Visible Effort: Visualizing and Measuring Group Structuration Through Social Entropy

Sorin Adam Matei, Robert Bruno, and Pamela L. Morris

1 Introduction

A large amount of research supports the benefits of group collaboration in terms of positive outcomes, individual satisfaction, and powerful cognitive effects (Johnson & Johnson, 1999; Slavin, 1996). The practice of computer-mediated collaboration (CMC) comes in many forms and many definitions for its meaning have been proposed. However, much research still needs to be done to understand the nature of the processes that take place during CMC. For example, despite recurring claims that online collaboration is innately egalitarian (either in terms of access or outcomes) and potentially superior due to some form of "collective intelligence" that spontaneously emerges without much coordination (Kelly, 1995; Rheingold, 2002), there is mounting evidence that online interaction follows traditional patterns of human interaction (Lampe, Ellison, & Steinfield, 2006; Matei & Ball-Rokeach, 2001).

We hold that effective group collaboration using CMC needs division of labor, coordination, and clear goals. Moreover, CMC groups that are rooted in norms or

A version of this paper has been presented at the Global Communication Forum, Jiao Tong University, Shanghai, September 29–30, 2010.

S.A. Matei (✉)
Brian Lamb School of Communication, Purdue University, West Lafayette, IN 47907, USA
e-mail: smatei@purdue.edu

R. Bruno
Communication Studies Department, City University of New York-LaGuardia, 718-704-7780, Long Island City, NY 11101, USA
e-mail: rbruno@lagcc.cuny.edu

P.L. Morris
Department of Communication Studies, University of Wisconsin-La Crosse, 1725 State St., La Crosse, WI 54601, USA
e-mail: pmorris@uwlax.edu

S.A. Matei et al. (eds.), *Transparency in Social Media*,
Computational Social Sciences, DOI 10.1007/978-3-319-18552-1_6

local cultures and that foster specific ethical guidelines are more likely to be productive. Conversely—and quite significantly—individual effort, inputs, and outputs are regularly observed to be *unevenly distributed* with naturally-occurring coordination and/or power hierarchies accompanying these uneven distributions. Barabási (2003) and Huberman (2001) have documented this uneven distribution for linkages between websites while Anderson (2006) and Shirky (2008) have done the same thing for online interactions related to e-commerce and online content consumption.

It is therefore of great importance that online collaboration be supported by new tools and be studied with appropriate methodologies that determine in what manner such uneven distribution of effort functions or how it can be modeled to facilitate maximum individual and group effectiveness. At the same time, egalitarian work paradigms can and should be employed in an informed, measured and intelligent manner. This is especially important in view of numerous claims that egalitarian collaborative systems are the preferred future organizational form (Brafman & Beckstrom, 2006), which would foster some form of "wisdom of crowds" (Lease, 2007; Powazek, 2009; Tapscott & Williams, 2006).

Some practitioners speculate that online groups are particularly adept at solving large problems by breaking them down into smaller and roughly similarly sized tasks to be allocated to many uncoordinated participants (Tapscott & Williams, 2006). A related expectation is that the larger the group and the more equitable the social structure, the more likely the problem will be solved effectively (Brafman & Beckstrom, 2006). As an example, an often invoked broadly-distributed process such as open source software development has been labeled by Raymond (2001) as the "bazaar" process. Accordingly, he notes that the hugely successful Linux operating system is the product of "bazaar" style micro-negotiation and collaboration between unknown and equally qualified programmers who take turns in fixing each other's mistakes. Illustrating the power of distributed open source programming, he states, "Given enough eyeballs, all bugs are shallow" (p. 30).

The egalitarian assumption that surrounds online interaction can be interpreted in many ways. One could be that equality of access should not be confounded with that of outcome or consumption. This distinction could be very important if the undeniable fact that the Internet gave more people more access to educational, business, or entertainment resources than previous media is to be reconciled with the body of observable evidence, supported by sociological theory, which suggests that collaboration online is in fact highly structured, that the Web has leaders and followers, and that equality of contributions and consumption is rarely if ever present in spontaneously emerging online groups (Kuk, 2006; Shirky, 2008). In opposition to Raymond's perspective, Kuk found a correlation between structuring, participation inequalities and the most productive processes of open source software development.

Taking a cue from this evidence, we propose a method for measuring the amount of equality and the emergence of social structure in groups that participate in CMC. The method relies on measuring the level of social "entropy" of an online environment. Social entropy, which will be discussed at length below, captures the degree of equality, evenness, and diversity of collaboration in any given system or group. The measure is visualized within the wiki environment "Visible Effort" (Veffort.us) with color-coded page frames and graphs, which can be used by learning groups for self-monitoring their collaborative progress. You can visit the experimental site at Veffort.us.

The measure and visualization method proposed serve two goals. First, they are used for measuring and visualizing the degree of collaborative evenness and emergence of social structure in a collaborative online wiki environment. Second, they can be used for steering the collaborative processes to attain specific goals (Matei & Bruno, 2015; Matei, Oh, & Bruno, 2006). This can be accomplished either passively or actively. It can passively provide users feedback on the processes that take place in their online space or can actively provide site administrators, project leaders or instructors the information necessary to intervene and moderate collaborative efforts. The present paper will illustrate these capabilities by describing a specific quasi-experimental teaching activity in tandem with a detailed discussion of theoretical justification, methodological underpinning, and technological capabilities of the Visible Effort approach.

2 CMC and Uneven Online Interaction

A significant amount of empirical evidence indicates that CMC in online environments tends to be distributed in the shape of a highly skewed curve (Anderson, 2006; Huberman, 2001; Kittur, Chi, Pendleton, Suh, & Mytkowycz, 2007; Ortega, Gonzalez, & Robles, 2008). Examples include the well-known metric of 10 % of Wikipedia editors contributing almost 90 % of the online encyclopedia's articles (Matei & Bertino, 2014; Ortega et al., 2008); which use similar inequities of production along the lines of 20–80 % that occur within the practice of the open source software (OSS) and Linux movement (Matei & Bruno, 2015); and multiple manifestations of uneven social distributions on Yahoo user groups, assorted emailing lists, user-generated "question & answer" forums, and so on (Matei & Bruno, 2015). Although utilizing different measurement techniques and theoretical perspectives, other terms that have cropped up in recent years to describe this extreme inequality are "Zipf's Law," "Power Law," or "long tail" distributions (Anderson, 2006; Barabási, 2003; Huberman, 2001). These terms all point to the fact that online phenomena, be it amount of contributions to a user-generated site, traffic, overall attention or usage share are highly skewed (Huberman, 2001). Some may say that the figures are nominal, and this would not be untrue. Nielsen (2006) proposed for the online environment a so-called "90/10/1" rule, which is probably closer to the truth, as Wikipedia research confirmed (Matei & Bruno, 2015): 90 % of users are mere consumers of content, 10 % contribute some time, while 1 % are responsible for the bulk of the contributions. Yet, even this radically skewed distribution would not alter the core idea, namely, online interaction can and is skewed.

However, this phenomenon is not native to computer-mediated environments. Seminal studies of small discussion groups ranging in size showed that top contributors dominate the conversation to the tune of 40–50 % of the time, with the next participator coming in at a percentage in the teens, and all those that follow generally registering below 10 % of the total (Bales, 1950; Stephan & Mishler, 1952). This suggests that human interactions tend to follow a skewed output and input allocation curve. While part of such skewness can be tracked to power, privilege, and control issues, much of it can be put under the rubric of functional differentiation of roles and tasks (Bailey, 1990; Matei & Bruno, 2015). Any task-oriented group needs to allocate roles, rewards, responsibilities, and workloads. Allocation involves a coordination mechanism, attendant communication processes, implementation schedules, and so on. These work best when redundancies are minimized and activities are distributed according to the nature of the task and to individual qualifications. These processes result in uneven distribution of individual input and output. Thus, a significant part of group inequalities can be tracked down to the functional requirements of forming human groups.

While the reality of uneven online collaboration and its impact is an undeniable fact, its ultimate theoretical explanation is still insufficiently understood. To some online activists and media observers, who for the past decades have promoted the idea of cyberspace as a liberating and equalizing force (Barlow, 1994; Benkler, 2007; Hiltz & Turoff, 1978; Raymond, 2001; Tapscott & Williams, 2006), these findings might appear as phenomena of less importance than purported peer-production processes that encourage egalitarian participation (Benkler, 2007). Yet, this opinion might ignore an important argument. As groups increase in size, they meet the hard barriers of mounting transaction costs. When narrowly defined, such costs are the financial expenditures associated with social and economic exchanges. When broadly understood, transaction costs are the energy, time, or financial resources spent on maintaining a group's coordination and communication mechanisms (Coase, 1937; Surowiecki, 2004). In the absence of hierarchies and division of labor, group members need to constantly survey all the other members and communicate with them to keep the project going. This takes more and more attention and resources, which as the group increases in size can undermine its ability to subsist as a whole. The typical solution to this problem is to create specialized roles and coordination mechanisms, which allow some of the members to work on the intended group goal, while other members manage the collaboration process. It is also only fair to note that highly hierarchical and strictly compartmentalized groups, with tightly defined divisions of labor, can run into problems of their own. The most prominent is that of inefficient utilization of resources, poor allocation of effort, and inability to fully capture and redistribute local or tacit knowledge throughout the organization (Coase, 1937).

The dilemmas of human collaboration were neatly captured in the seminal work "Wisdom of Crowds" (Surowiecki, 2004). Although sometimes understood as an argument for flat organizations and egalitarian collaboration, the book makes a more complex point. It highlights the fact that task-oriented social groups work optimally when combined with a high degree of autonomous decision supported by flexible methods of aggregating and communicating information about group processes. Groups are, according to Surowiecki, more likely to come to right solutions when sufficient diversity of opinion, expertise, and interest is combined with social structures and communication tools that can aggregate these opinions and experiences and make them visible to the group in an effective way. Extending Surowiecki's phrase, we propose that for groups to be wise, they need division a labor, role allocations, and the communication tools and channels that allow them to become aware of their own inner working. Furthermore, self-awareness can be enhanced if information refers not only to the task and its completion rate, but also to the manner in which its outcome is produced. Given the uneven and socially structured nature of human tasks already discussed, it is especially important that information aggregation systems communicate in an effective manner how effort has been allocated, who has done what and to what effect. While this can be accomplished in many ways, the ideal situation would be one where such information reflects both global and individual facets of collaboration. In what follows we will present a methodological approach and online tool for monitoring and fostering group collaboration, especially in a learning environment. The tool provides information about the level of collaborative evenness and group structure through charts and colors that reflect group entropy levels. In addition, the tool is meant to facilitate our understanding of how uneven collaboration influences group effectiveness especially in a learning environment.

3 Measuring Collaborative Unevenness

3.1 Shannon's Entropy Theory

In previous work (Matei et al., 2006) we have proposed Shannon's Theory of Communication (Shannon & Weaver, 1949) as an approach and its companion measure, social entropy, as a possible measure for understanding collaboration within online and/or technological systems, especially wikis. Shannon used the social entropy index to capture the degree to which a communication system contains information (Shannon & Weaver, 1949). To accomplish this, Shannon employed a well-known physics measure, entropy, which is connected to the second law of thermodynamics, that states that all physical systems have a tendency to devolve to the point where the level of energy is zero and all their elements are equally likely to be in a random state. Shannon took the entropy measure from the physical to the communicative and as we will show below, to the social realm. His novel proposition was that communication can be conceived in terms similar to those of a physical system. In nature, when all elements of a system (e.g., atoms) occur randomly, their prevalence is approximately equal. The system is in a state of chaos and entropy is at a maximum. When physical particles get organized in more and more complex compounds, which privilege some elements at the expense of others, entropy decreases.

Communication can be seen as a system as well. Symbols, similar to atoms in the physical world, are the basic units. A communication system will probably contain no information and its entropy will be at a maximum when symbols are equally likely to occur. In other words, when the order of the symbols is decided by chance alone, there is no information (Shannon & Weaver, 1949). On the other hand, information-laden communication will utilize specific units of meaning more often than others, and entropy will decrease as symbols, just like physical particles, occur in a biased manner (Seife, 2007). Thus, if applying the entropy formula to a communicative system, the less organized it is, the higher the entropy and the less likely to contain information. The opposite is also true—the more organized the system, the higher the amount of information, and the lower the entropy.

3.2 *Social Entropy Theory*

Shannon's theory can be extended further, from communicative to social interaction. If we consider communication broadly, as the main mechanism by which social interaction takes place, all human affairs can be understood through the exchange processes that make them possible. Social interaction can be seen as an extended process of communication reliant upon a system of symbols and can be studied through the lens proposed by Shannon. Social systems whose members interact with each other in a nearly random manner, quasi-egalitarian, are more likely to lack a definite structure. Social systems that form a specific structure of interaction, where symbols are exchanged according to specific rules and patterns possess a more definite, structured form. Moreover, while in the first situation the exchanges will be completely even in terms of output/input ratios (everyone is equally likely to send symbols to everyone else), in the second case there will be a definite bias in terms of who will send information to whom.

From a mathematical or statistical perspective, social entropy measures to what degree specific system units (individuals) are more likely to contribute to or in the workings of the system than what chance alone would predict. The social entropy of a group is maximized when a group member is just as likely to communicate, share the effort or contribute an output unit as any other member. In statistical terms, for each of them, contribution would not be greater than what chance alone would predict. It would be purely random. On the other hand, as members take upon themselves or are assigned specific tasks and communicate in a patterned way by interacting in a preferential manner with other members, frequency and amount of output or contribution become non-random. Chance alone cannot predict these outcomes. Entropy, when measured as likelihood of individuals to contribute randomly, starts to decrease. When non-random behavior emerges, however, we have more than simple unevenness and deviation from what chance alone would dictate. Patterned interaction goes hand in hand with roles, rules and division of labor or functional differentiation. The group has become, in fact, structured. More concisely, a social group is more structured when its members are organized in a specific chain of communication and coordination, where some interact more than others, and less structured when members interact randomly (thus, theoretically, equally) to each other. Calculating the entropy of each social situation reveals in fact how structured the group is. Structure is inversely proportional to entropy.

4 Entropy: A Higher Level Structural Indicator

As previously mentioned, groups that are dominated by some of their members are also more likely to have a given structure. This structural characteristic can be captured in a direct way by social entropy: top heavy groups have lower, while egalitarian groups have higher, entropy levels. In this we take a cue from Shannon's original intent in proposing social entropy as a measure for how "informed" (organized) a social (communicative) reality is.

In extending Shannon's theory from information to other realms of inquiry, we continue a line of work with a distinguished past. For example, social and communication scientists, such as Hiltz and Turoff (1978), Schramm (1955) or Bailey (1990) have applied entropy theory and its attendant methodologies to specific social scientific problems, such as small group structuring, system theory, media landscape organization, diversity of media production, and so on. Economists, environmental scientists, or human geographers have also used entropy to characterize the social structure and diversity of industries, occupations, species, or populations (Bailey, 1990; Matei et al., 2006).

In our own work we have analyzed the emergence of social structures on Wikipedia utilizing articles as systems, contributors as system units and their amount of contribution as means for characterizing "system states" (Matei, Braun, & Petrache, 2009; Matei & Bruno, 2015). Calculating the degree to which contributions to Wikipedia articles are random or not, we observed that such contributions tend to be generated by a relatively small group of logged in contributors. Using entropy as a synthetic measure of contribution bias we found that article specific entropic contribution values tend to decrease and to reach a plateau after the 500th editorial intervention. Furthermore, even after this point, entropy keeps decreasing steadily, although at a slower pace. In other words, after the 500th editorial intervention the structure of collaboration within an average Wikipedia article is dominated by a relatively small number of users whose influence keeps increasing at small but steady pace. At the level of the entire Wikipedia space, entropy varies widely at the beginning of the project (2001–2002), reached a peak in 2005, and reached a steady state by 2006 (Matei & Bruno, 2015). Overall, 1 % of Wikipedia users generate 77 % of content. This reflects findings of similar research, such as of Ortega et al. (2008), who found that less than 10 % of Wikipedia members contribute up to 90 % of content, a trend that has dominated Wikipedia for the last several years.

Rooted in this scholarly tradition and building upon our own research, we propose that social entropy could be used to measure how structured or unstructured a group is. More specifically, we reformulate Shannon's theory of information to suggest that:

1. Information and "structure" go in the opposite direction of entropy;
2. Information and structure, especially in the social realm, are intrinsically connected; and,
3. Structure (of a language, symbol system, or group organization) can be measured with one synthetic indicator, namely entropy.

We emphasize the connections between social entropy and structure because groups are more than mere aggregations of people who share the same space. A group is the structure of ties between its individuals. Individuals that occupy specific roles in this structure communicate, contribute or interact in a specific way. The distribution of outputs in the group will follow the curve of abilities, productivity, task and power allocation specific to each role. Employing Shannon's entropy measure to describe group efforts, communicative patterns and collaborative patterns, we expect that as a group becomes more structured (i.e., roles emerge, tasks are assigned or assumed, power and information starts flowing from specific nodes to other nodes), imbalances in the distribution of communication or work will appear.

In other words, as the group starts to form and its structure to emerge, group units (individuals) start behaving in a predictable and non-random way. This predictable pattern entails a specific amount of unevenness. It is important to mention that "specific" has no normative meaning in our research. We have no a priori preference for any given level of unevenness, nor do we think that unevenness is demanded by "natural," individual characteristics. Rather, we propose that unevenness, while ever present, is a dynamic group process. Any group member can theoretically occupy any level of contribution or interaction. For each group and type of structure, some of which can be flatter while other more hierarchical, there is a "specific" level of unevenness and social entropy that needs to be observed and explained, not predicated.

5 Visible Effort: A Technology for Moderating Wiki Collaboration

In what manner can social entropy be employed for building and employing online collaborative tools? We use entropy in a collaborative tool, built on top of a wiki platform that communicates in a direct and active way, i.e., how even or uneven the collaborative efforts of any given group is at any specific point in time. Specifically, Visible Effort measures and displays entropy levels and, as discussed above, group structure. Entropy and structural information are funneled directly back into the collaborative process, or delivered to the group moderators or administrators (who can monitor and direct the process in a proactive manner).

The Visible Effort tool, used with a wiki, has the ability to measure and monitor on a continuous basis the degree to which a group is structured. If needed, it can also be used to maintain collaborative work within certain levels of equitability and evenness. Thus the tool serves a double purpose. On the one hand, it can be used as a monitoring tool, for understanding how collaboration is structured. On the other, it can be employed for adjusting collaboration along particular parameters desired by the instructor or site administrator.

Visible Effort is powered by a Mediawiki extension. Mediawiki is a content management system, originally designed to power Wikipedia, through which content can be edited by any user, including non-registered ones. All changes are permanently stored, and access to information that was edited or added is instantaneous. In addition, all pages come with "talk" areas, which allow discussions and interactions about the editing process. This makes it well adapted for collaborative work, especially of a textual nature.

The fact that all contributions of all users are preserved, regardless of whether they still exist in the current version of the text or not, facilitates an ongoing analytic process that can tell, for each point in time, how even or structured the process of collaboration is. This is accomplished by counting the number of characters that each user has contributed to the document. This count may also include credit for images or other types of content, depending on the option chosen by the administrator. There are two counts that may be utilized. The *gross* contribution uses the total number of words the user has contributed over the document's entire life, whether those words have survived into the current version or not. The *net* contribution is the count only of contributed words that exist in the current, or latest, version of the document. Once calculated, these values are stored by Visible Effort for each revision of the document, so that users can view the contribution scores for any past version of the document.

To process any particular revision for word counts, Visible Effort retrieves the wiki-markup pages for the current and the immediately preceding revision, converts them to plain text, and stores them in files. A UNIX utility is used to compare the files on a word-by-word basis. A difference value calculated for each specific version is assigned to each user and saved in the wiki database. These values are then used for calculating entropy values. Entropy values are then used to shape the page layout using easily comprehensible conventions. The goal is to provide "at-a-glance" information about the collaborative process. As collaboration becomes more (or less) even, background colors change and the graph indicates the size of the collaborative group and who has done the most work so far. In this way the cognitive effort involved in comprehending the project's collaborative status is dramatically minimized.

Key visual elements of the collaborative space (page) are formatted using visual cues that communicate the project status through a diversity of measures. The visual elements include text frames of specific colors and interactive displays (charts). Of these, the most important is the frame that surrounds the page, which changes colors/shades according to the entropy value of each page version that is displayed at a particular point in time—the colors darken or condense as the level of entropy increases. This communicates, at a glance, to the instructors and to the users how even (or structured) the collaboration process currently is. When the color is the lightest, the collaborative effort should be assigned to only one member of the team, thus entropy is 0. When the color is the darkest, there is perfect equality (evenness/high entropy = 100). In addition, there is a chart that visually reflects the distribution of effort for each collaborator as well as tabular information that reflects the number of words or characters contributed by each individual. The system allows electing to visualize or not visualize the entropy levels of each given page, according to the manager or instructor's preferred strategy. Administrators can use the entropy level as a direct indicator for the users, who would be able to see how even or balanced the collaborative effort is. Or, they can hide the information from the users, who would work blindly. Managers or instructors would only send textual and verbal messages to participants about their level of contribution or, given the data provided by VE, they could alter or improve the assignment while it is underway.

6 Use Scenario

Online collaborative learning is in many situations a very effective educational tool. For example, researchers continue to examine the possibility of how distance-learning within virtual worlds, like Second Life, fosters socialization (Kehrwald, 2008), while providing virtual spaces for exploration and creativity that enhance the collaborative learning experiences. Such community learning spaces foster interaction and intrinsic motivation while discovering new knowledge (Faiola & Smyslova, 2009). Moreover, the notion of intrinsic motivation has significant implications for researchers interested in understanding what occurs when the learning activity and environment elicit motivation in students. This is seen when the goals and rewards of learning are meaningful or when the learning assists the learner in obtaining valued accomplishments (Brandt, 1995; Chance, 1992).

Yet, it is inarguable that within groups some individuals have more to offer, others less, and teachers are intimately aware of this reality. If left to his own devices, Stephen, a motivated academic star, may do more than his fair share in the project. Clearly, though meant to benefit the group, teachers would be misguided in stifling his contributions in an attempt to bring them down to the level of the others. Likewise, Sally, a reserved, shy student, may have something of value to contribute to the group even though her participation efforts might not seem overtly active or significant (Lave & Wenger, 1991). Such a perspective is congruent with constructivist learning theory (Vygotsky, 1978), which emphasizes the social nature of knowledge and learning. It is expected that individuals will learn more when interacting with others, because they will be able to construct knowledge socially. Furthermore, collaboration need not be perfectly egalitarian to be successful.

Groups lead by the best students, who contribute above average, tend to perform better (Webb, Nemer, Chizhik, & Sugrue, 1998).

Our current usage of Visible Effort is situated in this constructivist context. Visible Effort aims to foster smart user choices and interactions, as well as instructor interventions, all guided by knowing how even or uneven the collaborative process is. At present, the extension is used in a number of research activities that aim to better understand the advantages and disadvantages of collaborative learning. Another motive behind our research is that while the positive effects of structured collaborative learning (Johnson & Johnson, 1999; Slavin, 1996) have been well known for some time, previous research seemed to explain this in view of individual attributes (Dillenbourg, Baker, & Blaye, 1996). Synthetic group measures that capture general level of structure in the manner proposed by our interpretation of Shannon's social entropy have rarely been used in research on group learning. In addition, equality of effort seemingly has always been the assumed goal. While our research agenda makes no specific point whether this preference for a normative state of equality arose from value-laden positions or not, we do propose that complete evenness of effort would rarely be an ideal operational state of interaction.

To test this proposition, we have devised a quasi-experimental program, which utilizes VE. Students are tasked to create group reports and term glossaries that are incorporated in class assignments. The main goal is to empirically identify the degree to which collaborative evenness promotes learning or not. Learning is measured as acquisition of knowledge related to specific concepts and theories discussed in the group reports and glossaries. Our main contention is that learning outcomes improve as groups become differentiated. As members start contributing according to their level of knowledge and learning needs, a specific social structure of learning emerges. This structure offers each student a given role and comfort zone. Consequently, students will contribute in different ways, according to their needs, abilities, and motivations. The groups they participate in will be characterized by a specific level of collaborative differentiation and unevenness that will go hand in hand with a specific level of learning effectiveness. We further hypothesize that the relationship between learning outcomes, group structure and collaborative unevenness is curvilinear. If collaborative unevenness and its companion level of group structure reach the level where some of the group members constantly dominate the collaborative process or where too many members "free ride," learning is disrupted. Group processes are increasingly hindered by discussions and conflicts about optimal level of contribution, reward allocation, and equity. Collaboration slows down or even ceases. However, the inverse is also problematic. On the other extreme, collaboration can also become too even, wherein top performers may not be allowed to stand above the others and consequently raise performance of the group whole. In this context, we are interested in finding out to what degree making the level of collaborative evenness and group structure known to the group members through the visual cues provided by Visible Effort can maintain the group within optimal collaboration values.

Our ultimate goals are, thus, three: (1) to determine the range of collaborative unevenness within which collaboration and learning are optimal, (2) to uncover the inflection point where collaborative unevenness and group structure ceases to promote learning in an online environment, and (3) to understand to what degree visual feedback can be used for moderating group CMC behavior.

7 Aligning the Conceptual with the Actual

Bruno explored these issues in a study of learning gain through wiki interaction (2010). He focused mostly on the social dynamics, rather than on the visualization effects. He observed small groups of individuals (<10 members) working on finding answers to questions posited by the researchers. A total of 170 undergraduate students were organized into 23 groups and tasked to answer questions posted on the Visible Effort wiki. The questions were mainly focused on widely available information about university campus buildings, history, or traditions. The answers were purely textual. Respondents could add new data, edit, or replace existing information. Respondents were asked factual questions about the topics both before and after the activity. Comparing answers in pre and posttests, Bruno (2010) calculated a net knowledge gain score, which was averaged for each group. Comparing learning gain with group inequality, quantified as normalized entropy (observed entropy/maximum possible entropy) he detected a curvilinear relationship. The highest and lowest levels of entropy hindered knowledge gain at the group level. Knowledge gain was maximized at a level between these two extremes, although it should be said that it was toward the high end of the spectrum, where interaction was much more even than in non-experimental situations (Bruno, 2010).

8 Significance

Bruno's findings highlight the ability of online interaction spaces to foster learning and the need to monitor the level of structuration/entropy to better understand the optimal levels of interaction. Of course, this is best done in spontaneously occurring situations, rather than in experimental settings. Furthermore, research needs to be conducted on the impact of the interaction visualization of present level of structuration/entropy by the participants in the ongoing social interaction. Would such visualization help interactors better understand their current status in the collaboration process? Would the higher performing actors become more or less motivated by their presence in the contributor elite? Will the other contributors be positively or negatively impacted by their relatively lower position in the interaction hierarchy?

Above and beyond these questions, the Visible Effort wiki can be utilized by teachers or knowledge managers in a unique manner. It offers immediate individual and group-level participation feedback that can be passively or actively utilized. It is not proposed to take the place of other tasks any teacher must undertake in the way of student and group assessment, but is simply another powerful implement in the toolkit. This paper only scratches the surface of what is possible. If proven effective, the theoretical and technical applications of the ideas discussed here could conceivably be applied in countless ways, collaboratively utilizing countless emerging technologies. The tool might also be extended to other platforms, such as online management and writing environments (Google Docs, Microsoft Office Live or Zoho), where collaboration can be supported by group work on free standing documents, not directly connected to a wiki, and for any type of assignment. In identifying and isolating what constitutes optimal student collaboration, many different kinds of group projects with different intended goals and outcomes could be carried out—not only those of a cognitive nature. And of course, in that sense, many different forms of learning could also be conducted and measured through similar means. What is not in doubt is the significant benefit theory-driven technologies, such as the Visible Effort wiki, would offer students, instructors, and business organizations.

Acknowledgements The authors would like to thank the programmers Bill Schroeder and Shammi Diddla, as well as David Eisert and Terri Patterson who worked on the software code of the MediaWiki extension. Visible Effort is supported by a competitive grant awarded by the Teaching and Learning Technologies Unit at Purdue University.

References

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. New York: Hyperion.

Bailey, K. D. (1990). *Social entropy theory*. New York: SUNY Press.

Bales, R. (1950). A set of categories for the analysis of small group interaction. *American Sociological Review, 15*(2), 257–263.

Barabási, A. (2003). *Linked: How everything is connected to everything else and what it means for business, science, and everyday life*. New York: Plume.

Barlow, J. P. (1994, August 11). *Jack in, Young Pioneer*. Retrieved December 12, 2014, from http://w2.eff.org/Misc/Publications/John_Perry_Barlow/HTML/jack_in_young_pioneer.html.

Benkler, Y. (2007). *The wealth of networks: How social production transforms markets and freedom*. New Haven, CT: Yale University Press.

Brafman, O., & Beckstrom, R. (2006). *The starfish and the spider: The unstoppable power of leaderless organizations*. New York: Portfolio.

Brandt, R. (1995). Punished by rewards? A conversation with Alfie Kohn. *Educational Leadership, 53*(1), 13–16.

Bruno, R. J. (2010). *Social differentiation, participation inequality and optimal collaborative learning online* (Order No. 3444483). Available from Dissertations & Theses @ CIC Institutions; ProQuest Dissertations & Theses A&I. (858608194). Retrieved from http://search.proquest.com/docview/858608194?accountid=13360.

Chance, P. (1992). The rewards of learning. *Phi Delta Kappan, 73*, 200–207.

Coase, R. (1937). The nature of the firm. *Economica, 4*(16), 386–405.

Dillenbourg, P., Baker, M., & Blaye, A. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science*. Amsterdam: Pergamon Press.

Draper, J. V., & Blair, L. M. (1996). Workload, flow, and telepresence during teleoperation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1030–1035. Minneapolis, MN: IEEE Society.

Faiola, A. & Smyslova, O. (2009). Flow experience and telepresence in second life: Correlating pleasure, immersion, and interaction in virtual communities. In G. Salvendy and J. Jacko (Ed.), *Proceedings of the 13th International Conference on HCI*. San Diego, CA: Springer.

Hiltz, S. R., & Turoff, M. (1978). *The network nation: Human communication via computer*. Reading, MA: Addison-Wesley.

Huberman, B. (2001). *The laws of the web*. Boston: The MIT Press.

Johnson, D., & Johnson, R. (1999). Making cooperative learning work. *Theory into Practice, 38*(2), 67–73.

Kehrwald, B. (2008). Understanding social presence in text-based online learning environments. *Distance Education, 29*(1), 89–106.

Kelly, K. (1995). *Out of control: The new biology of machines, social systems and the economic world*. Reading, MA: Addison-Wesley.

Kittur, A., Chi, E., Pendleton, B., Suh, B., & Mytkowycz, T. (2007). *The Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie*. Presented at the 25th ACM SIGCHI, San Jose, Retrieved from www.users.cs.umn.edu/~echi/papers/2007-CHI/2007-05-altCHI-Power-Wikipedia.pdf.

Kuk, G. (2006). Strategic interaction and knowledge sharing in the KDE developer mailing list. *Management Science, 52*(7), 1031–1042.

Lampe, C., Ellison, N., & Steinfield, C. (2006). A Face(book) in the crowd: Social searching vs. social browsing. *Proceedings of CSCW-2006*, 167–170. New York: ACM Press.

Lave, J., & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. Cambridge, England: Cambridge Press.

Lease, E. (2007, January 22). *Wise crowds with long tails* [Web log message]. Retrieved from http://techessence.info/node/85.

Matei, S. A., & Bertino, E. (Eds.). (2014). *Roles, trust, and reputation in social media knowledge markets: Theories and methods*. New York: Springer.

Matei, S. A., & Bruno, R. (2015). Pareto's 80/20 Law and social differentiation: A social entropy perspective. *Public Relations Review, 41*(2), 178–186. http://doi.org/10.1016/j.pubrev.2014.11.006

Matei, S. A., Braun, D., & Petrache, H. (2009). The structure of social collaboration on Wikipedia. *Paper presented at Wikimania 2009*, Buenos Aires, August 26–28.

Matei, S., & Ball-Rokeach, S. J. (2001). Real and virtual social ties: Connections in the everyday lives of seven ethnic neighborhoods. *American Behavioral Scientist, 45*(3), 550–564.

Matei, S., Oh, K., & Bruno, R. (2006, Nov). Collaboration and communication: A social entropy approach. Paper presented at the *National Communication Association Annual Convention, San Antonio*. Retrieved from http://matei.org/ithink/wp-content/uploads/2008/12/communicationcollaborationentropyfinal.pdf.

Nielsen, J. (2006, October). *Participation inequality: Encouraging more users to contribute* [Web log message]. Retrieved from http://www.useit.com/alertbox/participation_inequality.html.

Ortega, F., Gonzalez, J., & Robles, G. (2008). On the inequality of contributions to Wikipedia. In *Proceedings of the 41st Hawaii International Conf on Systems Science*. Washington, DC: IEEE Computer Society.

Powazek, D. (2009, May). The wisdom of community. *A List Apart Magazine,* 283. Retrieved from www.alistapart.com/articles/.

Raymond, E. (2001). *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary*. Cambridge, MA: O'Reilly.

Rheingold, H. (2002). *Smart mobs: The next social revolution* (1st ed.). Cambridge, MA: Basic Books.

Schramm, W. (1955). Information theory and mass communication. *Journalism Quarterly, 32*, 131–146.

Seife, C. (2007). *Decoding the universe*. London: Penguin.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shirky, C. (2008). *Here comes everybody*. New York: Penguin Press.

Slavin, R. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43–69.

Stephan, F. F., & Mishler, E. G. (1952). The distribution of participation in small groups: An exponential approximation. *American Sociological Review, 17*(5), 598–608.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York: Doubleday.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics*. New York: Portfolio.

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Webb, N., Nemer, K., Chizhik, A., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal, 35*(4), 607–651.