

“Writing up rather than writing down”: Becoming Wikipedia Literate

Heather Ford
Ushahidi
hford@ushahidi.com

R. Stuart Geiger
UC-Berkeley School of Information
102 South Hall, Berkeley CA 94703
stuart@stuartgeiger.com

ABSTRACT

Editing Wikipedia is certainly not as simple as learning the MediaWiki syntax and knowing where the “edit” bar is, but how do we conceptualize the cultural and organizational understandings that make an effective contributor? We draw on work of literacy practitioner and theorist Richard Darville to advocate a multi-faceted theory of literacy that sheds light on what new knowledges and organizational forms are required to improve participation in Wikipedia’s communities. We outline what Darville refers to as the “background knowledges” required to be an empowered, literate member and apply this to the Wikipedia community. Using a series of examples drawn from interviews with new editors and qualitative studies of controversies in Wikipedia, we identify and outline several different literacy asymmetries.

Categories and Subject Descriptors

H.5.3 [Information Systems]: Group and Organization Interfaces—computer-supported collaborative work

Keywords

Literacy; Wikipedia; new literacies; educational technology; ethnography

1. INTRODUCTION

Working as a teacher, advocate and researcher in the literacy community for nearly 30 years, Richard Darville wrote a series of articles about how people learn and what it means to become literate. His message, an extension of the current shift towards a “social practices” conception of literacies, reframes literacy as fundamentally a way for people to be able to take part in some social activity. As he argues, “It’s not that there’s a skill (a merely cognitive process), and then a social context. Literacy is social all the way down.” [1] One of the most important consequences of such a framework is that learning is seen as inherently contextual and localized, with a wide variety of background knowledges and skills necessary to read and write in a particular genre. This is particularly the case in institutional genres – texts about and in the workplace, law, healthcare, even the news which Darville has spent the majority of his career studying. He argues that because texts are strongly linked to social practices in institutions, they:

are often put together so that their sense is impenetrable for novice readers. All texts assume particular “background knowledge.” All aim at some “implied reader.” Texts effectively exclude learners when they take someone very

different as their implied reader, or simply assume knowledge that a learner doesn’t have. These are roadblocks to “becoming literate” that cannot be overcome even by reading more closely, or between the lines. [6]

Our understanding of literacy draws extensively from Darville’s subsequent concept of “organizational literacy”, which refers to the ability to read and write in such a way that takes into account how a particular document has been or will be circulated, interpreted, and evaluated within an organization. Darville’s most compelling example is that of job applicants: unlike those without organizational literacy, “experts” to the process know how to articulate their skills and qualifications in such a way that it will flow through various levels of a human resources bureaucracy, receiving favorable evaluations each time. For example, a job applicant with high organizational literacy knows how to read a job listing and intuit how their potential application will be read.

While many of these skills are also involved in the extensive studies and theories around organizational learning, organizational literacy is not equivalent to organizational know-how or familiarity, largely because it is specific to the practices around documents and text. However, this is a tenuous distinction, primarily because our framework embraces the multifaceted ways in which reading and writing are often the primary way in which members of an organization interact with each other. This is particularly the case in highly mediated organizations like Wikipedia, and a number of Wikipedia researchers have exploited this fact to perform a wide variety of analyses. Similarly, Wikipedia as an organization is highly structured and learning such structures is a critical aspect of socialization, as has documented in rich detail.

Editors have complex ways of doing and documenting encyclopedic work, particularly through the use of templates and other institutional forms, leaving traces that are invaluable for each other but difficult for new users to comprehend. In fact, as Geiger and Ribes demonstrate in their “trace ethnography” [8, 9], almost every socially relevant action that takes place in Wikipedia is logged and often categorized. Due to the sheer size of the encyclopedia project, Wikipedians have developed sophisticated ways and tools to track these logs and revision histories to efficiently organize ad-hoc groups to collectively make specific decisions based on general norms, principles, and procedures. This decentralized mode of governance is what has made it possible for the all-volunteer Wikipedian community to collectively build and maintain the project. However, this mode is often impenetrable for new editors who lack the organizational literacies required to interpret and author texts and traces.

For example, nominating an article for deletion is a well-defined procedure in Wikipedia’s article deletion bureaucracy, requiring that an editor leave a series of traces that, first and foremost, serve to inform others that the article is being nominated for deletion. Wikipedians with and without administrative privileges routinely

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WikiSym '12, Aug 27–29, 2012, Linz, Austria.

Copyright © 2012 ACM 978-1-4503-1605-7/12/08... \$15.00

patrol new pages, and any editor can “tag” an article for speedy deletion if they believe it fits one of the project’s dozens of criteria for speedy deletion. Administrators are then able to unilaterally delete tagged articles that meet certain criteria – including obvious copyright infringement, lack of notability, or vandalism. Nominating an article for deletion or contesting a deletion are simultaneously rhetorical, normative, organizational, and technical acts. Yet in Wikipedia, like in many socio-technical systems, these aspects are compressed into the creation of a single standardized trace – which consequently makes the nomination able to be seen, aggregated, conceptualized, and discussed. A proper response is also codified into this regime of tracing: to contest a deletion, one must not only craft the proper argument and click “submit”, but more importantly do so in such a way that the statement is made visible to Wikipedians who utilize specialized modes of tracing to manage the article deletion process. This is but one example that illustrates how new editors to Wikipedia are often at a substantial disadvantage when interacting with veterans.

We focus on these asymmetries, arguing that a variety of activities and practices collectively constitute a new kind of literacy that has been generally passed over in most discussions of both the problems new Wikipedia editors face as well as the broader “digital literacy” literature. We present three cases that illustrate the various problems new users have when entering Wikipedia. Each demonstrates a different way in which the ability to interpret and author texts are prerequisites for full participation in a massively distributed and highly mediated community like Wikipedia. Darville’s argument is quite applicable to these cases, as such literacy asymmetries are not the kind that can be overcome by “reading more closely” or “between the lines.” Without an understanding of the people and the processes that are attached to these texts, a newbie can become confused not only about what the texts are saying, but also how to use them, or even “how to talk back” [6]. We try not to make any claims in the above examples about whether a deletion decision was legitimate or illegitimate. What is more important is to describe and analyze the misunderstandings that occurred, pointing to the kinds of “background knowledges” required to be a more effective editor.

2. LITERATURE REVIEW

When Wikipedia is discussed in terms of literacy it is usually in response to attacks against the use of the encyclopedia by students in learning practice. This means that the majority of writings about Wikipedia and literacy have focused on the efficacy of Wikipedia as a teaching tool [3, 4] rather than the question of where are the key challenges in becoming Wikipedia literate, or the role of culture, power and social context in becoming Wikipedia literate. Wikipedia literacy is often discussed as a set of universally applicable skills and competencies without a keen understanding of the unique literacy challenges of people in different socio-economic contexts.

This trend mirrors developments in the field of literacy with the so-called “new literacies” that encompass the skills and competencies required in the age of digital technologies. On the one hand, there are scholars who have proposed a universal set of skills they believe are required and also somehow inherent to the skillset of the so-called “Digital Natives” [5, [10], [1] [15]]. On the other, are scholars like Nishant Shah who recently argued [13] for a more political, contextual approach to understanding digital literacy that goes beyond the presentation of “Digital Natives” as largely young, white, male, affluent, English speaking and located in contexts of ubiquitous connectivity. It was

problematic to Shah to use this image as the prototype around which digital natives in the “rest of the world” were imagined and so he sought to showcase the challenges and experiences of “Other Digital Natives” in places like India, South Africa, Egypt and Brazil [2]. In the field of literacy studies, Brian Street describes these differences in terms of disciplinary distinctions [14]. The first group is distinguished by what Street describes as “culturally narrow approaches” that have “predominated” the field, with psychologists and educators focusing “on discrete elements of reading and writing skills”. Anthropologists and sociolinguists have concentrated on what he calls “literacies” – “the social practices and conceptions of reading and writing” and the rich cultural variations of these practices [14].

3. Methodology

Like Street et al, we take an ethnographic approach to Wikipedia literacy, attempting through a detailed in-depth account of actual practice in different cultural settings to understand the central role of power relations in literacy practices. We choose examples from both the US and Kenya that highlight a particular facet of misunderstanding from Darville’s adult literacy practice. The Kenyan examples, including interviews, come from a larger ethnographic study on the relationship between Swahili and English Wikipedia in east Africa. Material for the United States Messer-Kruse example comes from the writings surrounding the case in the online and radio news media. In each case, we analysed texts including the edit and talk pages of the case in question, as well as related interview transcripts and writings in the blogosphere and news media. We then used Richard Darville’s writings on literacy that are aligned significantly with this political, contextual approach to outline the cause of the misunderstanding and then propose a new theory of Wikipedia literacy that draws from the socially grounded experiences of Wikipedians in different cultural contexts.

4. MISREADINGS, MISUNDERSTANDINGS

4.1 Misreading organizational texts

Darville: “misreading is not about being unable to get words off the page, but rather being unable to participate effectively in the social action and relations that are carried in texts and documents - or even to see what that action and those relations are.” [5]

Bowling Green State University Professor Timothy Messer-Kruse recently wrote an article in *The Chronicle of Higher Education* about his failed experience editing the Wikipedia article about the 1886 Chicago Haymarket trial and riot – where a bomb thrown during an anarchist rally sparked America’s first Red Scare [11]. Messer-Kruse holds that the mainstream historical record is faulty, and has spent much of his academic career trying to correct what he argues most historians have gotten wrong. Attempting to edit the article with the new facts that he had uncovered in researching a book about the subject, Messer-Kruse’s edits were continuously reverted for a variety of reasons, most importantly because of his inability to convince other editors that the majority of scholarship regarding the riots had been incorrect and that his research had uncovered new facts that accurately contradicted those claims.

Replying to Messer-Kruse’s article, one of the page’s editors countered his claims of unfair treatment, writing that he had not correctly cited sources and had not presented himself well. In one of the edits for which he was reverted, he had removed and replaced a cited quotation “instead of providing a

counterargument” [15]. In the “edit summary” field provided for editors to provide a note about any particular edit, Messer-Kruse noted that he had removed the claim that “most police officers were killed by “friendly fire”” in the article since, he wrote, “Neither of the two footnotes to this information support it – in fact, one flatly contradicts it.” Removing a cited phrase, however, required more than a simple edit summary according to Wikipedia editor, *operalala*. It seemed to require the presentation of sources supporting his claim and perhaps a detailed and reasoned argument on the article’s talk page. *operalala* also suggested in the aftermath that Messer-Kruse might have done better to use an anonymous screen name, since his use of his last name as his username (“MesserKrus”) appeared to the other editors of the page that he might be merely marketing his own work when he attempted to cite the book that he had written on the subject.

operalala: “This makes it easier to cite your own published work without alarming other editors. If they see a “ProfSmith” citing Prof Smith all over the place, their first reaction will be to think that this is a narcissistic self-published wingnut.” [15]

This example reflects a misunderstanding between Messer-Kruse and other editors born out of the professor’s inability to see the assumptions and workings of the article and its related talk page as embedded in particular organisational structures and alliances, and with individual notions of the Wikipedia principles of “verifiability” and “reliable sources”. Messer-Kruse expressed an understanding of the article as simply a location of the most up-to-date facts about the Haymarket riots. In fact, the article and its related pages housed the inner-workings of a group of individuals with a variety of background knowledges. Those “knowledges” dictate how editors with new information should present that information, how editors should present themselves, as well as where and which tone is required on edit summaries, talk pages and articles. In Darville’s terms, this might be seen as an example of the misreading of organizational texts – the article, talk pages and edit summaries that required particular tones, styles and codes. Unable to decode these organizational texts, Messer-Kruse was unable to participate effectively in the social action relating to the article or even, as Darville notes, to see what action and those relations are.

4.2 “Agentless accounts”

Darville: “When the agents of actions are deleted from texts, readers must ‘fill them in,’ using a background knowledge of how actions are done and who would do them.” [7]

In an interview with a relatively new English Wikipedia editor, Kipsizoo, we talked about a challenge that he had helped coordinate in Nairobi where volunteers would take photographs of administrative buildings in the capital and upload them to Wikimedia Commons so that they might be used in Wikipedia articles. Kipsizoo described how the volunteers’ attempts were met with almost immediate deletion and how he wasn’t even sure whether a person or a bot had actually deleted them.

Kipsizoo: “So we were trying to create some stubs in the English Wikipedia and the minute we created those stubs they were all nominated for deletion. We felt bad. We needed those photos to be uploaded to Commons so that they would be used in articles...”

HF: “Did you try talk back to him?”

Kipsizoo: “Yes, we did but we opted to try create it in Swahili rather. Cause the guys are so harsh... We clearly indicated that

these are just stub articles. I think it was a bot or something. It wasn't a real guy... maybe.” (Interview, 4 August, 2011)

Kipsizoo refers here to the complex English Wikipedia deletion process, but confuses the speedy deletion (CSD) and articles for deletion (AfD) processes, indicating that his edits were “nominated for deletion”, a feature of the AfD and not CSD process. Speedy deletion does not require consensus (a significant part of the AfD process) and is difficult to argue against since the pages are deleted and are invisible to regular editors. Administrators are not required to let editors of new pages know when their article has been deleted, and even if they do, new editors often do not visit their talk pages or are embarrassed when their articles are deleted and do not see the opportunity to “talk back”.

The above exchange indicates that Kipsizoo is unaware of the differences between the deletion processes that would be essential to knowing how to fight back. If he had been aware of the entire deletion process – how CSDs are controlled by patrolling administrators and can be opposed by leaving a message on that article’s talk page or adding the “hang on template”, Kipsizoo might have been more successful. As it is, Kipsizoo is not even sure whether a real person who deleted the articles or a bot. His confusion probably emerges from the immediacy of the deletion. Since he did not know how to read the deletion traces for any sign of human life, they appeared somehow automated. Suggesting that the process was driven by a bot, Kipsizoo was “filling in” using a background knowledge of other processes he knows about, for example spam software on email platforms that appear configurable, and without an opportunity to appeal against.

Part of the problem in this particular case is that without knowing that the person who deleted the article was indeed a human, Kipsizoo could not know the kinds of challenges that this particular type of human typically faces. Perhaps if he had understood the kinds of work done by so-called “vandal fighters” who must face thousands of illegitimate articles that are created by spammers or marketers daily, he might have known how to distinguish his particular edits from those bad-faith edits. Without an understanding of the speedy deletion process of which this experience was part, it became impossible for Kipsizoo to have his edits endure. In this case, it did not matter how well he was able to write using MediaWiki syntax. Without access to the communication channel that would enable him to dispute the deletion, he was left feeling unwanted and disempowered, deciding instead to take what he perceived as second best: to go and create the articles in Swahili Wikipedia.

4.3 “Writing up rather than writing down”

Darville: “What counts is how matters can be written up (to enter them into the organizational process), not how they can be written down (to relate experience or aid memory)”[15]

Almost anyone can start a Wikipedia article but not any subject will do. Wikipedia calls for articles to subscribe to its notability guidelines, whereby an article warrants a new article if it has received significant coverage from independent reliable sources [16]. When subjects do not have significant coverage from independent secondary sources, they become vulnerable to challenge and attack. Writing a new article, Wikipedia editors must understand, not just how to write a few lines describing a subject, but how to argue that the new subject deserves its own page. This becomes especially important in contexts where few traditionally “reliable sources” exist or when reliable sources that

do exist are located outside the purview of most editors (such as small newspapers from a country in the global South).

When Wikipedia editor, Abbasjnr attempted to create an article on “Nairobi Java House” which he described in an interview as “Kenya’s equivalent of Starbucks”, the article was tagged for speedy deletion twice in the 30 minutes in which the article was being edited. Abbasjnr started the new article with a list and description of each Java House location and provided references to the Nairobi Java House website and to an article in a Kenyan newspaper about the founder’s child sex abuse scandal. Without a reference to establish notability, however, the article was tagged for deletion using the code, “A7” that refers to an “Article about a company, corporation, organization, or group, which does not indicate the importance or significance of the subject”. Later the article was re-written by an administrator with a reference to a 2007 Guardian newspaper article featuring the cafe as the only place in Nairobi to find export-grade Kenyan coffee [10].

Abbasjnr, then new to editing on English Wikipedia, lacked a holistic perspective of how individual articles within the English Wikipedia corpus came to be. In Darville’s terms, it wasn’t just a matter of writing down the salient features of Nairobi Java House. It meant writing up, in Wikipedia’s particular style and according to particular rules regarding notability and the referencing of notable characteristics that were made even more complex given the sparseness of such resources about organisations in the east African nation. On the administrator’s side, User:Glenfarclas, the editor who speedily deleted the page, had no clues to assist him to make a more discretionary decision about the article, or to have a conversation with the editor about it, given that as User:SJ later wrote on the article’s talk page that “countries outside of the US and Europe have vastly fewer newspaper articles written about them per organization or project”, which might perhaps have enabled the administrator to engage more constructively with Abbasjnr.

5. CONCLUSION

Literacy is a means of exercising power in Wikipedia. Keeping traces obscure help the powerful to remain in power and to keep new editors from being able to argue effectively or even to know that there is a space to argue or who to argue with in order to have their edits endure. Much of the debate around what some claim are unfairly deleted articles surrounds whether the right decisions were made by the parties involved, without recognizing that the most important feature of these cases is that misunderstandings are occurring that are preventing the development of a richer, more comprehensive global encyclopedia.

We need to gain a deeper understanding of the kinds of experiences faced by editors in particular social, cultural, economic and political contexts. In doing so, we will learn what kinds of “background knowledges” are required to be an empowered member of the Wikipedia community. These stories can help to expand designers’ perception of users beyond the prototypical, so that they may design clearer, more understandable roadmaps to replace the veritably invisible traces that characterize current Wikipedia processes such as those of deletions. By enabling new Wikipedians to more quickly understand how the process of deletion works, who are the people behind it, what motivates them, as well as how and where they can be engaged, Wikipedia could ramp up the time it takes to become an effective editor and to prevent some of the rapid fallout of new editors. Finally, Wikipedia literacy needs to engage with the social and cultural aspects of article editing, with training materials and

workshops provided the space to work through particularly challenging scenarios that new editors might find themselves in and to work out how this fits within the larger organizational structure.

The threat is that, if this kind of understanding is not cultivated, that newbies will not stay long enough to persevere and/or will use alternative narratives to fill in what they think happened to silence them. Justified or not, having claims that Wikipedians are merely opposed to the perspectives of those dissimilar from them (whether that is that they are from Africa or the academy) is not helping the encyclopedia grow in areas that it is currently weak.

6. ACKNOWLEDGMENTS

Our thanks to all the Wikipedians who talked with us about this topic, Hivos and OSI for supporting Ford’s research as part of the “Understanding Sources” grant to Ushahidi, and UC-Berkeley and the Wikimedia Foundation for supporting Geiger’s work.

7. BIBLIOGRAPHY

- [1] Aviram, A. and Eshet-Alkalai, Y. Towards a Theory of Digital Literacy. *European Journal of Open, Distance and E-Learning* 1 (2006).
- [2] Ben-David, A., McKay, K., Flor, L., et al. *Digital Alternatives with a Cause*. Bangalore: Hivos, 2011.
- [3] Bruckman, A. and Forte, A. Scaling Consensus: Increasing Decentralization in Wikipedia Governance. *Proc HICSS 2008*, (2008).
- [4] Bruckman, A., Bryant, S.L., and Forte, A. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. *ACM* (2005), 1-10.
- [5] Darville, R. The Language of Experience and the Literacy of Power. In M. Taylor and J. Draper, eds., *Adult Literacy Perspectives*. Culture Concepts, Toronto, 1989.
- [6] Darville, R. Literacy as practices, teaching as alignment: A message in a bottle. *Literacies* 10, (2009).
- [7] Forte, A., Kittur, N., Larco, V., Zhu, H., Bruckman, A., and Kraut, R.E. Coordination and beyond: social functions of groups in open content production. (2012), 41-426.
- [8] Geiger, R.S. and Ribes, D. The Work of Sustaining Order in Wikipedia: The Banning of A Vandal. *Proc CSCW 2010*, (2010).
- [9] Geiger, R.S. and Ribes, D. Trace Ethnography: Following Coordination Through Documentary Practices. *Proc HICSS*, IEEE (2011).
- [10] Jenkins, H. What Wikipedia can teach us about the new media literacies. *The Journal of Media Literacy* 54, Media Convergence and the Collective Community (2007).
- [11] Messer-Kruse, T. The Undue Weight of Truth on Wikipedia. *The Chronicle of Higher Education*, 2012.
- [12] Rice, X. Nairobi cafes buzzing as real coffee finally comes home. *The Guardian*, 2007.
- [13] Shah, N. In Search of the Other: Decoding Digital Natives. *DML Central*, 2011. <http://dmlcentral.net/blog/nishant-shah/search-other-decoding-digital-natives>.
- [14] Street, B., ed. *Cross-Cultural Approaches to Literacy (Cambridge Studies in Oral and Literate Culture)*. Cambridge University Press, Cambridge, 1993.
- [15] Tapscott, D. *Growing Up Digital: The Rise of the Net Generation*. McGraw-Hill, New York, 1998.
- [16] Wikipedia. Identifying reliable sources. *Wikipedia*, 2012. http://enwp.org/wiki/Wikipedia:Identifying_reliable_sources

