A., Rodriguez-Torres, R., & Vaes, J. (2003). Emotional prejudice, essentialism, and nationalism. *European Journal of Social Psychology*, 33, 703–717.
- Lindsay, R. C. L., & Adair, R. G. (1990). Do ethically recommended research procedures influence the perceived ethicality of social psychological research? *Canadian Journal of Behavioral Science*, 22, 282–294.
- 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.
- Lyon, D. (1994). *The electronic eye*. Oxford: Polity.
- Lyon, D., & Zuriek, E. (1996). *Computers, surveillance and privacy*. Minneapolis: University of Minnesota Press.
- McDermott, M., Öpik, L., Smith, S., Taylor, S., & Wills, A. (2002). *The Experiment: Report of the independent ethics panel*. <http://www.experimentethics.org.uk>.
- McGarty, C., & Smithson, M. (2005). Independence and non-independence: A simple test for comparing groups using multiple measures and the binomial test. *European Journal of Social Psychology*, 35, 171–180.
- 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 and Row.
- Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. Nye & A. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 57–84). Newbury Park, CA: Sage.
- Morgan, A. H. (1979). Comments on S. H. Lovibond, X. Mithiran and W. G. Adams "The effects of three experimental prison environments on the behaviour of non-convict volunteer subjects". *Australian Psychologist*, 14, 285–286.
- Navas, M., Morales, J. M., & Moya, M. C. (1992). Categorization and prototypicality as determinants of leader effectiveness. *International Journal of Psychology*, 27, 289–290.
- Norusis, M. J. (1985). *SPSSx: Advanced statistics guide*. New York: McGraw-Hill.
- Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Oxford: Blackwell.
- Orford, J. (1992). *Community psychology*. Chichester: Wiley.
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington.
- Organ, D. W. (1997). Organizational citizenship behavior: Its construct clean-up time. *Human Performance*, 10, 85–97.

- Pfeffer, J. (1981). *Power in organizations*. Boston, MA: Pitman.
- Postmes, T. (2003). A social identity approach to communication in organizations. In S. A. Haslam, D. van Knippenberg, M. J. Platow, & N. Ellemers (Eds.), *Social identity at work: Developing theory for organizational practice*. Philadelphia, PA: Taylor & Francis.
- Postmes, T., Baray, G., Haslam, S. A., & Morton, T. (in press). Personal and social identity formation through interaction: Individuality in small groups, social movements and categories. *Individuality and the group: Advances in social identity*. London: Sage.
- Postmes, T., & Spears, R. (1998). Deindividuation and anti-normative behaviour: A meta-analysis. *Psychological Bulletin*, 123, 238–259.
- Postmes, T., Spears, R., Lea, M., & Reicher, S. D. (2000). *SIDE issues centre stage*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Rees, L. (2002). *The Nazis: A warning from history*. London: BBC Worldwide.
- Reicher, S. D. (1982). The determination of collective action. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 41–84). Cambridge: Cambridge University Press.
- Reicher, S. D. (1996). Social identity and social change: Rethinking the context of social psychology. In P. Robinson (Ed.), *Social groups and identities: Developing the legacy of Henri Tajfel*. Oxford: Butterworth-Heinemann.
- Reicher, S. D. (2001). The psychology of crowd dynamics. In M. Hogg & S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 175–197). Oxford: Blackwell.
- Reicher, S. D. (2004). The context of social identity: Domination, resistance and change. *Political Psychology*, 25, 921–946.
- Reicher, S. D., Drury, J., Hopkins, N., & Stott, C. (2001). A model of crowd prototypes and crowd leadership. In C. Barker, M. Lavalette, & A. Johnson (Eds.), *Leadership and social movements*. Manchester: Manchester University Press.
- Reicher, S. D., & Emler, N. (1985). Delinquent behaviour and attitudes to formal authority. *British Journal of Social Psychology*, 24, 161–168.
- Reicher, S. D., & Haslam, S. A. (2003). Social psychology, science, and surveillance: Understanding *The Experiment*. *Social Psychology Review*, 5, 7–17.
- Reicher, S. D., & Haslam, S. A. (2004). Why not everyone is a torturer. Retrieved May 10, from: <http://news.bbc.co.uk/1/hi/magazine/3700209.stm>.
- Reicher, S. D., & Haslam, S. A. (in press). On the agency of individuals and groups: Lessons from the BBC Prison Study. In T. Postmes & J. Jetten (Eds.) *Individuality and the group: Advances in social identity*. London: Sage.
- Reicher, S. D., & Hopkins, N. (1996). Seeking influence through characterising self-categories: An analysis of anti-abortionist rhetoric. *British Journal of Social Psychology*, 35, 297–311.
- Reicher, S. D., & Hopkins, N. (2001a). Psychology and the end of history: A critique and a proposal for the psychology of social categorization. *Political Psychology*, 22, 383–407.
- Reicher, S. D., & Hopkins, N. (2001b). *Self and nation: Categorization, contestation and mobilisation*. London: Sage.
- Reicher, S. D., & Hopkins, N. (2004). On the science of the art of leadership. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership, power and identity* (pp. 197–209). London: Sage.
- Reicher, S. D., & Levine, M. (1994). Deindividuation, power relations between groups and the expression of social identity: The effects of visibility to the outgroup. *British Journal of Social Psychology*, 33, 145–163.
- Reicher, S. D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology*, 6, 161–198.
- Reynolds, K. J., & Platow, M. J. (2003). Why power in organizations really should be shared: Understanding power through the perils of powerlessness. In S. A. Haslam, D. van Knippenberg, M. J. Platow & N. Ellemers (Eds.), *Social identity at work: Developing theory for organizational practice*. Philadelphia, PA: Psychology Press.

- Reynolds, K. J., Turner, J. C., Haslam, S. A., & Ryan, M. K. (2001). The role of personality and group factors in explaining prejudice. *Journal of Experimental Social Psychology*, 37, 427–434.
- Robinson, W. P. (Ed.). (1996). *Social groups and identities: Developing the legacy of Henri Tajfel*. Oxford: Butterworth-Heinemann.
- Schwarzer, R. (2001). Stress, resources, and proactive coping. *Applied psychology: An international review*, 50, 400–407.
- Sherif, M. (1956). Experiments in group conflict. *Scientific American*, 195, 54–58.
- Sherif, M. (1966). *Group conflict and co-operation: Their social psychology*. London: Routledge and Kegan Paul.
- Sieber, J. E., Iannuzzo, R., & Rodriguez, B. (1995). Deception methods in psychology: Have they changed in 23 years? *Ethics and Behavior*, 5, 67–85.
- Smith, E. R., & Mackie, D. M. (2000). *Social psychology* (2nd ed). New York: Worth.
- Smith, H. J., Tyler, T. R., & Huo, Y. (2003). Interpersonal treatment, social identity and organizational behaviour. In S. A. Haslam, D. van Knippenberg, M. J. Platow, & N. Ellemers (Eds.), *Social identity at work: Developing theory for organizational practice*. Philadelphia, PA: Psychology Press.
- Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research. *Journal of Personality and Social Psychology*, 44, 1075–1082.
- Smithson, M. (2000). *Statistics with confidence*. London: Sage.
- Spears, R., Oakes, P. J., Ellemers, N., & Haslam, S. A. (Eds.). (1997). *The social psychology of stereotyping and group life*. Oxford: Blackwell.
- Spears, R., & Smith, H. J. (2001). Experiments as politics. *Political Psychology*, 22, 309–330.
- Staub, E. (1989). *The roots of evil: The origins of genocide and other group violence*. Cambridge: Cambridge University Press.
- Tajfel, H. (1978). Interpersonal behaviour and intergroup behaviour. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 27–60). London: Academic Press.
- Tajfel, H. (Ed.). (1982). *Social identity and intergroup relations*. Cambridge: Cambridge University Press.
- Tajfel, H., Flament, C., Billig, M. G., & Bundy, R. F. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1, 149–177.
- 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.
- Tedeschi, J. T. (Ed.). (1981). *Impression management theory and social psychological research*. New York: Academic Press.
- Terry, D. J., Carey, C. J., & Callan, V. J. (2001). Employee adjustment to an organizational merger: An intergroup perspective. *Personality and Social Psychology Bulletin*, 27, 267–290.
- Tetlock, P. E. (1985). Accountability: The neglected social context of judgement and choice. *Research in Organizational Behavior*, 7, 297–332.
- 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. (1981). Some considerations in generalizing experimental social psychology. In G. M. Stephenson & J. H. Davis (Eds.), *Progress in applied social psychology* (Vol. 1, pp. 3–34). Chichester: Wiley.
- 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. (1985). Social categorization and the self-concept: A social cognitive theory of group behaviour. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–122). Greenwich, CT: JAI Press.
- Turner, J. C. (1991). *Social influence*. Milton Keynes: Open University Press.

- Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 6–34). Oxford: Blackwell.
- Turner, J. C. (2005). Explaining the nature of power: A three-process theory. *European Journal of Social Psychology*, 35, 1–22.
- Turner, J. C., & Brown, R. J. (1978). Social status, cognitive alternatives and intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 201–234). London: Academic Press.
- 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., Reynolds, K. J., Haslam, S. A., & Veenstra, K. (in press). Reconceptualizing personality: Producing individuality through defining the personal self. In T. Postmes & J. Jetten (Eds.) *Individuality and the group: Advances in social identity*. London: Sage.
- Tyler, T. R., & Blader, S. L. (2000). *Cooperation in groups. Procedural justice, social identity, and behavioral engagement*. Philadelphia, PA: Psychology Press.
- Underwood, P. W. (2000). Social support: The promise and reality. In B. H. Rice (Ed.), *Handbook of stress, coping and health* (pp. 367–391). Newbury Park, CA: Sage.
- Wright, S. C. (1997). Ambiguity, social influence, and collective action: Generating collective protest in response to tokenism. *Personality and Social Psychology Bulletin*, 23, 1277–1290.
- Wright, S. C., & Taylor, D. M. (1998). Responding to tokenism: Individual action in the face of collective injustice. *European Journal of Social Psychology*, 28, 647–667.
- Wright, S. C., Taylor, D. M., & Moghaddam, F. M. (1990). Responding to membership in a disadvantaged group: From acceptance to collective protest. *Journal of Personality and Social Psychology*, 58, 994–1003.
- Zimbardo, P. (1969). The human choice: Individuation, reason and order versus deindividuation, impulse and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska symposium on motivation* (Vol. 17, pp. 237–307). Lincoln, NE: University of Nebraska Press.
- Zimbardo, P. (1989). *Quiet rage: The Stanford prison study video*. Stanford, CA: Stanford University.
- Zimbardo, P. (2001). Opposing terrorism by understanding the human capacity for evil. *Monitor on Psychology*, 32, 48–50.
- Zimbardo, P. G., Maslach, C., & Haney, C. (1999). Reflections on the Stanford prison experiment: Genesis, transformations, consequences. In T. Blass (Ed.), *Obedience to authority: Current perspectives on the Milgram paradigm* (pp. 193–237). Mahwah, NJ: Erlbaum.