

Gibson, Stephen ORCID logoORCID:
<https://orcid.org/0000-0002-5648-7669>, Blenkinsopp, Grace,
Johnstone, Elizabeth and Marshall, Aimee (2018) Just following
orders? The rhetorical invocation of 'obedience' in Stanley
Milgram's post-experiment interviews. *European Journal of Social
Psychology*, 48 (5). pp. 585-599.

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/2597/>

The version presented here may differ from the published version or version of record. If
you intend to cite from the work you are advised to consult the publisher's version:
<http://onlinelibrary.wiley.com/doi/10.1002/ejsp.2351/full>

Research at York St John (RaY) is an institutional repository. It supports the principles of
open access by making the research outputs of the University available in digital form.
Copyright of the items stored in RaY reside with the authors and/or other copyright
owners. Users may access full text items free of charge, and may download a copy for
private study or non-commercial research. For further reuse terms, see licence terms
governing individual outputs. [Institutional Repository Policy Statement](#)

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorks.ac.uk

Just following orders? The rhetorical invocation of ‘obedience’ in Stanley Milgram’s post-experiment interviews.

Stephen Gibson, Grace Blenkinsopp, Elizabeth Johnstone, and Aimee Marshall

York St John University

Author Note

Stephen Gibson, School of Psychological & Social Sciences, York St John University; Grace Blenkinsopp, School of Psychological & Social Sciences, York St John University; Libby Johnstone, School of Psychological & Social Sciences, York St John University; Aimee Marshall, School of Psychological & Social Sciences, York St John University.

Aimee Marshall is now at the School of Social Sciences, Business and Law, University of Teesside.

We acknowledge the assistance of Cynthia Ostroff, Mary Caldera and Stephen Ross of Yale University’s Manuscripts and Archives service in accessing material from the Stanley Milgram Papers archive. Permission to quote extracts from condition 02 of the Milgram obedience experiments is granted by Alexandra Milgram. The research was supported by grants from the Nuffield Foundation (SGS/36502) and the Leverhulme Trust (RF-2015-431). We would like to thank Graham Hamilton for his comments on an earlier version of this paper.

Correspondence concerning this article should be addressed to Stephen Gibson, School of Psychological & Social Sciences, York St John University, York, UK, YO31 7EX. E-mail: s.gibson@yorks.j.ac.uk

Abstract

Recent research has begun to challenge the received idea that Milgram's 'obedience' experiments are demonstrations of obedience as typically understood (i.e. as social influence elicited in response to direct orders). One key warrant for explaining the studies in terms of obedience has been the post-experiment interviews conducted with participants. The present study uses data from archived audio recordings of these interviews to highlight the extent to which participants used rhetorical strategies emphasising obedience when pressed by the interviewer to account for their behaviour. Previous research that has used these accounts as reports of underlying processes misses the extent to which they performed particular social actions in the context of their production. It is concluded that the standard social psychological version of 'obedience' *is* present in the experiments after all, but in a rather different way than is typically assumed – rather than an empirical finding, obedience is a participants' resource.

Keywords: discourse, Milgram, obedience, rhetoric, social influence

Just following orders? The rhetorical invocation of ‘obedience’ in Stanley Milgram’s post-experiment interviews.

Claiming obedience as a strategy for avoiding responsibility for one’s actions is inextricably associated with the Nuremberg trials of perpetrators of the Nazi Holocaust following the Second World War. The claim to have been ‘just following orders’ functions to mitigate one’s culpability insofar as it positions oneself as having being a mere subsidiary actor in a chain of authority. In legal terms, it is known as the defence of *superior orders* (e.g. Gaeta, 1999), and implies that those at higher levels in the chain of command are those who should really be held to account. In social psychology, the study of obedience is of course most famously associated with Stanley Milgram. Milgram – who was explicitly aiming to shed light on some of the psychological mechanisms that had led to the Holocaust – conducted a series of experiments at Yale University in the early 1960s which purported to show how orders from an authority figure could lead large numbers of US citizens to cause harm to an innocent person (Milgram, 1963, 1965a, 1974). These studies, together with Hannah Arendt’s (1963/1977) thesis on the banality of evil, were key elements in the crystallization of a conventionalized view of the Holocaust as having being the result primarily of ordinary people being caught up in a bureaucratic process and simply following orders (Reicher, Haslam & Miller, 2014).

In recent years this perspective has come under increasing scrutiny from both historians and psychologists (e.g. Lang, 2014; Overy, 2014; Haslam & Reicher, 2007). Milgram’s studies were always controversial (e.g. Baumrind, 1964; Miller, 1986; Orne & Holland, 1968), but recent years have seen a new wave of critique which has brought into question the extent to which the studies can be seen as demonstrations of obedience at all. The present study seeks to extend this critique by drawing attention to hitherto under-

explored aspects of Milgram's own data – his interviews conducted with participants who had just completed the experiment. Analysis of these data point to two key issues:

First, it challenges the use of the post-experiment interviews as sources of objective data on responsibility. Milgram (and others) have used these data to argue that 'obedience' is the product of a psychological shift whereby individuals no longer see themselves as responsible for their own actions. Detailed analysis of the interviews suggests instead that participants were engaged in the contextually-contingent activity of accounting for their conduct in the situation immediately preceding the interview. Any attempt to use these data which neglects this misses the extent to which participant accounts of responsibility serve a particular function in the specific social context of their production.

Second, while recent research has cast doubt on Milgram's experiments as demonstrations of obedience in terms of following orders, the present analysis will show that this version of 'obedience' *does* make an appearance in the experiments, albeit in a rather different way than is typically assumed. Instead of providing a conceptual tool for *analysts* to explain Milgram's experimental findings, it will instead be shown that 'obedience' can be re-cast as a *participants' resource* for making sense of – and accounting for – their behaviour in the experiment. Before outlining these findings more fully, however, it is worth re-visiting some of the key elements of Milgram's experiments, and their recent re-evaluation, in more detail.

The 'obedience' experiments

The 'official' account of Milgram's experiments is well-known, but it is useful to provide a brief overview here in order to juxtapose it with the emerging re-evaluation of the studies to be discussed below. Milgram (1974) reports 18 different experimental conditions, but it is arguable that a core group of these have come to constitute the most widely

disseminated version of the experiments. This is due in no small part to the extent to which a particular condition dominates the narrative of Milgram's (1965b) film of the experiments, which is widely used on undergraduate psychology courses and which has informed the broader circulation of Milgram's ideas beyond academia (e.g. in Michael Almereyda's 2015 film *Experimenter*).

In these versions of the experiment, which correspond with conditions 2, 5, 6 and 8 in Milgram's (1974) fullest account of the studies, a naïve participant arrived at a laboratory to take part in what was apparently an experiment on the effects of punishment on learning. The participant was greeted by an experimenter, and shortly afterwards another participant arrived. Unbeknownst to the naïve participant, this second person was in fact a confederate. Through a rigged allocation process, the naïve participant was given the role of teacher, and the confederate the role of learner. The experimenter explained that the teacher was required to administer a memory test to the learner, and to deliver punishments in the form of electric shocks when the learner made a mistake. The shocks started from 15 volts, and increased in 15-volt increments to a maximum of 450 volts. The learner was strapped into a chair in an adjoining room, and electrodes were attached to his arm in order to deliver the shocks. Once the procedure began, it became clear quite quickly that the learner was struggling with the task and that it would be necessary to move up the shock levels. When the shocks reached 75 volts, the learner began to grunt, and these exclamations escalated until the 150-volt shock was administered, at which point the learner demanded to be released. If participants continued administering shocks beyond this point, the learner's protests continued to increase in intensity along with the increasing shock levels. At 300 volts the learner refused to answer any more, and shortly thereafter the protests ceased, creating the impression that he had perhaps lost consciousness. Of course, the learner did not receive any real shocks and the protests were all pre-recorded on tape and played back at the appropriate shock level.

If participants hesitated or refused to continue at any point, the experimenter had at his disposal four sequential prods that were to be used in order, and begun again for each separate attempt at defiance:

Prod 1: Please continue, *or*, Please go on.

Prod 2: The experiment requires that you continue.

Prod 3: It is absolutely essential that you continue.

Prod 4: You have no other choice, you *must* go on.

(Milgram, 1974, p. 21, italics in original)

Only if participants still resisted after receiving the fourth prod did the experimenter draw the session to a halt. Milgram used his procedure to create a binary dependent variable whereby any participant who continued to the end of the shock generator, and thus administered all the shocks, was classified as obedient, whereas any participant who defied the experimenter and brought the session to a close was classified as disobedient. Under these conditions, Milgram found obedience levels of between 50% and 65%. However, in recent years the experiments have been subjected to a renewed level of critical scrutiny that has challenged the extent to which they were successful demonstrations of obedience at all.

Obeying orders?

Recent scholarship has drawn attention to a range of ethical (e.g. Nicholson, 2011; Perry, 2012), theoretical (e.g. Gibson, 2013a; Reicher & Haslam, 2011) and methodological (e.g. Gibson, 2013b; Russell, 2011) issues that challenge the received view of Milgram's experiments. For present purposes, one key emergent finding from this work is of particular note: the suggestion that the experiments are not demonstrations of obedience as typically understood (i.e. as social influence elicited in response to a direct order), but rather that they

in fact stand as demonstrations of the overwhelming *failure* of attempts to exert influence through direct orders.

Specifically, it has been suggested that Milgram's experimenter did not typically issue orders (Gibson, 2013a), and that most of the standardized prods fail to constitute orders as such. For example, 'Please continue' appears to be a polite request rather than an order. Indeed, Reicher and Haslam (2011, p. 167) have suggested that it is only the fourth prod that constitutes an order or command, and have therefore argued that, 'The question of whether or not people obey this fourth prompt is decisive in establishing the validity of those interpretations of Milgram's studies that see them as a demonstration of how people follow orders.' Recently, three lines of convergent evidence have suggested that the fourth prod was actually *ineffective* at getting participants to continue administering shocks. Specifically, it has been shown that hardly any participants continued after receiving the fourth prod in a selection of Milgram's experimental conditions (Gibson, 2013a); that no participants continued after receiving the fourth prod in a recent partial replication of Milgram's paradigm (Burger, Girgis & Manning, 2011); and that the fourth prod was not particularly effective in an experimental analogue designed to unconfound the ordering of Milgram's sequential prods (Haslam, Reicher & Birney, 2014). It thus appears that, whatever Milgram's experiments show, they *do not* show that people have a propensity to obey direct orders. This raises the question of why, for over fifty years, the experiments have been understood in these terms. There are a number of reasons for this, not least amongst which is the extent to which they provided an explanation which chimed with other attempts to understand the processes that led to the Holocaust, and in particular with Hannah Arendt's (1963/1977) 'banality of evil' thesis.

The agentic state and the banality of evil

The suggestion that the Holocaust was the product of a large number of people simply doing their jobs and unthinkingly following orders has proved to be remarkably durable. As noted above, this is largely due to the twin intellectual influences of Milgram's experiments and Arendt's (1963/1977) 'banality of evil' thesis. However, as many scholars have argued, the idea of the banality of evil is something of a simplification based on a mis-reading of Arendt's account of the trial of Adolf Eichmann (e.g. Lang, 2014; Reicher, 2014). In recent years, however, several scholars have set about systematically challenging the received view (e.g. Brannigan, 2013; Cesarani, 2004; Lang, 2014). Haslam and Reicher (2007, 2008) summarize these critiques, with some scholars suggesting that Eichmann was engaged in a project of self-presentation to minimize his culpability during the early part of his trial. Several historians have argued that Eichmann was much more committed to the Nazi project than the 'banality of evil' thesis suggests, and – crucially – Cesarani (2004, p. 11) has argued that 'It is a myth that Eichmann unthinkingly followed orders'. This revised view is echoed in the wider literature on the Holocaust (e.g. Fenigstein, 2015; Mastroianni, 2015; Overy, 2014), which has cast doubt on the idea that the Nazi atrocities were the result of people simply doing their duty and blindly following orders. These debates are far from settled, but they nevertheless highlight the extent to which the 'banality of evil' thesis is no longer tenable as a straightforward understanding of the Holocaust.

This critique of the received view of Eichmann has important echoes in the way in which the 'obedience' experiments have been interpreted. Milgram's (1974) own theoretical account of his findings was based on the idea of the agentic state. Milgram suggested that obedience involved a psychological shift in which individuals no longer feel themselves to be in control of their own actions, but rather are merely an *agent* for the commands of an

authority figure. In this respect, Milgram's explanation complements the 'banality of evil' thesis, and indeed this is hardly surprising given that Milgram was explicitly influenced by Arendt's work during the formulation of his theoretical account (Blass, 2004). Milgram argued that,

Arendt's conception of the *banality of evil* comes closer to the truth than one might dare imagine. The ordinary person who shocked his victim did so out of a sense of obligation – a conception of his duties as a subject – and not from any particularly aggressive tendencies.

(Milgram, 1974, p. 6, italics in original)

This highlights the extent to which Milgram was keen to contrast his account with personality-based explanations for authoritarian behaviour (e.g. Adorno, Frenkel-Brunswik, Levinson & Sanford, 1950; and see Elms & Milgram, 1966). Importantly, Milgram's interpretation relied on the self-reported responses of participants in post-experiment interviews. In the early experimental conditions, participants completed a measure in which they attributed responsibility for administering shocks against the learner's will. Milgram used an instrument resembling a pie chart, termed the 'responsibility clock' (see Milgram, 1974, p. 204) which enabled participants to divide responsibility between the experimenter, the learner and themselves.

Milgram (1974) reports that participants who had continued with the experiment until the highest level on the shock generator (i.e. 'obedient' participants) attributed less responsibility to themselves compared to participants who had defied the experimenter. This finding is used to support the suggestion that 'obedient' participants no longer perceived themselves as being fully in control of their actions, but were rather merely acting out the demands of the authority figure. Taken together, Milgram's work and the idea of the banality of evil exerted a powerful hold on attempts to make sense of the Holocaust, to the extent that

it has been suggested that they ‘straightjacketed research for two decades’ (Cesarani, 2004, p. 15).

However, Milgram’s theoretical explanation has long been regarded as flawed, even by the most sympathetic of commentators (e.g. Blass, 2004; Miller, 1986). In particular, Mantell and Panzarella (1976) cast doubt on the extent to which the responsibility data actually support Milgram’s theory. In this context, it might be argued that Milgram’s responsibility data and the agentic state theory have been conclusively discarded and further consideration of these issues is not warranted. However, the received view of Milgram’s findings still enjoys pre-eminence in textbook accounts (Griggs & Whitehead, 2015a, b), and indeed the issue of responsibility is still very much a live one in recent work on obedience. For example, Caspar, Christensen, Cleeremans and Haggard (2016) have recently used measures of electrophysiological activity to argue that coercive conditions lead to a reduced sense of agency. They conclude that the defence of ‘just following orders’ may not simply be an excuse but instead reflects people’s experiential reality. There has even been a recent attempt to rehabilitate the concept of the agentic state, with Fennis and Aarts (2012) suggesting that a reduction in perceptions of personal control increases susceptibility to social influence, and explicitly framing this in terms of Milgram’s theory.

Of most direct relevance to the present study, however, is Burger, Girgis and Manning’s (2011) analysis of participants’ post-experiment comments following a partial replication of Milgram’s paradigm (Burger, 2009). It is worth considering this study, and its relationship with Milgram’s own analysis, in more detail, and in particular the assumptions made by both Milgram and Burger et al concerning the status of participant accounts of responsibility.

Accepting and denying responsibility

It was noted above that advocates of the ‘banality of evil’ thesis have been challenged on the grounds that they may have been taken in by Eichmann’s attempts at self-presentation. Milgram dismisses similar explanations for his own participants’ denial of responsibility:

‘The most frequent defense of the individual who has performed a heinous act under command of authority is that he has simply done his duty. In asserting this defense, the individual is not introducing an alibi concocted for the moment but is reporting honestly on the psychological attitude induced by submission to authority’ (Milgram, 1974, p. 146).

Given that in other respects, Milgram’s research has been used to show people’s inability to appreciate the impact of social forces on their behaviour, it is perhaps surprising that he was so ready to treat these post-experiment accounts as straightforward verbal reports of underlying processes. We can understand this as an example of selective reification (Potter, 1996), in which the analyst selectively identifies some accounts as flawed, while choosing to reify – to treat as literally true – others. Similar assumptions can be identified in Burger et al’s (2011) research.

Burger et al (2011) undertook a content analysis of participant responses in Burger’s (2009) partial replication of the Milgram experiment. In one set of analyses, Burger et al focussed on whether participants had made comments at any stage during the experimental session or post-experiment debriefing interview that indicated that they took responsibility for the potential consequences of their actions during the experiments. In both the experimental sessions and the interviews, defiant participants were more likely to state that they were personally responsible than participants who continued past the 150 volt shock level (the highest level in Burger’s procedure). Burger et al argued that ‘[t]he results support the notion that a sense of personal responsibility contributed to the participants’ decision to

continue or end the procedure during Burger's (2009) replication of Milgram's studies' (Burger et al, 2011, p. 463).

However, this explanation relies on the same set of assumptions about the status of participant accounts as does Milgram's use of his post-experiment interviews. Indeed, Burger et al explicitly address the issue of the believability of the comments:

'Because the spontaneous comments came without prompting and because participants were unlikely to have anticipated that these comments would be analyzed by the investigators, the comments provide a unique and relatively uncensored glimpse into what participants were thinking as they moved through the experimental procedures' (Burger et al, 2011, p. 461).

We suggest that such an assumption is not warranted, and indeed that it results from a longstanding neglect of the constructive and functional nature of language in social psychology. Instead, a position informed by discursive-rhetorical psychology highlights the extent to which language should be understood as an action-oriented medium where people do not simply describe the world, but rather they construct it.

Discursive-rhetorical psychology

Since the 'turn to language' in the 1980s, a number of approaches within social psychology have sought to take seriously the argument that language is not a relatively straightforward and transparent medium for the communication of thoughts from one mind to another, but rather that it is a fundamentally action-oriented medium which is geared towards *doing things*. For example, conversation analysts have highlighted the virtues of detailed analysis of interaction on a line-by-line basis to show how social order is maintained in the most mundane of contexts (e.g. Hutchby & Wooffitt, 1998). By contrast, discourse analysts influenced by Foucauldian perspectives have sought to map the contours of power and

ideology within language (e.g. Parker, 1992). Other scholars, drawing on Gilbert and Mulkay's (1984) characterization of their study of scientists' discourse as occupying 'a middle ground' (p. 17) position, have attempted to plot a course between micro and macro approaches to the analysis of discourse, and the present study proceeds in a similar spirit. Thus, for present purposes we want to draw together strands from Potter and Wetherell's (1987) classic outline of discourse analysis, Edwards and Potter's (1992; Potter & Edwards, 2001) subsequent development of discursive psychology, and Billig's (1996) rhetorical psychology perspective. For ease of reference we term this approach discursive-rhetorical psychology (DRP), and it is worth briefly outlining four key propositions that we derive from it.

Discourse as constructive. DRP conceptualises discourse as inherently constructive. Rather than merely reflecting the way the world is, it is instead through language that we construct the world. Analytically, therefore, describing something as a fact, for example, is not to be assessed in terms of the extent to which the claim *really is* factual, but rather it is to be analysed for how specific terms are deployed in such a way as to create the impression of factuality (Potter, 1996).

Discourse as functional. When people are speaking, they are not simply reporting on the world as it is, but rather their words are geared towards action. Language is used to perform a whole manner of social actions, such as requesting, blaming, excusing, identifying, promising, and so on. Moreover, the action-orientation of a formulation is not determined by its grammatical form. Something that is ostensibly a statement (e.g. 'It's rather warm in here') can – in an appropriate context – be treated as a request (e.g. if, on hearing the utterance, someone goes to open a window).

Discourse as situated. DRP directs analytic attention towards the specific contexts in which discourse is used. Rather than treating accounts as reflective of psychological or social reality, DRP conceptualizes them as tied to the business of the immediate social context.

Discourse as rhetorical. Finally, DRP emphasises the extent to which discourse is fundamentally rhetorical insofar as people are constantly engaged in activities designed to persuade others of something. This extends far beyond self-evidently adversarial encounters to encompass the full range of mundane and everyday situations in which an overarching concern with accountability can be identified. By this is meant that in the course of everyday life people are constantly orienting to the possibility that they may be held to account for their actions in one way or another.

Previous research on Milgram's studies has applied this approach to the experimental interaction between the experimenter and participant, with a particular focus on the way in which participants sought to argue their way out of the experiments, and the way in which the experimenter adapted the standardized procedure in an attempt to keep participants in the experiment (Gibson, 2013a, b, 2014, 2017a; also see Hollander, 2015; Hollander & Maynard, 2016). In this previous research, it has been argued that the experimental interaction between experimenter and participant should be understood as a rhetorical encounter; here, we suggest that the interactions in the post-experimental interviews should be seen in similar terms. Whereas Milgram – and others, such as Burger et al (2011) – have used such interviews as a source of information about underlying processes, we conceptualize these encounters as occasions on which participants were constructing discourse in order to perform particular actions in the specific social context in which they found themselves. Moreover, participants will be in a position of having to account for their behaviour, and as such, a failure to conceptualise the interviews as rhetorical encounters will misconstrue the nature of these encounters. Such an analysis is overdue: The acceptance of the flaws in Milgram's account

has not led to further consideration of the data used by Milgram to develop his agentic state theory. Our guiding research questions were thus, how is obedience invoked in the interviews, and what is it doing on those occasions when it is invoked? The present analysis therefore addressed the question of how this classic understanding of obedience – as following orders – is invoked *by participants themselves* as they account for their behaviour in the post-experiment interviews.

Method

Data

The data are drawn from condition 02 of Milgram's experiment, first reported by Milgram (1965a), and also included in his book-length account of the studies (Milgram, 1974). This condition, known as the voice-feedback condition, was selected for use in the present analysis as it is the only condition that uses both the procedure described above and the responsibility clock instrument. Forty participants – all males – took part in the condition, 25 (62.5%) of whom proceeded to the highest point on the shock generator and were thus classified as obedient by Milgram. The condition was conducted in late August 1961, beginning just a few days after the judges in the Eichmann trial had retired to consider their verdict.

In common with most of Milgram's experimental conditions, audio recordings from the voice-feedback condition are held in the Stanley Milgram Papers archive at Yale University. Thirty-nine experimental sessions from the voice-feedback condition are present in the archive. For the purposes of the present analysis, the post-experiment interviews were transcribed in full using a simplified form of Jeffersonian transcription notation to capture the content of the interviews as well as some basic interactional features such as pauses, overlap and false starts (see Appendix for conventions).

The interviews were conducted immediately after the completion of the experimental session, either by John Williams, still in his role as experimenter, or by Milgram himself. At the time the interviews were conducted, participants had not been informed that the experimental set-up was a hoax, and indeed they were not told this until much later when they received a written report on the experiments (Perry, 2012). In neither of the publications in which the condition is reported (Milgram, 1965, 1974) are details of the interview schedule reported, however the interviews typically involved the participants being asked to explain the experiment in their own words, and being asked to improvise a 'story' based on a picture in which punishment is being administered to what appears to be a misbehaving schoolboy. Participants were also asked to complete a number of measures, of which all but the responsibility clock were completed via paper and pencil. At various points throughout this process, the interviewer could ask participants to expand on their answers, and particular lines of questioning appear to have been improvised depending upon the actions of participants in the experimental sessions (e.g. participants who laughed could be asked about the reasons for this). In modern parlance, the interviews appear to have followed what might be termed a 'semi-structured' format, although we might be cautious about applying such terms retrospectively.

Analytic Procedure

Analysis involved the identification of terms that can be said to relate to the standard social psychological definition of obedience – the defence of superior orders. Thus, references to *following orders/instructions/commands, doing as one is told*, and similar formulations were treated as part of the same interpretative system, or repertoire. Potter and Wetherell (1987, p. 149) defined interpretative repertoires as 'recurrently used systems of terms used for characterizing actions, events and other phenomena.'

The interview transcripts were read repeatedly to identify instances of the obedience repertoire, and more broadly to ascertain how the discussions of responsibility and related issues unfolded. At first, the analysis aimed at over-inclusion, so that borderline cases were retained. These included cases where the interviewer (rather than participants) invoked orders/commands, and where more oblique references were made to the experimenter being ‘in charge’ of the experiments. Subsequent readings refined this analysis to the point where only those occasions where the repertoire was used by participants clearly in reference to themselves were included in the analytic corpus. The resulting data were again read repeatedly with a view to identifying in more detail how participants invoked the superior orders defence. At this stage, the micro-interactional features of the interviews were attended to, with a particular focus on the use of specific discursive devices and techniques (Edwards & Potter, 1992). An over-arching principle was the concern for the four key elements of the DRP approach outlined above: the analysis sought to explore how the superior orders defence was constructed, what functions it performed, how it was situated in its specific context of articulation, and how it formed part of particular rhetorical strategies designed to mitigate participants’ responsibility for their actions. In short, we were concerned with the situated construction of superior orders for the performance of particular rhetorical actions.

Analysis & Discussion

In presenting the analysis, two major findings will be highlighted: First, it will be shown how the obedience repertoire was used by participants as a strategy in accounting for shocking the learner against his will. Second, attention will be drawn to the role of the interviewer in these exchanges, and in particular to what might best be described as the use of a strategy of ‘active’ interviewing to hold participants to account. Ultimately, the post-experiment interviews did not access participants’ feelings of responsibility in any

straightforward, neutral or objective sense, but rather they made responsibility a matter of contestation and accountability.

Obedience as accounting strategy

A total of 13 (12 obedient; 1 defiant) participants drew on the superior orders repertoire in the post-experiment interviews. As Milgram (1974) himself noted, many participants did indeed attribute responsibility to the experimenter, and/or sought to downplay their own responsibility in various ways. However, even a cursory listen to the recordings of the interviews indicates that participants were not passively responding to the use of the responsibility clock, but rather they were in a situation in which they were providing an account for their behaviour. The first extract illustrates the use of the responsibility clock, featuring discussions concerning the percentage of responsibility to be allocated to the experimenter (E), teacher (T) and learner (L). The extract begins after the experimenter has explained the responsibility clock and asked the participant (teacher) to position the arms of the clock in order to indicate his allocation of responsibility. During the opening lines of the extract, the participant is providing an account of the reasons for his allocation:

Extract 1 (0218; 450v.)

- 1 T: I don't think he had much
- 2 responsibility because he had
- 3 nothing to say into- into the
- 4 matter. So his would be very
- 5 small. I say yours was more so
- 6 because you're experimenting
- 7 and you knew, you told me

8 exactly what to do and I was
9 doing what you told me.
10 E: Mmhm
11 T: So it would maybe something
12 like that.
13 E: So you'd give the learner about
14 ten percent
15 T: Yeah.
16 E: And you'd give yourself about
17 thirty percent.
18 T: Thirty, yeah.
19 E: Yeah, and you'd give me, er,
20 well you'd give me the rest
21 about sixty percent.
22 T: Yeah
23 E: [I see]
24 T: [Because] I was certainly
25 following instructions

The participant attributes more responsibility to the experimenter on the grounds that the experimenter knew what he was doing (line 7: 'you knew') and that he told the participant what to do (lines 7-9; lines 24-25). It is worth considering the precise way in which the participant invokes the obedience repertoire here. The reference to the experimenter knowing what he was doing functions to attribute epistemic authority to the experimenter such that the participant not only positions himself as doing what he was told, but doing so because the

experimenter was in a better position than he was to understand what was going on. The subsequent formulation ‘you told me *exactly* what to do’ (lines 7-8) also functions to attribute responsibility to the experimenter insofar as it implies that no leeway was available for the participant to use his judgement – every action was precisely determined by the experimenter.

Following the use of the responsibility measure, the participant provides a further brief account. He gives the experimenter the largest share of the responsibility, and the experimenter glosses this as ‘about sixty per cent’ (line 21). This is not simply treated as a factual statement by the participant, who instead orients to the need for an account and again uses the obedience repertoire, this time with the added firmness of ‘I was *certainly* following instructions’ (lines 24-25) to leave no room for doubt. This formulation also removes the experimenter’s agency insofar as it is simply ‘instructions’ rather than ‘you’ or ‘your instructions’ that the participant has followed. Insofar as the denial of one’s own responsibility in this situation necessitates an attribution of responsibility to the experimenter, who is also acting as interviewer, this highlights a further nuance of the situation in that attributing blame in such a way can also be seen as a face-threatening and accountable action.

This account is in many respects a canonical use of the obedience repertoire in that there is no orientation to the possibility that following instructions might itself be accountable. Instead, the simple claim to have been doing the bidding of someone with greater knowledge of the situation is sufficient to warrant the attribution of greater responsibility to the experimenter.

The use of the obedience repertoire was not restricted to discussions that arose from the use of the formal measure of responsibility. In many interviews, matters of responsibility were also raised by the interviewer in a more general sense. In the following example, Milgram (SM) has been asking the participant about his impressions of the level of pain experienced by the learner:

Extract 2 (Participant 0201; 450v.)

- 1 SM: Did you- did you hear him
2 hollering?
3 T: Yes.
4 SM: H- what did you, how did you
5 feel about that?
6 T: I felt rather, gui-eh-lty.
7 SM: Why?
8 T: Cos I was still er giving him the
9 shocks.
10 SM: Uhuh. Well er, why did you
11 give him the shocks then?
12 T: Why did I give him the shocks?
13 SM: Yeah.
14 T: Because the er, other fellow told
15 me to continue, exactly as the,
16 the word said.

In this extract the participant uses the obedience repertoire in response to the interviewer pressing him as to why he had continued. The interviewer's initial questions are seemingly innocuous but function to position the participant as having blatantly disregarded the learner's 'hollering'. The first question addresses the seemingly straightforward perceptual matter of whether the participant had heard the learner's protests (lines 1-2), and the second concerns how he felt about the 'hollering' (lines 4-5). These questions create a

moral ambiguity: If you can hear someone in pain, and you feel guilty because you are causing the pain, then the appropriate and obvious thing to do is to stop performing the action that is the cause of the pain, and yet the participant did not. The interviewer's subsequent question addresses this incongruity by asking why he continued to administer the shocks (lines 10-11) and it is at this point that the participant draws on the obedience repertoire to account for his actions (lines 14-16). Again, there is no orientation here to the possibility that doing as he was told was in any sense problematic in itself.

In addition to straightforward invocations of following orders and similar formulations, participants could construct analogies with other situations in order to normalise their behaviour. For example, in extract 3, we see a participant explaining why he has attributed a greater proportion of responsibility to himself than to the experimenter:

Extract 3 (Participant 0219; 345v.)

- 1 T: so I do believe that
2 (2)
3 T: eh, in my case here I would take
4 the greater punishment of him
5 E: The [greater responsibility.]
6 T: [The greater respons]ibility to
7 him.
8 E: Yeah, but not as much as I
9 [(((inaudible)))]
10 T: [But not as] much as you
11 because, you be the experimenter
12 let us say, if you work in a place

- 13 you're the boss
- 14 E: mm
- 15 T: So you tell me what to do.

The participant here attributes more responsibility to the experimenter than to himself, and justifies this using an analogy with a workplace in which the experimenter is 'the boss'. The category-defining features of a 'boss' involve the power to make decisions, with the concomitant responsibility for their consequences. In this respect, we might understand the participants' analogy as a form of rhetorical anchoring (Billig, 1991; Moscovici, 1984) insofar as it is used to frame a novel and unfamiliar situation (i.e. the experiment) in terms of a more familiar setting with clearly defined – and consensually understood – roles (i.e. a workplace). In this respect, the assertion that 'you tell me what to do' (line 15) not only functions as an account for the participants' behaviour, but also as a justification for attributing greater responsibility to the experimenter. Thus once again not only is the obedience repertoire used to manage accountability, but doing as one is told is not itself oriented to as accountable.

However, in some cases participants could orient to obedience as accountable even as they were drawing on the obedience repertoire to manage their accountability for having administered shocks to the learner:

Extract 4 (Participant 0206; 450v.)

- 1 T: Oh I- I felt somewhat
- 2 uncomfortable giving him the
- 3 shocks to be honest with you.
- 4 E: You did?

5 T: Yeah.

6 E: And you continued to do so

7 though.

8 T: Well you told me to d- d- do so.

9 E: Yeah well I s-

10 T: I mean I hesitated several times

11 and you can- tell me to keep

12 going. As a matter of fact er

13 long before I reached the end

14 of that board I was tempted to

15 get a- get away from it.

16 E: You were?

17 T: But er-

18 E: Why didn't you?

19 T: Because you were there

20 telling me to carry on.

21 (2)

22 E: Even erm, o- over this man's er,

23 our learner's complaints you, er

24 continued to go on, with me

25 giving the er, the command.

26 T: (?I was just?) following

27 instructions maybe I'm a- an

28 Eichmann eh heh. But er, I

29 really didn't- didn't want to

- 30 continue.
- 31 E: Maybe you were a- maybe you
- 32 were a what?
- 33 T: Eichmann. Eh heh.
- 34 E: Eichmann.
- 35 T: I really didn't want to continue
- 36 even er when I reached the end
- 37 of the board there I was hoping
- 38 I could start over again on the
- 39 lower voltage. And er, o- only
- 40 because you e- e- sternly told
- 41 me to carry on I did eh heh.
- 42 E: I see.
- 43 T: But er, I really didn't want to.

In this extract, the interviewer repeatedly puts the participant in a position where he needs to account for his actions. The participant attempts to maintain as positive a self-presentation as possible by constructing a disjuncture between his outward behaviours and his internal processes. This is insufficient on its own to satisfy the interviewer, who then proceeds to hold him to account for not bringing his behaviour into line with his internal state. In this context, obedience is a useful accounting strategy for the participant insofar as it treats the role of a third party (i.e. the person issuing orders) as key to his continuation and as providing an explanation for the disjuncture between internal state and outward behaviour. It is useful to explore in detail how this process unfolds.

The participant's initial assertion of discomfort in administering the shocks comes in response to a preceding discussion in which the experimenter – now in the role of interviewer – had asked the participant to explain the purpose of the experiment in his own words (a typical opening question in these interviews). At this stage, therefore, there is no indication that the participant is to be pressed on his reasons for continuing and as such his claim bears the hallmarks of confession (line 3: 'to be honest with you') in which he is bringing to the experimenter's attention the fact that the experiment made him feel uncomfortable, while doing so in such a way as to mark it as dispreferred (Edwards & Fasulo, 2006). Such an assertion might ordinarily be oriented to as requiring an account in itself (i.e. because it can be heard as being critical of the experimenter), but of course the experimenter does no such thing and instead turns the rhetorical tables on the participant with his statement on lines 6-7. Note that this is not formulated as a direct accusation, or even as a question. In taking the grammatical form of merely observing that the participant had continued despite the feeling of discomfort, the experimenter nevertheless draws attention to an inconsistency between the participant's stated feelings (discomfort) and his actions (continuing with the task that is the cause of this discomfort).

The participant immediately responds to this, beginning with an appositional (line 8: *well*). Such words are commonly found at the beginning of speaking turns, and function as turn-entry devices, or what we might describe as devices for seizing the floor (Sacks, Schegloff & Jefferson, 1974). In this context the appositional serves as a form of 'yes but' formulation (Pomerantz, 1984) in that it acknowledges what has just been said but rejects the implication (Schiffrin, 1985; see also Gibson, 2017a). The participant is thus meeting the implied contradiction by placing responsibility for continuation on the experimenter ('you told me to d- d- do so'). The experimenter attempts to respond in kind (line 9: 'Yeah well') but is cut off as the participant expands on his account. In doing so, the participant works up

his reluctance to continue both in behavioural (line 10: 'I hesitated several times') and psychological (lines 14-15: 'I was tempted to get a- get away from it') terms. This serves to construct a description of himself as a reluctant continuer who would not have gone on without the instructions of the experimenter. A number of discursive devices are used to build this impression of himself. For example, the reference to hesitating *several times* constructs his resistance as persistent. Similarly, he prefaces his assertion that he was tempted to end the experiment by constructing it as a fact (line 12: 'As a matter of fact') and as something that had occurred to him 'long before I reached the end' (line 13). Thus, he is able to position himself as having wanted to discontinue, and as having displayed outward indications of his lack of willingness, but as having continued because he was told to do so. This strategy serves to maintain something of a moral position for the participant as a basically *good person*, while placing the responsibility for continuation firmly with the experimenter.

In response to the participant's claim to have been tempted to end the experiment, the experimenter then asks simply 'Why didn't you?' This again places the responsibility back with the participant, who again asserts that the reason he didn't stop is due to the experimenter 'telling me to carry on' (line 20). The experimenter's response juxtaposes the 'commands' to continue with 'our learner's complaints' (line 23), to which the participant again invokes the experimenter's 'instructions'. It is not entirely clear from the recording, but it sounds as though the participant begins this phrase with 'I was just', and the use of 'just' functions to minimise the gravity of his actions in line with the canonical 'just following orders' phrase. In any case, his subsequent suggestion that he may be 'an Eichmann' (lines 27-8) clearly shows that he is orienting to the potential parallels between his line of argument and the classic view of Eichmann developing at the time. Following this, he uses extreme case formulation (Pomerantz, 1986) to re-assert his lack of willingness to go on

(lines 28-30: ‘I *really* didn’t- didn’t want to continue.’) This continues on lines 35-43 as the participant works up descriptions of his inner states (lines 35 & 43: ‘I really didn’t want to’; line 37: ‘I was hoping’) and again contrasts these with the instructions of the experimenter. It is notable that in this final formulation of the experimenter’s prodding, the participant upgrades the previous glosses of the experimenter’s instructions. Now, instead of simply *telling* him to go on, the experimenter is constructed as having *sternly* told him to continue (line 40). Obedience is thus used as a strategy for managing the apparently – and self-confessed – inconsistency between internal desire (not wanting to continue) and outward behaviour (continuing to administer the shocks).

It is worth dwelling on the invocation of Eichmann a moment. As noted above, this experimental condition occurred at a time when Eichmann’s trial was salient and as such the moral issues it raised were an important contemporary concern. In orienting to these issues, the participant employs a formulation resembling a rhetorical strategy that Antaki and Wetherell (1999) have termed a *show concession*, in which an assertion is followed by an apparent concession, which is in turn followed by a restatement of the original assertion. This acknowledges, while attempting to dismiss, a potential weakness in the participant’s argument. In this instance, the original assertion (‘I was just following instructions’) and its restatement (‘But er, I really didn’t- didn’t want to continue’) are not directly equivalent, but both function to position the participant as having continued despite not wanting to do so. This concession allows the participant to appear reflexive by acknowledging Eichmann as being a problematic figure, and in so doing to avoid alignment with Eichmann insofar as this is acknowledged by the invocation itself and by the laughter particles. Importantly, this displays the participant’s orientation to the accountability of the obedience repertoire itself, over and above the accountability of the act of delivering shocks against someone’s will.

The participant's strategy appears to constitute a good example of Milgram's claims about those who went all the way on the shock generator being more inclined to pass responsibility to the experimenter. However, in the immediate situation in which they found themselves, passing responsibility was one of the best available options for maintaining as positive an impression of themselves as possible. Constructing a mis-match between desire and behaviour, and using obedience as a device to account for this, represented – in the circumstances – the best strategy for presenting the self as fundamentally moral: *I may have done this thing that I might have been better off not doing, but at least I didn't want to do it.*

‘Active’ interviewing

We have already seen how participants were held to account for their behaviour in the experimental session, and in this sense it is not the case that they were simply completing a straightforward measure of their perceptions of responsibility. Invocations of the obedience repertoire can thus usefully be understood as accounting strategies for impression management in the face of the interviewer's probes. In this second section of the analysis, we extend this line of argument further to focus in more detail on the interviewer's role in the production of the responsibility data. Specifically, we will suggest that the interviewing style adopted by Milgram and Williams was essentially an adversarial one that has some similarities with more recent notions of ‘active’ interviewing (Holstein & Gubrium, 1995). We do not wish to overstate this comparison by anachronistically projecting back a contemporary concept, but we would suggest that the way in which Milgram's interviews were conducted is much closer to what is currently understood as active interviewing than to any notion of the non-interventionist interviewer who attempts to maintain neutrality and minimise bias. Indeed, there are also similarities here with early discourse analytic advice on

interviewing, in which researchers were encouraged ‘to make the interview a much more interventionist and confrontative arena than is normal’ (Potter & Wetherell, 1987, p. 164).

In the first extract in this section, participant 0201 is being interviewed by Stanley Milgram, who creates the impression that he had not been aware of the precise details of what had occurred in the experimental session. He uses this ambiguous status to pose ‘naïve’ questions which function to hold the participant to account. The extract – which occurs shortly after the section of the interview presented above as extract 2 – begins with Milgram asking the participant why he didn’t stop administering the shocks at the point at which the learner demanded to be released:

Extract 5 (0201; 450v.)

1 SM: Why didn’t you stop at that

2 point?

3 (3)

4 T: Why didn’t I stop?

5 SM: Yeah.

6 T: Because the man told me to

7 continue.

8 SM: He told you to continue even

9 when the er, learner told you to

10 stop?

11 T: That’s correct.

12 SM: Why did you listen to that man

13 rather than the man that was

14 getting hurt.

- 15 T: Apparently he knows what he's
16 doing.
- 17 SM: Well the man in there knows
18 that he's getting hurt.
19 (7)
- 20 T: Well I was er, right in the
21 middle I didn't know which
22 way to turn I-
23 (4)
- 24 T: I felt sorry for the man outside
25 and
26 (2)
- 27 T: I thought the experiment er,
28 depended on my going ahead.
- 29 SM: Uhuh. Well what about, him,
30 didn't he have some er, I mean
31 rights as a person?
32 (3)
- 33 T: Certainly.
34 (2)
- 35 SM: Well, was one of his rights the
36 er, the right to stop being, given
37 an electric shock when he didn't
38 want anymore?
39 (5)

- 40 T: Oh yes, yeah. Yep.
- 41 SM: Well why did you give it to him
- 42 then?
- 43 (3)
- 44 T: eh heh nobody told me to stop.
- 45 SM: Uhuh. But he told you to stop.
- 46 (4)
- 47 T: That's true but he's just- but he
- 48 was just the er
- 49 (6)
- 50 T: subject sh- shall we say.

The participant's initial account for failing to stop draws on the obedience repertoire (lines 6-7: 'Because the man told me to continue.'). Milgram then works up his seeming naivety concerning the events of the experiment by asking for clarification. By creating the impression that the events of the experimental session were somewhat unexpected, Milgram places the onus on the participant to account for his behaviour in failing to draw proceedings to a halt. There is the implication that the experimenter might himself be held to account for this failure too, but Milgram's pressing of the participant is notable in that it challenges his account for continuing in several ways. First, Milgram contrasts the experimenter with 'the man that was getting hurt' (lines 13-14). This makes relevant a normative response that if someone is being hurt then others are obliged to do what they can to assist, and in so doing puts the participant in a position where an account is required for breaching this norm. Second, after the participant responds to his previous question by mobilising the experimenter's apparent expertise (lines 15-16: 'Apparently he knows what he's doing'),

Milgram contrasts the ‘apparent’ knowledge of the experimenter with the learner’s knowledge of his experiences of ‘getting hurt’ (line 18). He then focuses on the learner’s rights in order to suggest that the participant has disregarded the learner’s basic ‘rights as a person’ (line 31). Asking loaded questions, Milgram first gets the participant to agree that the learner did indeed have some basic rights (lines 29-30), and subsequently that one of these rights entailed not being given an electric shock against his will (lines 25-28). Both these questions are loaded in that they are formulated in such a way as to make disagreement difficult, and indeed the participants’ agreements come after noticeable delays in both cases. Having elicited these agreements, Milgram then asks the participant why he continued administering the shock (lines 41-42) – in effect why he behaved in such a way as to deny the rights which he had readily agreed were due to the learner. This again shows how the interviewer could construct incongruity between belief and action in order to hold participants to account. The participant’s response (line 44: ‘nobody told me to stop’) positions him as simply doing what he was told such that he would have stopped only if instructed to do so. When Milgram points out that *the learner* told him to stop (line 45), the participant invokes the power relations inherent in the situation (lines 47-50: ‘he was just the ... subject’), with the term ‘subject’ in particular drawing on contemporaneous norms concerning the role and status of research participants. Again, we see obedience being invoked as a way of downplaying responsibility when being held accountable. Milgram’s questioning is not designed for the straightforward elicitation of information, but rather it functions to hold the participant to account, to challenge his motives, and even his morality, in continuing with the experiment past the point at which the learner had demanded to be released.

In other interviews, the impetus to hold participants to account led the interviewer to explicitly contest the participants’ attributions of responsibility. In the next extract we see the

experimenter challenging the participant's claim that the experimenter should be allocated all the responsibility. However the participant, seemingly so submissive to the experimenter's authority just a few moments earlier, refuses to change his position:

Extract 6 (0214; 450v.)

- 1 T: Er, well, basically I would have
2 to say that er, you were
3 responsible as the examiner,
4 I was merely following your
5 instructions and, of course, he
6 had [no choice I mean]
- 7 E: [And yet you-]
- 8 T: he did his best. At least at the
9 beginning anyway.
- 10 E: Uhuh.
- 11 T: So I would have to put the
12 blame on you.
- 13 E: The whole blame?
- 14 T: I would say so.
- 15 E: Well er as a matter of fact you
16 could have er, could have got
17 up, said no. See I mean, and er
18 and he volunteered so er,
19 everyone has to share, er, well
20 n- not everyone has to but, let's

- 21 indicate on this pie.
 ((24 lines omitted in which the experimenter explains the responsibility
 clock))
- 45 T: Well er, as you mentioned I
46 could have possibly gotten up
47 myself and would not have
48 (?assumed?) any responsibility
49 but, of course that never
50 occurred to me I went along
51 with the instructions so I'm,
52 (?doing, with them?) all the way
53 up.
- 54 E: All right and that gives me all
55 the responsibility?
- 56 T: I gu-eh-ess it does.

In the first part of the extract, the participant uses the obedience repertoire (line 4: 'I was merely following your instructions') as part of his rationale for attributing all of the responsibility for the learner being shocked against his will to the experimenter. However, the experimenter does not take this at face value as a satisfactory response (line 13: 'The whole blame?'). When the participant re-states his position the experimenter begins to offer an alternative account. Beginning with an appositional ('well'; see above) on line 15, he notes that the participant could have refused to continue. He frames this as something that should have been clear and obvious to the participant (line 15: 'as a matter of fact'), and subsequently highlights the learner's voluntary participation (lines 17-18) in order to further

demonstrate that other judgements may be just as plausible. He then suggests that ‘everyone has to share’ responsibility (line 19), before quickly correcting himself (line 20). This self-correction can be heard as an orientation to having been too directive, and in moving on to using the responsibility clock/pie he moves away from attempting overtly to challenge the participant. However, in continuing to attribute total responsibility to the experimenter, the participant now orients to the experimenter’s previously stated objections to his position (line 45: ‘as you mentioned’) and proceeds to reject them by formulating an account of why he did not discontinue the experiment himself. He does this by stating that the possibility ‘never occurred’ to him (lines 49-50), which insofar as it prioritises a lack of recognition of refusal as an option serves as an attempt to retain a moral identity as someone who didn’t continue out of malice but rather simply by virtue of not realising that ceasing was an option. This is followed with a further use of the obedience repertoire (lines 50-51: ‘I went along with the instructions’), to restate the extent to which he was simply doing as he was told, but in such a way as to downplay the experimenter’s agency. Whereas previously the participant had constructed himself as having followed *your instructions*, following the experimenter’s provision of possible alternatives, it is simply *the instructions* themselves that were followed. As in extract 1, such a formulation attends to the delicate business of having to blame one’s interlocutor in order to minimise one’s own responsibility. Moreover, in placing the blame on the ‘the instructions’, the participant allows for the possibility that in other contexts – ones in which no such instructions were present – he would have behaved differently. His actions thus spring not from any deficiency of his own moral identity, but rather from the external contingencies of the situation, again anticipating Milgram’s own rejection of personality-based explanations.

This extract neatly demonstrates the two conflicting imperatives in the post-experiment interviews, only one of which has been fully appreciated by theorists, including

Milgram himself. In using the responsibility clock, Milgram was aiming to collect objective data concerning participants' feelings of responsibility. However, what has not typically been understood is the extent to which the interviews were also occasions for holding participants to account. The interviews are thus best conceptualized as rhetorical encounters, with interviewers probing participants as to why they behaved as they did, and in doing so giving the impression that they *ought* to have acted differently; the participants accounted for their behaviour as best they could in the circumstances, but their responses need to be conceptualised in precisely these terms – they were produced in, and for, these specific circumstances.

Concluding remarks

The present analysis constitutes an extension of the rhetorical perspective on the obedience experiments, which has to date focussed on the interactions in the experimental session itself. In this previous research, it has been argued that the experimental interaction between experimenter and participant should be understood as a rhetorical encounter (Gibson, 2013a, 2014, 2017a); here, we suggest that the interactions in the post-experimental interviews should be seen in similar terms. When discussing the issue of responsibility, the interviewer subtly – and sometimes not so subtly – probed in such a way as to make it clear to participants that they were accountable for their actions. It is in this context that participants invoked obedience, and suggested that they were not responsible for their actions, and thus rather than conceptualising their responses as reflecting underlying psychological shifts, we should instead see them as situated social actions. In this final concluding section we consider how this shift in perspective has the potential to re-orient our understanding of the experiments, and of obedience more broadly.

First, the shift to treating obedience in terms of a rhetorical device available for participants to deploy, rather than as an underlying psychological mechanism, highlights the value of studying social influence processes not only in straightforward causal terms, but as participants' concerns that can be invoked, oriented to, and accounted for (Gibson & Smart, 2017; Smart, 2013). Allied to this is the shift from trying to work out what underlying psychological process was occurring in participants' heads during the obedience experiments, to treating psychological matters themselves as things to be worked up and accounted for in the course of the experiments and post-experiment interviews. This can be seen, for example, in the repeated mobilization of a disjuncture between underlying thought/feeling and outward behaviour/action in the interviews. Interviewers could construct such an inconsistency as part of the process of holding participants to account; participants could in turn mobilize the inconsistency in order to assert that they may have done something that was problematic, but at least they hadn't *wanted* to. This enabled participants to present themselves as not being the type of person who would inflict pain on another human being, and in reifying these accounts Milgram was essentially choosing to believe statements that buttressed his general preference for situational explanations over ones based around personality. Such selective reification (Potter, 1996) neglects the extent to which these accounts were bound up with self-presentational concerns in the immediate social context, and a more agnostic position on the referentiality of accounts overcomes this. Such a perspective follows directly from the discursive psychological injunction to re-specify objects typically conceptualized as internal mental constructs as action-oriented social constructions (Edwards & Potter, 1992). None of this is to deny that there may be *something* going on 'under the skull', but it points to the difficulty of inferring this from discourse. Indeed, there is a fundamental circularity at work in the traditional use of these accounts as reports: people are understood as not responsible for their actions because they say they are not responsible; yet the only evidence that people

are not responsible is these verbal accounts themselves. Assurances that participant accounts offer a ‘spontaneous ... and relatively uncensored glimpse into what participants were thinking’ (Burger et al., 2011, p. 461), or that they were not ‘alibi[s] concocted for the moment’ (Milgram, 1974, p. 146) thus miss the point and risk reproducing a naïvely realist perspective on the relationship between thought and language. Given the moral issues involved in taking a perspective that effectively functions to absolve perpetrators of responsibility (Haslam & Reicher, 2006), at the very least a higher threshold of proof is required to demonstrate the psychological shifts purported in traditional accounts of Milgram’s experiments.

Second, this analysis adds to recent debates concerning the nature of the phenomena captured in Milgram’s experiments, and indeed the nature of obedience itself. It is clear from recent research that, whatever the experiments *do* show, they are not demonstrations of people following orders in the usual sense of that term (Burger et al., 2011; Gibson, 2013a; Haslam, Reicher & Birney, 2014). The experimental procedure required much more delicate rhetorical work than merely the issuing of blunt orders in order to ensure participants continued. This implies that the extent to which people administered the shocks depends on much more subtle features of the experimental context than the standard interpretation would imply. The issue of whether this means that the experiments are not demonstrations of *obedience*, however, depends on the conflation of obedience with following orders/commands. The standard social psychological definition of obedience certainly requires orders/commands for something to be understood in terms of obedience, but the key question is arguably whether this was ever a particularly good definition of obedience (Gibson, 2017b). In everyday language we might, for example, speak of obeying the law, but we do not need direct orders to do so. Rather, the act of obedience depends on shared social conventions and norms. At a more theoretical level, we need only consider Foucault’s (1979)

classic work on the way in which disciplinary power is exercised in contemporary societies to see that the most successful form of obedience is that which can be elicited *without* the need for the administration of anything so blunt as a direct command. Far better that people come to assume that their conduct is always potentially available for sanction and regulate their own behaviour accordingly. There is thus a need to refine and re-think the discipline's conceptualization not only of the Milgram experiments, but of obedience more broadly. Moreover, there is a separate analytic project to be undertaken exploring what counts explicitly as obedience in particular contexts, and how this is constructed by social actors across a range of situations. What can be said, however, is that the classic social psychological version of obedience as superior orders *is* present in Milgram's experiments after all, but in a rather different way than has typically been assumed. Rather than a process of social influence which facilitates an agentic shift in participants, it is instead present as a participants' resource for use in accounting for their behaviour. Indeed, there is a fundamental irony in that Milgram's work is held as the exemplar *par excellence* of the extent to which people underestimate the importance of social context as a determinant of behaviour, and yet Milgram and generations of subsequent theorists have neglected the extent to which the situation in which participants found themselves immediately *after* the experiment may have shaped their accounts.

References

- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper & Row.
- Antaki, C., & Wetherell, M. (1999). Show concessions. *Discourse Studies*, 1, 7-27. doi: 10.1177/1461445699001001002
- Arendt, H. (1977). *Eichmann in Jerusalem: A report on the banality of evil*. London: Penguin. (original work published 1963).
- Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's "Behavioural study of obedience". *American Psychologist*, 19, 421-423. doi: 10.1037/h0040128
- Billig, M. (1991). *Ideology and opinions: Studies in rhetorical psychology*. London: Sage.
- Billig, M. (1996). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.
- Blass, T. (2004). *The man who shocked the world: The life and legacy of Stanley Milgram*. New York: Basic Books.
- Brannigan, A. (2013). *Beyond the banality of evil: Criminology and genocide*. Oxford: Oxford University Press.
- Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *American Psychologist*, 64, 1-11. doi: 10.1037/a0010932
- Burger, J. M., Girgis, Z. M., & Manning, C. C. (2011). In their own words: Explaining obedience to authority through an examination of participants' comments. *Social Psychological & Personality Science*, 2, 460-466. doi: 10.1177/1948550610397632
- Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion changes the sense of agency in the human brain. *Current Biology*. Advance online publication.

- Cesarani, D. (2004). *Eichmann: His life and crimes*. London: William Heinemann.
- Edwards, D., & Fasulo, A. (2006). "To be honest": Sequential uses of honesty phrases in talk-in-interaction. *Research on Language and Social Interaction*, 39, 343-376. doi: 10.1207/s15327973rlsi3904_1
- Edwards, D., & Potter, J. (1992). *Discursive psychology*. London: Sage.
- Elms, A. C., & Milgram, S. (1966). Personality characteristics associated with obedience and defiance toward authoritative command. *Journal of Experimental Research in Personality*, 1, 282-289.
- Fenigstein, A. (2015). Milgram's shock experiments and the Nazi perpetrators: A contrarian perspective on the role of obedience pressures during the Holocaust. *Theory & Psychology*, 25, 581-598. doi: 10.1177/0959354315601904
- Fennis, B. M., & Aarts, H. (2012). Revisiting the agentic shift: Weakening personal control increases susceptibility to social influence. *European Journal of Social Psychology*, 42, 824-831. doi: 10.1002/ejsp.1887
- Foucault, M. (1979). *Discipline and punish: The birth of the prison*. London: Penguin.
- Gaeta, P. (1999). The defence of superior orders: The statute of the International Criminal Court versus customary international law. *European Journal of International Law*, 10, 172-191. doi: 10.1093/ejil/10.1.172
- Gibson, S. (2013a). Milgram's obedience experiments: A rhetorical analysis. *British Journal of Social Psychology*, 52, 290-309. doi: 10.1111/j.2044-8309.2011.02070.x
- Gibson, S. (2013b). 'The last possible resort': A forgotten prod and the *in situ* standardization of Stanley Milgram's voice-feedback condition. *History of Psychology*, 16, 177-194. doi: 10.1037/a0032430

- Gibson, S. (2014). Discourse, defiance, and rationality: “Knowledge work” in the “obedience” experiments. *Journal of Social Issues*, 70, 424-438. doi: 10.1111/josi.12069
- Gibson, S. (2017a). Developing psychology’s archival sensibilities: Revisiting Milgram’s ‘obedience’ experiments. *Qualitative Psychology*, 4, 73-89. doi: 10.1037/qup0000040
- Gibson, S. (2017b). Obedience without orders: Expanding social psychology’s definition of obedience. Manuscript submitted for publication.
- Gibson, S., & Smart, C. (2017). Social influence. In B. Gough (Ed.), *The Palgrave handbook of critical social psychology* (pp. 291-318). London: Palgrave.
- Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora’s box: A sociological analysis of scientists’ discourse*. Cambridge: Cambridge University Press.
- Griggs, R. A., & Whitehead III, G. I. (2015a). Coverage of Milgram’s obedience experiments in social psychology textbooks: Where have all the criticisms gone? *Teaching of Psychology*, 42, 315-322. doi: 10.1177/0098628315603065
- Griggs, R. A., & Whitehead III, G. I. (2015b). Coverage of recent criticisms of Milgram’s obedience experiments in introductory social psychology textbooks. *Theory & Psychology*, 25, 564-580. doi: 10.1177/0959354315601231
- Haslam, S. A., & Reicher, S. (2006). Debating the psychology of tyranny: Fundamental issues of theory, perspective and science. *British Journal of Social Psychology*, 45, 55-63. doi: 10.1348/014466605X80686
- Haslam, S. A., & Reicher, S. (2007). Beyond the banality of evil: Three dynamics of an interactionist social psychology of tyranny. *Personality & Social Psychology Bulletin*, 33, 615-622. doi: 10.1177/0146167206298570

- Haslam, S. A., & Reicher, S. D. (2008). Questioning the banality of evil. *The Psychologist*, 21, 16-19.
- Haslam, S. A., Reicher, S. D., & Birney, M. E. (2014). Nothing by mere authority: Evidence that in an experimental analogue of the Milgram paradigm participants are motivated not by orders but by appeals to science. *Journal of Social Issues*, 70, 473-488. doi:10.1111/josi.12072
- Hollander, M. M. (2015). The repertoire of resistance: Non-compliance with directives in Milgram's 'obedience' experiments. *British Journal of Social Psychology*, 54, 425-444. doi: 10.1111/bjso.12099
- Hollander, M. M., & Maynard, D. W. (2016). Do unto others...? Methodological advance and self- versus other-attentive resistance in Milgram's "obedience" experiments. *Social Psychology Quarterly*, 79, 355-375. doi: 10.1177/0190272516648967
- Holstein, J. A., & Gubrium, J. F. (1995). *The active interview*. London: Sage.
- Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Cambridge: Polity.
- Lang, J. (2014). Against obedience: Hannah Arendt's overlooked challenge to social-psychological explanations of mass atrocity. *Theory & Psychology*, 24, 649-667. doi: 10.1177/0959354314542368
- Mantell, D. M., & Panzarella, R. (1976). Obedience and responsibility. *British Journal of Social & Clinical Psychology*, 15, 239-245. doi: 10.1111/j.2044-8260.1976.tb00030.x
- Mastroianni, G. R. (2015). Obedience in perspective: Psychology and the Holocaust. *Theory & Psychology*, 25, 657-669. doi: 10.1177/0959354315608963
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal & Social Psychology*, 67, 371-378. doi: 10.1037/h0040525

- Milgram, S. (1965a). Some conditions of obedience and disobedience to authority. *Human Relations*, 18, 57-76. doi: 10.1177/001872676501800105
- Milgram, S. (Producer). (1965b). *Obedience* [DVD]. Available from PennState Media Sales, <http://www.mediasales.psu.edu>
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Miller, A. G. (1986). *The obedience experiments: A case study of controversy in social science*. New York: Praeger.
- Moscovici, S. (1984). The phenomenon of social representations. In R. M. Farr & S. Moscovici (Eds.), *Social representations*. Cambridge: Cambridge University Press.
- Nicholson, I. (2011). "Torture at Yale": Experimental subjects, laboratory torment and the "rehabilitation" of Milgram's "obedience to authority". *Theory & Psychology*, 21, 737-761. doi: 10.1177/0959354311420199
- Orne, M. T., & Holland, C. C. (1968). On the ecological validity of laboratory deceptions. *International Journal of Psychiatry*, 6, 282-293.
- Overy, R. (2014). "Ordinary men," extraordinary circumstances: Historians, social psychology, and the Holocaust. *Journal of Social Issues*, 70, 515-530. doi: 10.1111/josi.12075
- Parker, I. (1992). *Discourse dynamics: Critical analysis for social and individual psychology*. London: Routledge.
- Perry, G. (2012). *Behind the shock machine: The untold story of the notorious Milgram psychology experiments*. London: Scribe.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of

- preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.
- Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies*, 9, 219-229. doi: 10.1007/BF00148128
- Potter, J. (1996). *Representing reality: Discourse, rhetoric and social construction*. London: Sage.
- Potter, J., & Edwards, D. (2001). Discursive social psychology. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 103-118). Oxford: Wiley.
- Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: beyond attitudes and behaviour*. London: Sage.
- Reicher, S. (2014). In praise of activism: Rethinking the psychology of obedience and conformity. In C. Antaki & S. Condor (Eds.), *Rhetoric, ideology and social psychology: Essays in honour of Michael Billig* (pp. 94-108). Hove: Routledge.
- Reicher, S., & Haslam, S. A. (2011). After shock? Towards a social identity explanation of the Milgram 'obedience' studies. *British Journal of Social Psychology*, 50, 163-169. doi: 10.1111/j.2044-8309.2010.02015.x
- Reicher, S. D., Haslam, S. A., & Miller, A. G. (2014). What makes a person a perpetrator? The intellectual, moral, and methodological arguments for revisiting Milgram's research on the influence of authority. *Journal of Social Issues*, 70, 393-408. doi: 10.1111/josi.12067
- Russell, N. J. C. (2011). Milgram's obedience to authority experiments: Origins and early evolution. *British Journal of Social Psychology*, 50, 140-162. doi: 10.1348/014466610X492205

- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 696-735. doi: 10.2307/412243
- Schiffrin, D. (1985). Conversational coherence: The role of *well*. *Language*, 61, 640-667. doi: 10.2307/414389
- Smart, C. A. (2013). *Understanding social influence differently: A discursive study of livery yards* (Unpublished doctoral dissertation). The Open University, UK.

Appendix: Transcription conventions.

- Speakers are identified as E (Experimenter); T (Teacher; the naïve participant); L (Learner); SM (Stanley Milgram).
- Commas are used to indicate brief pauses of less than a second.
- Longer silences are indicated with a number in brackets. The number indicates the length of the silence, in seconds, rounded to the nearest whole second.
- Square brackets are used to indicate overlapping talk.
- A dash is used to indicate a sharp cut-off of the preceding sound.
- Question marks are used to indicate a questioning intonation, rather than a grammatical question as such.
- Full stops are used to indicate a stopping intonation, rather than the end of a sentence as such.
- Laughter particles are indicated by 'Eh heh'.
- Double parentheses are used to indicate comments from the transcriber.
- Question marks enclosed in single parentheses are used to indicate words that were not clearly audible, but which represent the transcriber's best guess at what was said.