Trust, Control, and Formalization in Open-Collaboration Communities: A Qualitative Study of Wikipedia

Dariusz Jemielniak
Kozminski University

Paper presented at the Academy of Management conference in Boston, MA (USA) 5-7 August 2012

(work in progress)

Abstract

Trust, control and organizational routines are often perceived as key concerns of organization studies.

Theoretically, this study presents an investigation into the enactment of interpersonal and institutional trust, credentials and identity in computer supported cooperative work. It adds to the ongoing debate on participative designs, online communities and open collaboration by demonstrating via a study that trust in people can be substituted for by not only increased control, but also by trust in procedures and legalistic remedies, especially in organizational designs where face to face interactions are limited or nonexistent. This paper relies on a longitudinal, participative, netnographic study of Wikipedia community.

Keywords

control in virtual communities; credentials in CSCW; open collaboration trust enactment; participative designs; Wikipedia
Introduction

Computer supported cooperative work is a dynamically growing field of organization studies. In the eve of distributed work and virtual team systems, it becomes one of the most promising, and yet understudied phenomena of this field (Gibson and Cohen 2003, Hackman 2011). For obvious technological reasons and in lack of regular face-to-face interactions, identity enactment, authority claims, the creation of trust and the way control of credentials is exerted differ significantly from traditional organizational environments (Oxley et al. 2010; Green et al. 2011).

Wikipedia, with the current number of registered accounts exceeding 30 million in all languages, and over 85 thousand editors making more than 5 edits every month, as well as with the current number of existing wiki pages exceeding 27 million, is beyond any doubt the largest computer supported cooperative project run by non-professionals in the history of humankind.

Yet, so far, there have been no studies of the application of trust and credentials in computer supported cooperative work on the example of Wikipedia, in particular conducted through qualitative, long-term research. Therefore, the purpose of this paper is to present the results of an ethnographic investigation of the enactment of interpersonal and institutional trust, credentials and identity in Wikipedia community. The paper continues the ongoing debate on participative designs, online communities and open collaboration by demonstrating via a study that trust in people can be substituted for by not only increased credentials control, but also by trust in procedures and legalistic remedies, especially in organizational designs where face to face interactions are limited or nonexistent.
Trust in organizations

Trust has been a common topic of organization studies both on the inter-organizational level (Hoffmann et al. 2010; Newell et al. 2000) and the intra-organizational level (McEvily 2011; Gambetta 1988; Thau et al. 2007). The profound influence of new technologies on trust relations has also been a topic for considerable academic speculation and investigation (Latusek et al. 2010; Latusek et al. 2007).

Interpersonal trust can be simply defined as a confident and positive expectation of the actions of another person (Dirks et al. 2004). On interpersonal level, it relies on a suspension of uncertainty about the other person’s behavior and can be defined as “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party” (Mayer et al. 1995, p. 712). On organizational level, trust has also been traditionally depicted as a social tool to use to reduce risk and uncertainty by relying on values and cooperative norms, and thus being often symmetrical in nature. "Trust is an important lubricant of any social system. It is extremely efficient; it saves people a lot of trouble to have a fair degree of reliance on other people's word" (Arrow 1974, p. 23). In contrast, the typical substitute of trust, namely control, is described as usually based on an asymmetrical relation and including the element of power. From a functionalist perspective, both trust and control serve the same purpose, i.e., reproducing and reinstating interactional order and also reducing transactional costs (Lane et al. 2000).

This dichotomy of trust and control has been criticized as presenting an overly polarized view of organizational reality (Ezzamel et al. 1992). For instance, trust has been portrayed as the opposite of distrust, rather than the opposite of control (Latusek et al. 2007; Vlaar et al. 2007). The dual, and not dichotomous, character of trust and control relation has been proposed
(Möllering 2005; Sitkin et al. 1993). Similarly, Reed (2001) makes the point that the interplay of trust and control is in fact a unique way of expert power structuring, and indeed, both are complementary rather than exclusive. Their dual and at least potentially interchangeable nature is thus generally accepted (Das and Teng 1998; Inkpen and Currall 2004; Khodyakov 2007; Costa and Bijlsma-Frankema 2007; Meyer and Rowan 1977).

The issue of trust is particularly important in the new knowledge society and in its new virtual environments, which present new challenges both for trust and control (Knights et al. 2001). These questions are particularly valid and of concern in open collaboration community environments, where participation is voluntary, incentives or punishments are limited, membership dispersed, and identities only virtual.

**Trust and open-source communities**

The development and social production of knowledge requires cooperative relationships within organizations. However, these relationships are hampered by the dominant assumption that people are mainly motivated by self-interest (Nahapiet et al. 2005) and may not necessarily engage in activities that bring them no material gain/reward over the long term. The identified modes of collaboration, typical for open collaboration communities, prove this assumption to be false, however, as many people are willing to cooperate and contribute without material benefit and are actually able to successfully develop products and ideas that compete with their commercial counterparts on equal grounds (Benkler 2006). There are many examples of such successful endeavors. One is Linux, the most popular server operating system, which was developed and distributed as the model of free and open source software by software engineers all around the world. The other example, arguably even more interesting for organization studies, is Wikipedia, the largest online encyclopedia also
developed as an open-source “copyleft” model. However, contrary to Linux, Wikipedia was and is created mostly by non-professionals. As Benkler points out (2011), its design is atypical for most other organizations in the sense that Wikipedia not only does not rely on incentive compensation or punishments, but also does not have hierarchical control in the traditional sense and still works perfectly well.

While there have been many studies on the trustworthiness of Wikipedia articles (Dondio et al. 2006; Kwan et al. 2009), the mechanism of trust and credential creation within the Wikipedia editor community has rarely been a subject of serious academic inquiry (Krupa et al. 2009), in particular through any in-depth participative and qualitative study. Other studies do indicate, however, that online communities with no offline interaction among their members (which, in most cases, does apply to Wikipedians) enact trust differently than groups with actual face-to-face contacts (Matzat 2010). When this mentioned lack of traditional credentials control is taken into account, Wikipedia becomes a particularly interesting topic for the study of trust enactment. Also, Wikipedia is a model example of an organization that relies on empowerment, a flat structure, and decentralized authority; in such an environment, trust plays a crucial role (Jones et al. 1998; De Jong et al. 2010).

Johnson argues that in our contemporary society, credentials have become more important in the enactment of trust, than have personal identity and relationship status. She sees the lack of confirmed credentials in the online environment as a major obstacle to regular and effective social activities (Johnson 1997). Similarly, according to other authors, the lack of transparency in terms of revealing Wikipedia contributor identity also questions the actual validity of the outcome, i.e. the contents of the articles (Waters 2007). Santana and Wood, when criticizing Wikipedia, even make the melodramatic point that (2009, p. 143) “the anonymous production and use of information prevents human users from achieving the
deepest possible meanings in life, and violates as well the ethical principle of integrity of information. In addition, anonymous providers need not exercise moral responsibility for there is no accountability.” I believe this statement is not only obviously flawed (considering both the popularity and proven quality of Wikipedia), but also testifies to the authors’ ignorance of the Wikipedia community and the nature of cooperative work. Yet, the authors do have a point, namely, that quasi-anonymous online communities do need to address the problem of accountability in their own way, which so far in the case of Wikipedia at least, has not been a topic for major academic discussion in the context of open collaboration, non-expert communities. Thus, it is vital for the studies of computer supported cooperative work to examine and understand and perhaps define the mechanisms of trust and control in the new world environment that is now often deprived of typical incentives and punishments and hierarchies, but still does allow for much more fluid, even obscured, identities. This issue is particularly important, because open collaboration projects are growing in popularity and becoming a great part of mainstream organizational design (Neus et al. 2005).

In Internet environments especially, strong assurance control systems are often portrayed as a solution to the nagging problem of trust (Cheshire 2011): Identity and credential verification provide a threshold/gate check, as it were, so as to allow for less control later. Interestingly, as shown in this paper, this scenario is not the case for Wikipedia. Even though its community has undergone some trust crises, it decided against introducing closer entry control systems. It has seemed however, to have decided to substitute for the typical trust-control functions by adding bureaucratic scripts. In other words, the community discarded the idea of increasing credential checks, as well as trust-building mechanisms for the users, in favor of trusting organizational routines and processes instead (in particular, in the case of Wikipedia, these include behavioral guidelines, essays, policies, as well as the technologies and norms of dispute resolution). Apparently, a crisis of interpersonal trust was resolved by means of strong
institutional trust and organizational para-legalistic remedies, even though in traditional, non-virtual organizations these would actually often increase and deepen interpersonal distrust (Sitkin et al. 1993).

Thus, the purpose of this paper is to analyze the community of Wikipedia and its approach to trust and credentials control through a precise analysis of one of the largest crises on both trust and control that the community experienced (the so-called “Essjay controversy”, in which one of the most distinguished editors on Wikipedia was found to be a liar about his credentials), and the internal debates within the community that followed that crisis.

**Research methods**

This paper is the result of a five-year qualitative study on Wikipedia. It started in 2006 and is still continuing with this paper the first written report on that study. The research is ethnographic (Hammersley et al. 1995), and benefits heavily from active participation of the researcher in the community, even though it is not confessional in terms of its narrative style (Van Maanen 1988; Schultze 2000). Such ethnographic studies of computer supported cooperative work have been growing in recognition over the last 15 years (Forsythe 1999; Baba 2012), and are particularly useful for the analysis of less structured and categorized problems (Denzin et al. 2007). The study is performative rather than ostensive (Latour 1986), in the sense that it aims at helping the reader understand the examined community’s perceptions and that community’s culture better, rather than at applying a preconceptualized theoretical model onto the existing culture.
Since the beginning of the project, there has hardly been a day when I\(^1\) did not log into Wikipedia, often spending an hour online, editing, discussing Wikipedia, and as chatting with other Wikipedians on IRC (more recently, also on Facebook). I have participated in several Wikipedia and Wikimedia projects, totaling a five-digit total edit count. On one of the projects, I was elected an administrator and later also a “bureaucrat” (a user with a technical ability to grant administrative rights and change usernames, among other aspects). I also became a member of the ombudsmen commission for all Wikimedia projects (a body responsible for evaluating privacy policy violations, and overseeing “checkuser” conduct) for one term, and serving as a mediator. Currently, I am also one of some 40 elected “stewards” for all Wikimedia projects (this role gives highest access and responsibilities on all Wikipedias).

The research methods used for the project have been participant observation and case study (Flyvbjerg 2001; Denzin et al. 2003). It should be noted, however, that participant observation in the case of netnographic research (Kozinets 2010) combines observation and discourse analysis (Grant et al. 1998), since virtually all interactions and behaviors are textual, and relies heavily on an analysis of narratives (Jemielniak et al. 2010). English Wikipedia is quite possibly the largest culture in the world based virtually on textual asynchronous interactions only.

The presented study belongs to the field of organizational ethnography (Schwartzman 1993; Kostera 2007), and even though it has been, to some extent, grounded-theory inspired (Glaser et al. 1967), especially in terms of analyzing the gathered material into useful categories, the research relies on anthropological reflection (Czarniawska-Joerges 1992; Clifford et al. 1986) more than it does on coded categorizations. For the purpose of this paper only, many relevant

\(^1\) First-person use as a choice herein is determined by the methodological considerations, see e.g. (Denzin et al. 2007; Golden-Biddle et al. 1997)
discussion-, talk-, and comment pages were carefully analyzed and interpreted (a total of roughly 500,000 words of field material)\(^2\).

This paper offers an analysis of a scandal from 2007 which shook Wikipedia’s community and was widely commented on in the media (suffice it to say that the exact phrase “Essjay controversy” in Google search still brings over 28,000 hits). The scandal resulted in several lengthy internal discussions and debates on editor trust and credibility. While the event is now long passed and the scandal quite gone, it is worthwhile to interpret the reaction and decisions of the community, since they reveal several social mechanisms and basic assumptions regarding Wikipedia. This case has also been mentioned in several academic publications (Bruns 2008; Lih 2009), but has not as yet been a subject of any detailed analysis, in particular from an organizational point of view, even though that it may shed new light on the enactment of trust and control in the online communities context.

**Essjay controversy**

A Wikipedia user called “Essjay” started to edit for Wikipedia in 2005. From the beginning, he claimed on his user page that he was a tenured Professor of Religion, with a Ph.D. in theology, a Doctorate in Canon law, and held other degrees in his field, and that he was regularly teaching theology to undergraduate and graduate students at a private university in the U.S. He was indeed quite active in the articles on religion, and in just five months, he earned enough respect from the community that he became an administrator (with a tremendous support of over 98% of voters). He then increased his engagement in Wikipedia

\(^2\) For the convenience of the reader, all references to Wikipedia rules are given in the format typical for internal Wikipedia links. Getting a full URL is very easy, since the syntax is stable. For example, whenever I refer to [[Wikipedia:Policies and Guidelines]], I reference the link http://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines
and was active in the community, mediating conflicts, fighting vandals, and doing well what any admin (as administrators are usually called on Wikipedia) should do. This success lead to his being elected a bureaucrat and a checkuser (a member of the Wikipedian elite who is trusted to check sensitive information about other users’ IP addresses), as well as being the Chair of the Mediation Committee.

He was such a dedicated and efficient editor in fact that as a paragon example of Wikipedia’s collaboration with the academic world, he was invited in July 2006 to participate in an interview for *The New Yorker* (Schiff 2006), via a recommendation from the Wikimedia Foundation. His experience on Wikipedia was wide enough to help him get hired by Wikia in January 2007 (a commercial company offering Wiki-based hosting solutions and created also by Jimmy Wales, the co-founder of Wikipedia). Soon after the scandal broke: Daniel Brandt, a social activist dedicated to exposing misconduct on Wikipedia (and harshly criticizing Wikipedia, sometimes using borderline conspiracy theories) discovered that Essjay was not who he said he was. Brandt wrote to *The New Yorker* and posted his suspicions on his own website.

As it turned out, upon achieving employment at Wikia, Essjay had published an online bio note on the company’s website in which he admitted to being a 24-year old former paralegal clerk and an account manager for a Fortune 20 company (a statement that many also doubted later). When some of the editors saw the discrepancy between this note and the one posted on Wikipedia, they started expressing their doubts about his identity. In early February, he admitted to have created a false persona, but for security reasons:

One of the things that tends to happen as you become, let us say, "popular" on Wikipedia is that you attract the attention of an unsavory element. There are a number of trolls, stalkers, and psychopaths who wander around Wikipedia and the other Wikimedia projects looking for people to harass, stalk, and otherwise ruin the lives of (several have been arrested over their activities here). (...) Many people have
tried many things to keep their identities secret: They worry over every little detail they may have released, or refuse to answer anything about themselves, making it very difficult to form any personal ties. Quite unfortunately, it simply isn't possible to keep your details quiet: You will eventually say something that will lead back to you, and the stalkers will find it. My approach was different: I decided to be myself, to never hide my personality, to always be who I am, but to utilize disinformation with regard to what I consider unimportant details: age, location, occupation, etc. (…) I was actually under the impression that the stalkers and psychopaths were the only people who actually believed the story; a quick examination of the time I've spent here should lead to the conclusion that there's no way I could be who the statistics said I was. (Essjay talk page, 06:07, 2 February 2007).

This explanation was viewed as sufficient for many editors. The fact that Daniel Brandt had had a record of attacking Wikipedia, as well as stalking and exposing the real identity of some admins, worked well in Essjay’s favor; he seemed to be an object of malicious attacks, and it appeared quite reasonable to have had his identity protected. Later Essjay confessed he was a regular, weekly subject of numerous death threats, torture monologues, and legal threats, which definitely won him some additional sympathy. In my roles on Wikipedia, I also was threatened a couple of times (even though I honestly don’t believe in the volume of death threats that Essjay reported). In my view, certainly many other Wikipedia activists did relate to and sympathize with the situation that Essjay described.

Possibly to show solidarity with this singled out Wikipedian, on 23 February 2007, Jimmy Wales nominated Essjay to the Arbitration Committee, a highly prestigious body that is responsible for resolving conflicts and disputes in the community and has high decision powers. A couple of days later, The New Yorker published an update to their earlier article, clarifying that Essjay had misled the interviewer about his credentials. The note also quoted Jimmy Wales who was said to disregard the whole situation as problematic and insist on Essjay’s right to use a pseudonym. This rendering immediately led to a discussion on Jimmy
Wales’ talk page, simultaneous to the one appearing on Essjay’s. Various editors now were accusing Essjay of serious misconduct and lying.

Still, many highly regarded Wikipedians continued to defend him. For example, one admin and also a bureaucrat wrote on Essjay’s talk page:

Just wanted to express my 100% support for everything you do around here. I think you were totally entitled to protect your identity. Don't let all the fuss get you down! (WjBscribe, 16:49, 1 March 2007)

The point of view of the defenders can be well summed up by the comment a different admin (and also a check user) made on Jimmy Wale’s talk page:

Since qualifications don't matter here, who cares? WP:V and WP:RS are required both from PhDs and junior high kids if they're editing articles. As far as personal integrity is concerned, in cyberspace, nobody knows you're a dog; people make up personae right and left around here (jpgordon 21:36, 28 February 2007)

The references to WP:V and WP:RS refer to “verifiability” and “reliable sources”, both fundamental principles that apply to all information introduced into Wikipedia articles. They determine precisely what kind of statements and references are allowed and are meant to eliminate editors’ original research and opinions. The user basically made the point that, since all information has to be verifiable and based on trustworthy materials, it does not matter what the usurped credentials of the editors are. He also pointed out that online identities are not particularly trustworthy anyway (“nobody knows you’re a dog”), and that creating virtual identities is a common practice on Wikipedia.

Still, after a while, the tide started to change. While Jimmy Wales was reported as accepting Essjay’s apology and considering the whole matter settled (Cohen 2007), he was still not able to fully participate in the debate, because he traveled in Asia and had limited Internet access (which also, quite likely, limited his ability to delve into the issue in detail).
One of the other administrators and bureaucrats succinctly declared on Jimmy Wales’ talk page what many others had expressed:

We enjoy our fantasy of an exclusively merit-based system, but it really is nothing more than a fantasy. If someone says he has a Ph.D., no amount of protesting the egalitarianism of the project will change the effect this claim has on other editors' opinions of him (or, in this case, the opinions of the New Yorker's readership, who are doubtless accustomed to put much stock by advanced degrees). He could easily have chosen a set of fake characteristics which did not carry such strong preconceptions if he wished to be anonymous. He has introduced a biasing factor -- I cannot say whether it was deliberate, but we cannot pretend it has no effect. (Dan, 21:45, 28 February 2007)

In the wake of a now public scandal, many editors started to look more closely into Essjay’s editing history and did not like what they found. For example, in one of his early edits, he insisted on using “Catholicism for Dummies” as a source by writing “This is a text I often require for my students, and I would hang my own Ph.D. on its credibility”. In another discussion, he referred to his personal experience as a monk to validate a point he was making about Psalms. He claimed he had been the head of his department and that he wrote a letter to a few other professors in which he explicitly represented himself as “a tenured professor of theology”. Based on these credentials, he was referred to as “one of Wikipedia's foremost experts on Catholicism,” a comment which he was happy to quote himself, when offering his advice in disputes. All in all, on at least several occasions, he used his faked credentials to add weight to specific disputes on Wikipedia.

Many Wikipedians decided to express their disappointment with Essjay’s conduct. Suffice it to say, the “Request for Comment” page dedicated to the whole case contained nearly 40,000 words of different opinions, statements, and views. While some were supporting Essjay’s right to create a fake identity, the vast majority were critical and occasionally even hostile (a fact also pointed out by some Wikipedians who were appealing for more civility).
Among the many disappointed editors, a member of the Wikimedia Foundation Board of Trustees urged Essjay (on Essjay’s talk page) to step down:

Creating a pseudonym is one thing. Creating an elaborate fake persona with fake credentials, and using it in arguments, letters, and interviews is another. I am deeply troubled by this behavior, consider it highly unethical, and would like to ask you to seriously consider stepping down from your official Wikimedia roles. At the very least, I believe you owe the community an apology for this behavior. You have damaged both the reputation of the project, and your own. I am deeply saddened and disappointed. (Eloquence, 22:13, 1 March 2007)

Finally, Jimmy Wales posted a statement (after archiving his talk page, so that the statement would be immediately visible at the top), and he explained:

I only learned this morning that EssJay used his false credentials in content disputes. I understood this to be primarily the matter of a pseudonymous identity (something very mild and completely understandable given the personal dangers possible on the Internet) and not a matter of violation of people’s trust. I want to make it perfectly clear that my past support of EssJay in this matter was fully based on a lack of knowledge about what has been going on. (Jimbo Wales 06:42, 3 March 2007)

He also asked Essjay to resign his positions on Wikipedia. On the very same day, Essjay decided to do just that, deleted his talk page, and left.

**What Did Essjay Do Wrong?**

While for the rest of the world, Essjay’s wrongdoing was one of creating fake credentials, it seems that for a huge majority of Wikipedians, such “wrongdoing” was not a problem at all. This view may be related to the fact that many people come to edit on Wikipedia to gain the status and authority they lack in regular life (Jemielniak et al. 2012). Since creating an alter ego is a widely accepted behavior in online communities (Boellstorff 2008; Bailenson et al.
2006) and since the Wikipedic “user page” is considered to be one’s own private turf and is less regulated, even though creating a totally fictional profile is not widely condoned, that activity is not frowned upon neither.

Also, privacy is a very serious concern. As mentioned before, many admins are subjects of different kinds of threats. I was cyber-stalked and bullied by a person who did not like the content of their biographic article on Wikipedia, to the extent that he (a martial arts instructor, with a history of violent behavior, as well as legendary persistence in suing everybody he considered to be his enemy, including Internet disputants) tried to find out where I live and also tried to contact my employer with his (luckily, delusional) accusations. After this incident, I removed all identifying data from my Wikipedia pages, but I know of at least several other similar situations that have also happened to administrators. Unsurprisingly, Internet trolls, as well as frustrated non-notable freaks with aspirations to garnering encyclopedic “fame” often end up interacting with the admins. Some people take it worse than others do.

In this light, Essjay’s first reply chimed in well with what the whole community is particularly sensitive to, namely, protecting your own identity, which is perceived as your fundamental right. Essjay’s initial explanation was perfectly legitimate. After all, he only changed several facts about his occupation, age, etc., to make it more difficult for stalkers to find him. Taking into account that his masquerade was revealed by somebody considered to be hostile toward Wikipedia in general and selected admins in particular, it is not surprising that the community, including some of its influential figures, supported Essjay in the beginning.

For Wikipedians, the problem with what Essjay did was not that he just created a false persona. What infuriated many members of the community was what he used that persona for and to did with it. He referred to his fake credentials in discussions, inflating his arguments,
and adding weight to his point of view in interactions using that persona, which by design were meant to be meritocratic and not based on formal authority. Wikipedia guidelines clearly state that edits have to stand (or fall) on their own merit. Thus even using one’s true credentials in a discussion may be considered a display of bad manners (an editor is supposed to clearly present his/her arguments as such, rather than trump a disputant’s comments with a reference to a diploma or degree).

With such “assuming good faith” being one of the fundamental guidelines on Wikipedia, as well as recognizing the high position of power occupied at the time by Essjay, most of his disputants thought they indeed must be mistaken, when Essjay disagreed with them in articles in his fake specialization area. This use of the fake persona created a feeling of major trust violation. Indeed, Essjay’s use of fake credentials quite likely persuaded other editors not to perform references checks, since they simply trusted his professional judgment.

**The post-Essjay disputes**

Jimmy Wales, who initially was declared to be “anti-credentialist” (possibly because of the failure of Wikipedia’s predecessor, Nupedia, which was intended to be developed by experts, but turned out instead to be a flop), announced to Associated Press in March 2007 that Wikipedia would start requiring contributors, who claim serious credentials, to identify themselves, rather than use pseudonyms (Read 2007). It seemed that this issue of credentialing and use of personas had been decided finally.

Yet, typically for Wikipedia, which relies on a participatory design and draws heavily from shop floor democracy roots (Rayton 1972; Greenwood et al. 1991), a decision had to be made by the community itself and not by any of its representatives. And still, the very community,
which was appalled by Essjay’s conduct, remained reluctant to accept Jimmy Wales’ proposal.

On 7 March 2007 Jimmy Wales proposed simply that credential verification could rely on other Wikipedian testimonies, and be completely voluntary (it would only apply to users who wanted their credentials checked). Still, the debate within the community was fierce for about two weeks (producing over 130,000 words of discussion). The vast majority of commentators considered credential verification a horrible idea. They were afraid that this verification “will become a bragging rite” (JoeSmack), and that “it makes a class of privileged editors, which was totally against the spirit of Wikipedia. Issues should be solved using cited, verifiable facts, not credentials” (pschemp). Many users pointed out that Wikipedia policies and guidelines (requiring, among other things, verifiability, reliable sources, or no original research) actually made the problem of credentials irrelevant. Also, they insisted that introducing official credential verification might actually work against the rules of current efficacy, since it would allow users with formal qualifications to gain the upper hand in disputes, when in normal meritocratic circumstances, these individual would not have such a high standing. In general, a visible majority believed that the policies already regulated discussions and allowed argumentative practices adequately, and allowing the use of credentials and credential verification would disturb the system both in terms of weakening meritocratic discussion, and the of stratification of users.

There were also other minor concerns, such as the possibility of fake credentials and the indispensable right to remain anonymous. Some editors perceived this proposal as an opportunity to attract more experts to Wikipedia (a long outstanding concern), but they were criticized, occasionally even with slightly hostile comments, e.g. “There we go. stewardship. Appeal to authority already, and this hasn't even been implemented yet.” (Corvus).
On 14 March 2007, one of the disputants, Netesq, created an alternative proposal, namely to “Ignore all credentials”. As he explained:

the idea of verifying credentials on Wikipedia is a bad one, and THERE SHOULD BE A POLICY in place stating that any information which people choose to set forth on their Wikipedia user pages is inherently unreliable and *NOT* subject to any sort of verification.

This proposal gained some interest and even support from Jimmy Wales himself, who commented:

I would support the elevation of that to a more fundamental position of respect within the community. I actually suspect that the problem of people inappropriately relying on credentials in an argument is much more likely in cases where people are faking the credentials. Actual PhDs are not normally pompous jerks; they are actually well trained intellectuals who have devoted themselves to a life of rational discussion and debate, and they know quite well that "I have PhD so STFU" is not a valid argument. It's the fake PhDs who are likely to try that nonsense (Jimbo Wales, 07:19, 16 March 2007)

However, this solution, too, was contested. Wikipedians made a point that Wikipedia already suffers from very low expert retention and that emphasizing that credentials would have no value on Wikipedia did not exactly help. In about 8,000 words, the disputants continued to deliberate on the pros and cons of the proposed solution.

Almost at the same time, one of the administrators (WikiLeon) created another essay, explaining that “Credentials are irrelevant”. A simultaneous discussion on this essay took up 17,000 words, in general repeating the arguments for the other two. It was already quite clear that a consensus would be very difficult to reach.

A couple of other proposals emerged in the meantime, including e.g., a suggestion of a full credentials ban.
Finally, a poll was organized to make a decision, and the results showed that the community flatly rejected the idea of credential verification. Jimmy Wales’ proposal received only 5 votes in favor and 29 against. Other proposals to regulate (or ban) credentials were also refused. This vote outcome has indicated, basically, staying with the status quo, where credentials could be listed, but could not be officially confirmed, so should be always treated with caution and not taken into account in meritocratic discussions.

One of the few stable results that emerged from Essjay case was the creation of an information page (poll support of 14 in favor and 13 against) that expressed a communal consensus about the rules of conduct for an honest Wikipedian. As of 27 November 2011, that page states:

One of the officially accepted norms on Wikipedia is honesty. According to this norm description, an honest Wikipedian does not intentionally misrepresent their identity or credentials. The choice of anonymity and pseudonymity is part of Wikipedia, but it is not a license to fabricate real world credentials. It is strongly recommended that you decline to share details you wish to keep secret rather than to invent alternatives. Fabrication of credentials will lose an editor his or her credibility and damage the credibility of the project as a whole (Wikipedia:Honesty).

The whole analysis of the Essjay case sheds some new light on the definition and nature of trust and control present in all online communities.

**Trust on Wikipedia**

Trust has often been portrayed as reciprocal in nature; it is being built (or destroyed) in a mutual relationship (Fox 1974). However, more recent publications show that the process may in fact be asymmetrical (Schoorman et al. 2007); trust in someone does not necessarily invoke that someone trusting back. In the case of Wikipedia, trust in other editors is often
asymmetrical and can be reduced to a simple expectation that the other editor argues with
good faith and in the honest belief that his/her reasoning is truthful.

Yet, the norm of reciprocity indeed is in and of itself a trust-building mechanism and may be
perceived as seminal to both social group and group norm formation (Gouldner 1960). One
important difference seen in the regular social context on Wikipedia is that interpersonal trust
is being built from scratch through online interactions and article editing with no stigmas
present that (Goffman 1963) relate to dress code, appearance, social class, gender, or race
(since editors have the liberty to represent themselves freely in this respect), although it is
quite likely, that they have other stigmas, based on text (vocabulary, punctuation, grammar,
etc.).

New virtual identities on Wikipedia can be created very easily, and this ease makes
Wikipedia, along with other online communities, particularly attractive for users who are
seeking the chance to interact without the usual stigmas they have previously seen.
Additionally, any unblocked editor has the right to leave and come back under a different
name or persona, which makes the creation of identities potentially transient and also
temporary (even though it should be noted that the more an editor edits, the more their
identity becomes stable, since the edits build up trust and standing for a single particular
identity, and this investment is of course non-transferable to a new identity).

Curiously though, creating a couple of different accounts and identities and using them
simultaneously (called “sock-puppetry”), especially if done to create an illusion of
independent support, is considered to be one of the most serious crimes a Wikipedian may
commit and often results in a lifetime ban. The act also definitely disqualifies an editor from
community service roles (Welham et al. 2009). Even if these edits are relatively innocent,
editing from two or more different virtual identities is considered fraudulent, much more so
than just inventing a single online persona. This view is so because within the community, the virtual identity is the only one that should matter. One can create any identity representation one chooses, but a person should not use more than one identity at the same time. There is, of course, also a practical side to the issue; a person operating multiple avatars could create an illusion of support for their own ideas and votes and thus inflate their own discursive argument falsely.

Yet, in a world where everybody can easily create a fake persona and where everybody can (and is allowed to) choose or change the presentation of self in this online community, trust can only be earned through actual participation in Wikipedia and a history of edits and interactions. Still, considering just the sheer size of the project (250,000 accounts created every month, 300,000 editors editing Wikipedia every month, 5,000 editors’ making more than 100 edits every month, more than 700 active administrators, etc.), for most members of this community it is clearly impossible to employ more sophisticated trust-evaluation strategies. Thus, interpersonal trust, in theory, should be replaced by control.

In some online communities, a so-called “swift trust” serves as a useful surrogate. Meyerson and Weick (1996) introduced this notion to describe a phenomenon that typical for virtual teams formed around a common project, clear tasks, and a defined project life-span. Swift-trust relies on team members’ suspending their doubts about others. Still, it also requires the members to reasonably expect beneficial outcomes for the project, as well as some stability within the team (most, if not all people should remain part of the team), and thus, this process quite understandably cannot emerge on Wikipedia. It is useful, therefore, to observe how traditional organizational control is actually exerted in this community.
Control on Wikipedia

Control on Wikipedia is much more limited in scope than it is in other online communities. While credential control systems are believed to be a common answer to the problem of low trust in most Internet interactions (Cheshire 2011), the case of Wikipedia clearly proves this speculation at least partially wrong. Even during its most extreme trust crisis, the community firmly rejected the idea of credentials verification. I believe this view ruled because one of the fundamental (even though not verbatim) assumptions of the Wikipedia community is to provide a clean slate start for all participants. This focus is an embodiment of true democracy the community believes; everybody is on equal grounds with everybody else. Status and trust have to be earned within the community, and attempts at shortcuts (such as leveraging one’s high social standing outside the community) are generally discouraged. Control, in the sense of credentials assurance, is thus non-existent, and even such public relations disasters as the Essjay case, do not serve as enough of a justification to change the status quo.

Traditional, bureaucratic control establishes a clear hierarchy and top-down authority and may be perceived by some as a tool of management domination (Braverman 1974). While this control provides structure and stability through designated procedures, it is often criticized contemporarily as being less volatile and likely soon to be replaced by more fluid organizational designs (Lewin et al. 1994), even though some point out that changes in liquid modernity may force bureaucracies to migrate control from the structural sphere to a more ideal sphere of influence (Cobb et al. 2001).

The need for such control on Wikipedia is significantly reduced by the fact that Wiki-media technology makes checking all changes by any editor, with all of a history track, very easy to accomplish. Naturally, since all experienced editors are aware that all their actions are traceable and that any user is allowed to perform a silent control check on any editor, control
becomes largely internalized. However, since all users of Wikipedia (even the unlogged-in users) are allowed to control the content of any edit changes, there is no group of controllers that is singled out as the management. Even though Wikipedia has administrators who can block other users to prevent disruptive behavior, users without administrative privileges often control those by requesting explanations of particular actions or by discussing the shape of an article’s parts. Wikipedia policy on administrators clearly states that “…administrators should never develop into a special subgroup. Rather, administrators should be a part of the community like other editors, with no special powers or privileges when acting as editors.” To reflect this unprivileged status of administrators, on Wikipedia, they are often referred to as “janitors”, and their role is strongly emphasized as ancillary; becoming an admin in the popular Wikipedic adage by no means should be treated as a “big deal” (which means that, at least in the official discourse of the community, everybody is expected not to ever consider becoming an administrator as moving up in the Wikipedia hierarchy).

Thus, while there is wide awareness of the possibility of control, everybody is actually in control as well as controlled which eliminates many of the degrading effects of the controlled-the controllers division (Braverman 1974). Control stays both concertive (Barker 1993) and reciprocal.

Similarly, the forms of control that are speculated to emerge in a hierarchical organizational designs, such as technocratic control (Burris 1989), relying on one’s expertise (and partly on professional credentials), are significantly reduced or even become non-existent on Wikipedia, because of the very basic rejection of credentials checks, as well as because of the playful character of “work” (Hunter 2010).

Organizational control is actually most visible and present only in the instance of unknown editors. The community has easy ways of observing such edits from unregistered users (called
“IPs”) and to some extent also newly registered ones (in fact, many Wikipedia editors specialize in watching the recent changes coming from unregistered or new users). This review leads to partial exclusion of unregistered users from the community to the extent that in one of the advisory essays on Wikipedia behavior, editors are carefully reminded that “IPs” are actually human too:

Many users believe that unregistered users’ sole contributions to Wikipedia are to cause disruption to articles and that they have fewer rights as editors compared with registered users (…) As current policy stands, unregistered users have exactly the same rights as registered users to participate in the writing of Wikipedia. (…) Remember this when dealing with unregistered users. They are not a lower category of users. They are not a special subset that we tolerate. They are not locust swarms intent on destroying your article ([[Wikipedia:IPs are human too]]).

However, once a user is registered, makes some edits, creates his or her own user page and has some discussions on the talk page (non-existence of these two pages serves as an immediate and visible warning to experienced editors that they are dealing with a rookie), control gradually decreases.

Overall, the scope of control exertion on Wikipedia seems to be relatively narrow, just as it is in the case of trust. Therefore, it is important to consider what the alternatives to this classical dichotomy might be.

**Interpretation and implications**

Active participation in Wikipedia helps build a local (intra-community) identity, status, and reputation (Anthony et al. 2007). Resigning from the means of control (such as assurance, credential checks, hierarchical control), as well as accepting the low-trust environment
(anybody can present themselves any way they want, trust is developed locally and fluid),

apart from its obvious disadvantages, also produces significant organizational benefits.

Rejecting credentials control allows for a full democratic participation of different people
with different backgrounds. Paradoxically, since credentials are not checked, claiming one’s
expertise ends up to be just an empty gesture, as all arguments still have to make valid points
to be considered. However, making a point “valid” according to Wikipedia standards is not
something everybody is born to do. In fact, even though Wikipedia rules say that you can
“ignore all rules” and “be bold” (meaning that knowing all or even any procedures is not
necessary to successfully edit on Wikipedia), rules of conduct in discussions on Wikipedia
remain highly formalized. There are 9 official guidelines for working with others totaling
14,000 words, 20 official behavioral guidelines totaling 45,000 words, and many more semi-
official suggestions and essays, as well as official procedures displayed for dispute resolution,
mediation, arbitration, requesting comments, blocking, to say nothing of general principles
and article editing standards, and more. This array leads to an unusual observation.

Apparently, the need for trust or control is substituted with precise behavioral scripts and
formalization of discussion rules.

While Essjay’s case showed that the Wikipedia community is prone to crises of trust in its
contributors; it also proved that more formal rules of discussion, dispute resolution, and
decision-making (Jemielniak 2011) can provide a necessary alternative to the introduction of
tighter controls and credentials checks. Thus, “legalistic remedies”, often perceived as
ineffective substitutes of trust in organizations (Granovetter 1985; Sitkin et al. 1993), turned
out to be actually working in a democratic, participative online environment.

Clear article editing rules do allow for non-expert administrators to exercise judgment in any
meritocratic disputes. Wikipedia has been able to successfully reduce the problem of truth to
the problem of sources. In the same way, it has reduced the problem of credentials, as well as
low trust, to the problem of following the rules of editor behavior and proper discussion.
Thanks to the creation of such precise and elaborate guidelines, requiring all disputants to
ground their points of view well, Wikipedia is probably one of the very few communities in
the world where a teenager can actually win a meritocratic debate with a person holding a
Ph.D. in a discussed field (Kapiszewski 2011). The fact that knowing all the guidelines and
rules is not mandatory also adds more flexibility to the discussions as long as beginners make
a point that is valid in the eyes of others; not adhering to the letter of the rules is not
considered to be a problem. Not sticking to guidelines is a powerful formal argument,
however, which experienced users may address, situation allowing.

Shifting the weight from a trust-control duo to formalization (or in other words, trust in the
procedures and not the individuals and their credentials) is not a universal remedy and has
other, somewhat undesirable implications. For example, although in theory Wikipedia is not a
bureaucracy ([[Wikipedia:Buro]]), accumulation of instruction creep is to be avoided
([[Wikipedia:CREEP]]). All editors are expected to use common sense ([[Wikipedia:Common
sense]]) and ignore the rules when necessary. Still it is difficult not to notice that the number
of rules, guidelines, and policies has grown dramatically over years (Butler et al. 2008).
Increasing bureaucratization of Wikipedia also serves, among other goals, the purpose of
reinforcing and protecting the status of editors. The sheer amount of rules increases the
threshold for newcomers and also introduces a natural, if unorganized hierarchy of those in
the know and those outside that same know.

Additionally, rules are not listed or collected in one, single consistent manner. Wikipedia
distinguishes between policies, guidelines, and essays, but that difference is blurry with no
clear differences between these in the sense that violating some policies and guidelines often
does not bring any undue consequences (for example, the verifiability policy is violated hundreds of times every day), while violating rules that need to be expressed in “just an essay” in some cases may lead to blocks (for example, [[Wikipedia:Tendentious editing]]).

Finally, experienced editors often use cryptic abbreviations and Wikipedic lingo in their discussions. Sentences such as “PROD per nom”, “fails to meet WP:GNG”, “fails WP:N”, or “SD – copyvio” are perfectly understandable to the site’s natives, but will discourage beginners even from expressing their views easily or regularly.

All these issues add to newcomer confusion, and a discursive advantage for experienced editors. Clearly, both the number of regulations and the number of guidelines, as well as their chaotic and inconsistent presentation and hermetic slang, even if unintentional, serve the purpose of increasing the threshold level for less experienced and new editors, and as many other seemingly functional behaviors and concepts, such as e.g. time enactment, play also a symbolic role (Jemielniak 2009). When trust in people as well as control are substituted by only trust in procedures (also known as institutional trust, see e.g.: Shapiro 1987; Sitkin et al. 1993), apparent proficiency in those procedures as such can have a trust-building value.

Anybody can claim they have Ph.D., and with so many different academic fields, it is difficult to verify one’s expertise. However, refinement in Wikipedic rules and language is immediately transparent to natives of this online culture. Editors displaying good understanding of Wikipedia policies are naturally considered more trustworthy, at least not to make the mistakes typical for those who do not have much Wikipedia article editing experience.

It is also in understanding bureaucratic rules (and their cryptic abbreviations) that one can determine at least to some extent one’s status in a community. Therefore, it is only natural
that introducing a possibility to build one’s status through external credibility and expert status could quite easily disturb this delicate social system.

What is really most striking in the studied case of Wikipedia and the Essjay crisis is seeing that this online community believes that if the rules of conduct are really precise and appropriate and developed in a participative manner, the members will prefer to trust them and their execution, rather than turn to the development of more interpersonal trust or introduce precise credentialing and general control systems. Even though, as some authors indicate (Sitkin et al. 1993) legalistic solutions may have limited effectiveness in restoring trust, this is not the case of Wikipedia. The fully democratic and participative character of its rules and procedures may be decisive in their communal acceptance and preference over para-managerial control or local trust development. As it seems now, procedures and organizational routines do substitute for the need for trust and control, even though this process is possibly limited only to virtual communities, which are able to develop their own rules and structures and also have naturally limited regular trust-building face-to-face interactions.

Yet, the conclusion here that sometimes formalization of discursive practices may successfully substitute for the need for control in low-trust environments is interesting and certainly worth further research. Even if the implications of this phenomenon are limited only to online open-source communities, they are still significant because they indicate that traditional control methods can often serve as an ersatz solution to trust deficiency, but only when there are no clear rules of conduct and behavior and as long as those rules are created in an a-hierarchical, self-managed community, and created by the community itself.

With virtual workplaces and virtual teams becoming more often parts of regular organizational structures and with new product innovation development and deployment
growing in popularity among many corporations, these study findings call for further investigation into how such solutions can integrate well into the nature of collaboration practices in the online world and beyond.

References


Jemieleniak, Dariusz, and Andreea D. Gorbatai. (2012). Power and Status on Wikipedia: working paper at collaboration seminar at Bermann Center for Internet and Society


Welham, Jamie, and Nina Lakhan. (2009). Wikipedia sentinel quits after 'sock-puppeting' scandal". *The New Zealand Herald*. 6 June,

(retrieved on 26 November 2011).