Commentary

After shock? Towards a social identity explanation of the Milgram 'obedience' studies

Stephen Reicher[1] and S. Alexander Haslam[2]*
[1]University of St Andrews, St Andrews, Fife, UK
[2]University of Exeter, Exeter, Devon, UK

Russell's forensic archival investigations reveal the great lengths that Milgram went to in order to construct an experiment that would 'shock the world'. However, in achieving this goal it is also apparent that the drama of the 'basic' obedience paradigm draws attention away both from variation in obedience and from the task of explaining that variation. Building on points that Russell and others have made concerning the competing 'pulls' that are at play in the Milgram paradigm, this paper outlines the potential for a social identity perspective on obedience to provide such an explanation.

The theatre of 'obedience'

The Milgram studies are great drama as well as great science. Therein lies much of the secret of their success. Few can be left unmoved by reading transcripts of the unfolding sessions – or still better, by watching these on grainy black-and-white film. Few can escape being gripped as the participants agonize over their predicament before finally, seemingly inexorably, pressing the buttons on Milgram's bogus shock generator.

Certainly, Milgram's films and the narrative style of his classic 1974 text have contributed to his work becoming one of the best-known pieces of psychological research ever conducted – not only in the academic world but also in the larger culture (Blass, 2004; Novick, 2000). In the words of Ross and Nisbett, Milgram's work is 'part of our society's shared intellectual legacy' (cited in Miller, 2009, p. 20). But this positive impact comes with an accompanying danger: that we may become so absorbed by the phenomena we are witnessing that we fail to see the background against which they are set. All we take in is the evidence that people obey, not the particular circumstances in which this obedience occurs. Hence, we slip too easily from the evidence that ordinary people can obey even the most extreme of instructions to the conclusion that people cannot help but obey destructive authorities.

Moreover, if we do notice anything of the background, it is not the structure of the immediate social situation, but rather the broader culture into which the participants

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

have been socialized. We therefore ask whether the results would differ with different participants or in alternative cultural contexts (e.g., Blass, 1999; Burger, 2009b; Mantell, 1971; Twenge, 2009). Would women obey as much as men? Would other nationalities obey as much as Americans? Do we obey now as much as we did then? Indeed, it is noteworthy that Burger's recent replication of Milgram's research – which was the focus of an extended discussion in *American Psychologist* – is subtitled 'Would people still obey today?' Such a question, of course, implies that people did obey in the past. It distracts our attention from interrogating the conditions under which people do or do not obey yesterday, today, or indeed on any day. In other words, the focus on behaviour gets in the way of the focus on the processes that govern obedience. Or rather, to be more accurate, an implicit 'culturalist' theory of obedience gets in the way of an explicit theory of how the precise constitution of social context impacts upon obedience. In either case, what Ross observed over 20 years ago remains equally true today: we do not really know why people did what they did in Milgram's experiments (Ross, 1988; see also Miller, 2004, 2009; Russell, 2011).

The important achievement of Russell's (2011) archival research is precisely to reveal the great care that Milgram took in order to produce an experimental demonstration that would 'shock the world' [to borrow from the title of Blass's engaging narrative history of Milgram's life]. Two details of Milgram's inspired craft stand out and, together, are highly revealing. First, Russell shows how much effort went into the precise appearance of the shock generator. Second, he shows how the labelling of the shock levels was equally critical. Thus, Milgram replaced the original designation of the highest levels of shock as 'LETHAL' with the ominous, but more ambiguous 'XXX' in all his published studies. His concern was that to be too explicit about the harm being [seemingly] inflicted might mean that *not enough* people [if any] would be willing to administer the highest level of shock.

On the other hand, Russell reveals that Milgram dropped a version of the study that led *too many* people [in fact, everybody] to go to this extreme. In this version, the participant received no feedback at all from the 'learner' – not even a grunt or a cry – and followed the experimenter's instructions to the end without any apparent signs of strain. Total obedience, obedience without tension, became banal. It lacked the drama that Milgram recognized as essential for engaging his audience.

Overall, then, it is clear that the studies had to be very carefully calibrated in order to achieve the forms and levels of obedience that Milgram sought. It is equally clear that small changes in the nature of the set-up could have very dramatic consequences for the amount of obedience that was found. In sum, by deconstructing the theatricality of Milgram's paradigm, Russell lets us see beyond appearances in order that we may better focus on the key missing questions about why – through what *process* – contextual variations affected the hold of the experimenter upon the participant.

The essentialization of 'obedience'

Before we try to tackle issues of process, we have to deal with another question that is posed by Russell's analysis. This concerns the extent to which Milgram is himself responsible for the essentialization of his studies. Here, one could take from Russell's article a sense that Milgram set out as a film-maker as much as a scientist with the primary goal of dazzling [and perhaps discomfiting] his audience. Certainly, we see that he sought to 'make his mark' by producing 'a phenomenon of great consequence' (2010, p. 7; see also Millard, 2010). This may be true, but it is important to note that,

having achieved this, Milgram then wanted to 'worry [the phenomenon] to death' (Russell, 2011, p. 146; after Blass, 2004). In this regard, one of the most impressive aspects of Milgram's work is the sheer number of variants to the obedience paradigm that he investigated [20 are described in the 1974 book]. These produced levels of obedience that ranged from 0% to 92.5% as a function of context [0% to 100% if we add the 'zero feedback' condition]. Moreover, in the preface to the second edition of *Obedience to Authority* (Milgram, 1983) Milgram bemoans the fact that people distorted his research by focussing on the occurrence of obedience rather than considering *variation* in its occurrence.

In this, one can perhaps conclude that Milgram – like Hannah Arendt with whom his ideas have become inextricably linked – was a victim of his own success. For just as Arendt's (1963) phrase 'the banality of evil' [which, aside from the subtitle, appears only once at the very end of her book *Eichmann in Jerusalem*] has been misinterpreted, simplified, distorted, and used to obscure the richness and subtlety of her thinking (see Haslam & Reicher, 2007; Newman, 2001; Reicher, Haslam & Rath, 2008), so Milgram's image of the banal participant delivering electric shocks has served to occlude the complexity and variability of his findings.

While Milgram's own intentions and responsibility are difficult territory, what seems undeniable is that the broad impact of his work has been reduced to the claim that 'people obey orders' or, more strongly, that 'people can't help but obey orders'. Indeed, as testimony to this, while writing this piece, we asked 100 high-school students who were studying psychology to write a few sentences summarizing the key message of Milgram's studies. Ninety-four percent of respondents were able to provide a reasonable description and, of these, almost all [93%] were of the form 'people obey the instructions of people in authority', with about half [56%] adding something to suggest that such obedience is unconditional [e.g., '. . . even when this involves harming others'].

The need for a compelling theory of obedience [and disobedience]

All this makes it all the more important to return to our earlier question concerning the processes underlying contextual variations in obedience. One thing is for sure. The process account that Milgram introduced in his seminal 1974 book does not suffice. Even those who are most sympathetic to Milgram's work in general find his notion of an agentic state singularly unilluminating (e.g., Blass, 2004). The notion of the agentic state suggests that, in the face of authority, people lose sight of their own goals and values and cede responsibility to those in authority – so that they become fixated on how well they carry out their allotted instructions rather than on what those instructions are. Quite apart from the fact that there is little or no evidence that people do in fact enter such a state, there is certainly nothing to suggest that variation in this state can explain variation in levels of obedience across different experiments (Blass, 2004; Mantell & Panzarella, 1976).

But apart from its empirical deficiencies, there is something strangely one-sided and static about the notion of an agentic state. For it reduces the obedience paradigm to a relationship between the participant and the experimenter. And it reduces this relationship to a state in which the participant is completely subjugated to the experimenter and becomes a mere cipher (see Reicher & Haslam, 2006, for a similar critique of role-based explanations of tyranny). Yet, as Milgram himself makes explicit in his early papers (e.g. Milgram, 1963, 1965a,b), the whole dilemma and the whole drama of the paradigm lies in the fact that participants are torn between different relationships, different obligations, different moralities. Mostly, these pit obligations to

the experimenter against obligations to the learner. Sometimes they bring in additional 'teachers' as well. Who one decides to heed therefore depends on the relative 'pull' of these different constituencies.

This sense of being torn between competing demands is evident from the film and transcripts of the studies. Participants do not – as the notion of an agentic state would suggest – simply concentrate on being exemplary functionaries. Rather, they argue and remonstrate with both the experimenter and the learner. They reflect the words of the learner as a challenge back to the experimenter; they reflect the words of the experimenter as a rationalization back to the learner. Far from being passive, they actively try to make sense of their predicament and to find a way through it. Indeed, in this regard, even when they accede to the requirements of their role as teachers, there is often a richness to their responses that the anodyne term 'obedience' fails fully to convey (Haslam & Reicher, 2007).

Towards a social identity explanation of obedience [and disobedience]

There are a number of findings that dictate the need for a dynamic and dilemmatic explanation of obedience levels. First, it is quite clear that the levels of obedience in different variants of Milgram's paradigm depend upon participants' exposure to the voices of different constituents. If the participant only hears the experimenter and is insulated from any other voice [as in the abandoned pilot], there is 100% obedience. As soon as the slightest voice of protest is heard from the learner – even if it is only pounding the wall at the 300 volt mark [1974, Experiment 1] – obedience falls precipitately to 65%, and as the victim gets closer and can be heard, seen, and then felt, so obedience falls to 62.5%, 40%, and then 30% [1974, Experiments 2, 3, 4]. As a corollary, when experimenter gives instructions and then leaves, so that his voice becomes less prominent, obedience falls to 21% [1974, Experiment 7]. When two fellow teachers argue against continuing the study [importantly, not when one dissents but when both do so and hence create a consensus of opposition] obedience falls to 10% [1974, Experiment 17]. And, even more starkly, when two experimenters issue contradictory demands [thereby undermining the capacity for either to represent a scientific consensus] obedience falls to 0% [1974, Experiment 15].

When one looks more closely at the studies, it also becomes apparent that the point at which disobedience occurs is intimately tied to the point at which dissenting voices are heard. Packer (2008) has shown that the critical decision points are at 150 volts and, to a lesser extent, at 315 volts. Not only are these the points at which most disobedience took place [36.9% and 10.6% of total noncompliant responses, respectively], but also overall obedience rates across studies covaried with rates of disobedience at 150 volt but no other point. The relevance of this is that, apart from Experiment 1, 150 volts is the first point where the learner explicitly demands to be released from the study (exclaiming 'Ugh!!! Experimenter! That's all. Get me out of here. I told you I had heart trouble. My heart's starting to bother me now. Get me out of here, please. My heart's starting to bother me. I refuse to go on. Let me out'; Milgram, 1974, p. 56). The 315 volt point is where the learner formally withdraws his consent to participate in the study [responding with '(*an intensely agonized scream*) *I told you I refuse to answer*. I'm no longer part of this experiment'; 1974, p. 57, original emphasis].

Of course, it is one thing to hear a voice and another to heed it. And while participants cannot heed something they cannot hear, we still need to ask how, when exposed to different voices, they decide to listen to the experimenter as opposed to the learner

[or else to fellow teachers]. This is something Russell (2011) alerts us to in detailing Milgram's sensitivity to both (1) the authenticity and quality of his accomplices [e.g., the importance of having a learner who 'is mild and submissive; not at all academic'; 2011, p. 20], and (2) the 'binding factors' [such as the experimenter's status] that help put the participant under the experimenter's influence.

What is more, amongst Milgram's rich and thoughtful reflections in his early papers, he refers in passing to the importance of *group formation* as a key psychological process in binding participants to the experimenter. In particular, he suggests that: 'the changing set of spatial conditions leads to a potentially shifting set of alliances over the several experimental conditions' (1965a, p. 64). In other words, when the experimenter and participant are in a room together and separate from the learner, the participant is far more likely to categorize himself together with the experimenter than when all three share the same room. This, we suggest, is a critical insight. It clearly needs to be developed in order to explain when groups are formed and what groups are formed. But, with such development, it can be used more generally to explain the generation of obedience [or disobedience] in the Milgram paradigm as a reflection of varying patterns of *social identification* (or shared social self-categorization; Turner, 1991).

More specifically, there are several factors in addition to spatiality which, we believe, affect the extent to which participants see themselves as sharing a common social identity with the experimenter, and hence impact on their levels of obedience. In broad terms, these can be divided into *who the experimenter is* and *what the experimenter does*. To take these in turn, we start by assuming that, during the studies, participants define themselves – at least on one level – as parties to a scientific enterprise. It follows that they are more likely to identify with the experimenter and accept both his legitimacy and his expertise, to the extent that he represents the category 'science'. This condition is fulfilled when he is a devoted and prestigious scientist, associated with Yale University. As such, he can be seen to be *prototypical* of 'science' (Turner, Hogg, Oakes, Reicher & Wetherell, 1987) and, on this basis, participants listen and respond positively to what he has to say. When the experimenter is associated with a less prestigious institution [because the study is relocated from Yale to a commercial building in Bridgeport, 1974, Experiment 10], he is less prototypical and so obedience falls to 48%. And when the experimenter is just an 'ordinary man' who represents no-one but himself [1974, Experiment 13], obedience falls still further to 20%.

Now consider what the experimenter does. Milgram sought, as far as possible, to systematize the interaction between experimenter and participant by providing four prompts [or 'prods' as he terms them]. They are, in escalating order: (1) 'Please continue', or, 'please go on'; (2) 'The experiment requires that you continue'; (3) 'It is absolutely essential that you continue'; and (4) 'You have no other choice, you must go on'. The first three of these prods are either requests or else root the instruction in a scientific rationale [the needs of the experiment – which is the basis for shared social identity]. What distinguishes the final prompt is that it is the only one that takes the form of an unqualified command or order (Burger, 2009a). Here, the participant is simply told to continue whether he likes it or not. So, 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 [and of their inherent propensity to do so].

There is only one published transcript of reaction to the final prompt in Milgram's studies. In this, the participant – a Professor of Old Testament liturgy – responds 'if this were Russia, maybe, but not in America. *(The experiment is terminated)*'

(Milgram, 1974, p. 48). Fortunately, though, more systematic evidence is provided by Burger's (2009a,b) careful replication of the Milgram paradigm. Tellingly, he discovered that *on every single occasion* that the experimenter issued the fourth prompt, participants refused to continue. One way of interpreting this finding is to suggest that when the experimenter imposes himself over the participant, this serves to emphasize their lack of shared identity, it dismantles the group relationship between them, and hence produces disobedience (Haslam, Reicher & Platow, 2011).

There is a powerful irony at play here. For, as we have noted, Milgram's studies are widely remembered as showing that people obey orders. However, upon closer inspection, it appears that one thing that they show unequivocally is that, when requests are framed as orders, people do *not* obey.

Our explanation of this apparent anomaly is derived from a social identity/self-categorization account that stresses that obedience is predicated upon perceptions of shared identity. This, therefore, directs us to enquire into the various ways that the constitution of context impacts upon social identities and into the way in which identities impact [and are impacted] by moral obligations to the different parties in the study. Of course, this is a *post-hoc* interpretation based on indirect evidence. Accordingly, if we want to sustain such an argument, it is obviously necessary to conduct further studies specifically designed to test this hypotheses.

Exactly 50 years after the original studies, we are finally in a position to do so. Thanks to the innovative work of researchers such as Burger (2009b, 2010) and Slater *et al.* (2006), we now have methodological tools that allow us to investigate the Milgram paradigm without falling foul of the attendant ethical problems. Thanks to the work of Russell (2011), we have been re-directed to address the contextual conditions and causes of obedience. It is therefore time to move forward and produce an explanation of Milgram's findings that is every bit as vivid, rich, and compelling as the findings themselves.

References

Arendt, H. (1963). *Eichmann in Jerusalem: A report on the banality of evil*. New York, NY: Penguin.

Blass, T. (1999). The Milgram paradigm after 35 years: Some things we now know about obedience to authority. *Journal of Applied Social Psychology*, *29*, 955–978. doi: 10.1111/j.1559-1816.1999.tb00134.x

Blass, T. (2004). *The man who shocked the world: The life and legacy of Stanley Milgram*. New York, NY: Basic Books.

Burger, J. (2009a). *In their own words: Explaining obedience through an examination of participants' comments*. Paper presented at the Meeting of the Society of Experimental Social Psychology, Portland, ME, October 15–17.

Burger, J. (2009b). Replicating Milgram: Would people still obey today? *American Psychologist*, *64*, 1–11. doi: 10.1037/a0010932

Haslam, S. A., & Reicher, S. D. (2007). Beyond the banality of evil: Three dynamics of an interactionist social psychology of tyranny. *Personality and Social Psychology Bulletin*, *33*, 615–622. doi: 10.1177/0146167206298570

Haslam, S. A., Reicher, S. D., & Platow, M. J. (2011). *The new psychology of leadership: Identity, influence and power*. New York, NY and Hove: Psychology Press.

Mantell, D. M. (1971). The potential for violence in Germany. *Journal of Social Issues*, *27*, 101–112. doi: 10.1111/j.1540-4560.1971.tb00680.x

Mantell, D. M., & Panzarella, R. (1976). Obedience and responsibility. *British Journal of Social and Clinical Psychology*, *15*, 239–245.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*, 371–378. doi: 10.1037/h0040525

Milgram, S. (1965a). Liberating effects of group pressure. *Journal of Personality and Social Psychology*, *1*, 127–134. doi: 10.1037/h0021650

Milgram, S. (1965b). Some conditions of obedience and disobedience to authority. *Human Relations*, *18*, 57–76. doi: 10.1177/001872676501800105

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York. NY: Harper & Row.

Milgram, S. (1983). *Obedience to authority: An experimental view* (2nd ed.). New York. NY: Harper & Row.

Millard, K. (2010). *Stanley Milgram's story machine: Experimental psychology, improvisation and the reinvention of the screenplay*. Paper presented at "Screenwriting: History, Theory, Praxis" conference. Copenhagen, 9–11 September.

Miller, A. (2004). What can the Milgram obedience experiments tell us about the Holocaust? Generalizing from the social psychology laboratory. In A. Miller (Ed.), *The social psychology of good and evil* (pp. 193–239). New York, NY: Guilford.

Miller, A. G. (2009). Reflections on 'replicating Milgram'. *American Psychologist*, *64*, 20–27. doi: 10.1037/a0014407

Newman, L. S. (2001). *The banality of secondary sources: Why social psychologists have misinterpreted Arendt's thesis*. Unpublished manuscript: Syracuse University, Syracuse, NY.

Novick, P. (2000). *The holocaust in American life*. Boston, MA and New York, NY: Houghton Mifflin.

Packer, D. J. (2008). Identifying systematic disobedience in Milgram's obedience experiments: A meta-analytic review. *Perspectives on Psychological Science*, *3*, 301–304. doi: 10.1111/j.1745-6924.2008.00080.x

Reicher, S. D., & Haslam, S. A. (2006). Rethinking the psychology of tyranny: The BBC prison experiment. *British Journal of Social Psychology*, *45*, 1–40. doi: 10.1348/014466605X48998

Reicher, S. D., Haslam, S. A., & Rath, R. (2008). Making a virtue of evil: A five-step social identity model of the development of collective hate. *Social and Personality Psychology Compass*, *2*, 1313–1344. doi: 10.1111/j.1751-9004.2008.00113.x

Ross, L. D. (1988). Situationist perspectives on the obedience experiments. *Contemporary Psychology*, *33*, 101–104. doi: 10.1037/025376

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

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., *et al*. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE*, *1*, e39. doi:10.1371/journal.pone.0000039

Turner, J. C. (1991). *Social influence*. Milton Keynes: Open University Press.

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.

Twenge, J. M. (2009). Change over time in obedience: The jury's still out, but it might be decreasing. *American Psychologist*, *64*, 28–31. doi: 10.1037/a0014475