

The Role of the Moderator: Problems and Possibilities for Government-Run Online Discussion Forums

SCOTT WRIGHT

1 Introduction

Governments at all levels, and across many continents, have adopted online discussion forums as a means to promote democratic participation. These vary greatly in structure and may encourage a two-way link between government and people and/or help create a virtual public sphere (Wright 2002). Asynchronous forums might facilitate the kind of large-scale discussion often considered unrealistic. Thus, they have the potential to facilitate broader-based interactive policy making (Coleman and Götze 2001; Wright and Street 2007). New technologies do not, however, deterministically produce idealized conditions for discussion. There are many potential problems such as flaming and polarized debates. Moderators, it is sometimes suggested, are crucial to shaping the democratic potential of online discussion, because they help to mitigate many problems by actively intervening in the debates (Edwards 2002; Coleman and Götze 2001; Wright 2006b, 2007). There is, however, a great deal of confusion about exactly what roles moderators should, and do, perform (Barber 2003). This, in turn, leads to disputes about the nomenclature for such activities: are they moderators, facilitators, or censors?

Online Deliberation: Design, Research, and Practice.
Todd Davies and Seeta Peña Gangadharan (eds.).
Copyright © 2009, CSLI Publications.

This chapter develops two models that take account of the different roles which moderators perform in government-run online discussion forums. Two case studies, approximate to these models, are then presented. First, the Downing Street website. This featured a large-scale moderated discussion forum. Second, Citizen Space's E-democracy Forum, which was a smaller, policy-linked forum with interactive moderation. The case studies will highlight the practical positives and negatives of these models and lead to the generation of a series of policy suggestions about how the e-discussion agenda can be taken forward.

2 The Necessity of Moderation

Moderation is generally thought to positively influence the quality and usefulness of government-run online debates. For Kearns et al. (2002): 'The use of moderators is important in keeping citizen engagement focused and in consequently ensuring that such engagement adds value to services, to policy, and to citizens' (26). This is because: 'Free speech without regulation becomes just noise; democracy without procedure would be in danger of degenerating into a tyranny of the loudest shouter—or, in the case of e-democracy, the most obsessive, loquacious poster' (Blumler and Coleman 2001: 17-18). Barber (2003) likewise supports this position. He states: 'The question is not whether or not to facilitate, mediate, and gate-keep. It is *which* form of facilitation, *which* mediation, and *which* gatekeeper? The pretence that there can be none at all, that discourse is possible on a wholly unmediated basis, breeds anarchy rather than liberty and data overload rather than knowledge' (42). For Edwards (2002), 'the moderator can be characterized as a democratic intermediary' but must be independent of government in order to avoid a 'shadow of control' (5).

Blumler and Coleman (2001) argue for the creation of a civic commons in cyberspace, under the umbrella of 'a new kind of public agency' that would 'connect the voice of the people more meaningfully to the daily activities of democratic institutions' (16). This organization would be funded publicly but would be independent of government. Although such a proposal has many potential benefits, it is unclear how the summaries produced by such a body would be fed into the political process and what is meant, in practice, by the requirement that public bodies would 'be expected to react formally to whatever emerges from the discussion' (16). Secondly, it is questionable whether the rules and reports produced by an overarching body would be suitable across all government contexts.

3 The Fear of Moderation

In general, there is neither an acceptance of moderators as enhancing democracy, nor of what activities a moderator should perform. In fact, moderation does not come without potential costs if it is poorly structured, and can be very counterproductive (Coleman, Hall, and Howell 2002). Noveck (2004) has argued that: ‘To be deliberative, the conversation must be free from censorship’ and this ‘includes any distortion or restraint of speech that would hinder the independence of the discussion or cause participants to self-censor’ (22).¹ But I argue that we must be very careful not to automatically demonize the censorial role of the moderator: there are legitimate reasons for censoring the content of online discussion forums. This is because in the online world, constitutive (and/or self) censorship is arguably weakened by perceived anonymity.² Moral and social cues that shape speech acts are missing, and this gives people greater freedom to use profanity.

The fear remains, however, that the power to moderate the content of online forums will be abused. This could be done by setting overly restrictive rules or by ignoring ‘fair’ rules and deleting messages that are critical of the authority involved. It is, thus, necessary to draw a line between legitimate and illegitimate censorship. Determining what constitutes legitimate censorship is dependent on the context and is thus hard to define except in broad terms. Legitimate censorship could be defined as occurring when messages are deleted that do not meet specific, and open, rules for debate that have been discussed and agreed upon by a range of stakeholders. Illegitimate censorship occurs, then, when the rules are either too restrictive or are ignored by the moderator. To avoid value judgments, the analysis here will concentrate on whether or not the given rules were enforced.

The development and enforcement of moderation rules must be seen as fair—a complicated endeavor given that censorship can appear arbitrary. A great deal of trust has to be placed in the judgment of the moderator not to unduly censor messages. However, who should moderate discussions? Should it be independent, trained moderators, relevant policy experts, or ‘unbiased’ software? Because it is difficult to know when messages are being censored, these sorts of questions prove complicated to answer.

¹ Despite such an unequivocal statement against censoring discussions, the Unchat software, which was designed in relation to the values listed, contained a number of flexible censorship tools, potentially open to both individual participants and site administrators.

² Constitutive censorship relates to the latent, taken-for-granted rules by which discourse is structured (McGuigan 1996).

4 The Form of Moderation

Moderators can perform a range of duties. These are shaped by administrative aims, the technology used, the institutional context, the funding given to support the moderator, and by the moderator's decisions which determine the extent to which rules are followed. The list of potential roles presented below can be used interchangeably by moderators dependent on the specific aims and context:

- Greeter: making people feel welcome
- Conversation Stimulator: posing new questions and topics, playing devil's advocate in existing conversations
- Conflict Resolver: mediating conflicts towards collective agreements (or agreeing to disagree)
- Summarizer of debates
- Problem Solver: directing questions to relevant people for response
- Supporter: bringing in external information to enrich debates, support arguments
- Welcomer: bringing in new participants, either citizens or politicians/civil servants
- 'Cybrarian': providing expert knowledge on particular topics
- Open Censor: deleting messages deemed inappropriate, normally against predefined rules and criteria. Feedback is given to explain why, and an opportunity to rewrite is provided
- Covert Censor: deleting messages deemed inappropriate, but without explaining why
- Cleaner: removing or closing dead threads, hiving off subdiscussions into separate threads

Two broad models have been developed that take account of the potential forms of moderation. These are not fixed models. Moderation policies evolve and change in response to events. For example, if consensus is the goal, mediation strategies come to the fore (Morison and Newman 2001). They are, thus, intended as guides. A third possible model is unmoderated forums, e.g. Usenet. This is not explored here as no government-run forum was found to have adopted this policy.³

³ A further independent variable is the use of mechanized moderation on top of human moderation. Mechanized moderation is an electronic filter that blocks specific barred words. It is also possible for individual users to control the content that is made visible to them through 'kill files' on some forums. These block messages on particular topics or from specific people.

5 Content Moderation: The Downing Street Website

One model is human-based content moderation. The rules for moderation are set by the institution. This is a silent form of moderation because no feedback is given either to posters or to the institution. Silent moderation can create a conspiratorial atmosphere as messages are removed without explanation (Coleman, Hall, and Howell 2002). This is exacerbated if debates are not fed into the decision making process. Coleman (2001) describes this as 'tokenism', arguing that it is very counterproductive because: 'rational citizens seek outcomes from their participation and meaningful outcomes often depend upon there being a link between the virtual world of open discussion and the physical world of complex political relationships and institutions' (120).

The Downing Street website's online discussion forums differed from this model in two important ways. First, it used post-moderation: messages went straight onto the discussion forum before being moderated. Second, at least during the early stages of the forum, a mechanical filter was used in combination with human moderation.

The Downing Street website was redesigned on the February 10, 2000, and two discussion forums, 'Speaker's Corner' and 'Policy Forum', were added in an attempt to create a 'two-way link between government and people'.⁴ In contradiction to the aim, limited resources meant only a selection of posts received Official Responses: around 0.27% in the Speaker's Corner and these were primarily to questions about the discussion board itself (Wright 2002). This is unsurprising as the forums were moderated by the website team rather than by people with a direct policy making background.

The primary task for Downing Street's moderators was moderating the content of messages. Determining whether messages breach posting rules is subjective. One person might consider the word 'prat' acceptable while another might not. The degree of subjectivity can be limited by having clear and detailed guidelines. In this case, the rules were quite vague. Initially, the site carried only a warning not to swear because children may visit the site. This was subsequently strengthened, in line with government guidelines: 'Please do not make inappropriate postings, including those containing offensive, defamatory or libelous comments'.⁵ Nevertheless Kevin Webster, Chairman of the site's Independent Users Group, noted that the

⁴ This quote forms part of the original stated aim of the site.

⁵ See <http://www.cabinet-office.gov.uk/moderngov/download/modgov.pdf> (last accessed October 6, 2008).

Downing Street discussion forum had ‘become a haven for people to post offensive and meaningless messages’.

To give an idea of the number of abusive messages and thus highlight exactly why content moderation was necessary, a search for various offensive words was conducted on a random sample of seven forums from the second version of the discussion board.⁶ In total, there were 256 messages (out of 20,540 total, including all ‘missing’ messages) that featured one or more of the words (some messages contained literally hundreds of swear words, but are counted here as one). The findings suggest that the government was justified in moderating the content of the discussions as there were numerous offensive messages—although arguably the use of profanity was surprisingly small given the volume of messages involved.

The problem for Downing Street was drawing a line between ‘abuse’ and legitimate criticism of the government. The site was, after all, designed to make government more transparent and accountable. The moderators noted: ‘it is often a difficult line to tread to ensure that the debate is kept as open as possible, while removing inappropriate postings. The emotive topics which are discussed on this forum make that task particularly challenging’.⁷ Indeed, a number of mistakes were recognized: ‘The Magna Carta was deleted in error, I know that it has caused a lot of irritation and please accept my apologies for the mistake’.⁸

We have seen that censoring the content of online discussion is necessary if debates are not to be fractured by rude language. There is still a fear that moderators might abuse this power and censor messages that legitimately criticize the government. Such a fear would appear to have been at least partially upheld. There were numerous claims that Downing Street officials censored discussions inappropriately, particularly in the first incarnation of the forum. The IR35 discussion forum was particularly heated and many critical messages ‘magically disappeared overnight’.⁹ This created bad publicity for the government, leading to accusations of excessive control and censorship in the Times and technological naivety in The Observer.

⁶ The discussion board collapsed on May 12, 2000, after the message number field exceeded 32,767. The site designers took this opportunity to improve the software. Unfortunately, all the messages that were made to the forum before this crash were lost—except for the <http://no10.quivcustodiet.net/> (last accessed October 6, 2008) cache, which had copies of all the messages that were sent to the forum.

⁷ See <http://archive.cabinetoffice.gov.uk/moderngov/download/modgov.pdf> (last accessed October 6, 2008).

⁸ See http://no10.quivcustodiet.net/cgi-bin/show_archive2?fid=13&mid=31968 (last accessed October 6, 2008).

⁹ See posts in the IR 35 Forum 15/03/2000 (IR 35 refers to income taxation and the self-employed). Archived at: <http://no10.quivcustodiet.net/cgi-bin/archive2?fid=73> (last accessed October 6, 2008).

Two deleted messages are presented below, taken from the first incarnation of the forum at the time by the Quiscustodiet team:¹⁰

Shouldn't this website be independent of any particular party? It definitely tries to give the impression that the labour party and the government are the same thing whereas they are merely the current 'majority shareholders'. Is there an independent alternative to this site?

And:

What is the point of this website if the points made and questions posed are not responded to by government?

If we cannot expect some level of response we may as well make the point with a paint can on the nearest wall!! Government must do more than just provide the 'wall' and then pretend they are a listening open government because they opened a website.

In the light of the rules, it is difficult to explain why these messages were deleted. Removing them appears to be politically motivated: they criticize the government/website without the use of foul, racist, libelous, or offensive language.

Moderators do not just delete messages because of their content, however. They also perform housekeeping functions such as deleting stale threads. Such legitimate activities heavily shaped the discussions: 53.9% of messages sent to the Downing Street website were not visible at the end. Undoubtedly the majority of these deletions were legitimate. However, there was still the potential for this to lead to accusations of censorship: several moderation practices were not listed in the rules. Most importantly, if a message did not receive a reply within three days, it was automatically deleted and messages that replied to a deleted message were also deleted in an attempt to maintain the coherence of the discussions. Thus, in contradiction to the rules, many legitimate messages were deleted. This was further complicated by the use of a language filter that operated on the first version of the site.¹¹ It was initially set to block messages containing words such as 'bomb', 'anarchy', and 'fairy', but these were significantly reduced after complaints. Such practices explain much of the controversy about political censorship.

¹⁰ A systematic analysis of censorship during the first incarnation is not possible because of the way the data was stored.

¹¹ See http://no10.quiscustodiet.net/cgi-bin/show_archive2?fid=13&mid=33800 (last accessed October 6, 2008).

6 Interactive Moderation: The E-democracy Forum

A second model is interactive moderation. In this model, the communication is two-way and the moderator far more interventionist. This model is approximate to Edwards' conceptualization of the moderator as a democratic intermediary (Edwards 2002). The moderator brings both new citizens and political institutions into the discussion, encourages existing users to respond, moderates the content of messages and attempts to maintain civility, where possible, by persuasion and not censorship, frames the debate and sets subtopics, provides feedback to the institution, and participates in the debates.

The E-democracy forum, hosted on Citizen Space, was a small-scale, policy-linked discussion board that included 427 posts in 73 separate threads. A system of interactive pre-moderation was adopted, mitigating the problem of inappropriate posts being aired publicly before being removed. Moderators gave ongoing feedback and generally guided the discussions by providing topics for debate. Although the moderator initiated the most new threads, these topics also produced the most responses: 31.9% of all discussion.¹² The moderator was also successful in getting politicians to participate.

Moderators adopted stricter rules and regulations than found on the Downing Street website, yet fewer messages were censored: 26.3% (152 of 579). These were censored either because messages were repeats, used foul language, had inappropriate Web links, or were considered off-topic.¹³ The lower levels of censorship can probably be explained by this forum's lower profile and less emotive topic.

Despite the time and resources invested in premoderation, the results could be considered ambiguous. It was 'not possible to conclude that the provision of the consultation on the Internet significantly increased the number of people included or the spread of the e-democracy debate'.¹⁴

¹² This is only a rough guideline, because it does not take into account the number of initiated topics. Secondly, discussions within a topic tend to have a life on their own and move on from the initial post. Thus, people are often not responding to the initial post.

¹³ See http://web.archive.org/web/*/http://www.edemocracy.gov.uk/feedback/responses/edemocracy_discussion_final_summary.doc (last accessed November 15, 2008, original site, http://www.edemocracy.gov.uk/feedback/responses/edemocracy_discussion_final_summary.doc is no longer available).

¹⁴ See http://web.archive.org/web/*/http://www.edemocracy.gov.uk/downloads/your_response_report.doc (last accessed November 14, 2008, original site, http://www.edemocracy.gov.uk/downloads/your_response_report.doc is no longer available). The online provision was, however, 'particularly successful in distributing the consultation

Moreover, it is not clear what effect the forum had. A consultation report was developed which showed that messages were analyzed at some length, but the development of a formal policy has floundered (Wright 2006a).

7 Conclusion

This analysis has highlighted the problems and possibilities when moderating government-run online discussion forums. Interactive moderation can promote discussion and bring in new participants and can, thus, produce democratic/discursive benefits. The extent to which benefits outweigh financial costs depends on the aims and size of the discussion. The value of this model decreases for larger discussion forums.

On the downside, numerous problems were discovered, particularly with the large, content-moderated Downing Street website. Most prominent were allegations of censorship that dogged the forum. This was primarily a structural problem caused by poorly designed and poorly advertized rules. To resolve this, following Blumler and Coleman, and Edwards, I argue that the censorial power of the moderator would most fruitfully be enforced by an independent body following detailed (and openly available) rules set by the institution in negotiation with a range of stakeholders. This proposal differs from their models in which there is a link to policy making. I argue that it would be beneficial to separate the roles of the moderator into two clearly defined areas. Independent censors would be supplemented by civil servant facilitators. This would stop the facilitator (and government) being tainted by accusations of censorship. It would also mean that the facilitator would have direct experience of, and links with, the governmental body concerned—mitigating both the problems experienced on the Downing Street website, and with Blumler and Coleman's model wherein the summarizer does not have direct experience of the policy being discussed.¹⁵ The moderator can, thus, perform important democratic functions, but such practices are not without problems and must be carefully planned and thought through.

document' with 22,000 copies of the main document being downloaded and about 18,500 copies of the background and summary.

¹⁵ This is not to say that this model must always be used. The strategy must be adopted in relation to the aims—one can envision situations where having completely independent facilitators would be appropriate.

Acknowledgements

A longer version of this chapter was published in 2006: Government-run Online Discussion Fora: Moderation, Censorship and the Shadow of Control. *British Journal of Politics and International Relations* 8(4): 550-568.

References

- Barber, B. 2003. Which Technology and Which Technology? *Democracy and New Media*, eds. H. Jenkins and D. Thorburn, 33-47. Cambridge, MA: MIT Press.
- Blumler, J. G. and S. Coleman. 2001. *Realising Democracy Online: A Civic Commons in Cyberspace*. London: Institute for Public Policy Research.
- Coleman, S. 2001. The Transformation of Citizenship. *New Media and Politics*, eds. B. Axford and R. Huggins. London: Sage.
- Coleman, S. and J. Götze. 2001. *Bowling Together: Online Public Engagement in Policy Deliberation*. London: Hansard Society.
- Coleman, S., N. Hall, and M. Howell. 2002. *Hearing Voices: The Experience of Online Public Consultations and Discussions in UK Governance*. London: Hansard Society.
- Edwards, A. 2002. The Moderator as an Emerging Democratic Intermediary: The Role of the Moderator in Internet Discussions about Public Issues. *Information Polity* 7(1): 3-20.
- Kearns, I., J. Bend, and B. Stern. 2002. *E-participation in Local Government*. London: Institute for Public Policy Research.
- McGuigan, J. 1996. *Culture and the Public Sphere*. London: Routledge.
- Morison, J. and D. R. Newman. 2001. On-line Citizenship: Consultation and Participation in New Labour's Britain and Beyond. *International Review of Law, Computers and Technology* 15(2): 171-194.
- Noveck, B. S. 2004 Unchat: Democratic Solution for a Wired World. *Democracy Online: The Prospects for Political Renewal through the Internet*, ed. P. M. Shane, 21-34. London: Routledge.
- Wright, S. 2002. Dogma or Dialogue? The Politics of the Downing Street website. *Politics* 22(3): 135-142.
- Wright, S. 2006a. Electrifying Democracy: 10 Years of Policy and Practice. *Parliamentary Affairs* 59(2): 236-249.
- Wright, S. 2006b. Government-run Online Discussion Fora: Moderation, Censorship and the Shadow of Control. *British Journal of Politics and International Relations* 8(4): 550-568.
- Wright, S. 2007. A Virtual European Public Sphere? The Futurum Discussion Forum. *Journal of European Public Policy* 14(8): 1167-1185.
- Wright, S. and J. Street. 2007. Democracy, Deliberation and Design: The Case of Online Discussion Forums. *New Media and Society* 9(5): 849-869.