Normative behaviour in Wikipedia

Chris Goldspink¹, Bruce Edmonds², Nigel Gilbert¹

¹ Centre for Research in Social Simulation, Department of Sociology, University of Surrey. Guildford, GU1 7XH, UK.
² Centre for Policy Modelling, Manchester Metropolitan University, Aytoun Bldg., Aytoun Street, Manchester, M1 3GH. UK.

Email address of corresponding author: c.goldspink@surrey.ac.uk

Abstract

This paper contributes to the debate about governance behaviour in on-line communities, particularly those associated with Open Source. It addresses evidence of normative self-regulation by analysing the discussion pages of a sample of Wikipedia Controversial and Featured articles. The assumption is that attempts by editors to influence one another within these pages will be revealed by their use rules and norms as well as the illocutionary force of speech acts. The results reveal some unexpected patterns. Despite the Wikipedia community generating a large number of rules, etiquettes and guidelines, explicit invocation of rules and/or use of wider social norms appear to play a small role in regulating editor behaviour. The emergent pattern of communicative exchange appear not well aligned either with these rules or with the characteristics of a coherent community and nor is it consistent with the behaviour needed to reach agreement on controversial topics. The paper concludes by offering some tentative hypotheses as to why this behaviour may be as it is.

Wikipedia

Wikipedia grew out of an earlier Web encyclopaedia project called Nupedia founded by Jimmy Wales with Larry Sanger appointed as its first editor-in-chief. From its inception Nupedia was linked to a free information concept and thus the wider open source movement. Nupedia used traditional hierarchical organizational methods which were time consuming. This and an associated lack of openness have been argued to explain the failure of the Nupedia. Sanger (2005; 2006), however, argues that the expert model was sound but needed simplifying.

Sanger was introduced to the WikiWiki software platform in 2001 and saw in it a way to address the limitations hampering Nupedia. The inherent openness of the Wikiwiki environment was, however, seen as a problem so Wikipedia began as an experimental side project. Sanger notes that a majority of the Nupedia Advisory Board did not support the Wikipedia, being of the view ‘...that a wiki could not resemble an encyclopaedia at all, that it would be too informal and unstructured’ (Sanger 2005). However the intrinsic openness of Wikipedia attracted increasing numbers of contributors and quickly developed a life of its own. Almeida et al (2007) note that growth in articles, editors and users have all shown an exponential trajectory. From Sanger’s earlier comments it is clear that he had been surprised at the rate of
development and of the quality achieved by the relatively un-coordinated action of many editors.

Social Self-Regulation in Wikipedia

The use and enforcement of principles and rules has been an ongoing issue within the Wikipedia community with a division emerging between the founders and within the wider community about whether rules were necessary and how they should be policed. The power to police rules or impose sanctions has always been limited by the openness of the technology platform.

In the early stages Sanger argues the need was for participants more than for rules and so the only rule was ‘there is no rule’. The reason for this, he explains, was that they needed to gain experience of how wikis worked before over prescribing the mechanisms. However, ‘As the project grew and the requirements of its success became increasingly obvious, I became ambivalent about this particular “rule” and then rejected it altogether’ (Sanger 2005). However, in the minds of some members of the community, it had become ‘the essence’ of Wikipedia.’

In the beginning, complete openness was seen as valuable to encourage all comers and to avoid them feeling intimidated. Radical collaboration — allowing everybody to edit everyone’s (unsigned) articles — also avoided ownership and attendant defensiveness. Importantly it also removed bottle necks associated with ‘expert’ editing. That said, the handpicking of a few core people is regarded by Sanger as having had an important and positive impact on the early development of Wikipedia. Sanger argues ‘I think it was essential that we began the project with a core group of intelligent good writers who understood what an encyclopaedia should look like, and who were basically decent human beings’ (2005). In addition to ‘seeding’ the culture with a positive disposition, this statement highlights the potential importance of establishing a style consistent with the Encyclopaedia genre, a stylistic model that might shape the subsequent contributions of others.

Sanger argues that in the early stages ‘force of personality’ and ‘shaming’ were the only means used to control contributors and that no formal exclusion occurred for six months, despite there being difficult characters from the beginning. The aim was to live with this ‘good natured anarchy’ until the community itself could identify and posit a suitable rule-set. Within Wikipedia rules evolved and as new ones were needed they were added to the ‘What Wikipedia is not’ page. Wales then added the ‘Neutral Point of View’ (NPOV) page which emphasised the need for contributions to be free of bias. The combination of clear purpose and the principle of neutrality provided a reference point against which all contributions could be judged. Sanger regards the many rules, principles and guidelines which have evolved since as secondary and not essential for success.

How do newcomers learn these (ever increasing) rules and do they actually influence behaviour? Bryant et al (2005) suggest that there is evidence of ‘legitimate peripheral practice’, a process whereby newcomers learn the relevant rules, norms and skills by serving a kind of apprenticeship. These authors argue that this is evident in new editors of Wikipedia initially undertaking minor editing tasks before moving to more significant contributions, and possibly, eventually, taking administrative roles. Bryant et al. have a somewhat idealistic view of involvement, however, for here newcomers
have the same rights as long standing participants, and that the process for socialising newcomers is often bypassed.

In some Open Source environments (such as Open Source Software) it is possible to gain reputation which may be usable in the wider world. The commitment to the community is often explained by arguing that a desire for reputation increases compliance (for an excellent overview see Rossi April, 2004). However, in the Wikipedia environment there is no list of contributors to which an editor can point as evidence of their contribution (although they can self-identify their contributions on their user page). Contributions are, in essence, non attributable. In the case of Wikipedia identification with product, community, values or personal satisfaction appear to be more likely reasons for remaining involved

**Wiki technology**

Wiki technology is easy to learn and contributing is extremely simple. There are few technical impediments confronting novice users. Wiki platforms are intrinsically open supporting decentralised action unless modified to control or restrict access. A division of labour emerges as editors choose which pages interest them and which they want to focus on contributing to or maintaining.

Wikipedia has added a number of facilities which support the ready detection and correction of vandalism. These include a *reversion* facility that supports the rapid reinstatement of the page content. Lih (2004: 4) attributes significance to this feature noting that ‘This crucial asymmetry tips the balance in favour of productive and cooperative members of the wiki community, allowing quality content to emerge’. and Stvilia et al. (2004: 13) note that ‘By allowing the disputing sides to obliterate each others’ contributions easily, a wiki makes the sides interdependent in achieving their goals and perhaps surprisingly may encourage more consensus building rather than confrontation’. Stvilia, Twidale, Gasser, & Smith (2005) among others identify *discussion pages* as an important ‘… coordination artefact which helps to negotiate and align members perspectives on the content and quality of the article.’ Discussion pages provide an opportunity for managing minor disputes about content or editing behaviour and for movement towards agreement.

Cifferoli (2007) has argued that a significant consequence of these technical features is the way in which they alter transaction costs (Coase 1993; Williamson and Winter 1993). The efficacy of cultural control will be influenced by factors such as the homogeneity of the user group and that group’s propensity for endogenous norm formation, rates of turnover of the group, and the effect of external perturbation of the group or of the task on which they are working. This may also be subject to feedback effects: reduced norm compliance may lead to higher turnover and reduced commitment, further reducing norm compliance.

In conclusion then, Wikipedia is a volunteer open source project characterised by low ties between contributors, no formal obligations and very few means for the exercise of formal sanction. There is a low level of reciprocity with contributors under no obligation to maintain engagement. The wiki technology is open, inviting many to the task and imposing low costs on participation while reducing transaction costs. There is however high reliance on pro-social behaviour dominating if agency costs (borne by individuals) is not to lead to high turnover and possible governance failure. The
anonymity of Wikipedia precludes much significant reputation effects outside of the small group of co-editors who maintain extended involvement with an article. Wikipedians have produced a set of permissions, obligations, rules and norms which have been documented in guidelines and etiquettes as well as embedded in technical artefacts such as style bots (programs that automatically scan pages for stylistic anomalies and correct them). The need for and effect of these is however controversial. From a governance perspective there are relatively few means within Wikipedia by which formal control can be exercised using these rules and the community relies instead on the use of informal or ‘soft’ control.

Analysis of Self-Regulation mechanisms

In Wikipedia there are two classes of activity: editing; and conversation about editing. This paper is not concerned with the editing activity (although this is to be considered in future research) but with the conversations which help to coordinate it. Insight into this can be gained by examining the Discussion pages which accompany many of the articles. The activity on Discussion pages comprises a series of ‘utterances’ or speech acts between contributors about editing activity and the quality of product. On the face of it, these pages should provide a fertile source of data to support analysis of how governance operates in the Wikipedia, in particular ‘soft’ governance. Within these pages we expected to see attempts by editors to influence the behaviour of one another through the only means available to them – communicative acts. We anticipated that these may exhibit some regularity which would allow us to examine both the range and type of events that led to the explicit invocation of rules and norms and which revealed emergent influence patterns which were themselves normative. We wanted also to examine what conventions prevailed and how these compared and interacted with the goal of the community and its policies. A convention is defined here as a behavioural regularity widely observed by members of the community. Policies include explicit codes of conduct as well as guidelines (etiquettes) and principles.

Method

For the study we randomly selected a sample of Discussion pages associated with both Controversial and Featured articles. At the time of the study (May/June 2007) there were 583 articles identified by the Wikipedia community as controversial and approximately 1900 as featured. The analysis reported here is based on a sample of 19 Controversial and 11 Featured articles. The most recent three pages of discussion were selected for analysis from each Discussion associated with the article included in the sample. These were subjected to detailed coding using the Open Source qualitative analysis software WeftQDA. Both qualitative and quantitative analysis was performed. The latter was undertaken by re-processing the coded utterances such that each utterance constituted a case and each applied code became a variable associated with that case. This data was then analysed using SPSS and MLwin.

A number of coding schemes for natural speech were considered before choosing the Verbal Response Mode (VRM) taxonomy (Stiles 1992). VRM has been developed over many years and used in a wide range of communication contexts. Stiles defines it as ‘a conceptually based, general purpose system for coding speech acts. The taxonomic categories are mutually exclusive and they are exhaustive in the sense that every conceivable utterance can be classified.’ (Stiles 1992: 15). The classification
schema is attractive where there is a need (as here) to capture many of the subtleties of natural language use that derive from and rely on its intrinsic flexibility and ambiguity yet map them to a more formal system needed for analysis. Additional codes were applied to identify: valence, subject of communication, explicit invocation of norms or rules and the associated deontic and trigger, whether the receiver/s accepted the illocutionary force of the utterance and the ID and registration status of the person making the utterance. There were 3654 utterances coded in these thirty three documents.

Findings

Style of Communication

There was a statistically significant correlation between the article group (Controversial vs Featured) and broad style of communication. This was however very small at -0.078 (p < 0.01, 2-tailed). This difference was most apparent when examined at the level of specific styles. Both groups had approximately similar proportions of neutrally phrased utterances (approximately 64%). Nearly one quarter (22.5%) of all utterances in the discussion pages of Featured articles were positive compared to only 11% in those of Controversial ones. By comparison nearly one quarter (23.9%) of all utterances in Controversial discussion pages were negative compared to 14% for featured. The positive styles of ‘affirming’, ‘encouraging’ and ‘acknowledging’ were significantly overrepresented in the discussion pages of Featured articles but underrepresented in the Controversial ones. The reverse was the case for the negative styles of ‘aggressive’, ‘contemptuous’ and ‘dismissive’ in the controversial. Overall, the most common positive utterance was affirming (4.7%) closely followed by encouraging (4.7%) and acknowledging (4.3%). The most common negative utterance was dismissive (8.2%) followed by defensive (6.4%) and contemptuous (3.5%). All the Wikipedia discussions sampled reflected a strongly neutral-objective style (although it was apparent from the Qualitative study that the content was sometimes far from objective or balanced).

Validation

Within speech act theory (Searle 1969; Habermas 1976), validation refers to whether an utterance made by one speaker is accepted, rejected, ignored or let go unquestioned by the intended recipient/s. In the Wikipedia sample 50% of all utterances were accepted without question. A further 18% were explicitly accepted by at least one editor; 11% were explicitly rejected and a substantial 22% were ignored. 25% of positive style utterances were accepted by at least one editor compared to 18% of neutral and only 9% of negative. By comparison only 2% of positive utterances were rejected compared to 9% of neutral and 26% of negative. Positive utterances were more likely to be accepted without question (61%) compared to negative (21.7%) and neutral (54.4%). Negative comments were more likely to be ignored (44.1%) compared to neutral (18.2%) and positive (11.4%). From this we can conclude that positive utterances are more likely to be validated than negative, but that overall, a significant number are ignored or rejected.

Normative and rule invocation

A norm is coded where an editor specifically invokes a norm which is a) not the subject of an existing Wikipedia rule or b) the editor does not link to a specific
Wikipedia practice or rule (even if one exists) but rather refers to a wider social standard. A rule is coded where an editor explicitly cites a rule present in a Wikipedia guideline, etiquette or style guide. Overall 5.2% of all utterances involved norm or rule invocation. This meant that Wikipedia rules were invoked 122 times and general social norms a further 77 times in 3654 utterances. This overall number was contributed to disproportionately by three (outlier) articles in the sample. Without these outliers the rate of invocation was in the order of 2%. Rules were most commonly invoked in response to neutral style communication (63.9%) followed by 27% in response to a negative style. Only 9% of positive style utterances were responded to with a rule invocation. By comparison, norms were most commonly invoked in response to negative style utterances (53.2%) followed by neutral (44.2%) and then positive (2.6%). The difference in likelihood of invocation by style was statistically significant (p=.001).

A Wikipedia rule invocation was most likely to be triggered by the form of an article (44.9%), an edit action (22%), an article fact or a person’s behaviour (both 16%). A norm was most likely to be triggered by a person’s behaviour (35.6%), an edit action (23.3%), article form (21.9%), or article fact (19.2%). This pattern did not differ to a significant degree between the Featured and Controversial sites. Nearly three quarters (73.6%) of rule invocations had the implicit deontic of ‘it is obligatory’. Norms also were most likely to carry this deontic (61.3%). The second most likely deontic was ‘it is permissible that’ (9.7%).

While there was no statistically significant difference in the degree to which either norms or rules were invoked between the Featured and Controversial articles, there was a qualitative difference in the role norm and rule invocation played. In Controversial discussions, social norms and rules were most likely to be invoked against the behaviour of an editor who was of a different view while in Featured sites, norms and rules were somewhat more often used by the editor as a reflection on their own contribution – i.e. involved a level of self-check.

Registered vs non-registered users

Although there was no statistically significant difference in the likelihood for either registered or non-registered editors to invoke norms or rules, there was a statistically significant difference between registered and non-registered editors (p < 0.01) when it came to validation. Registered editors were more likely than non-registered to be explicitly accepted (18.7% of utterances compared to 13.9%), less likely to be rejected (9.9% compared to 13.7%), considerably less likely to be ignored (18.3% compared to 34.7%) or unquestioned (53.1% compared to 37.6%). Qualitatively, however, it was much more common that un-registered users would make suggestions before undertaking edits, particularly in the Features articles, so their behaviour was less likely to attract action or comment. Non-registered editors were more likely to make negative style utterances (24.3% compared to 18.5%) and less likely to make positive utterances (9.5% compared to 17.4%). This difference was significant (p < 0.01).

Influence through Illocutionary Force

The theory of speech acts distinguishes between the meaning of an utterance and its pragmatic intent. A typical utterance may have a form that differs from the intent. The
utterance ‘could you close the door?’, for example, has the form of a question but the intent of advisement: the speaker intends the listener to close the door. With the VRM coding frame used in this research each utterance is coded twice, once to capture the semantic form and again to capture the use of language to exert (illocutionary) force (Searle 1969). In VRM, the relationship of form to intent is expressed using the statement “in service of” (Stiles 1992). In this example the question ‘could you close the door’ is ‘in service of’ the advisement ‘close the door’. In standard presentation this is recorded as (QA).

Edification in service of Edification (EE) is the most frequent form of utterance in the Wikipedia sample – 37% of all utterances were of this mode. The Edification mode is defined as deriving from the speaker’s frame of reference, making no presumption about the listener and using a neutral (objective) frame of reference shared by both speaker and listener. This mode is informative, unassuming and acquiescent. It reflects attempts to convince by neutral argument. An example would be ‘That edit was made last week’.

The second most common mode is that of Disclosure in service of Disclosure (DD). Disclosure is defined as being from the speaker’s experience, making no presumption, but being framed using the speaker’s frame of reference. This is summarised as informative, unassuming but directive. Unlike EE mode, DD mode represents an attempt by the speaker to impose or have the listener accept the speaker’s frame. 12% of all utterances adopted this form. An example would be ‘I don’t know much about this topic’.

The third most common mode is Disclosure in service of Edification (DE). The DE mode represents an utterance which is from the speaker’s frame of reference but as if it is neutral or from a shared frame. 8% of all utterances used this mode. This is a somewhat neutral mode where the speaker offers clearly labelled personal knowledge as information. An example would be ‘I believe it occurred in 1987’.

The fourth most common mode is Advisement in service of Advisement (AA). AA mode represents speech from the speaker’s experience, which makes presumptions about the listener and adopts the speaker’s frame of reference. It can be summarised as informative, presumptuous and directive. An example would be ‘You should change this immediately’. Approximately 7% of utterances were in this mode. A further 12% of utterances have the directive pragmatic intent of advisement masked by a less presumptuous form – Edification or Disclosure ( ‘It should be changed immediately’ or ‘I think it should be changed immediately’).

Significantly, utterances associated with politeness were very rare in this sample.

Discussion of Findings

What is significant about the utterance strategies is that they typically involve an exchange of assertions delivered with a neutral – i.e. non-emotive style. There are very few explicit praises or put downs, and few niceties like explicit acknowledgements. Seldom do contributors refer to one another by ID – the exchanges are very impersonal. This does not tally with what one would expect if the Wikipedia etiquette (http://en.wikipedia.org/wiki/Wikipedia:Etiquette) had been institutionalised. The Featured articles conform a little more closely with what one
would expect than do the Controversial, but if we assume that the etiquette captures the community’s ideal, the utterances do not conform to that ‘ideal’ in either case. Similarly we see low levels of questioning or of reflection (i.e. feeding back the words of the speaker to check understanding or to come to a better understanding of the other’s intentions). This is inconsistent with the task needs – to reach consensus on controversial topics. The frequency with which utterances were ignored also suggested low engagement by participants in the discussion. All of this would seem to need some explanation.

The absence of any expression of acknowledgement of emotions and/or similarity of attitude (homophily) among many contributors suggests that Wikipedia lacks many of the qualities of verbal exchange that would identify it as strong community. It is more consistent with being a place to share coordination of a task. This could suggest that the goal is the primary orientating point. However, the lack of quality of discourse needed to achieve consensus is more indicative of a brief encounter between individuals who struggle (or are not fully committed) to find common understanding rather than of a community committed to a common goal (Becker and Mark 1997). This might suggest that the shared goal may be subordinate to more personal goals for a considerable proportion of contributors. Or it may be that the technology and environment will support no more than this.

The Wikipedia environment supports saboteurs who can use the opportunity afforded by the open and anonymous platform to use identity deception i.e. to mimic the language and style of an ‘expert’ or to present as a genuine editor while trying to pursue a personal or political agenda hostile to the aims or interests of the Wikipedia. We found no direct evidence of this behaviour in the Discussion pages we sampled even though the discussions about controversial articles provide particularly fertile ground for such sabotage. However, vandalism is not generally evident in Discussion pages only on articles, there were some interactions which could have been regarded as trolling but were not explicitly labelled as such. Nevertheless the threat of these behaviours could have an influence on the type of communication conventions which arise. Editors may, for example, display reserve and suspicion, withholding trust and taking conventional signals of authority and identity (Donath 1998) as unreliable.

Qualitatively there was considerable evidence that editors appeared to form judgements about the intent of others on relatively little information. There was, however, little evidence of the use of utterance strategies to better understand or check these judgements. Only occasionally would an editor modify his/her style significantly if challenged. Of the rule invocations 26% were accepted, a similar proportion were rejected or ignored and the remainder went unquestioned. This is consistent with norms being triggered by a limited range of cues which allow individuals to locate themselves and select identities appropriate to a context and which then remain essentially stable. The invocation of rules and norms appears to have little to no immediate effect on behaviour. It is not clear if it has an effect on future behaviour as this cannot be ascertained from the available data.

Conclusions and future work

In this study we set out to identify the mechanisms that underpin the emergence of systemic self-organisation in a volunteer on-line global institution. The findings have challenged some of our assumptions and expectations, in particular:
• The detailed and specific behavioural etiquette published in Wikipedia seems to have little influence on the overall character and style of interaction.

• The overall quality of interaction of editors falls short of the range and quality of communicative style characteristic of a community and also of that which would be expected, given the nature of the task.

• Most regulation is achieved without the need for frequent explicit invocation of rules or norms. Rather, behaviour seems to accord to a convention which editors quickly recognise and conform to (or bring to the Wikipedia) and which minimally accommodates what needs to be done to satisfy the task in a context of divergent personal goals.

• There was a lack of evidence of active negotiation of expectations and standards and convergence of behaviour towards a norm. Within the discussion pages there appeared to be little obvious norm innovation, evolution, adaptation or extension. This suggests that on first encounter with Wikipedia, editors read a set of cues as to what constitutes appropriate or acceptable behaviour and then more or less accommodate it. Alternatively the order observed may be largely attributable to the prior socialisation of participants with local norms and rules playing a very minor role.

• While there is a difference between controversial and featured sites this is small and the quality of the interaction cannot explain the difference in status. Similarly there appeared to be little in the subject matter of the two groups of articles that would explain the difference – both contained subject matter which was contestable and subject to significantly diverse opinion.

There is no clear basis to argue that the apparent order is a direct result of the use of deontic commands associated with social norms and environment specific rules. Despite the fact that the community has been a prolific rule generator, they appear to play a minor role. Contributors demonstrate a style which is broadly inconsistent with these rules and not a good fit with the task.

Overall though there is order and it appears to be emergent. The mechanisms that underpin this emergence have not been revealed by the analysis undertaken to date although some hypotheses can be tentatively suggested. The neutral-objective style may be a consequence of the anonymity and open nature of the environment – leading to a suspension of trust. It may propagate as newcomers copy the pattern through a process of behavioural cueing. It is possible also that the order is due to pro-social behaviour internalized and brought to the task. The volunteer nature of Wikipedia, and the level of commitment required, is likely to mean that long term editors reflect a pro-social disposition (Penner, Dovidio et al. 2005). In this context a little norm/rule invocation may go a long way if not by influencing immediate behaviour then by encouraging future compliance and/or by giving an incentive for non-compliers to leave. Such a view is quite different from that presumed by previous theories of social norms.

While the findings of the research to date are far from conclusive they do challenge many of previous assumptions and suggest a range of alternative hypotheses. Some of these will be able to be critically examined by further analysis of the current data.
and/or by data currently being collected through a controlled wiki experiment as well as data proposed to be collected in a case study in Second Life. The EMIL simulator is being designed to support a range of alternative assumptions and so should allow us to test alternative hypotheses and contribute to our understanding of this increasingly significant phenomena.

**Acknowledgements:** The work on which this paper is based has been supported by EMIL n EU funded project (Sixth Framework Programme - Information Society and Technologies - Citizens and Governance in the Knowledge Based Society). EMIL comprises six Partner institutions: National Research Council, Italy; University of Bayreuth, Germany; University of Surrey, UK; Universität Koblenz – Landau, Germany; Manchester Metropolitan University, UK; and AITIA International Informatics Inc. Budapest. We thank the partners for their contributions to the research.

**References**


Rossi, M. A. (April, 2004). Decoding the "Free/Open Source (F/OSS) Puzzle" - a Survey of Theretical and Empirical Contributions, University of Sienna: 42.


