

The Hidden Wikipedia

Wikipedia as a System of Acoustic Separation

5/1/2009

Harvard Law School

LLM paper

Ayelet Oz

Table of Contents

Prologue: two understandings of Wikipedia.....	4
I. Introduction.....	6
II. Newbies and insiders: the social and technological construction of Wikipedia.....	11
III. Acoustic separation, transparency and crowding out	17
(1) Acoustic separation	17
(2) Crowding out theory.....	19
(3) Enforcement, acoustic separation and crowding out.....	23
IV. Conduct rules and Decision rules on Wikipedia.....	26
(1) Conduct rules: Wikipedia's policies	26
(2) Vandalism and talk pages.....	31
(3) Blocks and off-wiki platforms.....	39
(4) Bans and blocks.....	48
The Community Sanctions Noticeboard.....	50
The Arbitration Committee.....	57
Acoustic separations of Images	65
VII. Challenging the acoustic separation? - Flagged revisions and the politicization of code	70

Flagged Versions	72
VII. Conclusion: beyond bug and feature	80

Prologue: two understandings of Wikipedia

The most detailed description so far of Wikipedia's conflict resolution mechanism was presented recently by David Hoffman and Salil Mehra, in their article *Wikitruth Through Wikiorder*.¹ Hoffman and Mehra describe Wikipedia's conflict resolution system by focusing on the decisions of Wikipedia's Arbitration Committee and analyze them through game-theory tools to show how they help solve cooperation problems of this large-scale online collaborative project.

Nevertheless, shortly after the article was distributed, one of the main contributors to the English Wikipedia commented on it and argued that it failed to give an accurate description of the dispute resolution system on Wikipedia. The user, Durova, noted in his personal blog that Hoffman and Mehra focused on arbitration as the only means of banning editors, while ignoring the variety of community-based remedies available on Wikipedia.²

Without resolving the factual debate, the following paper will show that Hoffman and Mehra's 'mistake' is not incidental or a case of poor research. It is but one example of a structural character of Wikipedia. Using different mechanisms of "acoustic separation", Wikipedia deliberately presents two distinct images of its dispute resolution system to two different audiences. One image is transparent to 'newbies' (law professors included) while

¹ See David Hoffman & Salil Mehra, *Wikitruth Through Wikiorder* (2009) (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1354424)

² See Durova, *Wikitruth without Wikiorder*, in DUROVA: THE WIKI WITCH OF THE WEST (March 19, 2009), <http://durova.blogspot.com/2009/03/wikitruth-through-wikiorder.html>.

another is presented to long-time Wikipedians, such as Durova. The characteristics of this separation, the discourse it initiates and its theoretical contribution to the study of large-scale online collaboration will be the focus of this paper.

I. Introduction

A common tale of internet communities starts with a small and peaceful community. As it expands, however, the peaceful community sees its initial audience 'driven out by spammers, flammers, trolls, and know-nothings'.³ The fear of being taken over by unwanted users has been a recurring concern of many large-scale collaborative projects.⁴

Different online communities have responded differently to this concern by creating different mechanisms to exercise control over their virtual space. Some prevent the posting of unwanted content in the first place, by requesting prior approval from the site owners or by using automated filtering systems. Others use ex-post mechanisms, which remove the unwanted content after it has been posted and sanction the unwanted user.

In both cases, the use of enforcement mechanisms triggers a similar dilemma: under-enforcement could result in a paralysis of the community and damage to the collaborative project. Over-enforcement, on the other hand, could be perceived as an abuse of power by the community leaders, and might lead to a backlash and outrage from the members of the community.

³ Lior Strahilevits, *Wealth Without Markets?*, 116 YALE LAW JOURNAL 1472, 1493 (2007).

⁴ See Chris Wilson, *The Wisdom of the Chaperones: Digg, Wikipedia, and the Myth of Web 2.0 Democracy*, SLATE (Feb 22, 2008, <http://www.slate.com/id/2184487>, last visited Apr 25, 2009); Mattathias Schwartz, *The Trolls Among Us*, THE NEW YORK TIMES (Aug 3, 2008, <http://www.nytimes.com/2008/08/03/magazine/03trolls-t.html>, last visited Apr 25, 2009).

One obvious solution to this dilemma is to set an accurate enforcement mechanism, which creates neither under- nor over-enforcement. Slashdot's decentralized moderation system, for example, could be understood as an attempt to create such a system.⁵ In this paper I will present another solution which is used, as I will demonstrate, by the online encyclopedia Wikipedia.

Wikipedia is an open-content encyclopedia, built on wiki technology. The first wiki was launched in 1995 by Ward Cunningham on the premise that publicly editable web-spaces are a promising way to achieve fast, productive online collaboration.⁶ Established in 2001, Wikipedia is among the most prolific collaborative authoring projects. The English-language Wikipedia contains nearly three million articles.⁷ Other active Wikipedias exist in over 100 other languages. Due to its extraordinary success and its centrality for contemporary theoretical understandings of online collaboration, my focus on Wikipedia will also illuminate the phenomenon of large-scale collaboration as a whole.

The solution used by Wikipedia to the problem of enforcement is based on mechanisms of structural 'acoustic separation'. I refer to this concept, originally used by Meir Dan-Cohen to describe the criminal justice system, in two ways: the first is similar to Dan-Cohen's original concept, and it refers to separation between different sets of 'conduct rules' and 'decision rules'. One example of this kind of separation is the common deviances from the

⁵ See Slashdot, *Slashdot Moderation* (<http://slashdot.org/moderation.shtml>), last visited Apr 30, 2009).

⁶ Bo Leuf and Ward Cunningham, *THE WIKI WAY: QUICK COLLABORATION ON THE WEB* (2001)

⁷ See Wikipedia, *The Size of Wikipedia*, http://en.wikipedia.org/wiki/Size_of_Wikipedia. (last visited Apr 1, 2009).

formal policies in everyday enforcement practices by administrators. The second meaning of 'acoustic separation' I will use relates to different images Wikipedia presents to different publics. The most obvious example of this second kind of separation is the use of 'flagged versions', aimed at presenting to the general public of Wikipedia visitors a different version of an article than the one viewed by Wikipedia's registered editors. Many other kinds of separations – for example, undermining of the importance of Wikipedia administrators in the official policies, in contrast to their actual power - could be placed along the spectrum in which a set of rules becomes an 'image' of Wikipedia.

Contemporary studies regarding Wikipedia have presented a limited picture of the practice of its internal regulation. Most scholars have focused on the expressed policies and guidelines as an accurate description, and have failed to take account of the ways in which these policies are exercised in everyday online reality, thus adopting a nearly-utopist vision of Wikipedia.⁸

The following paper will present a realist understanding of Wikipedia's internal governance structure, one which looks beyond the policies and into the actual everyday practices.⁹ Several scholars have created quantitative analysis of Wikipedia's practices;¹⁰

⁸ See Jonathan Zittrain, *THE FUTURE OF THE INTERNET AND HOW TO STOP IT* (2008); Andrew Lih, *THE WIKIPEDIA REVOLUTION: HOW A BUNCH OF NOBODIES CREATED THE WORLD'S GREATEST ENCYCLOPEDIA* (2009); Brian Butler, Elisabeth Joyce and Jacqueline Pike, *Don't Look Now, But We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia* (2008); Fernanda B. Viegas, Martin Wattenberg & Kushal Dave, *Studying Cooperation and Conflict between Authors with History Flow Visualizations* (2004).

⁹ See Roscoe Pound, *Law in Books and Law in Action*, 44 AM. L. REV. 12 (1910).

¹⁰ See Hoffman & Mehra, *supra* note 1.

very few have used methods of qualitative and ethnographic analysis along this path.¹¹ This paper aims to fill this gap.

I will suggest a detailed analysis of specific discussions on Wikipedia in order to illuminate its practices. By providing such close reading, I will be able to point out new evidence of the actual practices on Wikipedia that stand in contrast to the official Wikipedia policies, and were not presented before. This evidence will not only demonstrate that Wikipedia's policies are not necessarily an accurate description of the everyday practice of Wikipedia, but that there is a systematic difference between them. This close qualitative study, although limited in scope, challenges many of the current studies and aims to present a more complex image of Wikipedia's mechanisms of power.¹²

Moreover, I will show that the use of separation is not limited to a list of practices, but that it also sheds specific rules on the debate it creates, which are reflected in the fears, concerns and arguments Wikipedia users make in this regard. This discourse analysis will both strengthen the factual evidence I will provide – by suggesting that it is not incidental or

¹¹ See Joseph M. Reagle Jr's works on his website, <http://reagle.org>, last visited Apr 30, 2009.

¹² For a critique of the current literature see Henry Farrel & Mellissa Schwartzberg, *Norms, Minorities, and Collective Choice Online*, 22 ETHICS & INTERNATIONAL AFFAIRS 357 (2008). As examples for the current literature, Farrel & Schwartzberg refer, for example, to David R. Johnson & David Post, *The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367 (1996); LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE, VERSION 2.0 (2006); Cass Sunstein, *Neither Hayek nor Habermas*, 134 PUBLIC CHOICE 87 (2008); YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM (2006); and Zittrain, *supra* note 8.

anecdotal – and will also add another layer to understanding the internal structure of the separation and the discussions it initiates.

The paper is composed of several parts. In part II, I will introduce the social construction of Wikipedia and its technological infrastructure as they are described in previous studies of Wikipedia. This part will serve as the core basis for my construction of the acoustic separation. In part III, I will present the concept of acoustic separation and its use as a solution to the problem of crowding out, which is imminent to the literature of large-scale collaboration. I will further present its internal paradoxes. In part IV, I will move to a close description of several Wikipedia discussions, which will illustrate the use of acoustic separation between conduct rules and decision rules on Wikipedia and the discourse they create. Part V will account for the relationship between the technological infrastructure of Wikipedia and the possibility to challenge it by politicizing it. Part VI will offer a conclusion and final remarks.

II. Newbies and insiders: the social and technological construction of Wikipedia

Wikipedia's users are composed of three distinct groups or classes. The first group is the 'consumers' of Wikipedia, its non-participating readers. Wikipedia, which has an Alexa rank of 7, attracts nearly 700 million visitors yearly.¹³ The second group is composed of the 'newbies', who are the majority of Wikipedia's 10 million registered users.¹⁴ These are new or low-intensity Wikipedia editors, who do not take an active part in constructing Wikipedia's governance mechanisms but contribute to its articles in different degrees. The third group is the project's leaders, the true insiders. In this group I include 1,638 Wikipedia administrators, and an order of up to 10,000 (probably much fewer)¹⁵ highly-contributing users who engage in the governance discussions of Wikipedia.¹⁶

¹³ ¹³ See Wikipedia, *About* (<http://en.wikipedia.org/wiki/wikipedia:about>), last visited Mar 28, 2009).

¹⁴ See Wikipedia: *Wikipedians*. (<http://en.wikipedia.org/wiki/Wikipedia:Wikipedians>, last visited Apr 25, 2009).

¹⁵ See Aaron Schwartz, *Who Writes Wikipedia?*, RAW THOUGHT (Sep 4, 2006) (<http://www.aaronsw.com/weblog/whowriteswikipedia>, last visited Apr 25, 2009)

¹⁶ See Wikipedia, *List of Wikipedians by Number of Edits*, http://en.wikipedia.org/wiki/List_of_wikipedians_by_number_of_edits. (last visited Apr 1, 2009). For a qualitative analysis of the contribution of different classes of users to Wikipedia's articles see Aniket Kittur, Ed Chi, Bryan A. Pendleton, Bongwon Suh, Todd Mythkowitz, *Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie* (2007). For a review of Wikipedia insiders as the projects' leaders see Joseph M. Reagle Jr., *Do as I Do: Leadership in the Wikipedia* (2005, <http://reagle.org>, joseph/2005/ethno/leadership.html, last visited Apr 30, 2009).

The composition of these groups changes as individual users change their roles along the spectrum, and become more engaged with the Wikipedia community.¹⁷ As one's participation becomes more central and frequent, she adopts new goals, new roles, and uses different tools.¹⁸ The process in which newbies become insiders on Wikipedia was recently analyzed according to the theoretical model of Legitimate Peripheral Participation (LPP) used in communities of practice.¹⁹ According to LPP, newcomers become members of a community initially by participating in peripheral tasks that contribute to the overall goal of the community. These activities are typically simple and carry low risk to the community. Through these peripheral activities, the newcomers become acquainted with the tasks, vocabulary and organizing principles of the community, and gradually they become insiders.²⁰ LPP suggests that newcomers should have direct access to the practices of experts, and the broader context into which their own efforts fit, in order to make the membership in the community optimal.²¹

According to this framework, Bryant, Forte and Bruckman explain how the use of technology and emergent social norms on Wikipedia influence the transformation of

¹⁷ Susan L. Bryant, Andrea Forte, Amy Bruckman, *Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia* (2005)

¹⁸ *Id.*

¹⁹ JEAN LAVE AND ETIENNE WEGNER, *SITUATED LEARNING: LEGITIMATE PERIPHERAL PARTICIPATION* (1991).
See also ETIENNE WEGNER, *COMMUNITIES OF PRACTICE: LEARNING, MEANING, AND IDENTITY* (1998).

²⁰ *See* Susan L. Bryant, Andrea Forte, Amy Bruckman, *Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia* (2005).

²¹ *Id.*

members' participation over time.²² At the periphery of Wikipedia, newbies contribute by reading articles that interest them and fixing minor mistakes or omissions, based on their own personal knowledge.²³ As they become insiders, the encyclopedia as a whole becomes the center of their interest, rather than any single set of articles. They recognize the project's overarching goals, the appeal of community, and its contributions to society.²⁴ Because their goals are broader than assessing the quality of a particular article and fixing it, the scope of their activities also extends beyond editing. They use their 'watch list' to track changes and monitor large sections of Wikipedia,²⁵ and start reading the accompanying pages to the main articles – the discussion pages and the page history. It should be noted that although, theoretically, the same set of tools are available to both newbies and insiders, their different goals and more sophisticated understanding of the site render some tools more visible and relevant to insiders than to newbies.²⁶

The process which Bryant, Forte and Bruckman present suggests a gradual and 'natural' process of growth inside the community, in which the newbies slowly develop into insiders

²² *Id.* For an in-depth discussion of Activity Theory's theoretical roots, see Yrjv Engstrom, *Activity Theory and Individual and Social Transformation*, in PERSPECTIVES ON ACTIVITY THEORY 19 (Yrjv Engestrom, Reijo Miettinen and Raija-Leena Punamaki, eds., 1999). For other uses of Activity Theory use Kari Kuutti, *Activity Theory as a Potential Framework for Human-Computer Interaction Research*, in CONTEXT AND CONSCIOUSNESS: ACTIVITY THEORY AND HUMAN-COMPUTER INTERACTION 17 (Bonnie A. Nardi ed., 1996).

²³ Bryant, Forte & Bruckman, *supra* note 17, at p. 4.

²⁴ *Id, id.*

²⁵ *Id*, at p. 5.

²⁶ *Id*, at p 6.

and reveal more and more of the Wikipedia world. I would like to challenge this harmonic description by referring to the barriers – social and technological – that newcomers need to pass in order to observe the internal practices of insiders. As I will show, Wikipedia elicits the gaps between newbies and insiders in the use of technological tools in order to structure different spaces of discussion – some spaces welcome newcomers, while others are in fact limited to insiders only, thus allowing a space in which to create a separate discussion, which presents values and practices different than those presented to newbies.

In order to understand the construction of these separate spaces, it is important to see the technological boundaries in light of the social practices. Although the technological boundaries themselves might seem very modest, since most of the information is open, and theoretically available to anyone, when seen in the light of the social patterns described, these barriers are in fact enough to create a functional separation of spaces.

The main technological tool I will refer to is the separation the Wiki software entails between talk pages and main pages.²⁷ The talk, or discussion, pages, accompany every Wikipedia page. They allow editors to explain and justify their changes to the main article. Moreover, the internal social norms of Wikipedia require such explanation on the talk page, so that controversial changes made without any corresponding explanation on the discussion page could be reverted by others even without any explanation on the merits.²⁸ Writing on the talk pages is subject to different rules of etiquette than those governing article pages. For

²⁷ For an analysis of the use of talk pages to refer to Wikipedia policies see Travis Kriplean, Ivan Beschastnikh, David W. McDonald, Scott A. Golder, *Community, Consensus, Coercion, Control: CS*W or How Policy Mediates Mass Participation* (GROUP, 2007).

²⁸ Zittrain, *supra* note 8, at p. 134.

example, in talk pages contributors are not supposed to erase each other's comments and every person is supposed to sign his or her postings using four tildes (~~~~), which in the syntax of the Wiki software links to their personal user page and notes the date and time. The threading in talk pages takes place according to different social conventions and posting rules meant to ensure that conversation threads are easily recognized.²⁹ Users are encouraged to indent related postings so that the relationship between associated contributions will be visually clear and to separate the talk page to different discussion topics by using visually separate sections.

Talk pages are among the oldest and most important coordination mechanisms on Wikipedia. They provide a channel for a debate regarding the articles and create small groups of people engaged in conversation before, during and after editing pages. This is one explanation why heavily-edited articles and talk pages go hand in hand. While the average edits per page in Wikipedia is roughly 15, more than 90% of the articles with more than 100 edits have related talk pages. Articles with associated talk pages have, on average, nearly 6 times more edits and 5 times more users than articles without talk pages.³⁰

But talk pages are not limited to improving specific articles. These are also the places in which conflict is often resolved, and governance structures are debated.³¹ These uses of talk pages, which are conducted away from the main article pages – and from the eyes of the

²⁹ *Id.*

³⁰ *Id.*

³¹ See Fernanda B. Viegas, Martin Wattenberg, Jesse Kriss, Frank van Ham, *Talk Before You Type: Coordination in Wikipedia* (2007), p. 5.

newbies – allow for a relatively private place for insiders' discussion. The gap which is created between the most important rules for newbies, articulated on the edit pages themselves, where new users are likely to encounter them, and the insiders' discussion threads, taking place on discussion pages of policy pages or on designated noticeboards, allows for a unification of the social construction with the technological infrastructure. The newbie/insider separation is reproduced by the technological separation between the main article and the talk pages. By placing different discourses on different virtual spaces, Wikipedia creates a system of acoustic separation between new and senior users, allowing it to elicit different motivations from the two groups.

III. Acoustic separation, transparency and crowding out

(1) Acoustic separation

In his 1984 article, *Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law*,³² Meir Dan-Cohen asked the readers to imagine a world in which there is an 'acoustic separation' between two sets of normative messages: conduct rules – rules which are addressed to the general public and are designed to guide its behavior, and decision rules – rules which are directed to the officials who apply the conduct rules.³³ In a world in which there is such separation, each group is unaware of the other's set of rules. Thus, the acoustic separation enables the designer of the system to send different normative messages to different groups.³⁴

³² See Meir Dan-Cohen, *Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law*, 97 HARV. L. REV. 625 (1984).

³³ *Id.*, at 630.

³⁴ *Id.*, at 635. For a critique of Dan-Cohen's article see Richard Singer, *Criminal Law: On Classism and Dissonance in the Criminal Law: A Reply to Professor Meir Dan-Cohen*, 7 J. CRIM. L. & CRIMINOLOGY 69 (1986). For other works influenced by Dan-Cohen's concept see Carol S. Steiker, *Counter-Revolution in Constitutional Criminal Procedure? Two Audiences, Two Answers*, 94 MICH. L. REV. 2466 (1996); Patricia S. Abril, *'Acoustic Separation' and the Hispanic Small Business Owner*, 10 HARV. LATINO. L. REV. 1 (2007); Elizabeth Philips Marsh, *On Rollercoasters, Submarines, and Judicial Shipwrecks: Acoustic Separation and the Good Faith Exception to the Fourth Amendment Exclusionary Rule*, 1989 U. ILL. L. REV. 941 (1989); Paul W.

The potential benefits of acoustic separation are demonstrated by Dan-Cohen by addressing the question of duress as a defense to a criminal charge.³⁵ The dilemma regarding duress stems from the fact that on the one hand, the law maker would like to deter potential actors so they will try and make the right choice even when they face external pressures; on the other hand, it seems unfair to punish a defendant for succumbing to pressures to which even his judges might have yielded. These conflicting arguments allegedly pose a hopeless tradeoff between the competing values of deterrence and fairness. The impasse dissolves, however, in a system of acoustic separation, which instructs the judge not to punish defendants who acted under duress, while at the same time hiding this decision rule from potential defendants, who will therefore be guided exclusively by the relevant criminal proscriptions.

Kahn, *Panel II: Managing Violence: Law as Fetish, Alibi or Deceit: Managing Violence: Acoustic Separation, Memorials and Scapegoats*, 77 REV. JUR. U.P.R. 317 (2008).

³⁵Dan-Cohen, *supra* note 32, at 633.

(2) Crowding out theory

The problem of duress, and the use of acoustic separation to solve it, is similar in its structure to the problem of 'crowding out', or 'motivation crowding effect' as it is defined in the study of large-scale collaboration. The problem of crowding out stems from the internal influences and relationships among different motivational levers in a given system.³⁶ The system can try and elicit several kinds of different motivations in order to attract people to contribute to its project. Several of these motivations could be intrinsic motivations (such as notions of solidarity, empathy or fairness),³⁷ while others will be external motivations (for example the use of reward and punishment).³⁸

The problem of crowding out is created due to the internal relations between these multiply motivations. Certain motivational levers could be mutually excludable or present costly tradeoffs, which will reduce the efficiency of one motivator following the use of another. It can happen, for example, when the system design mixes different elements, such as market mechanisms with social norms.³⁹ The basic idea that rewards, and in particular monetary rewards, may crowd out intrinsic motivations, was first suggested by Richard Titmuss in his 1970 book, *THE GIFT RELATIONSHIP*. There, Titmuss argued (without

³⁶ See Bruno S. Frey and Reto Jegen, *Motivation Crowding Theory: A survey of Empirical Evidence*, 15(5) J. OF ECONOMIC SURVEYS 589 (2001).

³⁷ See Yochai Benkler, *Law, Policy and Cooperation* (2008), p. 12.

³⁸ *Id.*, at p. 18.

³⁹ *Id.*, at p. 19.

empirical research) that paying for blood will reduce people's willingness to donate it.⁴⁰ Roughly at the same time, psychology studies showed as well that under particular conditions, monetary rewards will undermine intrinsic motivations.⁴¹

The crowding-out theory opposes one of the most fundamental economic principals – that raising monetary incentives increases supply. A famous example for this phenomenon is leaving a fifty-dollar check on the table at dinner party at a friend's house.⁴² The economic motivation will subtract from the social motivation, thus creating a situation which will probably not result in another dinner invitation. Another famous example stems from the study of daycare centers. Daycare centers are often confronted with the problem that parents arrive late to pick up their children, which forces the daycare's workers to stay after closing time. A typical economic approach would suggest introducing a fine for collecting children late. But the study showed that as a matter of fact, the number of late-coming parents increased substantially with the introduction of the fine.⁴³ That is, because the introduction of

⁴⁰ Richard M. Titmuss, *THE GIFT RELATIONSHIP* (1970)

⁴¹ See Edward L. Deci, *Effects of Externally Mediated Rewards on Intrinsic Motivation*, 18 (1) J. OF PERSONALITY AND SOC. PSYCHOLOGY 105; Edward L. Deci, *Intrinsic Motivation, Extrinsic Reinforcement and Inequality*, 22(1) J. OF PERSONALITY AND SOC. PSYCHOLOGY 113; EDWARD L. DECI, *INTRINSIC MOTIVATION* (1975); EDWARD L. DECI & RICHARD FLASTE, *WHY WE DO WHAT WE DO: THE DYNAMICS OF PERSONAL AUTONOMY* (1995); Edward L. Deci & Richard M. Ryan, *The Empirical Exploration of Intrinsic Motivational Processes*, 10 *ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 39; EDWARD L. DECI & RICHARD M. RYAN, *INTRINSIC MOTIVATION AND SELF-DETERMINATION IN HUMAN BEHAVIOR* (1985).

⁴² See Benkler, *supra* note **Error! Bookmark not defined.**, at p. 92.

⁴³ See Uri Gneezy and Aldo Rustichini, *A Fine is a Price*, 29 *J. OF LEGAL STUD.* 1 (2000).

a penalty system to the system of trust transformed the relationship between the parents and teachers and reduced the parents' intrinsic motivations to come on time.⁴⁴ Due to the change in the perceived nature of the performed task, the structure of motivations changed as well.⁴⁵

Although most instances of crowding out available in the literature deal with the effect of monetary incentives on intrinsic motivations, another kind of crowding out effect stems from the influence of extrinsic regulation, which focuses on reward and punishment, on intrinsic motivations of trust, solidarity and empathy. This instance of crowding out explains why it is hard to paint a picture of a free and open community while using authoritative, administrative and bureaucratic sanctions. Empirical support for this notion was given by Akerlof and Dickens who showed that severe punishment may not reduce crime, as it decreases self-motivation to obey the law.⁴⁶ The more a community is regulated, the more its members see themselves bound only by the exact scope of regulation (or more specifically, its enforcement).⁴⁷ The use of reward and punishment turns the focus away from a commitment to the community's values.⁴⁸ Another example relates to the connection between

⁴⁴ See Frey & Jegen, *supra* note 36, p. 18.

⁴⁵ See RONALD BENABOU & JEAN TIROLE, SELF-CONFIDENCE AND SOCIAL INTERACTIONS (2000); George A. Akerlof & William T. Dickens, *The Economic Consequences of Cognitive Dissonance*, 72(3) AMERICAN ECONOMIC REVIEW 307 (1982).

⁴⁶ Akerlof & Dickens, *supra* note 45.

⁴⁷ See Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

⁴⁸ Zittrain, *supra* note 8, at p. 128.

constitutional norms and disciplinary norms.⁴⁹ Civic virtue – understood as intrinsic motivation – grows as the law conveys the notion that citizens are to be trusted, by respecting their rights and allowing them to participate in governance. Rules that attempt to discipline the citizens tends to crowd-out civic virtue and undermine the support of the citizens to the law and their willingness to cooperate with authorities and the law.⁵⁰

⁴⁹ See Frey and Jegen, *supra* note 36, at p. 20.

⁵⁰ See EUGENE BARDACH & ROBERT A. KAGAN, *GOING BY THE BOOK: THE PROBLEM OF REGULATORY UNREASONABLENESS* (1982); E. ALLAN LIND & TOM R. TYLER, *THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE* (1988); Tom R. Tyler & Kathleen M. McGraw, *Ideology and the Interpretation of Personal Experience: Procedural Justice and Political Quiescence*, 42 J. OF SOC. ISSUES 115 (1986).

(3) Enforcement, acoustic separation and crowding out

Combining traditional economic deterrence theory with the insights of the crowding out literature suggests two possible opposing results to softening the certainty of authority-based reward and punishment mechanisms: either chaos due to the lack of external motivations or a re-introduction of order based on intrinsic motivations, as people choose to respect particular limits in the absence of enforcement.⁵¹ A third result, is elicited by the use of acoustic separation, is that these two scenarios will be realized: some people would take advantage of the situation to break the law, while others will internalize the normative commitments and their intrinsic motivations will become stronger.

Similarly to the problem of duress, using acoustic separation can help solve the concerns of crowding out in such a hybrid system. Hiding one design lever from the recipients of the other can decrease the tradeoffs which are created due to the clash between the two motivations. By using the separation between conduct rules and decision rules, the system can present free, open and fair conduct rules, while practicing authoritative decision rules. Using acoustic separation – leaving in place the external restraints in regard to the first group, while presenting leniency to the second group – will target the first group with authoritative external regulation, without undermining the intrinsic motivation of the second group. (Simply using two different sets of rules for the two groups, without concealing the use of sanctions from the second group, will not necessarily achieve that result, due to such system's internal biases and unfairness which in turn will undermine its intrinsic motivations).

Yet, despite the functional benefits of acoustic separation, it also has obvious costs. It (crowds out) motivations based on trust and credibility.⁵² If things are not actually as they are presented to be, it undermines the notion of 'playing fair', which Reagle spotted as a central component of Wikipedia.⁵³ Therefore, the use of crowding out posits a risk as well as a solution.

This kind of critique was presented by Dan-Cohen himself in his article.⁵⁴ I would like to use his answer here as well. Similar to law, which cannot escape the moral dilemmas that stems from the use of acoustic separation,⁵⁵ so can't a large-scale project escape the 'motivational dilemma' this process creates. The use of a 'strategic behavior', such as creating an acoustic separation, is surely not free of tensions and places different values in conflict.⁵⁶ But its internal paradoxes should not make us give up this concept. On the contrary, understood as a paradox and not as an absolute solution, the concept of separation also illuminates the semiotics of arguments that go along with it.

As noted before, I will use this concept both as an empirical claim and as a discursive claim. First, I will show that this kind of structure is actually being used in Wikipedia. In order to attract new users, they are presented with the balanced, fair and moderate official

⁵² See Benkler, *supra* note 37, at p. 14.

⁵³ Joseph M. Reagle Jr., *Is the Wikipedia Neutral?* (2006, <http://reagle.org/joseph/2005/06/neutralty.html>, last visited Apr 30, 2009).

⁵⁴ Dan-Cohen, *supra* note 32, at p. 665.

⁵⁵ *Id.*, at p. 667.

⁵⁶ *Id.*, at p. 675.

policies of enforcement. At the same time, the senior editors, and especially the administrators, practice an authoritative system of hidden sanctions meant to maintain order on Wikipedia. The separation between these systems, which is supported by Wikipedia's technological infrastructure, allows Wikipedia to attract those different users and elicit different motivations, without facing the costs of crowding out between the external and internal motivations. At the same time, it also serves a discursive function. The separation is put forward as a utilitarian argument in constructing policies, and is systematically debated against using transparency, legitimacy and trust arguments.⁵⁷

⁵⁷ Compare: Duncan Kennedy, *A Semiotics of Legal Argument*, 42 SYRACUSE L. REV. 75 (1991).

IV. Conduct rules and Decision rules on Wikipedia

(1) Conduct rules: Wikipedia's policies

The internal governance of Wikipedia is framed around a body of policies, guidelines and processes, which grow rapidly (and are one of the fastest-growing sections of Wikipedia).⁵⁸ 'Policies' are standards that, with rare exceptions, all users should follow; 'guidelines' are similar to policies, but considered more advisory, and 'processes' describe the way of following certain policies or guidelines.⁵⁹ For example, 'civility' is a policy, 'assume good faith' is a guideline, and the 'deletion process' sets the procedure to be followed before deleting articles.

All kinds of norms stem from one of three sources: documenting actual beneficial practices on a policy page, after seeking consensus for the documentation on the policy talk page; proposing a change to a policy in the 'village pump' and getting the community's consensus; and rarely, declarations from Jimmy Wales or the Wikimedia Foundation Board (particularly regarding copyright issues, legal issues, or server load issues).⁶⁰

⁵⁸ See Viegas, Wattenberg, Kriss & van Ham, *supra* note 31, at p. 5.

⁵⁹ Wikipedia, *Wikipedia Policies* i(http://en.wikipedia.org/wiki/wikipedia_policies, last visited Apr 10, 2009).

⁶⁰ Wikipedia, *Sources of Wikipedia Policy*

(http://en.wikipedia.org/wiki/wikipedia_policies#sources_of_wikipedia_policy, last visited Apr 10, 2009).

There are five kinds of Wikipedia policies: behavioral policies, content and style policies, deletion policies, enforcement policies and legal and copyright policies.⁶¹ The use of sanctions against users is regulated especially through the behavioral policies, which set the standards for behavior on Wikipedia, and the enforcement policies, which determine the possible actions available in enforcing the behavioral policies. Among these policies, the hard core of unwanted behaviors, which could result in a block or a ban, is embedded in the policies regarding civility, edit warring, harassment, making legal threats, personal attacks, sock puppetry, and vandalism.⁶²

While reading Wikipedia's policies – the conduct rules – it is obvious that their main message is one of fairness and solidarity. They try to address the voluntary aspects of compliance and not to elicit external motivations of punishment. The attempt to appear open and fair is not surprising, and goes hand in hand with a similar analysis that was conducted by O'Mahony and Ferraro regarding the formal authority in the open-source project Debian.⁶³ By blending bureaucratic and democratic mechanisms together, O'Mahony and Ferraro showed that by preserving democracy and accountability for its members, Debians' governance system can retain their interest and commitment.⁶⁴

⁶¹ Wikipedia, *List of Policies*, http://en.wikipedia.org/wiki/wikipedia:List_of_policies (last visited Apr 9, 2009).

⁶² Wikipedia, *List of Policies: Behavioral*, http://en.wikipedia.org/wiki/wikipedia:List_of_policies#Behavioral (last visited, Apr 9, 2009).

⁶³ Siobhan O'Mahony & Fabrizio Ferraro, *The Emergence of Governance in an Open Source Community*, TECHNOLOGY, INNOVATION AND INSTITUTIONS WORKING PAPER SERIES (2007).

⁶⁴ *Id.*, at p. 10.

The openness of Wikipedia is apparent from the basic rule of "ignore all rules". This slogan, which is one of the first 'Wikipedia policies', intended to appeal to those made 'nervous and depressed' by regulation⁶⁵ and it was understood to frame Wikipedia as radically inclusive. This combination of openness and fairness remained central throughout the formative stages of the formalization of policies, with the 'Statement of Principles' Jimmy Wales posted on his user page on October 27th 2001.⁶⁶ The second statement among these declared that Newcomers are always to be welcomed:

"There must be no cabal, there must be no elites, there must be no hierarchy or structure which gets in the way of this openness to newcomers. Any security measures to be implemented to protect the community against real vandals (and there *are* real vandals, who are already starting to affect us), should be implemented on the model of 'strict scrutiny'. 'strict scrutiny' means that any measures instituted for security must address a compelling community interest, and must be narrowly tailored to achieve that objective and no other."⁶⁷

⁶⁵ Larry Sanger, *The Early History of Nupedia and Wikipedia: A Memoir*, SLASHDOT.COM (Apr 18, 2005, <http://features slashdot.org/article.pl?sid=05/04/18/164213&tid=95>, last visited Apr 25, 2009).

⁶⁶Wikipedia, *Jimbo Wales' user page: Statement of Principles*, (http://en.wikipedia.org/wiki/user:Jimbo_Wales/Statement_of_principles, last visited Apr 18, 2009).

⁶⁷ *Id.*

The policy regarding administrators further states that 'administrators should never develop into a special subgroup of the community'.⁶⁸ This repeated notion is encompassed in Jimmy Wales' commonly referred-to comment from February 2003 (soon after creating the class of administrators):

"I just wanted to say that becoming [an administrator] is *not a big deal*. I think perhaps I'll go through semi-willy-nilly and make a bunch of people who have been around for a while [administrators]. I want to dispel the auto of "authority" around the position. It's merely a technical matter that the powers given to [administrators] are not given to everyone. I don't like that there's the apparent feeling here that being granted [adminship] is a really special thing."⁶⁹

One tenet of fairness in Wikipedia's policies relays on the processes they create. The focus on fair process mirrors empirical studies that showed that fairness is judged by subjects of a system through reference to the way in which the system operates, rather than its results. Psychological theories of justice showed that people care more about the fairness of the

⁶⁸ Wikipedia, *Administrators' Policy* (http://en.wikipedia.org/wiki/wikipedia_administrator, last visited Apr 19, 2009).

⁶⁹ Wikipedia, *Administrators policy page: No Big Deal.* (http://en.wikipedia.org/wiki/wikipedia_administrator#No_big_deal, last visited Apr. 15, 2009). See also Wikipedia, *What Adminship is Not* essay, (<http://en.wikipedia.org/wiki/Wikipedia:ANOT>, last visited Apr 15, 2009).

procedures through which the outcomes are derived than about the outcomes themselves. Therefore, the notion of fairness is based on neutrality, lack of bias, honesty, efforts to be fair, etc.⁷⁰

⁷⁰ Tyler, p. 7.

(2) Vandalism and talk pages

The most severe Wiki-crime is vandalism. It is defined as 'any addition, removal, or change of content made in a *deliberate* attempt to compromise the integrity of Wikipedia.' The vandalism policy page further lists a set of behaviors that are considered vandalism and others that do not constitute vandalism. Yet, although vandals obviously cause damage to the collaborative project of Wikipedia, the process governing the use of sanctions against vandals is very restrained. The central tool for the enforcement of the anti-vandalism policy is to leave an appropriate warning message on the potential vandal's talk page. The warnings are templates used to warn vandals, listed according to the nature and severity of the vandalism. The justification for the use of warnings is that though some of the vandals are incorrigible returning vandals who may be blocked quickly, others may be stopped by a simple warning and will go on to become productive contributors to Wikipedia.

There are five levels of warnings. They begin with Level 1, which assumes good faith and usually includes a 'welcome to Wikipedia' message or some variant of it and move up to Level 4im, which assumes bad faith and results in an immediate block.⁷¹ For example, the Level 1 warning for vandalism states:

Welcome to Wikipedia. Although everyone is welcome to make constructive contributions to Wikipedia, at least one of your recent edits did not appear to be constructive and has been reverted. Please use the

⁷¹ Wikipedia, *Template Messages* (<http://en.wikipedia.org/wiki/Wikipedia:WARN>, last visited Apr 27, 2009).

sandbox for any test edits you would like to make, and read the welcome page to learn more about contributing constructively to this encyclopedia.

Thank you.⁷²

On the other hand, the Level 4 warning for vandalism, to be used as the fourth warning, states:

This is the **last warning** you will receive for your disruptive edits. The next time you vandalize a page, you **will** be blocked from editing Wikipedia.⁷³

Even after a user is blocked, she may ask for another chance, in which case the "2nd chance" template will be used.⁷⁴ The template uses the following language:

⁷² Wikipedia, *Vandalism1 Template* (<http://en.wikipedia.org/wiki/Template:Uw-vandalism1>, last visited Apr 15, 2009).

⁷³ Wikipedia, *Vandalism 4im Template* (<http://en.wikipedia.org/wiki/Template:Uw-vandalism4im>, last visited Apr. 15, 2009).

⁷⁴ Wikipedia, *2nd Chance Template* (http://en.wikipedia.org/wiki/Template:2nd_chance, last visited Apr 15, 2009).

"This request for unblocking has been declined due to your history of vandalism and/or disruption to this encyclopedia. However, we are willing to give you another chance provided that you can earn back the trust of the Wikipedia community. To be unblocked you need to demonstrate that you are willing and able to contribute positively to Wikipedia."⁷⁵

The blocked user can demonstrate her ability to contribute to Wikipedia by creating substantial "mock-edits" of articles on her own talk page. Following those mock edits, she can re-request unblocking by calling an administrator to review her proposed edits. If the administrator is convinced that the proposed edits will improve Wikipedia as an encyclopedia, the user will be unblocked. Moreover, if the user needs help while working on her proposed edits, she may add a "help me" comment to her talk page and an administrator will come to her aid.⁷⁶

It's obvious, as noted, that the vandalism policy is aimed to be a design lever for openness and fairness. Yet, moving to the decision rules, as they are apparent from reading the policy's talk pages, reveals a picture of broad discretion to administrators, unlimited by the procedures listed on the policy's main page. In this case, the acoustic separation between the conduct rules and the decision rules is maintained by refusing to move the practice from the talk page to the main policy page.

⁷⁵ *Id.*

⁷⁶ *Id.*

The discussion I will focus on started with one suggestion out of many to allow immediate blocking of anonymous IPs who engage in vandalism.⁷⁷ This is a common problem and a repetitive suggestion for a policy change, since a major strain of vandalism comes from bored students editing Wikipedia from schools. This kind of vandalism typically consists of the students making articles about their teacher or listing their friends as "gay". Yet, the problem with blocking these IP addresses is that most schools use only one network access point for the entire school, so the edits of one student can trigger a block that would affect the entire school.⁷⁸

In this specific discussion, several users, among which the most dominant was Kotniski, presented a common "tough on vandals" approach:

"I don't know why we want to waste any more time than we have on these people. 'Harsh' isn't remotely the applicable adjective. If someone comes here to damage our project, just show them the door without fuss – that's less punishment than they deserve."⁷⁹

⁷⁷ Wikipedia, *Vandalism Policy Talk Page, Policy Change Thread* (http://en.wikipedia.org/wiki/Wikipedia_talk:Vandalism#Policy_change, last visited Apr 20, 2009).

⁷⁸ See Lih, *supra* note 8, at p. 178.

⁷⁹ *Supra* note **Error! Bookmark not defined.**

Another user, Philcha, added that he already deviates from the official policy, and does not use 'those wishy-washy templates', since they 'don't give the impression that they mean business.'⁸⁰ The administrator Kafziel answered these calls by referring to more than a bit of administrative elitism, while revealing some of the actual decision rules used by administrators:

"It's always non-admins that come up with this stuff because they don't understand how things work. And, yes, a certain level of tolerance for vandals *is* required for adminship. Proposals like this are always drawn up by frustrated users who don't understand that it's not the policy that dictates the actions of admins, it's the actions of admins that create the policy. You can't legislate change; you can't force us to do anything".⁸¹

Following this remark, which exposes the gap between the conduct rules of the official policy and the decision rules as described by Kafziel, Kotniski edited the Vandalism policy main page so it no longer included the requirement of warnings, "to emphasis that warnings are not necessary in every case."⁸² He explained his edit at the talk page by presenting the

⁸⁰ *Id.*

⁸¹ *Id.*

⁸² See the edit to the *Wikipedia Vandalism Policy* page on [http://en.wikipedia.org/w/index.php?title=Wikipedia:Vandalism &diff=prev&oldid=265484844](http://en.wikipedia.org/w/index.php?title=Wikipedia:Vandalism&diff=prev&oldid=265484844) (last visited Apr 16, 2009).

standard critique of acoustic separation: the fact that it is based on hiding practices from the subjects of the rule and thus stand in stark contrast to the values of openness and candor.⁸³

"I don't think this changes the status quo – everyone seems to agree that blocks without warnings are not uncommon. [...] In one of your edit summaries you say, quite rightly, 'administrators know when to make exceptions'. But these pages are not for administrators, but for *ordinary users*. It needs to be made clear to *them* that there is no need to go through a long process of warning when an active vandal is on the prowl. At the moment I'm sure many people are deterred by the way the instructions are phrased. [...] [I]t seems to go against the community spirit to deliberately keep information out of the policy pages because the masses can't be trusted to interpret it correctly. If it's sometimes right for the project for admins to block vandals without warnings, then it's something right for the project for non-admins to report them without warnings."

Kafziel, in reply, expressly opposed Kotinstki's proposal, and explained the need for separation by referring too to the standard benefits of acoustic separation: the need to engage certain motivations in users. In his answer, he argued that the best results will be achieved by presenting different conduct rules to the ordinary users:

⁸³ See Dan-Cohen, *supra* note 32, at p. 666.

"[D]oes the possibility exist that an admin will block a vandal without a warning? Yes. I've done it many times. Do we need to mention that possibility in the policy or the noticeboard? No. Because, quite frankly, it's none of your concern. And I don't mean that in some elitist way; it's none of *anyone's* business, admin or not, if I choose to wait for a full set of warnings. Nobody can *make* me block someone I don't feel like blocking, and I don't need people showing up and hounding me about it (which has happened in the past). So don't worry about it. Do what you're supposed to do, and if we see an exceptional case then we'll step in. The thing is, whenever we mention the possibility of an exception to *any* policy, everyone thinks *their* case is the exception. [...] Users should be keeping an eye on articles, not other users. Once you file a report, if the article stops being vandalized, that's all you need to know. [...] If and when there is an exception to be made (and that may never happen, as far as you know) the individual admin will make it on a whim. Some admins never block anybody, ever, for any reason. I might decide not to block a user simply because I just ate a nice breakfast, and French toast makes me happy. We don't need to add that to the policy. [...] [I]t needs to be clear to [ordinary users] that there *is* a need to go through a long process, because most users are not familiar enough to decide when exceptions should be made. [...] They waste our time by reporting content disputes, test edits, and honest mistakes. [...] People who think following the guidelines is too

much trouble are always free to work on something else. They often do, and they are rarely missed."

Following this discussion, the main 'vandalism' policy page was not changed, and the normative message of warnings as a conduct rule, although it is different from the decision rule as explained by Kafziel, remained in place. This discussion is a first example both of the existence of a gap between conduct rules and decisions rules, as well as to the discursive practices of separation, composed of a set of standard pro/con acoustic separation arguments.

(3) Blocks and off-wiki platforms

The central technical affordance used for sanctioning users is blocking. Blocking is carried out by an administrator, who is a Wikipedia editor that holds certain restricted technical abilities (currently, the English Wikipedia has 1,635 administrators).⁸⁴ Besides blocking, administrators can also protect frequently vandalized pages, move pages when there are name conflicts, exclude bulk vandalism from the recent changes list, and edit the front page.⁸⁵

When a blocked user attempts to edit a page, he is forwarded to a message stating that he is currently unable to edit pages on Wikipedia. The message also includes a short description of the meaning of a block and the remedies available to her (e.g. appealing the block).⁸⁶ The block is also recorded in the user's block log. (It should be noted that the blocking policy expressly prohibits using blocks for the sole purpose of recording warnings

⁸⁴ See Wikipedia, *Administrators Policy*, *supra* note 68. For an analysis of the administrators' nomination decisions on Wikipedia see Moira Burke & Robert Kraut, *Mopping Up: Modeling Wikipedia Promotion Decisions* (2008); B. Collier, M. Burke, N. Kittur and R.E. Kraut, *Retrospective versus Prospective Evidence for Promotion: The Case of Wikipedia* (2008).

⁸⁵ See Burke & Kraut, *supra* note 84, at p. 1.

⁸⁶ MediaWiki, *Blocked Text*, (<http://en.wikipedia.org/wiki/MediaWiki:Blockedtext>, last visited Apr 25, 2009).

or other negative events in a user's block log. This practice, typically involving very short blocks, is regarded as punitive and humiliating.)⁸⁷

According to the blocking policy, administrators should also notify the blocked user by leaving a message on their talk page (there are several templates which are meant to facilitate the notification, based on the reason for the block,⁸⁸ e.g. posting content which includes defamation,⁸⁹ a returning case of vandalism,⁹⁰ or when the user is a suspected sock-puppet.)⁹¹ The duration of the block should reflect the likelihood that the user will repeat the inappropriate behavior, the severity of the behavior and the past conduct of the user. Although the duration of the block varies according to the circumstances, incidents or disruptive behavior typically result in 24 hour blocks.⁹² Administrators must supply a clear and specific reason for blocking a user, and blocks are expected to be made based on

⁸⁷ Wikipedia, *Blocking Policy: When Blocking May Not Be Used* (http://en.wikipedia.org/wiki/wikipedia:BLOCK#When_blocking_may_not_be_used, last visited Apr 15, 2009).

⁸⁸ See Wikipedia, *User Block Templates* (http://en.wikipedia.org/wiki/Category:User_block_templates, last visited Apr 27, 2009)

⁸⁹ Wikipedia, *Blocking for Defamation Template*, (<http://en.wikipedia.org/wiki/Template:Defban>, last visited Apr 27, 2009).

⁹⁰ Wikipedia, *Blocking a Returning Vandal Template* (<http://en.wikipedia.org/wiki/Template:Returnvandal>, last visited Apr 27, 2009).

⁹¹ Wikipedia, *Blocking a Suspected Sockpuppet Template* (<http://en.wikipedia.org/wiki/Template:SSPblock>, last visited Apr 27, 2009).

⁹² Wikipedia, *Blocking Policy: Duration of Blocks*, (http://en.wikipedia.org/wiki/wikipedia:BLOCK#Duration_of_blocks, last visited Apr 27, 2009).

reviewable evidence and reasonable judgment, so that all the factors supporting the block could be subjected to independent peer review. This is the reason why administrators should not discuss blocking issues off-wiki (for example, in the #admins channel on IRC).⁹³

Removing a block or abridging a block is most common when an administrator decides that the block should be reversed, following a change of circumstances or an appeal. An appeal is in fact a motion that the block will be lifted by an uninvolved administrator. Usually, the unblocking administrator should not unblock the user before discussing the matter with the blocking administrator.⁹⁴ If consensus between the two cannot be reached, the user can appeal to the Arbitration Committee, whose decisions are binding. It should be noted, though, that in practice, the Arbitration Committee rarely hears cases involving short-term blocks, so this remedy is only available for unlimited blocks or bans.⁹⁵

However, in regard to the use of the blocking power, as in the case of sanctioning vandals, the decision rules are different from the conduct rules. One example is a recent ban appeal discussed on the Administrators' Noticeboard on April 2nd, 2009. The administrator KillerChihuahua banned the user Ferrylodge from editing the Sarah Palin article for one week, following a long and bitter history among the two. The sensitive relationship between

⁹³ Wikipedia, *Blocking Policy: Dealing with Off-Wiki Block Requests*, (http://en.wikipedia.org/wiki/wikipedia:BLOCK#dealing_with_off-wiki_block_requests, last visited Apr 27, 2009).

⁹⁴ Wikipedia, *Blocking Policy: Unblocking* (<http://en.wikipedia.org/wiki/wikipedia:BLOCK#Unblocking>, last visited Apr 27, 2009); Wikipedia, *Appealing a Block* (http://en.wikipedia.org/wiki/wikipedia:Appealing_a_block, last visited Apr 27, 2009).

⁹⁵ Wikipedia, *Appeal to the Arbitration Committee* (http://en.wikipedia.org/wiki/wikipedia:Appealing_a_block#Appeal_to_the_Arbitration_Committee., last visited Apr 27, 2009).

them is apparent, for example, from the comment Ferrylodge posted on his talk page following the block:

'KillerChihuahua, speaking bluntly, you are the admin at [Wikipedia] most desperately in need of admin training. [...] Your behavior continues to be disgraceful. [...] Your block is an abuse of admin tools. I suggest that you seek assistance from another admin. [...] I consider it an honor to be blocked by you, if that will serve to differentiate my editing habits from yours. I plan to delete this talk page section at the end of the day, and would kindly ask you to NEVER post at this talk page again, except to the extent absolutely required by Wikipedia rules, or to unblock. And I sincerely hope that means I will never see you posting at this talk page again. EVER. Be gone. Thanks."⁹⁶

Following Ferrylodge's appeal, several users engaged in discussion, trying to reach a consensus whether or not the block should be removed. Some of them commented that Ferrylodge has been a helpful editor, and that the block should be lifted, while others depicted him as a user who will continue a debate "at a length and degree that exhausts patience of nearly everyone who comes into contact with him" and therefore should be blocked. Yet, after several pages of discussion (which is required in order to reach consensus

⁹⁶ Wikipedia, *The user Ferrylodge's Talk Page*, (http://en.wikipedia.org/wiki/User_talk:Ferrylodge#One_week_ban_from_Sarah_Palin, last visited Apr 27, 2009).

according the blocking policy), the co-administrator enforcing the probation on the Sarah Palin article, SB Johnny, stepped into the discussion and moved it to informal and non-public mediation over email, in sharp contrast with the principles of transparency and consensus:

"KillerChihuahua, Ferrylodge, check your emails and back down please. It's pretty clear that neither of you trusts the other, and this is getting a bit unseemly (on both your parts). You're both here to improve the encyclopedia, not make drama, right? This is not a conversation, and if you can't have a conversation with each other, just talk to me, and I'll try to figure out why you can't just talk to each other. [...] I think we've gone over the edge into "schoolyard debate" at this point, so good taste suggests closure of this thread."

He later updated Ferrylodge's talk page, commenting that:

"KillerChihuahua agreed to lift the ban on the talk page, but I'm imposing an indef ban on the actual article, as we discussed via email. If it's OK with you, I think it would be better if we continued to correspond via email, agreed? For the benefit of those who are watching from the galleries, I want to work with Ferrylodge to try to see why some folks aren't willing to Assume Good Faith on his part, and see if there's something to be done to improve that situation. I've also asked both

KillerChihuahua and Ferrylodge to try and avoid dancing near each other's toes for a while."⁹⁷

This discussion shows again the gap between decision rules and conduct rules. It also suggests a stronger mechanism of acoustic separation than the one previously presented. Unlike the separation between the talk pages and the policy pages, here the acoustic separation is executed by moving the use of sanctions even further away from the open space and to the hidden platform of the personal email, thus making the decision rules unavailable to anyone but the parties themselves – leaving only the conduct rules transparent.

A second example for this stronger kind of acoustic separation is to be found in a discussion from November 18th 2007. The discussion started with the following comment by Durova, who placed an indefinite block on a user without giving any explanation:

"I have placed an indefinite block on this account as a disruptive sockpuppet. Due to the nature of this investigation, our normal open discussion isn't really feasible. Please take it to arbitration if you disagree with this decision. Thank you."⁹⁸

⁹⁷ *Id.*

⁹⁸ See the following edit on the Administrators' noticeboard (http://en.wikipedia.org/w/index.php?title=Wikipedia:Administrators%27_noticeboard/Incidents&diff=next&oldid=172323794, last visited Apr 15, 2009).

As suggested by the discourse of separation, the unsupported block initiated strong discontent and opposition from other users. For example, the user Newyorkbrad commented:

"I am sorry but some degree of further explanation here will be necessary. I appreciate your desire to keep your investigation confidential but a first review of the contribution history here reveals no disruptive edits of any nature and no warnings of any kind. A response on an urgent basis is requested."⁹⁹

But despite the outrage and although several threats were made that the matter will go to the Arbitration Committee, Durova kept her refusal to give more information. She answered the critics by asking them to have confidence in her. This reference to the functional benefits of separation, its contribution to better benefits, is again a part of the standard discourse I portrayed.

"This has been a tough call, but in my opinion a necessary one. I am very confident my research will stand up to scrutiny. I am equally confident that anything I say here will be parsed rather closely by some disruptive banned sockpuppeteers. If I open the door a little bit it'll become a wedge issue as people ask for more information, and then some rather

⁹⁹ *Id.*

deep research techniques would be in jeopardy. As I've said this before, take me to arbitration if you want to challenge this. I think I've said that enough time clearly – I opened this thread for exactly that purpose. More than half a dozen administrators have already seen this research."¹⁰⁰

This response initiated further opposition, but as a matter of fact, the other administrators did not act and no other administrator was willing to unblock the user until, two hours later, Durova herself unblocked the user and posted the following announcement:

"==Unblock with apologies==

When I make a mistake I like to be the first to step forward and correct myself. It's very surprising that a few facts didn't come to light sooner, given the amount of time my report circulated and the people who had access to it. This is, in fact, not a first account. But it's a legitimate situation. I request early closure and archiving of this thread to protect that person's privacy.¹⁰¹

¹⁰⁰ *Id.*

¹⁰¹ *Id.*

This discussion doesn't only serve as another example for the gap between the practice and the official policies and for the discursive power of the arguments of separation, but further provides a third mechanism of acoustic separation. Following this comment, the thread was completely deleted from the noticeboard – a very exceptional act in noticeboards, in which it is usually not acceptable to delete anything, since they serve as a place of discussion, similar to talk pages. The discussion is currently available only through the history pages, and in the noticeboards' archive. Furthermore, the blocked account's name was permanently deleted even from the history pages, so that it could not be retrieved at all. This is an extreme multiple use of mechanisms of acoustic separation: using the administrators' noticeboard, which is practically hidden from new users due to the different patterns of use; deleting the thread from the noticeboard, and permanently deleting the user's name. All these are meant to conceal the actual practices of enforcement away from the public eye and leave transparent only the policies themselves.

(4) Bans and blocks

A ban, unlike a block, is not a technical limitation but a formal revocation of editing privileges, usually in the scope of an article ban or topic ban (though a ban may extend to the entire encyclopedia). Bans are social constructs, which are not in themselves enforced by the Wiki software and do not disable a user's ability to edit any page. However, users who violate a ban are usually blocked, in order to enforce the ban.¹⁰²

A decision to ban a user from Wikipedia can arise from various sources: community bans, a remedy inflicted by the Arbitration Committee, or use of Jimmy Wales' authority to ban users.¹⁰³ The vast majority of banned users are banned by the community.¹⁰⁴ Jimmy Wales banned 12 users¹⁰⁵ and the Arbitration Committee 42 users (30 users for indefinite duration and 12 users for fixed periods),¹⁰⁶ while the community banned 209 users.¹⁰⁷ But

¹⁰² Wikipedia, *Banning Policy*, (http://en.wikipedia.org/wiki/wikipedia:Banning_policy, last visited Apr 27, 2009).

¹⁰³ The Wikimedia Foundation also has the authority to ban users, but it has rarely exercised this authority on the English Wikipedia. See Wikipedia, *Banning Policy: Decision to Ban* (http://en.wikipedia.org/wiki/wikipedia:Banning_policy#Decision_to_ban, last visited Apr 15, 2009).

¹⁰⁴ Wikipedia, *List on Banned Users* (http://en.wikipedia.org/wiki/wikipedia:List_of_banned_users), last visited Apr 27, 2009).

¹⁰⁵ See Wikipedia, *List of Banned Users: Banned by Jimbo Wales*, (http://en.wikipedia.org/wiki/wikipedia:List_of_banned_users#banned_by_Jimbo_Wales, last visited Apr 27, 2009).

¹⁰⁶ See Wikipedia, *List of Banned Users: Banned by the Arbitration Committee* (http://en.wikipedia.org/wiki/wikipedia:List_of_banned_users#banned_by_the_Arbitration_Committee, last visited Apr 27, 2009).

even the number of community bans is significantly outnumbered by the number of blocks – which de-facto ban a user from Wikipedia, but without the accompanying public consensus. According to the list of active blocks,¹⁰⁸ nearly 1,000 users are blocked every day, among which around 600 are permanent blocks. These numbers show that unlike, for example, Jonathan Zittrain's description of Wikipedia, blocks are definitely not 'rare and usually temporary', but an everyday practice which is hidden from the general public.¹⁰⁹ As a matter of fact, the vast majority of enforcement on Wikipedia is conducted through administrative blocks – a technical affordance – and not through the community's social constructs.

One sphere of conflict between the relative power of the administrators (practiced through blocks) and that of the community as whole (practiced through bans) is in the definition of a community ban. Community bans are defined on the banning policy page as bans placed by the community by consensus on editors 'who have exhausted the community's patience'. A ban can be placed in one of three ways: one way is that a topic ban may be placed when a user has been repeatedly disruptive in a certain area of Wikipedia. The community then engages in a discussion at a relevant noticeboard and reaches a consensus of uninvolved editors to ban him. A second way is when an administrator has blocked a user indefinitely and no uninvolved administrator is willing to unblock him. The third way is

¹⁰⁷ See Wikipedia, *List of Banned Users: Banned by the Wikipedia Community* (http://en.wikipedia.org/wiki/wikipedia:List_of_banned_users#banned_by_the_Wikipedia_community, last visited Apr 27, 2009).

¹⁰⁸ See Wikipedia, *Blocked IP Addresses and Usernames*, (<http://en.wikipedia.org/w/index.php?title=Special:BlockList>, last visited Apr 27, 2009).

¹⁰⁹ Zittrain, *supra* note 8, at p. 136.

when the community discusses the block on a relevant noticeboard, and reaches a consensus not to unblock the user after he was unilaterally blocked by an administrator.¹¹⁰

Due to the different emphasis given to community power and administrators' power by the different definitions, the issue of community bans is highly contested. One extreme view is that a community ban is in effect only where no administrator is willing to unblock the blocked user, thus leaving the community ban as a de-facto ban by administrators. The opposite line of argument is that a community ban is in effect only when the overwhelming majority of the Wikipedia community members who examined the situation (after it was posted publicly) support the ban. Understood through the lens of acoustic separation, the first view aims to strengthen the mechanisms of separation, while the latter is meant to increase the transparency of the use of sanctions to wider parts of the community. This conflict has reached its peak (and was decided on the side of separation) in the creation of the Community Sanction Noticeboard.

The Community Sanctions Noticeboard

The Community Sanctions Noticeboard was split off from the Administrators' Incidents Noticeboard in February 2007, so as to promote the discussion of community-related issues. The rationale of the creating the Community Sanction Noticeboard was explained as follows:

¹¹⁰ Wikipedia, *Banning Policy: Community Ban* ([#Community ban](http://en.wikipedia.org/wiki/Wikipedia:Banning_policy), last visited Apr 20, 2009).

"This page was created with the intent that it was to become a place for the community to come together and discuss matters similar to those that are normally discussed on the administrative and incident noticeboards. It was felt that because 'administrators' was used in the title of the other two pages, it excluded the rest of the community."¹¹¹

The user Durova, who was the main promoter of the Community Noticeboard, further explained its purpose in making the practice of community bans more transparent to the community as a whole by differentiating it from the administrators' noticeboard:

"A comment here on the increasing number of community ban discussions: although [the administrator's] noticeboard is open to everyone, its title does tend to scare away the unmopified crowd (it certainly had that effect on me before I assumed janitorial responsibilities). So since community bans – and potentially community enforced mediation as well – are in principle for the entire community, perhaps we should initiate a new noticeboard for community-specific action. I'm thinking something parallel to this [...]. Call it Wikipedia: Community noticeboard.

¹¹¹ See Ryulong's post on the first nomination for deletion discussion: (http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard#Wikipedia:community_sanction_noticeboard, last visited Apr 1, 2009).

The creation of the Community Sanctions Noticeboard was meant to increase the power of the community to decide on bans vis-à-vis the administrators. One of its main achievements should have been to make the process of banning more transparent, thus lifting some of the acoustic separation and the undermining the unlimited use of sanctions enabled by blocking. This familiar argument, a standard part of the discourse of separation, is apparent from EdJohnston's comment:

"[T]he consensus' may be up in the air somewhere, but occasionally people have the urge to make it tangible. The [Community Sanction Noticeboard] helps make Wikipedia less Kafkaesque, in the sense that people want to know where to turn or how decisions get made. Do you need to be an insider-insider before you have any idea what is going on?".¹¹²

The user Rocksanddirt's also added a standard argument that the opposition directed at the Community Sanctions Noticeboard was a 'double standard' which contradicts the policy as it is communicated on the official pages:

¹¹² See EdJohnston's comment on the nomination for deletion page (http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion /[Wikipedia:Community_sanction_noticeboard](http://en.wikipedia.org/wiki/Wikipedia:Community_sanction_noticeboard) #Wikipedia:Community_sanction_noticeboard, last visited Apr 1, 2009).

"There is a community consensus that a 'community ban' exists. There is modest consensus on what you have to do to get such a 'community ban' implemented. [...] [T]here is no real consensus on what the form of the discussion should be (how long, how many uninvolved editors, how many uninvolved administrators, closing and solution implementation). Therefore if we continue with the fantasy that the community runs the store, bans implemented (even in a haphazard and arguable way) are valid. [...] [W]e are often told that 'adminship is no big deal' by the powers that be on [the English Wikipedia]. We are also told it is a community run and governed organization [...]. Now we are being told in this case (and others) that *any administrator who objects to a community sanction invalidates that sanction*. Essentially, a **veto** of community decisions. That is an enormous power given to administrators. In my view it invalidates both 'adminship is no big deal' and that Wikipedia is a community governed organization".¹¹³

But despite the good intentions, the Community Sanctions Noticeboard became highly contested almost immediately after its creation. Therefore, it was nominated for

¹¹³ See Wikipedia, *Requests for Arbitration: Ferrylodge* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge/Workshop#That_the_ban_and_block_that_resulted_from_the_CSN_case_are_invalid, last visited Apr 15, 2009).

deletion on May 2007, only three months after it was created. Several users described the noticeboard on the nomination for deletion page as a place to 'get a vote on officially banning users' and compared it to the failed and infamous 'quickpolls' system.¹¹⁴ (The quickpolls system was a trial system for creating quick decisions about problematic users which consisted of polls among Wikipedia users. Quickpolls was in operation for only three months, from March to June 2004, before it was shut down).¹¹⁵ The discussion of the nomination for deletion was polarized between administrators and the power of the community. The administrator Nick commented by referring to the professionalism of administrators vis-à-vis regular users:

"I would never block someone based on the discussions present on the Community Sanctions Noticeboard, I resent a few editors behaving quite uncivilly and trying to force administrators into blocking users. The people requesting the sanctions aren't going to be liable for any fallout from Community ban, it's going to be administrators, and unless we have an overwhelming majority of administrators here who are happy to (in my view, recklessly) block a user per a request here, then it's totally useless. Now, why wouldn't I carry out a block/ban request from this board – Simple – The board doesn't disregard conflict of interest, if doesn't find all evidence, it can and frequently is very one sided and it's often used to

¹¹⁴ See Ryulong's comment, *supra* note 111.

¹¹⁵ See Wikipedia, *Quickpolls* (<http://en.wikipedia.org/wiki/Wikipedia:Quickpolls>., last visited Apr 19, 2009).

carry out vendettas against certain editors. There's enough little cliques on Wikipedia that would like to cause trouble and would like to be able to ban certain users, this board tries to give them ability to do so. For that reason, it must be deleted and any future board that works in a similar manner should be blocked. [...] If a user is to be banned, they are at least entitled to a fair and impartial investigation into their behavior."¹¹⁶

Other users, on the other hand, argued that administrators must follow community consensus. The user's badlydrawnjeff commented, sarcastically: "God forbid the community has a say in the process regarding removing people's editing [privileges]. How *dare* we infringe on the power of the administrators?"¹¹⁷

The May 2007 discussion did not reach a consensus on the deletion of the Community Sanctions Noticeboard, and resulted only in several suggestions for reform. But as the noticeboard continued to suffer from excessive speed in banning of users, simple voting, and forking of discussions, a second nomination for deletion was initiated in October of 2007. The nomination was justified by Picaroon in the following arguments:

¹¹⁶ See Nick's comment on the nomination for deletion page (http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard #Wikipedia:Community_sanction_noticeboard, last visited Apr 20, 2009).

¹¹⁷ See badlydrawnjeff's comment on the nomination for deletion discussion (http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard #Wikipedia:Community_sanction_noticeboard, last visited Apr 19, 2009).

"some [Community Sanctions Noticeboard] bans are discussed for less than a day; many are assented to with simple 'support' votes (no wonder 'votes for banning' is a common nickname); discussions have less than half dozen participants (how this could possibly be construed as consensus for something as major as revoking a user's editing privileges, indefinitely, escapes me); discussions to ban are started by people who are in edit warring with the person they propose the banning of, and this isn't even noted".¹¹⁸

In order to fully understand this nomination for deletion (which resulted in the end of the noticeboard), and the complex power relations that create the mechanisms of acoustic separation on Wikipedia, we need to look away for a short while from the discussion itself, and consider the discussion of the Arbitration Committee in the matter of Ferrylodge, who was banned by the community following a discussion on the board. At the same time, reading this case in the context of the nomination for deletion of the Community Sanctions Noticeboard sheds a new light on the case and on the mechanisms of acoustic separation used by the Arbitration Committee.

¹¹⁸ See Picaroon's summary on the second nomination for deletion discussion ([http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard_\(second_nomination\)#Nomination.](http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard_(second_nomination)#Nomination.), last visited Apr 19, 2009).

The Arbitration Committee

Wikipedia's Arbitration Committee imposes binding solutions to disputes that communal discussion or administrators are unable to resolve.¹¹⁹ The Committee started hearing cases in 2006, and it now includes 15 active arbitrators.¹²⁰ The cases heard by the Committee follow an extremely open form of participation: in addition to the parties, others may submit evidence or arguments about the conflict at issue, though the final result remains subject to a vote of the arbitration panel (it should be noted that the use of voting stands in some contrast to the general anti-voting policy in Wikipedia). In its rulings, the arbitration committee states the Wikipedia principles it finds applicable to the case and determines the facts of the conflict according to the evidence.¹²¹

The most severe sanction the committee can inflict is to permanently ban a user from Wikipedia. It may also impose different editing restrictions such as account restrictions (the user is limited to using only one account when editing), civility restrictions (the user is placed under probation and an administrator may block him he makes an uncivil edit), different kinds of probation, limiting his ability to revert changes or banning him from editing articles

¹¹⁹ Wikipedia, *Arbitration Committee* (http://en.wikipedia.org/wiki/wikipedia:Arbitration_Committee), last visited Apr 25, 2009).

¹²⁰ Wikipedia, *Arbitration Committee: Current Members* (http://en.wikipedia.org/wiki/wikipedia:Arbitration_committee#Current_members), last visited Apr 25, 2009).

¹²¹ See Hoffman & Mehra, *supra* note 1.

on certain topics.¹²² The Arbitration Committee is the main institute which places editing restrictions. Sixty-seven restrictions were placed so far by the Committee and only twelve were placed in consensus by the community (one was placed voluntarily).¹²³ The Committee may also place a user under mentorship of another editor, if the user agrees to the mentorship in order to avoid more severe remedies.¹²⁴ Besides particular restrictions placed on specific users, the Committee may also place general sanctions to be inflicted by administrators on any editors who misbehave while working on a particular sensitive topic.¹²⁵ Such general sanctions were placed, for example, on articles relating to pseudoscience, homeopathy, or Falun Gong.¹²⁶

When reading a main page of an Arbitration Committee case, the ruling seems very clear-cut. But an opinion of the Arbitration Committee is actually combined of eight different Wiki pages: besides the main page of the case (the Request for Arbitration page), there is an Evidence page, a Workshop page, a Proposed Decision page, and four respective talk pages

¹²² Wikipedia, *Editing Restrictions: Types of Restrictions* (http://en.wikipedia.org/wiki/wikipedia:Editing_restrictions#types_of_restrictions), last visited Apr 25, 2009).

¹²³ Wikipedia, *Editing Restrictions: Users Subject to Restrictions* (http://en.wikipedia.org/wiki/wikipedia:Editing_restrictions#Users_subject_to_restrictions), last visited Apr 25, 2009).

¹²⁴ Wikipedia, *Mentorship* (<http://en.wikipedia.org/wiki/wikipedia:Mentorship>), last visited Apr 10, 2009).

¹²⁵ Wikipedia, *General Sanctions* (http://en.wikipedia.org/wiki/wikipedia:General_sanctions), last visited Apr 10, 2009).

¹²⁶ Wikipedia, *General Sanctions: Active Sanctions* (http://en.wikipedia.org/wiki/wikipedia:General_sanctions#Active_sanctions), last visited Apr 12, 2009).

(one for each main page). Every case (which is laid out on a Wiki page) opens up with the following note:

"Please do not edit this page directly unless you wish to become a participant in this case. Only add a statement after the case has begun if you are named as a party; otherwise, your statement may be placed on the talk page, and will be read in full. Evidence, no matter who can provide it, is very welcome at [the evidence page]. Evidence is more useful than comments. Arbitrators, the parties, and other editors may suggest proposed principles, findings and remedies at [the workshop] page. That page may also be used for general comments on the evidence. Arbitrators will then vote on a final decision in the case at [the proposed decision page]. Once the case is closed, editors may add to the log of blocks and bans as needed, but closed cases should not be edited otherwise."¹²⁷

The division of the Arbitration Committee case into different pages is another mechanism of acoustic separation, meant to present a clear picture of the case to casual readers, who will read only the main case page, and leave the internal conflicts of the case transparent only to insiders, who will usually only be those who take part in the case.

¹²⁷ See, e.g., the case of Ferrylodge: Wikipedia, *Requests for Arbitration: Ferrylodge* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge), last visited Apr 19, 2009).

The Ferrylodge case was discussed from October 15th, 2007 to November 29th, 2007, just as the second nomination for deletion of the Community Sanctions Noticeboard was taking place. The case carries in total 146 printed-out pages, among which only 11 pages are the main Wiki page – meaning that the vast majority of the information regarding the case requires an in-depth reading of the different case pages (the Evidence page holds 33 printed pages, the Workshop page holds 37 printed pages, the Proposed Decision page holds 12 printed pages, and 63 printed pages are the accompanying talk pages of the case (almost half of the total information is to be found, therefore, on the talk pages).

In order to emphasize the gap between the image of the case presented to regular users and the one presented to insiders, I will first describe the case by referring to the main pages alone. The first part of the main case page lists the involved parties in the case:¹²⁸ the user Ferrylodge,¹²⁹ the administrator FeloniousMonk,¹³⁰ and the administrator KillerChihuahua.¹³¹ Following the introduction, the page moves to the opening statement of Ferrylodge (who was unblocked for purpose of his appeal).¹³² Ferrylodge describes how, upon request of

¹²⁸ Wikipedia, *Requests for Arbitration: Ferrylodge: Involved Parties* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge#Involved_parties, last visited Apr 5, 2009).

¹²⁹ Wikipedia, *Ferrylodge's User Page* (<http://en.wikipedia.org/wiki/User:Ferrylodge>, last visited Apr 5, 2009).

¹³⁰ Wikipedia, *FeloniousMonk's User Page* (<http://en.wikipedia.org/wiki/User:FeloniousMonk>, last visited Apr 5, 2009).

¹³¹ Wikipedia, *KillerChihuahua's User Page* (<http://en.wikipedia.org/wiki/User:KillerChihuahua>, last visited Apr 5, 2009).

¹³² See Ferrylodge's block log (<http://en.wikipedia.org/w/index.php?title=Special:Log&type=block&page=User:Ferrylodge>, last visited Apr 27, 2009).

KillerChihuahua, FeloniousMonk banned him from Wikipedia on September 21st 2007 (a month before the case started). Ferrylodge starts with his own personal narrative:

"I have never been to [the Arbitration Committee]. Also, no one has ever brought an [Request for Comment] on me. I started editing Wikipedia in 2004, and became more active in 2006. I accumulated 5523 edits to Wikipedia articles, and edited 540 pages."¹³³

He then moves to describe the uncivil behavior of KillerChihuahua towards him, following which he asked her to stop posting at his talk page at all. After his request, she filed a proposal to ban him from Wikipedia on the Community Sanction Noticeboard.¹³⁴ 24 hours later, FeloniousMonk banned him from Wikipedia for 'harassing' KillerChihuahua. Following Ferrylodge's opening statement, KillerChihuaua added her own opening statement, which focuses on describing Ferrylodge's disruptive conduct.

A second layer of the case is to be found on the Evidence page.¹³⁵ Anyone, whether directly involved or not, can add evidence to this page. In the Ferrylodge case, evidence was

¹³³ Wikipedia, *Request for Arbitration: Ferrylodge* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge#Statement_by_Ferrylodge, last visited Apr 15, 2009).

¹³⁴ See the edit on the Community Sanctions Noticeboard (<http://tinyurl.com/c2et4k>, last visited Apr 27, 2009).

¹³⁵ See Wikipedia, *Requests for Arbitration: Ferrylodge: Evidence* (http://en.wikipedia.org/wiki/Wikipedia:requests_for_arbitration/Ferrylodge/Evidence., last visited Apr 3, 2009).

presented by sixteen different users.¹³⁶ Most of the evidence relates to Ferrylodge and KillerChihuahua's behavior and editing patterns, as well as to process that took place on the Community Sanction Noticeboard. An example for the kind of evidence put forward is the user B's detailed analysis of the editing patterns of the users Odd nature and FeloniousMonk, aiming to prove that Odd nature was a 'sock puppet' of FeloniousMonk.¹³⁷ Ferrylodge presented evidence meant to prove that he did not have any chance to answer the accusations made against him at the Community Sanction Noticeboard and therefore he was banned without 'due process'.¹³⁸ The user Sbowers3 presented a detailed analysis of 350 of Ferrylodge's edits to show that he was a productive editor and therefore should not be banned.¹³⁹ Another line of submitting evidence was taken by the user Pleasantville, who asked to submit confidential evidence straight to the arbitrators' mailing list, without it being transparent to Ferrylodge's.¹⁴⁰ This request created some concern regarding the fairness of the process, but was nevertheless approved.

A third layer was added on the Workshop page. On this page, the Arbitrators, the parties, and other editors, may draft proposals for the ruling and post them on the page for

¹³⁶ *Id.*

¹³⁷ *Id.*

¹³⁸ *Id.*

¹³⁹ *Id.*

¹⁴⁰ *Id.*

review and comments, as well as miscellaneous information.¹⁴¹ The Workshop page includes motions and requests by the parties, questions to the parties, and analysis of the evidence; as well as proposed principles, findings of fact, remedies and enforcement mechanisms for the Arbitration's committee rulings.¹⁴² For example, among the suggested principles in the Ferrylodge case were 'polls are evil' and 'bans are last resort'.¹⁴³

After considering the Evidence page and the Workshop page, the Arbitrators place their proposals for voting on the Proposed Decision page, and later vote.¹⁴⁴ Only Arbitrators or clerks are allowed to edit this page, although others may comment on the talk page. The Proposed Decision page includes the proposed Wikipedia principles on which the decision should be based, the proposed findings of fact, the proposed remedies and the proposed enforcement mechanism. It should be noted that unlike the other pages previously described, which are filled with lengthy discussions, the Proposed Decision page, like the main case page, reveals no discussion or reasons for the votes of the Arbitrators, leaving them unknown and poorly reasoned. All it includes are selected propositions followed by the votes. The talk page of the Proposed Decision page also lacks any reasoning for the decision and is limited to various discussions by users.

¹⁴¹ See Wikipedia, *Requests for Arbitration: Ferrylodge: Workshop* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge/Workshop), last visited Apr 8, 2009).

¹⁴² *Id.*

¹⁴³ *Id.*

¹⁴⁴ Wikipedia, *Requests for Arbitration: Ferrylodge: Proposed Decision* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge/Proposed_decision), last visited Apr 9, 2009).

In the Ferrylodge case, four principles were discussed: 'community bans', 'right of appeal', 'Wikipedia is not a soapbox', and 'the issue of administrators'.¹⁴⁵ As for community bans, the Committee affirmed the authority of the community to ban users at its discretion, yet noted that "[w]here such bans affect established contributors, a period of discussion and consensus-building should precede the ban (or, in emergency situations, follow it)". The Committee added that any user considered to be banned by the community may appeal her ban to the Committee, which may overturn any ban that has improperly imposed.

The arbitration committee found, as a matter of fact, that Ferrylodge had a long history of disruptive editing on topics regarding pregnancy and abortion, but had edited reasonably on unrelated topics. It further found that his ban was based on a very brief discussion from which a consensus was not clear, so the community ban was invalid and Ferrylodge could be unblocked by any administrator at the conclusion of the case.¹⁴⁶ Nevertheless, The Committee decided that Ferrylodge should be subject to an indefinite editing restriction on any article which relates to pregnancy or abortion.¹⁴⁷ At the end of its decision, the Committee "[urged] the community to develop a coherent policy regarding the method by which community bans are to be imposed."¹⁴⁸

¹⁴⁵ *Id.*

¹⁴⁶ See Wikipedia, *Requests for Arbitration: Ferrylodge* (http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Ferrylodge, last visited Apr 27, 2009).

¹⁴⁷ *Id.*

¹⁴⁸ *Id.*

Following the Arbitration Committee's ruling, several other "Requests for clarification" were posted on the case's talk page during December 2007, February 2008 and April 2008, regarding Ferrylodge's involvement in several articles which did not make it clear whether the ban is effective.¹⁴⁹ These requests led to more long discussions, which ended with a decision of the Committee that the term "article" in its ruling should be constructed narrowly as referring to the main article alone, without talk pages, wiki projects, templates and such.¹⁵⁰

Acoustic separations of Images

When we limit ourselves to the main case pages alone, the case seems to be another random instance of user misconduct. But reading the talk pages of the case and unveiling the acoustic separation they hold, we can acknowledge the case's true relationship to the end of the Community Sanctions Noticeboard. The administrator B argued that the case is actually about the practices of the noticeboard, which are unacceptable:¹⁵¹

¹⁴⁹ *Id.*

¹⁵⁰ *Id.* See also Wikipedia, *Request for Arbitration: Ferrylodge: Talk Page* (http://en.wikipedia.org/wiki/Wikipedia_talk:Requests_for_arbitration/Ferrylodge, last visited Apr 20, 2009).

¹⁵¹ See Wikipedia, *B's User Page* (<http://en.wikipedia.org/wiki/User:B>), last visited Apr 5, 2009).

"I am pleased to see that Ferrylodge is appealing this ban. This is an appalling example of the lynch mob mentality of the community sanction board. The initial ban proposal was made at 23:32, 20 September 2007. The block was made at 18:23, 21 September 2007. That's less than 24 hours later and horribly inappropriate. [...] If I ask the question, 'should user X be banned' and ten people who don't like user X show up and answer yes, that doesn't mean that the community has reached a consensus – it only means that the people who don't like user X did a good job of showing up quickly. That's not enough to infer anything. That's why our court system has a jury make decisions rather than trying cases at a town hall meeting. A community-discussed ban shouldn't be overturned without good reason, but that safeguard needs to be there, otherwise [...] bans are determined by the tyranny of the heckler. [...] Ferrylodge has been blocked for an unrelated [three revert rule infringement] when the ban proposal was made and was unable to respond until well into the lynching [...] [T]he fact that there was no opportunity for rebuttal and meaningful discussion of KillerChihuahua's assertions before the final course of action was *de facto* decided makes it difficult to have respect for this process. The person who actually imposed the block, FeloniousMonk, is himself a participant in the discussion. To sum it all up, the process was horribly bad. I strongly encourage [the Committee] to reverse this ban."¹⁵²

¹⁵² See Wikipedia, *Requests for Arbitration: Ferrylodge, Talk page, supra* note 150. See specifically the statements by Ali'i, CitiCat, KWSN, and Nick.

On the other hand, pro-community users, such as Andrew C, argued that B's comment tries to use the case in order to influence the future of the noticeboard, and therefore the Arbitration Committee should have no authority at all to review the process held at the Community Sanctions Noticeboard:

"This shouldn't be about the [Community Sanctions Noticeboard] process. It may be broken, but the community is taking that under consideration at a deletion discussion. I do not believe anyone here is asking the [Arbitration Committee] to review the [Community Sanctions Noticeboard], so I ask that the [Arbitration Committee] take that into consideration."¹⁵³

Following these comments, the talk pages of the Ferrylodge case became a central sphere of discussion of the Community Sanctions Noticeboard. At the same time, the nomination for deletion discussion continuously referred to the Ferrylodge case. The two spheres became intertwined – in a way nearly invisible to users who did not directly engage in the discussion. None of it was transparent to the readers of the case's main pages. Following the nomination for deletion discussion, the Community Sanctions noticeboard was

¹⁵³ *Id.*

closed, and the ban discussions were returned to the Administrators Noticeboard for Incidents.¹⁵⁴

This story illustrates several of the mechanisms for acoustic separation, as well as their discursive functions. First, the decisions of the Arbitration Committee are structured in a fragmented way, in order to make it difficult for outsiders to understand the full meaning of its rulings. It thus creates an acoustic separation not only between two sets of rules, but between two images of Wikipedia. The common user, reading the case alone, sees the discussion as a neutral dispute resolution system for dealing with hard cases. Only the deeper reading of the case allows the reader to understand it in the context of the tension between the community and the administration.

Moreover, the story of the Community Sanctions Noticeboard and its failure is a story of failed opposition against the mechanisms of acoustic separation. The noticeboard made visible the practices of the community bans, as opposed to the secret practices of blocking, thus attempted to reduce the prominent power of administrators to place de-facto bans. But the appearance of the discussion created a backlash and a reference to the Community Sanctions Noticeboard as a place for a 'lynch mob'. Therefore, it was shut down, allowing the vast majority of sanctions (bans as well as blocks) to be acted out non-publicly by administrators. This is, therefore, a failed attempt to make the use of sanctions transparent and to reduce the acoustic separation between conduct rules and decision rules. This failed

¹⁵⁴ See *Wikipedia: Miscellany for Deletion: Community Sanction Noticeboard (Second Nomination)*, [http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard_\(second_nomination\)](http://en.wikipedia.org/wiki/Wikipedia:Miscellany_for_deletion/Wikipedia:Community_sanction_noticeboard_(second_nomination)), last visited Apr 27, 2009).

attempt, and the choice to conceal the procedures used by administrators, strengthens the centrality of the acoustic separation in Wikipedia's structure.

VII. Challenging the acoustic separation? - Flagged revisions and the politicization of code

From the discussion so far, the correlation between the social use of acoustic separation and the technological infrastructure of the Wiki software is apparent. Similar to other internal barriers in cyberspace, which regulate human interactions,¹⁵⁵ the mechanisms of acoustic separation utilize the 'digital divide' among different groups of users by placing the information targeted at new users on the main article pages and edit page, while placing contrary information on the talk pages, in relatively unknown noticeboards, inside pages of the Arbitration Committee cases or even off-wiki in alternative platforms.

This correlation between the technological infrastructure and the social practices of separation might seem to suggest one of two simplistic mechanisms of the connection between society and technology: either that it is the technological infrastructure that largely determines the social and discursive organization, or that technology is a conscious and intentional product put in place to support a detailed plan of social engineering to improve collective performance.¹⁵⁶ Yet, I will argue that Wikipedia manages to escape both understandings by its informed recognition of the social importance of its technological infrastructure and a 'politicization of code'.

¹⁵⁵ MARK STEFIK, *THE INTERNET EDGE*,: SOCIAL, TECHNICAL, AND LEGAL CHALLENGES FOR A NETWORKED World 14-15 (2000).

¹⁵⁶ See Butler, Joyce and Pike, *supra* note 8, at p. 3.

Lawrence Lessig noted several years ago, in his book *CODE*, that cyberspace demands a new understanding of how regulation works, one which looks both at norms and at the technological infrastructure.¹⁵⁷ Lessig further shows that 'architecture', 'law' and 'social norms' are not separate kinds of regulation, but that they are interrelated as the architecture influences the way the law works, and the legal rules attribute to certain social norms. Therefore, he argues, following William Mitchell, code should become a crucial focus of political contest,¹⁵⁸ and code programmers should be controlled by the political involvement of the community:¹⁵⁹ 'how the code regulates, who the code writers are, and who controls the code writers – these are the questions on which any practice of justice must focus in the age of cyberspace'.¹⁶⁰ The continuing discussion and conflict regarding the structure of technology is also the main characteristic of a 'recursive public' in the meaning given to this concept by Chris Kelty.¹⁶¹ Such a public makes, maintains, and modifies its technical and legal layers. Simultaneously, its notion of community and social constructions are created and reinforced by the same technological and legal layer.¹⁶²

¹⁵⁷ See Lessig, *supra* note 8, at p. 5.

¹⁵⁸ WILLIAM MITCHELL, *CITY OF BITS: SPACE, PLACE, AND THE INFOBAHN*, 112 (1996).

¹⁵⁹ Lessig, *supra* note 8, at pp. 78-79.

¹⁶⁰ *Id.*

¹⁶¹ CHRIS KELTY, *TWO BITS: THE CULTURAL SIGNIFICANCE OF FREE SOFTWARE* (2008).

¹⁶² *Id.*, at p. 29.

Flagged Versions

As a forward, it's important to note that even today the outsiders of Wikipedia – the visitors who only read the main article pages - are exposed to almost completely undisturbed and accurate pages, which portray Wikipedia as a reliable encyclopedia, what contributes to its outstanding popularity.¹⁶³ Most cases of vandalism, for example, are not transparent to outside users. Malicious edits are usually corrected in a matter of minutes by members of the community. The median time to revert a 'mass deletion' of a page is 2.9 minutes, and an 'obscene' mass deletion is reverted in a median of 2 minutes.¹⁶⁴ This efficiency is achieved by a combination of social norms - the first thing a Wikipedia editor should do when he spots vandalism is to revert the vandalized page – and technical tools: both the Recent Changes page and software applications that were specifically created for dealing with vandalism (e.g. Huggle, Twinkle or Vandal Fighter).¹⁶⁵ Some techniques used by these applications include highlighting edits from anonymous editors on the Recent Changes list (based on the fact that 97% of all cases of vandalism are created by anonymous IPs)¹⁶⁶ or identifying suspicious chunks of text ('poop' and the like). A unique Wiki project, the Counter-Vandalism Unit

¹⁶³ See Viegas, Wattenberg & Dave, *supra* note 8.

¹⁶⁴ See Viegas et al., *supra* note 31.

¹⁶⁵ See Wikipedia, *Counter-Vandalism Tools Category* (http://en.wikipedia.org/wiki/Category:Wikipedia_counter-vandalism_tools), last visited Apr 2, 2009).

¹⁶⁶ See Wikipedia, *Vandalism Studies WikiProject* (http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Vandalism_studies), last visited Apr 27, 2009).

(CVU),¹⁶⁷ specializes in producing tools to assist in removing vandalism and gathering data on the best ways to deal with vandals.¹⁶⁸ The result is that the work of vandals is virtually untraceable for the general public.¹⁶⁹

Nevertheless, a more extreme attempt to present a fault-proof page to outsiders is currently being discussed on the English Wikipedia, the 'flagged revisions' feature, which is currently being used by the German, Polish and Russian Wikipedias.¹⁷⁰ This feature shows the general public – visitors who are not logged on - a 'validated' version of an article and not its last version, in order to make it more credible and 'presentable' to the wider public.¹⁷¹ The approved versions are known as 'sighted versions' or 'flagged versions', and are checked to be free from spam, libel, obvious un-encyclopedic content, or spelling errors and to contain references to reliable sources.¹⁷² The main argument for the proposal (which should not be a surprise by now), is functional: to provide a better encyclopedia to the general public. As user Abd noted:

¹⁶⁷ Wikipedia, *Counter-Vandalism Unit WikiProject* (<http://en.wikipedia.org/wiki/Wikipedia:CVU>., last visited Apr 27, 2009).

¹⁶⁸ *Id.* See also MetaWiki, *Anti-Vandalism Ideas* (http://meta.wikimedia.org/wiki/Anti-vandalism_ideas., last visited Apr 27, 2009).

¹⁶⁹ See Lih, *supra* note 8, at p. 177.

¹⁷⁰ See *id.*, at p. 148; Wikipedia, *Proposed Wikipedia Policy: Flagged Revisions/Sighted Versions* (http://en.wikipedia.org/wiki/Flagged_revisions/Sighted_versions., last visited Apr 27, 2009).

¹⁷¹ *Id.*

¹⁷² *Id.*

"Wikipedia is and will remain 'the encyclopedia' that anyone can edit. However, this feature creates a layering, a layering that has for a long time been necessary but not achievable without this feature. Without flagged revisions [...] there must be *constant* defense not only against vandalism, but also against the insertion of unreliable information or the removal of text accepted by consensus. Immediate editing by anyone is in conflict with reliability, unless we could guarantee that all edits are immediately reviewed. [...] So how can we have immediate editing *and* reliability? It's simple: immediate editing is confined to a layer below the top-level default layer. Anyone may *propose* an edit. [...] If we have a publisher which automatically, without review, publishes what someone writes, we'd consider that tantamount to 'self-publishing'. We have been that publisher, which is one reason why we are so attractive to link spammers and POV-pushers and vandals. With a bit more sophistication in the structure, we could actually become a truly reliable source. [...] Wikipedia has guidelines and policies that require reliable sources, but the enforcement of this and the unevenness of our consensus process for applying and interpreting this has made reliability an unrealized ideal.¹⁷³

¹⁷³ See Wikipedia, *Flagged Revisions Proposed Policy Talk Page* (http://en.wikipedia.org/wiki/Wikipedia_talk:Flagged_revisions#Why_flagged_revisions_is_essential to fixing the problems of scale and reliability., last visited Apr 27, 2009).

The use of flagged versions will create an extreme kind of content-based 'acoustic separation' between non-users and users, by presenting two separate encyclopedias to these two groups. The opposition to the change stems mostly from the fact that new editors' contributions will not be visible, and thus their motivation to contribute will decrease.¹⁷⁴ An example for this argument is the user Computer's comment:

"[T]his is the most important change to Wikipedia since its birth, and is going to REVOLUTIONISE Wikipedia, a (if not *the*) powerhouse of the *edit-this-page* direct-user-contribution culture – we are facing a fundamental change to how users' contributions will be treated, according to the **user status** instead of the **quality** or **reasoning** of the contribution. Before, the software [...] either accepts the contribution of a user [...], or do not accept at all (protected pages). This gives an impression of openness, transparency and simplicity. Users feel **Trusted** and in control. Their contribution will only be removed [...] if their fellow equal Wikipedians **reasonably** challenge its quality. [...] This extension changes all these by treating all users as **untrustworthy** by default, and hides their initial contributions from public (non-Wikipedian) view [...]. Before the

¹⁷⁴ Id. See also BBC News, *Editorial Row Engulfs Wikipedia*, (BBC News, Jan 26, 2009), <http://news.bbc.co.uk/1/hi/technology/7851400.stm>, last visited Apr 27, 2009).

contribution gets 'sighted', the user feels excluded, uncertain [...] and helpless [...]. Freedom or quality? Time to choose."¹⁷⁵

Discussions about the use of flagged revisions have been going on the English Wikipedia for several years now, and it was a major issue at Wikimania 2006 conference.¹⁷⁶ But the discussion did not advance to a practical level until January 21st 2009, when Jimmy Wales announced on his talk page a partial activation of the flagged revisions tool, in a comment titled 'why I am asking Flagged Revisions to be turned on now':

[a Washington Post article about an edit on Wikipedia announcing Edward Kennedy's death] would have been 100% prevented by Flagged Revisions. [...] We have a tool available now that is (a) consistent with higher quality (b) will allow us to allow more people to edit it a wider range of circumstances and (c) will prevent certain [harms due to the edits of biographies of living persons]. We now have a community poll indicating approximately a 60/40 support for the future. This is a very wide margin, with 20% separation between the pro's and con's. [...] The proposal is for a time-limited test. To the Wikimedia Foundation: per the poll of the

¹⁷⁵ Wikipedia, *supra* note 173.

¹⁷⁶ See Wikipedia, *Pushing to Validation* (http://en.wikipedia.org/wiki/Wikipedia:Pushing_to_validation), last visited Apr 27, 2009).

English Wikipedia community and upon my personal recommendation, please turn on the flagged revisions feature as approved in the poll.¹⁷⁷

Not surprisingly, three minutes later, the first user, DS, opposed Wales's announcement.¹⁷⁸ Others have joined his opposition and noted that a 60% majority does not qualify as a clear consensus.¹⁷⁹ Others have supported the announcement, and the discussion went on and on, for hundreds and hundreds of comments (which are added to the ongoing discussion of the matter in other forums as well, such as the proposal page itself). Currently – based on an update by Jimmy Wales from March 28th, 2009, the subject is awaiting a discussion on the Wikimedia board meeting.¹⁸⁰ In the meeting the board will also regard a poll conducted on April 1st, 2009, which suggested a two-month trial of a limited version of flagged revisions, called 'flagged protection', referring to using flagged revisions only in regard to specific sensitive articles and not across the board.¹⁸¹ In the poll, 259 users

¹⁷⁷Jimbo Wales, *Why I Am Asking Flagged Revisions to Be Turned on Now*, *Jimbo Wales' User Talk Page* (http://en.wikipedia.org/wiki/User_talk:Jimbo_Wales/Archive_44#Why_I_am_asking_Flagged_Revisions_to_be_turned_on_now, last visited Apr 27, 2009).

¹⁷⁸ *Id.*

¹⁷⁹ *Id.*

¹⁸⁰ Jimbo Wales, *Update on BLP / Flagged Protection / Flagged Revs*, *Jimbo Wales User Talk Page* (http://en.wikipedia.org/wiki/User_talk:Jimbo_Wales/Archive_46#Update_on_BLP_2F_Flagged_Protection_2F_Flagged_Revs, last visited Apr 27, 2009).

¹⁸¹ Wikipedia, *Flagged Protection* (<http://en.wikipedia.org/wiki/Wikipedia:FLP>, last visited Apr 27, 2009).

supported the proposal, 61 users opposed, and 4 remained neutral (80% support, 19% oppose, 1% neutral).¹⁸²

Looking at the poll results and the centrality of Jimmy Wales's announcement in the decision to promote 'flagged revisions', it is easy to see the mechanisms of deliberation that led to the decision as meaningless, and to regard it merely as a legitimating cover for an authoritative act.¹⁸³ But such understanding will overlook the long and engaging discussion regarding this change that has been going on for several years. The ongoing discussion regarding the proposal to use 'flagged versions' shows that the social and political implications of changes to the Wiki software are discussed among the community and their importance is acknowledged. A similar debate has arisen, as I have showed, in regard to the use of separate noticeboards as place for different social practices.

This discussion suggests a new way to understand the place of the technological infrastructure in the social construction of Wikipedia, one which undermines both deterministic views I described. It combines the technological infrastructure, the social patterns, and the semiotics of discourse - and the combination of the three escapes determinism as to all of them. Therefore, it always holds the possibility of their future transformation.

¹⁸² Wikipedia, *Flagged Protection and Patrolled Revisions Poll* (http://en.wikipedia.org/wiki/Wikipedia_talk:Flagged_protection_and_patrolled_revisions/Poll), last visited Apr 27, 2009).

¹⁸³ Compare with Aaron Shaw, "*Calling Bullsh*t on the Facebook Governance Vote*", AARON SHAW'S WEBLOG, (April 24, 2009, <http://fringethoughts.wordpress.com/2009/04/24/calling-bullsht-on-facebook-governance/>), last visited Apr 27, 2009).

VII. Conclusion: beyond bug and feature

The obvious way to think about the internal barriers of Wikipedia, which separate, as I showed, the newbies from the insiders, is through a normative or functional lens. Such means of analysis would ask whether these barriers are a 'bug' – a mechanism for secrecy that undermines the legitimate use of sanctions, or a 'feature' – a functional mechanism that maximizes the attractiveness of Wikipedia to as many people as possible. This paper aimed to show that these barriers hold both. The concealing of sanctions serves both as a way to hide from the general public of users the ways in which the power is actually practiced, and at the same time it limits the use of sanctions to specific professional elites and prohibits 'lynch mobs'.

Moreover, I tried to show that a normative analysis could also hide the fact that this exact duality lies at the bottom of the discursive environment of these practices and navigates their different uses in different spaces in Wikipedia. Only by transcending the normative, it is possible to leave the discussion to the internal politics of the Wikipedia community itself, thus leaving it open to constant change, modification, and innovation.