THE ONLINE GARBAGE CAN:

CONTEXT AND ATTENTION IN ONLINE COMMUNITIES

Nicolai J Foss
iCRIOS--Department of Management and Technology
Bocconi University
Via G Roentgen 1
20135 Milano; Italy
Nicolai.foss@unibocconi.it

Lars Bo Jeppesen
Department of Strategy and Innovation
Copenhagen Business School
Kilevej 14, 3rd fl; 2000 Frederiksberg; Denmark

Francesco Rullani
Department of Economics and Business
LUISS Guido Carli
Viale Romania 32
00198 Rome, Italy
frullani@luiss.it

HOW CONTEXT AND ATTENTION SHAPE BEHAVIORS IN ONLINE COMMUNITIES: A MODIFIED GARBAGE CAN MODEL

Abstract

Online communities have emerged as important organizational forms, but there are many gaps in our understanding. In particular, researchers have mainly focused on individual-level drivers of behaviors in communities, while downplaying (formal, informal) context at various levels. We theorize that different dimensions of context (i.e., omnibus and discrete context) influence decision-making in online communities through mechanisms involving community members' attention. Specifically, context influences which problems members

perceive and which solutions they retrieve and apply, thereby shaping the process of matching solutions and problems. We derive four hypotheses about contribution behaviors in online communities and how such behaviors are influenced by context. The empirical setting for our study is the open-source software community. We find support for our hypotheses in a unique dataset that captures the behavior of 24,057 community members who used the SourceForge.net online platform from 2000 to 2002.

Keywords: Context, attention allocation, problem-solving view, online innovation community

INTRODUCTION

Online communities play an important role as productive and innovative systems in the modern economy.[1] These communities mobilize and leverage the forces of spontaneous self-selection into tasks and activities, and their organizing principles are increasingly being emulated within firm hierarchies as they engage in "crowdsourcing," "crowd involvement," "platforms," and "user communities" (e.g., Boudreau, 2012; Majchrzak & Malhotra, 2013; Nambisan et al., 2017). The drivers of behaviors in such communities have captured a great deal of research attention. Scholars have examined how behaviors are shaped by antecedents, such as certain forms of governance (e.g., Raymond, 1999; O'Mahony & Ferraro, 2007; Yoo, Henfridsson, & Lyytinen, 2010; West & O'Mahony, 2008), organizational dynamics (Butler, 2001; Faraj, Jarvenpaa & Majchrzak, 2011; Von Krogh, Späth, & Lakhani, 2003), modes of communication (Foss, Frederiksen, & Rullani, 2016) and, in particular, community members' incentives, abilities, and motivations (e.g., Lerner & Tirole, 2002; Wasko & Faraj, 2005; Lakhani & Wolf, 2005; Bagozzi & Dholakia, 2006).

However, these antecedents only capture part of what drives behaviors in online communities. The inherently "fluid" nature of these communities, which are characterized by self-organizing projects, voluntary contributions, and the near absence of formal hierarchies, organizational structures, and control, means that they represent a context for interaction that is quite different from that of traditional organizations (Johns, 2006; Lomi, Conaldo, & Tonellato, 2012). Organizational theory suggests that context is a strong force in shaping behaviors (Andreson, 1999) because it supplies stimuli in the work setting that attract, focus, and direct

[1] Open and collaborative problem solving and innovation have received attention from scholars conducting research under a variety of headings, such as commons-based peer production (e.g., Benkler & Nissenbaum, 2006), virtual organizations (Faraj, Jarvenpaa, & Majchrzak, 2011), open-source science (Lakhani *et al.*, 2007; Jeppesen and Lakhani, 2010), private-collective innovation (von Hippel and von Krogh, 2003; Stuermer, Spaeth, & von Krogh, 2009), community-based innovation (e.g., Lakhani & von Hippel, 2003), and a hybrid model of community product co-development between firms and external individuals (Dahlander & Wallin, 2006; Dahlander & Magnusson, 2005; Baldwin & von Hippel, 2009).

attention, which then influences behavior (March & Simon, 1958; Ocasio, 1997, 2011). However, there are many gaps in our understanding of how this particular and peculiar context influences behaviors in communities. Communities are increasingly important ways of organizing problem-solving and innovation. Moreover, as in firms (Andreson, 1999), context in communities can be influenced or even partially designed (Gulati et al., 2012; e.g., by activating or deactivating certain interaction channels, Foss et al., 2016; or purposefully using certain non-material artifacts rather than others, Rullani & Haefliger, 2013). Thus, in principle it is possible to shape behaviors in communities in desired directions. However, as we lack an understanding of the influence of context on behaviors in communities, it is unclear how this should be accomplished.

Accordingly, the aim of this article is to theorize and empirically examine the role of context in behaviors in online communities. A starting point for our theorizing is that, in many ways, these communities resemble the description of organized anarchy found in Cohen, March, and Olsen (1972: 1), as they have processes that "are not understood by [their] own members," and entail "fluid participation" and "self-selection to tasks." In this paper, we add to Cohen et al.'s (1972) formulation of the garbage can model by developing four hypotheses that link context and community behaviors, specifically contribution behaviors, in different ways.

Our first hypothesis is based on the notion that what Cohen et al. (1972) call "garbage" can be represented by an information stream monitored by an individual member of an online community. As such, what Cohen et al. (1972) call "garbage" corresponds to the broad macro or "omnibus" context (Johns, 2006) to which all members of the community are exposed. In this regard, we go beyond the original garbage can model by arguing that community members' contribution behaviors are positively related to their exposure to the community-wide information

stream.[2] In other words, if a community member is exposed to a great deal of community-wide information, that member will be more likely to offer solutions to problems emerging in the stream. We offer three interwoven mechanisms that can account for this main effect hypothesis. First, we recall the idea of the "energy" (Cohen *et al*., 1972) that community members have available for decision-making and acting. The larger the information stream surrounding a community member, the greater the "effervescence" (Collins, 2004) of the community environment and, in turn, the more the members' energy will be used for contributions. Second, the larger the pool of unpaired solutions and problems "floating" in the community information stream, the more problem-solution matching becomes likely. This provides a direct and costless type of contribution in which energy can be readily applied. Third, as energy comes from social exchanges among community members, it springs from identity sharing and a sense of belonging, which are manifested in reciprocity and pro-social behaviors within the community's reach. The latter activate members' contributions in response to evidence of others' contributions (e.g., of a larger information stream; Shah, 2006; Bagozzi & Dholakia, 2006).

Furthermore, we move beyond this initial addition to Cohen et al.'s (1972) model by developing three other hypotheses. In particular, we focus on more proximate or "discrete" contexts (Johns, 2006), which act as moderators of the influence from the more distant omnibus context. We specifically argue that the main effect relation is moderated by: 1) the number of individuals (i.e., sources of the information stream) present in the proximate community; 2) how those are organized in different "subcommunities" (and, therefore, how "siloed" or fragmented the community discussion is in terms of topics); and 3) the temporal dynamics of problem and solution

[2] Our empirical analysis deals with open-source software. In this setting, Cohen et al.'s (1972) garbage can take a plethora of forms, including suggestions; complaints; snippets of code; requests for features; debates on the ethos of the community; discussions of the elegance of certain programs; disputes over who should be project leaders; and information on certain topics, processes, members of the community, attitudes, coding styles, opinions, ideas, hints, issues, and views. Thus, it is anything that is stored in any form in a community's repository, and can be used by community members to formulate a problem, or to provide a solution or part of a solution.

matching. The latter is because decision-making temporarily depletes "energy," such that the main effect is negatively moderated by the completion of a major task (i.e., the release of a new product) in the community.

Thus, our theorizing adds at least four arguments on problem-solution matching (Lakhani & von Hippel, 2003; von Hippel & von Krogh, 2016) and attention allocation (Foss et al., 2016). First, it shows that these are linked to other mechanisms, such as energy dynamics, pro-sociality, and reciprocity. In this regard, we complement the "organized anarchy" representation of online communities offered by those who apply the lenses of the "garbage can" model to this phenomenon (e.g., Lomi et al., 2012). Moreover, we provide an original representation of the community interaction environment as structured in a series of "contextual layers" where one can clearly distinguish between an overall (omnibus) context and several proximity-based contexts, each layer acting in its own respect and in interaction with the others. We also show how attention plays a different role in each of these contexts, and how it is augmented, diminished, or reshaped by the interactions of the structure and pace of community communication and contextual elements like energy, pro-sociality, reciprocity, and problem-solution matching. Finally, our empirical exercise adds to our contribution regarding prosocial behaviors and reciprocity by employing a large observational dataset, a step that few other papers have proposed (e.g., Belenzon & Schankerman, 2015).

More specifically, we analyze a dataset that covers the contribution behaviors of 24,057 software developers ("community members") participating in the SourceForge.net online platform over a span of two years in the early 2000s. The setting and the choice of time period for our analysis offer a number of features relevant for gaining insights into project organization and developers' interactions in online communities. In particular, the setting and the choice of time period allow us to test our hypotheses on data that are largely free of company interference, which

could affect participants' behaviors, and of possible parallel communication on alternative social media, which would affect our ability to capture most solutions and problems surrounding each community member. Our hypotheses are supported by the data.

BACKGROUND AND HYPOTHESES

Context in Online Communities

Much research conceptualizes online communities as a novel kind of social context that is associated with particular social values as well as certain prosocial and intrinsic motivations (e.g., Benkler, 1996; Von Hippel & Von Krogh, 2003). Given such contextual attributes, this stream of research seeks to understand what causes individuals to join or exit online communities, contribute to those communities, and start or join projects in those communities (e.g., Wasko and Faraj, 2000, 2005; Von Krogh *et al.*, 2012; Belenzon & Schankerman, 2015). Research that explicitly grapples with the contextual characteristics of online communities and the resulting behaviors examines, for example, how the particular ethos of free,open-source software (FOSS) communities may trigger prosocial motivations in the community context (Coffin, 2006), the rules regulating online communities and other aspects of the formal organization of communities (e.g., Lessig, 2001; O'Mahony, 2003; Yoo, Henfridsson, & Lyytinen, 2010; Tiwana, Konsynski, & Bush, 2010; Majchrzak & Malhotra, 2017), and the organization and dynamics of projects in online communities (e.g., Giuri, Ploner, Rullani, & Torrisi, 2010, 2008; David & Rullani, 2008; Kane, Johnson, & Majchrzak, 2014; Kyriakou, Nickerson, & Sabnis, 2017).

Context matters because it represents "situational opportunities and constraints" (Johns, 2006: 386).[3] It influences what individuals can do and how motivated they are to take these

[3] Johns (2006) also argues that contexts reside at two levels: the omnibus and the discrete contexts. The former refers to the macro, general features of the context, while the latter refers to specific, proximal influences on individuals. The macro-contextual features mainly influence behaviors through the mediation of the discrete context. Stimuli embedded in the more proximal context usually have stronger behavioral implications than those embedded in the more embedded context. We make use of this distinction in the following.

actions. However, and of particular interest here, context also shapes the individuals' *attention*—that is, the acts of noticing, encoding, and interpreting to make sense of issues (Ocasio, 1997)[4]—as it transmits stimuli and organizes those stimuli in various ways. The allocation of attention in organizational contexts has been a subject of study since Simon (1947) argued that organizations shape decision premises by defining and allocating the stimuli that channel attention (see also Ocasio, 1997).

In comparison to formal, hierarchical organizations, online communities possess rather rudimentary organizational structures. Participation is voluntary, which does not align well with the direct, formal organizational-control mechanisms associated with behavioral control. Organizational structures are fluid and mainly manifest in projects that are formed by entrepreneurial community members into which other members can self-select (Foss *et al*., 2016). In fact, in key respects, online communities resemble what Cohen et al. (1972: 1) call "organized anarchies," that is, "collections of choices looking for problems, issues and feelings looking for decision situations..., solutions looking for issues..., and decision makers looking for work."[5]

We would expect context to influence attention in communities differently than it does in more traditional hierarchical forms of organization because of the former's similarity to organized anarchies. In fact, Lakhani and von Hippel (2003) argue that contribution behaviors in communities often manifest as offering well-tried solutions that contributors already have in hand. The model works because of the significant heterogeneity among community members, whose

[4] Attention is introduced in the context of online communities in Hansen and Haas (2001), Huberman et al. (2009), Haas, Criscuolo, and George (2015), and Piezunka and Dahlander (2015).

[5] They further explain that such organizational forms are characterized by "problematic preferences… The organization operates on the basis of a variety of inconsistent and ill-defined preferences. It can be described better as a loose collection of ideas than as a coherent structure; it discovers preferences through action more than it acts on the basis of preferences" (Cohen *et al*., 1972: 1). Also, organized anarchies function on the basis of "unclear technology," which means that they have to function "on the basis of simple trial-and-error procedures, the residue of learning from the accidents of past experience, and pragmatic inventions of necessity" (Cohen *et al.,* 1972: 1). Finally, organized anarchies have "fluid participation. Participants vary in the amount of time and effort …; involvement varies from one time to another. … boundaries of the organization are uncertain and changing; the audiences and decision makers for any particular kind of choice change" (Cohen *et al.,* 1972: 1).

knowledge may offer solutions to many different problems (Rullani and Haefliger, 2013). As a result, contributors rarely spend more than a few minutes providing a solution to the problem presented. The broader implication is that, as in organized anarchies, solutions have a life of their own, which may be unrelated to the problems they might resolve.

The community information stream and member contributions

A key characteristic of electronically mediated communities is the low cost of communicating information. In principle, all community members can monitor the problem-solving process and volunteer information. We expect the sheer size (or "richness") of the community information stream to influence members' contribution behaviors for three key reasons: 1) the "energy" of community members, 2) the matching of problems and solutions, and 3) reciprocity triggered by pro-social behavior. We consider each in turn.

First, Cohen *et al*. (1972) posit that organizational members allocate "energy" to organizational decision-making. Energy is the "feeling that one is eager to act and capable of acting" (Quinn & Dutton, 2005: 36). In other words, it is an affective emotional arousal associated with positive moods or longer-lasting affective states (Quinn & Dutton, 2005). Membership in a decentralized community that is united in attempting to meet some overall objectives (typically prosocial in nature) is likely to bring about such "effervescence" (Collins, 2004). Thus, the more visible activity—the more "excitement"—there is in the community, the higher the level of effervescence and, therefore, the more energy that members dedicate to community activities, including the development of ideas and solutions.

Second, in allocating this energy, community members may follow the "vigilant attention" scheme (Ocasio, 2011) in which they observe the information stream and wait for a problem and a solution that can be matched—or stored for future possible matches. Research shows that if a text has some element of familiarity, an observer can retrieve knowledge that is immediately accessible,

a process that has been described as “knowledge activation” (Higgins, 1996). Community members who volunteer contributions may do so when they spot suitable solutions to problems they identified in the flow of information to which they have been exposed (Haas et al., 2015). Therefore, the cost of providing that solution is low. In other words, contributions triggered by energy may be driven by good matches between the solutions the community’s members already have available (i.e., those found and stored in the recent or distant past), and the problems and opportunities emerging from the information flow. Therefore, members will activate themselves when their solutions are “called upon” by the appearance of a problem posted in the community’s information stream. An increase in the number of solutions offered and problems revealed increases the likelihood that a community member will identify an opportunity for matching a solution to a problem.

In addition, although they do not specifically refer to communities, Von Hippel and Von Krogh (2016: 207) suggest that “in informal problem solving, a need and a solution are often discovered together and tested for viability as a ‘need–solution pair’.” Because of their fluid, informal structure and speed of communication, communities are particularly likely to manifest such simultaneity in problems and solutions in addition to traditional problem solving by means of problem formulation followed by solution search and problem solving by means of solutions searching for problems (as in Cohen *et al.*, 1972). A rich information stream not only energizes members but also turns that energy into the discovery of need-solution pairs.

Third, as argued above, energy springs from the richness of the social exchange in the community (Cropanzano and Mitchell, 2005), where relations evolve into a sense of belonging to the community and its goals (Bagozzi and Dholakia, 2006), and generate pro-social behavior geared, at least, toward other community members, thereby pushing community members to reciprocate (Shah, 2006; Lakhani and von Hippel, 2003). A rich information stream may ground

and then bolster norms of reciprocity, which may be further supported by the ease of communicating and sharing information among community members, and by the energy that can be achieved in a vibrant online community with a rich and plentiful information stream.

These three elements and their relations lead to the following baseline hypothesis:

> ***Hypothesis 1****: The likelihood of an individual contributing to an open and collaborative innovation community is positively affected by the size of the information stream to which the individual is exposed in that community.*

Attention allocation as a moderator of the relationship between the information stream and contributions

Although the costs of communicating information may be very low in an online community, the costs of monitoring ongoing activity in the community as well as the costs of absorbing, processing, and combining information may be substantial in terms of time and effort. Thus, Lakhani and von Hippel (2003) find that 98 percent of the time that community members spend on an online community's website is devoted to monitoring and only 2 percent is spent on providing contributions by answering questions. Relatedly, David and Rullani (2008) find that online "lurking"—monitoring an online community's activity without actively participating in it—is important in terms of time. Such "onlooker" behavior allows members to gather information regarding what is being debated in the community, a process through which members encounter specific opportunities and problems that motivate them to provide feedback (Rullani & Haefliger, 2013; von Krogh *et al.*, 2003) and, thereby, to become "de-lurked" (Soroka & Rafaeli, 2006). However, the extent to which this and other contribution behaviors happen depends on the costs discussed above and, therefore, on the attention that community members devote to monitoring the information stream. Factors that influence the amount of attention that community members can devote to the community as well as how that attention is allocated across diverse topics of interest in the community will also affect members' behaviors.

Selective attention and pro-social motivations. While members monitoring a community's stream of information are in "vigilance mode" (Ocasio, 2011), attention remains scarce, which means that information overload may become a problem (Lavie, 1995).[6] Community members adapt by making their attention more selective (Lavie, 1995). Members' attentional challenges (i.e., allocating attention in an economizing, selective manner) are related to the number of different and not always consistent information sources to which they are exposed. Moreover, a large set of information sources may generate more noise in the communication process because of a loss of quality and the presence of generally irrelevant information. In such a situation, information about problems and solutions may be more confusing (Krishnamurti, 2005), making matching more difficult. Thus, the positive association between a richer information stream and members' behaviors is likely to be reduced as community members reach their capacity to pay attention to, interpret, and digest information from a plethora of different sources (i.e., other community members). In other words, it becomes difficult for community members to follow and engage in discussions when information comes from many other members. This, in turn, will reduce their contributions.

In addition, the depletion of attention and confusion may deenergize the focal community member when other members, each offering a different perspective, are engaged in the focal member's main discussion arena (i.e., his or her largest project). This challenges the focal member's attention the most. The strength of pro-social motivations and reciprocity as motivators may also be reduced when proximate communities become larger, so that direct observation of pro-sociality may come across as very blurred and the sense of belonging to a shared identity may

[6] Simon (1971: 40-41) notes the general trade-off between information load and attention: "[I]n an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it."

be weakened (Giordani et al., 2018). This reasoning suggests a negative moderation effect of (main) project size on the relationship hypothesized in Hypothesis 1:

> ***Hypothesis 2:*** *The positive effect of the size of the community information stream on contribution behaviors is negatively moderated by the number of other community members involved in the focal community member's main discussion arena (i.e., his or her largest project).*

Fragmentation and problem-solution matching. Organization theory suggests that organizational structures serve to address problems of information overload. Thus, departmentalization allows for cognitive specialization (Simon, 1947; March & Simon, 1958). Smaller units imply fewer stimuli that compete for scarce attention. Organizational structures, in turn, serve to focus and situate attention in certain ways (Koput, 1997; Sullivan, 2010). Research finds that individuals who selectively monitor and read large quantities of information will *a priori* allocate more attention to categories of information that they perceive as more relevant (Reynolds & Anderson 2007), while they may exclude other (possibly relevant) domains. The potentially negative aspects of cognitive specialization are a lack of flexibility and a lack of awareness of developments in other areas. Therefore, situations that call for flexibility and coordination among units may greatly tax attention (Quinn & Dutton, 2005).

In communities, all information on projects, tasks, and issues, and all general discussions are typically open to everyone. However, communities also exhibit "departmentalization." More specifically, discussions are usually structured into different virtual spaces, such as mailing lists or forums, or threads on specific topics. Forums function as discussion rooms that bridge project members and those external to a project who are interested in it. In such an environment, threads typically originate from requests from the public, announcements from project members, or discussions regarding features of the code or the principles guiding software development (e.g., free software versus open-source

software). In contrast, mailing lists are typically used as discussion rooms solely for project members, such that discussions among developers are clearly divided from those involving users.

This structure facilitates lurkers' monitoring as well as participation by members interested only in specific discussions. It also creates boundaries between debates, and pushes community members to focus on certain discussions and to only work on certain tasks while disregarding others. Like departments in formal organizations, community subgroups tend to develop their own informal codes and norms. Learning becomes concentrated within rather than across departments or subgroups, which gives the few community members able to move from one subgroup to another a crucial boundary-spanning advantage (Dahlander & Frederiksen, 2012). As individuals are in vigilance mode (Ocasio, 2011), their attention becomes focused on the problems and solutions that are posted in the discussion stream they are primarily following. Problems formulated using specialized language further reduce opportunities for members to easily span different discussion streams, thereby restricting the number and variety of solutions and problems they can find, and decreasing opportunities for problem-knowledge matching. This means that attempts to departmentalize in the online context, regardless of whether they are planned, tend to reduce the number of actual matches that community members find between solutions. These efforts also reduce the number of problems "floating" in the information stream to which members are exposed. Hence:

> ***Hypothesis 3:*** *The positive effect of the size of the community information stream on contribution behaviors is negatively moderated by the degree of fragmentation in the community's discussion stream.*

Development stage and energy. While energy propels decision-making in organized anarchies (Cohen *et al*., 1972), such as communities, energy is depleted when attention and effort are utilized for problem-solving tasks. Certain key stages of development may involve extensive

technical discussions, which require individuals to increase their attention and contribute more during those phases. In other stages, such as after the launch of a final version of a product, members' attention may decline. More specifically, completion of a major task (i.e., a product release or a new version of a product) may temporarily diminish the contributions of community members for a period of time, such that they will be less attracted to the community discussion. As they will pay less attention to the discussion overall, they will not easily recognize problems or opportunities to which they might contribute. Therefore:

> ***Hypothesis 4*****:** *The positive effect of the size of the community information stream on contribution behaviors is negatively moderated by the individual's participation in the completion of a major task in that community.*

EMPIRICAL ANALYSIS

Data and model

To test our hypotheses, we relied on data provided by SourceForge.net (http://sourceforge.net, henceforth *SF.net*), the largest existing repository of open-source software projects (e.g., Belenzon & Shankerman, 2015; Foss *et al*., 2016) where participants can jointly develop new open-source software. We use a fixed effects logistic regression on the activity of 24,057 software developers over a period of 840 days (using 30-day windows) from September 2000 to December 2002 (i.e., a maximum of 28 windows). We used data from the early 2000s because firms played a much less pronounced role in the open-source software community at that time. This minimizes the interference of commercial strategies in our processes of interest.

Dependent variable

Our aim is to uncover what triggers a community member's activity in general, regardless of what that activity may be. To capture this, we defined three status variables that represent the wide spectrum of activities in which a community member can engage on *SF.net*: acting as a central member of existing projects, founding new projects, and acting as an external contributor to

other projects. We then identified actions that could capture these activities (i.e., posting news items and job requests; founding a new project; and contributing patches, bug reports, support, and new feature requests, respectively). We aggregate these actions into one unique dummy variable, *community member contributes*, which distinguishing between "no activity" and "some activity." This approach is conservative—at worst, we may underestimate contributions to the community, as some contributions may be in forms other than those captured by our variable.

Independent variables

Similar to Cohen et al.'s (1972) concept of "garbage," we defined our "information stream" as the various kinds of problems and solutions. In our specific case, this information stream takes the form of mailing-list messages, bug reports, patches, task lists, questions in forums, code and snippets, etc. (Rullani & Haefliger, 2013). In other words, this stream is highly heterogeneous. Therefore, finding a point measure of it would be nearly impossible. Instead, we detect the propensity of those who surround the focal individual (i.e., members working on *i*'s projects, "colleagues") to produce and diffuse information of any sort. We use this as a proxy for the amount of information *i*'s "colleagues" introduce in the environment surrounding *i*. The question is then the following: How can we observe this propensity?

-------- *Insert Table 1 and 2 here* --------

The data provide information about surveys launched on the *SF.net* website by members registered on *SF.net*. Table 1 provides an example, while Table 2 lists the most important surveys. When individuals participate in a survey, they do so because they are interested in sharing their views and opinions, flagging problems, and proposing ideas and solutions. Therefore, survey participation is closely related to an individual's general propensity to provide information about various kinds of problems and solutions even without a direct match to a real context or issue (e.g., even when the flagged problem does not yet have a solution or when the provided solution

relates to an unspecified problem). Keusch (2015) reports that the literature on the use of web surveys points to a higher inclination to respond in the presence of general participation-fostering traits, such as an investigative and enterprising personality, curiosity, openness to experience, social engagement, and interest in testing new web services. Therefore, the number of answered surveys is a good measure of the respondents' *propensity* to produce and diffuse information of many different kinds and, thus, a good proxy for the size of the information stream in the community.

Endogeneity is a key issue in this research. To avoid it, the information stream surrounding each participant should be not only *contextual* but also exogenous. In other words, it must be *independent* of member *i*'s influence. Although surveys can be launched by anyone, a careful analysis of the data suggests that the majority of the survey answers were provided in response to prompts from *SF.net*'s maintainers. Thus, the influence of a random participant in a random survey is likely to be none or negligible. To be sure that this was actually the case, we also introduced the dummy variable *community member in SF.net staff*, which was set equal to 1 if community member *i* belonged to the *SF.net* maintenance project.

Thus, we expect the number of surveys answered by member *i*'s colleagues (*number of surveys answered by colleagues*) to capture *i*'s colleagues' propensities to produce and diffuse information. As such, it serves as a proxy for the size of the information stream surrounding *i*. At the same time, we expect this variable to be independent of *i*'s influence, which considerably reduces endogeneity concerns (we further account for endogeneity in the next section).

Other variables

Hypotheses 2 to 4 include moderators. We compute the *number of community member's colleagues (largest projects)* to account for H2, the *projects' number of forum threads* and *projects' number of mailing lists* for H3, and the *projects' number of file releases (average)* for

H4. See Table 3 for details on these variables and the control variables. Tables 4, 5, and 6 cover the main statistics for these variables and their correlations.

-------- *Insert Table 3, 4, 5 and 6 here* --------

As Hamilton and Nickerson (2003) explain, if the most important dimensions leading to endogeneity are controlled for in a regression, the resulting coefficient is unbiased. Therefore, including the right controls can further reduce endogeneity problems. In our context, there are three main sources of endogeneity. When community members increase their contributions over time (i.e., become more productive), they may also: (1) *attract* participants who produce and diffuse more information, (2) *stimulate* the production and diffusion of information around them, and (3) *move* to projects in which members produce and diffuse more information. We consider each of these possibilities in turn.

(1) *Attract*: More productive community members may attract more individuals who produce and diffuse more information. For example, producing and diffusing more information *and* moving closer to a productive member could be a suitable strategy for getting closer to a crucial node in the network, even without many skills or a commitment to programming. However, individuals pursuing such a strategy would need to detect the productive members before they could get close to them. We ensure that this process cannot be a factor by lagging the independent variables, as a positive coefficient for *number of surveys answered by colleagues* represents a positive effect of *today's* communicative environment on members' *future* level of activity. The endogeneity mechanism described above also works in the other direction and can, therefore, be ruled out. Finally, we include community member *i*'s level of *previous* productivity as a control in the regression using the proxy *community member's number of artifacts*.

(2) *Stimulate*: Increasingly productive community members may also become more communicative and, thereby, stimulate the production and diffusion of information around them.

To account for this possibility, we "cleaned" the measure of the size of the information stream surrounding community member *i* of the effect of the member's own communicative attitude by using the *number of surveys answered by colleagues* as the main regressor. Moreover, we ruled out any possible residual influence of member *i*'s change in his or her propensity to communicate by introducing the *community member's number of forum messages*.

(3) *Move*: The final source of endogeneity is the possibility that increasingly productive community members will self-select into projects in which participants produce and diffuse a lot of information. If members increase their contributions or launch new projects, and then join other projects of a similar kind, the lagged structure explained in point 1 ensures that the estimates of the coefficients of *number of surveys answered by colleagues* are not affected. The analysis relates *today's* exposure to an information-rich context to *future* contributions.

For some community members, the propensity to contribute begins to rise and, thereby, increases their *future* contributions because *in the past* they have strategically chosen projects with specific characteristics (related to information) that allowed them to realize their augmented propensity to contribute. An efficient way to control for this mechanism is to introduce past projects' characteristics that: (a) community members could view as "instruments" they could use to increase their contribution, and (b) are correlated with the amount of information produced and diffused by the projects' participants.

To identify the main controls to be introduced for this purpose, we rely on five classes of participant motivations that span the whole space of incentives detected in the literature: *signaling and career concerns* (Lerner & Tirole 2002); *own use* (Lakhani & von Hippel, 2003; Shah 2006); *learning* (von Hippel & von Krogh 2003); *social motivations* (Bagozzi & Dholakia 2006); and *psychological motivations* (Wasko & Faraj, 2005 Lakhani & Wolf, 2005). First, participants whose propensity to contribute is increased may search for more visible (i.e., central in the

network), large, and productive projects in order to diffuse their work, and to send stronger and more visible signals to the labor market. Second, they may search for more productive projects because they are looking for collaborators to work on software they need for their own use. They may also search for productive projects because they want to learn from other community members and ask questions about the specific problems they need to solve in order to increase their own contributions. Eventually, their choices might be driven by more "ideological" considerations (e.g., they are willing to mainly contribute to projects that are committed to a certain view of the open-source phenomenon) or simply by their willingness to "have fun" programming with other productive participants.

Clearly, all of these mechanisms are positively correlated with the propensity of *i*'s colleagues to produce and diffuse information. As such, they could be sources of endogeneity. Fortunately, the dataset is large enough to account for all of the project characteristics upon which these mechanisms are based, as a comparison of Table 3 with our description of the mechanisms indicates. Consider, for example, learning and ideology. Community members can search for productive projects because they want to learn from other members. To control for this, we used the variables *projects' number of file releases (average)*, which accounts for the projects' productivity; the *number of the community member's colleagues (all projects),* which roughly counts the number of members involved in the projects; and the *community member's number of forum messages*, which represents the number of messages the participant sends to the forums. The latter can also be viewed as a proxy for the participant's requests for information. In the case of *ideology*, another control is introduced, *number of projects with GPL*, which reflects the number of projects to which the community member belongs that have chosen the GPL as their first license (the most restrictive license) (Gambardella and Hall 2006; Lerner and Tirole 2005). This serves as a proxy for the ideological level of the context of the community member's actions.

Finally, most activities in open-source projects (and the interactions among a project's participants) occur online. Thus, our data account for all possible activities undertaken in projects and by their participants. This allows for the construction of a wide spectrum of controls that should considerably reduce endogeneity.

RESULTS

Results and Robustness checks

We estimated four logistic regression models with fixed effects (Table 7).

-------- *Insert Table 7 here* --------

The coefficients of *number of surveys answered by colleagues* are positive and significant in all four models, which suggests strong support for H1. The negative and significant coefficient of *colleagues' answered surveys x number of colleagues (largest project)* in Model 2 also supports H2. H3 is also supported, as the coefficients of *colleagues' answered surveys x number of threads in the forums* and *colleagues' answered surveys x number of mailing lists* are significant and negative. Similarly, the interaction between *number of surveys answered by colleagues* and *projects' number of file releases (average)* confirms H4, as it is significant and negative (the results remain unchanged when we use CVS commits, i.e., the number of changes made to the code basis of the same project).

We applied bootstrapping (Greene, 2008: 596) with 50 subsamples of 2,000 observations each to evaluate the role of large sample size (Wooldridge, 2009: 135). Although most other coefficients show much higher p-values and the number of p-values equal to 0.000 is reduced to less than 50%, our main coefficients remain significant for H1, H3, and H4. We are cautious about assessing H2, as the coefficients of *colleagues' answered surveys x number of threads in the forums* and *colleagues' answered surveys x number of colleagues (largest project)* bear the

expected negative signs but with p-values of 0.111 and 0.163, respectively.[7]

With regard to the magnitude of the results, we computed the odds ratio (OR) of *number of surveys answered by colleagues*. In all models, the OR oscillates between 1.100 and 1.139, which means that when one colleague answers one more survey in *t*, the odds that community member *i* will make a contribution the next month increase by a factor of 10% to 14%. For the sake of comparison, all of the other variables in the main regression (except *community members in SF.net staff* for obvious reasons) have ORs implying changes that rarely reach 10%, with the highest being 18%. Thus, increasing the number of average file releases in member *i*'s project by one or enlarging the pool of member *i*'s colleagues by one—events that seem much more relevant than one more answered survey—lead to the same range of change as that observed for *number of surveys answered by colleagues*. Altogether, this means that the economic significance of our results is adequate for our task.

Finally, the computation of the variance inflation factors reveals that multicollinearity may be an issue. We re-ran our regressions after deleting the variables that generate the problem progressively and in different combinations. This procedure confirmed our results. Moreover, this check highlighted the stability of our estimates—the greatest change in coefficients between these estimates and those in Table 7 is no larger than 0.08 when testing H3 with projects' number of mailing lists, and smaller than 0.04 for all the other regressions.

CONCLUDING DISCUSSION

Contributions

In this paper, we studied an online community through the lens of the garbage can model (Cohen et al., 1972) to gain new insights into the contextual antecedents of behaviors in online communities. Problems and solutions are parts of a broad stream of information within the

[7] The results are available from the authors upon request.

community setting (. We argued for a positive relationship between the information stream that a member is exposed to and his or her contribution behavior (H1). The association is likely to be the result of the joint work of three mechanisms, all led by the size of the information stream: more intense social exchanges among community members, which lead to *pro-social* behaviors (at least within the scope of the community) and the establishment of *reciprocity* rules; the presence of a larger pool of unpaired problems and solutions, which increases the number of opportunities for their *matching*; and a more lively environment, which increases effervescence (Collins, 2004) and, therefore, the *energy* that community members are willing to dedicate to their joint activities. In other words, rich information streams and people of "good will" are key if "organized anarchies" are to work properly.[8]

We also argued that because development in this setting is dependent on community members self-selecting into tasks while monitoring the community's development activities, the extent to which contributions are triggered by revealed opportunities and problems is moderated by the attention members can allocate to various discussion streams. Our results suggest that variables related to attention allocation moderate the link between the extent of garbage and contributions. More specifically, H2 suggests that the interruptions, noise, and conflicting views that community members encounter as part of their involvement in projects with many sources of information (i.e., many members) affect their energy as well as their ability to identify appropriate problem-knowledge matches and to preserve their pro-social inclination to engage in reciprocity. This hypothesis is confirmed, but only to some extent, as it does not completely pass all of the robustness checks we applied.

Relatedly, the negative moderation effect found in relation to H3 (i.e., the negative impact of fragmented discussion topics) suggests that the structure of a project's discussion affects

[8] We are indebted to a reviewer who pointed us toward pro-social and reciprocity-based explanations, and their complementarity with energy and problem-solution matching.

members' abilities to monitor relevant areas of the community discussion. In this case, dividing a discussion has a negative effect on the relationship between the size of the information stream and members' contributions, thereby emphasizing the presence of a tradeoff between the benefits of specializing a discussion and the costs of reducing community members' abilities to monitor and uncover appropriate matches for their knowledge. Again, the weak support suggests that this negative effect is limited, probably by the positive effect of being in larger arenas where more problems and opportunities are apparent. Finally, we find that "available energy" matters, as predicted by the garbage can model (i.e., H4). In other words, the timing of the focal development activity plays a clear role in the open community's ability to activate developers' attention and foster problem-knowledge matches.

Overall, we make four key contributions to the extant literature. First, we augment the "organized anarchy" metaphor by discussing how energy (Cohen et al., 1972) interacts with other mechanisms, such as problem-solution matching, pro-social behaviors, and reciprocity. Our data do not allow us to empirically investigate this nexus, but we offer a representation of three mechanisms that present a more nuanced picture of how online communities work in terms of the "garbage can" model.

Second, we further explain why community members contribute to open, collaborative, problem-solving, and innovation communities by pointing to the role of (layered) contexts in communities. This explanation complements extant explanations, which are mainly motivation-based (e.g., Wasko & Faraj, 2005; Lakhani & von Hippel, 2003), and the more general body of research on open collaborative innovation communities (Faraj, Jarvenpaa, & Majchrzak, 2011; Kane, Johnson, & Majchrzak, 2014; Majchrzak, & Malhotra, 2017; Kyriakou, Nickerson & Sabnis, 2017).

Our third contribution lies in our theorizing about attention allocation in online communities

and in the empirical support we uncover for our arguments. We show that attention may be augmented, diminished, directed, and shaped in online communities by a series of contextual factors related to how crowded the interaction space is, how it is designed and structured, and how the dynamics of community activities evolve.

The last contribution is empirical in nature: we offer a test for a series of pro-social- and reciprocity-based mechanisms in online communities, which are interwoven with problem-solution matching and energy (Cohen et al., 1972). This test employs a large observational dataset in which the influence of firms is limited. To the best of our knowledge, this is one of the few studies (e.g., Belenzon & Schankerman, 2015) to do so. As such, it offers empirical insights complementary to those provided by similar studies based on experiments (e.g., Gachter, von Krogh, & Haefliger, 2010) and agent-based simulations (e.g., Levine & Prietula, 2014).

Limitations

While our study adds new insight to the discussion of contribution behaviors in innovation communities, explanations based on motivation (i.e., Constant, Sproull, & Kiesler, 1996; Wasko & Faraj, 2005; Lakhani & Von Hippel, 2003; Jeppesen & Frederiksen, 2006), attention, and the matching of problems and solutions are highly complementary. Additional research is needed to determine the precise roles played by these (and perhaps other) factors. For instance, scholars may wish to establish the relative importance of these factors, their potential sequences, and the degree to which they are intertwined.

Similarly, we identified three main mechanisms—pro-sociality, matching, and energy—whose joint work is crucial for communities wishing to continually fuel their members' participation. However, we could not clearly disentangle and analyze each mechanism on its own, as their joint functioning and the lack of specific data for each of them limited our insight. Additional research is needed to tease out the components of this nexus that we have placed at the

core of our theorizing, the existence of which we tested empirically. Such efforts may follow the lines drawn by Bagozzi and Dholakia (2006), who developed surveys to directly capture community members' motivations, or the "revealed preferences" approach introduced by Belenzon and Schankerman (2015), who created observable point measures for specific motivations, like reciprocity. While the former approach can be true to the complex set of motivations among community members, it runs the risk of offering only a blurred view of how those motivations are translated into actual behavior. In contrast, the latter approach is able to directly detect behaviors, but runs the risk of producing point estimates that are quite restrictive and, therefore, unable to account for more complex phenomena (e.g., the authors define reciprocity as "*do ut des*" between OSS project pairs). Future research must be very careful in balancing adherence to constructs with measurement precision.

Future research should also explore the relationship between context and behavior, and apply sharper measures of the former. This is challenging, as context is usually made up of a plethora of elements that are connected to the individuals who populate it. Consequently, any statistical evaluation of causality is affected by endogeneity. In this study, we tackled this challenge by building a measure of context capable of reducing endogeneity. Moreover, by coupling it with specific controls, we took steps to keep endogeneity under control. However, this makes it impossible to have a direct measure of context. More research is needed to more directly capture the elements surrounding innovators and community members that tries to ameliorate the tradeoff between the exogeneity of the measure and its proximity to the construct.

REFERENCES

Afuah, A. N., & Tucci, C. L. 2012. Crowdsourcing as a Solution to Distant Search. *Academy of Management Review*, 37(3): 355-375.

Anderson, P., 1999. Complexity theory and organization science. *Organization Science*, 10: 216-232

Autio, E., Dahlander, L., & Frederiksen, L. 2013. Information Exposure, Opportunity Evaluation, and Entrepreneurial Action: An Investigation of an Online User Community. *Academy of Management Journal*, 56: 1348-1371.

Bagozzi, R. P., & Dholakia, U. M. 2006. Open Source Software User Communities: A Study of Participation in Linux User Groups. *Management Science*, 52(7): 1099-1115.

Baldwin, C. Y., & von Hippel, E. 2009. Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. Working paper 10-038 Harvard Business School, Boston, MA.

Baldwin, C., & von Hippel, E. 2011. Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. *Organization Science*, 22(6): 1399-1417.Bateman, P. Gray, P.H. & Butler, B.S. 2011, The Impact of Community Commitment on Participation in Online Communities, Research Note, *Information Systems Research*, 22(4):841-854

Belenzon, S., & Schankerman, M. 2015. Motivation and sorting of human capital in open innovation. *Strategic Management Journal*, 36(6): 795-820.

Bendor, J., Moe, T. M., & Shotts, K. W. 2001. Recycling the garbage can: An assessment of the research program. *American Political Science Review*, 95(1): 169-190.

Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale University Press.

Benkler, Y., & Nissenbaum, H. 2006. Commons-Based Peer Production and Virtue. *Journal of Political Philosophy*, 14(4): 394-419.

Bjelland, O. M., & Wood, R. C. 2008. An Inside View of IBM's 'Innovation Jam'. *Sloan Management Review*, 50: 32-40.

Bleiklie, I., Enders, J., & Lepori, B. 2015 Organizations as Penetrated Hierarchies: Environmental Pressures and Control in Professional. *Organization Studies*, 36: 873-896.

Boudreau, K. 2010. Open Platform Strategies and Innovation: Granting Access vs. Devolving Control. *Management Science*, 56(10): 1849-1872.

Butler, B. S. 2001. Membership Size, Communication Activity, and Sustainability: The Internal Dynamics of Networked Social Structures, *Information Systems Research* 12: 346-362.

Butler, B.S, P.J Bateman, P.H Gray & E.I. Diamant.2014. An attraction-selection-attrition theory of online community size and resilience. *MIS Quarterly* 38: 699-728

Cappelli, P. &. Sherer, P.D. 1991. The Missing Role of Context in OB: The Need for a Meso-Level Approach. *Research in Organizational Behavior*¸ 13. 55-111.

Coffin, J. 2006. Analysis of open source principles in diverse collaborative communities. *First Monday*, 11(6). http://firstmonday.org/issues/issue11_6/coffin/index.html.

Cho, T. S., & Hambrick, D. C. 2006. Attention as the mediator between top management team characteristics and strategic change: The case of airline deregulation. *Organization Science*, 17(4): 453-469.

Cohen, M. D., & March J. G. 1974. *Leadership and Ambiguity: The American College President.* Hightstown, NJ: McGraw-Hill.

Cohen, M. D., March, J. G., & Olsen, J. P. 1972. A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly*, 17(1): 1-25.

Cohen, M. D., March, J. G., & Olsen, J. P. 2012. "A garbage can model" at forty: A solution that still attracts problems. In Lomi, A., & Harrison, J. R. (Eds.), *The Garbage Can Model of Organizational Choice: Looking Forward at Forty* (Research in the Sociology of Organizations, Vol. 36): 19-30. Bingley, UK: Emerald Group Publishing Limited.

Collins, R.2004. *Interaction ritual chains*. Princeton: Princeton University Press.

Cropanzano R., Mitchell M.S. (2005), Social Exchange Theory: An Interdisciplinary Review, *Journal of Management*, 31(6), 874-900

Cyert, R. M., & March, J. G. 1963. *Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Wiley-Blackwell.

Dahlander, L., & Frederiksen, L. 2012. The Core and Cosmopolitans: A Relational View of Innovation in User Communities. *Organization Science*, 23(4): 988-1007.

Dahlander, L., & Magnusson, M. G. 2005. Relationships between open source software companies and communities: Observations from Nordic firms. *Research Policy*, 34(4): 481-493.

Dahlander, L., & Wallin, M. W. 2006. A man on the inside: Unlocking communities as complementary assets. *Research Policy*, 35(8): 1243-1259.

David, P. A., & Rullani, F. 2008. Dynamics of innovation in an "open source" collaboration environment: lurking, laboring, and launching FLOSS projects on SourceForge. *Industrial and Corporate Change*, 17(4): 647-710.

David, P. A., Waterman, A., & Arora, S. 2003. FLOSS-US: the free/libre/open source software survey for 2003. Preliminary draft quoted with permission of authors. Stanford Institute for Economic Policy Research, Stanford University, CA. Retrieved February 6, 2017, from http://www.stanford.edu/group/floss-us/report/FLOSS-US-Report.pdf

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. 2011. Knowledge Collaboration in Online Communities. *Organization Science*, 22(5): 1224-1239.

Felin, T., & Zenger, T. R. 2014. Closed or open innovation? Problem solving and the governance choice. *Research Policy*, 43(5): 914-925.

Felin, T., Foss, N.J., & Ployhart, R. 2015. Microfoundations for Management Research. *Academy of Management Annals*, 9: 575–632.

Foss, N. J., Frederiksen, L., & Rullani, F. 2016. Problem-formulation and problem-solving in self-organized communities: How modes of communication shape project behaviors in the free open-source software community. *Strategic Management Journal*, 37(13): 2589-2610.

Franke, N., & Shah S. 2003. How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, 32: 157–178.

Gächter, S., von Krogh, G., Haefliger, S. 2010. Initiating private-collective innovation: The fragility of knowledge sharing, *Research Policy*, 39(7), 893-906

Giordani P., Rullani F., Zirulia L. 2018. Endogenous Growth of Open Collaborative Innovation Communities, *Industrial and Corporate Change*, 27(4), 745-762.

Giuri, P., Ploner, M., Rullani F., & Torrisi, S. 2010. Skills, division of labor and performance in collective inventions: Evidence from open source software. *International Journal of Industrial Organization*, 28(1): 54-68.

Giuri, P., Rullani F., & Torrisi, S. 2008. Explaining leadership in virtual teams: The case of open source software. *Information Economics and Policy*, 20(4): 305-315.

Glynn, P.W., Greve, H.R., & Rao, H. 2020. Relining the garbage can of organizational decision-making: modeling the arrival of problems and solutions as queues. *Industrial and Corporate Change*, 29: 125-142.

Greene, W. 2008. *Econometric Analysis*, 6th Edition. Prentice Hall, NJ: Pearson.

Gulati, R., Puranam, P., Tushman, M. (2012). Meta-Organization Design: Rethinking Design in Interorganizational and Community Contexts. *Strategic Management Journal* 33, 571–586

Hansen, M. T., & Haas, M. R. 2001. Competing for Attention in Knowledge Markets: Electronic Document Dissemination in a Management Consulting Company. *Administrative Science Quarterly,* 46(1): 1-28.

Haas, M. R., Criscuolo, P., & George, G. 2015. Which Problem to Solve? Online Knowledge Sharing and Attention Allocation in Organizations. *Academy of Management Journal*, 58(3): 680-711.

Hart, S. L. 1992. An Integrative Framework for Strategy-Making Processes. *Academy of*

Management Review, 17: 327-351.

Hayek, F. A. 1945. The Use of Knowledge in Society. *American Economic Review*, 35(4): 519-530.

Heath, C., & Sitkin, S. 2001. Big-B versus big-O: What is Organizational About Organizational Behavior. *Journal of Organizational Behavior*, 2: 43–58.

Higgins, E. T. 1996. Knowledge activation: Accessibility, applicability, and salience. In Higgins, E. T., & Kruglanski, A. W. (Eds.) *Social Psychology: Handbook of Basic Principles*: 133-168. New York: Guilford Press.

Jeppesen, L. B., & Lakhani, K. R. 2010. Marginality and Problem-Solving Effectiveness in Broadcast Search. *Organization Science*, 21(5): 1016-1033.

Jeppesen, L. B., & Laursen, K. 2009. The Role of Lead Users in Knowledge Sharing. *Research Policy*, 38(10): 1582-1589.

Jeppesen, L. B., & Frederiksen, L. 2006. Why firm-established user communities work for innovation? The personal attributes of innovative users in the case of computer-controlled music instruments. *Organization Science*, 17(1): 45-64.

Johns, G. 2006. The Essential Impact of Context on Organizational Behavior. *Academy of Management Review*, 31: 386-408.

Kane, G. C., Johnson, J., and Majchrzak, A. 2014. Emergent Life Cycle: The Tension Between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities. *Management Science* 60(12): 3026-3048.

Keusch, F. 2015. Why do people participate in Web surveys? Applying survey participation theory to Internet survey data collection. *Management Review Quarterly*, 65(3): 183-216.

Koput, K. W. 1997. A Chaotic Model of Innovative Search: Some Answers, Many Questions. *Organization Science*, 8(5): 528-542.

Krishnamurthy, S. 2005. About Closed-Door Free/Libre/Open Source (FLOSS) Projects: Lessons

from the Mozilla Firefox Developer Recruitment Approach. *Upgrade*, 6(3): 28-32.

Kyriakou, H., Nickerson, J., & Sabnis, G. 2017. Knowledge Reuse for Customization: Metamodels in an Open Design Community for 3D Printing. *MIS Quarterly* 41(1): 315-332.

Lakhani, K. R., & von Hippel, E. 2003. How open source software works: "Free" developer-to-developer assistance. *Research Policy*, 32: 923-943.

Lakhani, K. R., & Lonstein, E. 2011. InnoCentive.com. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2014058.

Lakhani, K. R., & Wolf, R. G. 2005. Why Hackers Do What They Do: Understanding Motivations and Effort in Free/Open Source Software Projects. In Feller, J., Fitzgerald, B., Hissam, S., & Lakhani, K. R. (Eds.) *Perspectives on Free and Open Source Software*. Cambridge, MA: MIT Press.

Lavie, N. 1995. Perceptual Load as a Necessary Condition for Selective Attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3): 451-468.

Lerner, J., & Tirole, J. 2002. Some Simple Economics of Open Source. *Journal of Industrial Economics*, 50(2): 197-234.

Lerner, J., & Tirole, J. 2005. The Scope of Open Source Licensing. *Journal of Law, Economics and Organization*, 21(1): 20-56.

Lessig, L. 2001. *The Future of Ideas: The Fate of the Commons in a Connected World*. New York: Random House.

Levine S.S., Prietula M.J. 2014. Open Collaboration for Innovation: Principles and Performance. *Organization Science* 25(5):1414-1433.

Lomi, A., Conaldi, G., & Tonellato, M. 2012. Organized Anarchies and the Network Dynamics of Decision Opportunities in an Open Source Software Project. In Lomi, A., & Harrison, J. R., eds. The Garbage Can Model of Organizational Choice: Looking Forward at Forty. *Research*

in the Sociology of Organizations, 36: 363-397. Bingley, UK: Emerald Group Publishing Limited.

Lomi, A., & Harrison, J. R. 2012. *The Garbage Can Model of Organizational Choice: Looking Forward at Forty*. Bingley, UK: Emerald Group Publishing Limited.

Majchrzak, A., & Malhotra, A. 2017. Effect of Knowledge Sharing Trajectories on Innovative Outcomes in Temporary Online Crowds. *Information Systems Research*, forthcoming

March, J. G., & Simon, H. A. 1958. *Organizations*. New Jersey: Wiley-Blackwell.

Nambisan, S., Lyytinen, K., Majchrzak, A & Song., M. 2017. Introduction: Reinventing Innovation Management in a Digital World *MIS Quarterly*, 41(1): 223-238.

Ocasio, W. 1997. Towards an attention-based view of the firm. *Strategic Management Journal*, 18): 187-206.

Ocasio, W. 2011. Attention to Attention. *Organization Science*, 22(5): 1286-1296.

O'Mahony, S. 2003. Guarding the commons: how community managed software projects protect their work. *Research policy*, 32 (7): 1179-1198

O'Mahony, S., & Ferraro, F. 2007. The Emergence of Governance in an Open Source Community. *Academy of Management Journal*, 50: 1079-1106.

Piezunka, H., & Dahlander, L. 2015. Distant Search, Narrow Attention: How Crowding Alters Organizations' Filtering of Suggestions in Crowdsourcing. *Academy of Management Journal*, 58(3): 856-880.

Quinn, R. W., & Dutton, J. E. 2005. Coordination as Energy-in-Conversation. *Academy of Management Review*, 30: 36-57.

Raymond, E. 1999. *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. Sebastopol, CA: O'Reilly & Associates.

Rullani, F., & Haefliger, S. 2013. The periphery on stage. Properties, functions, and dynamics of the

periphery in the Free/Open Source Software community. *Research Policy*, 42(4): 941-953.

Shah, S. 2006. Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Management Science*, 52(7): 1000–1014.

Shane, S. 2000. Prior Knowledge and the Discovery of Entrepreneurial Opportunities. *Organization Science*, 11(4): 448-469.

Simon, H. A. 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Macmillan.

Simon, H. A. 1962. The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106(6): 467-482.

Simon, H. A. 1971. Designing Organizations for an Information-Rich World. In Greenberger, M. (Ed.) *Computers, Communication and the Public Interest*: 37-52. Baltimore, MD: John Hopkins Press.

Smith, M. A. & Kollock, P. 1999. *Communities in Cyberspace*. London, UK: Routledge.

Soroka, V., & Rafaeli, S. 2006. Invisible Participants: How Cultural Capital Relates to Lurking Behavior. Paper presented at the 15th International World Wide Web Conference WWW 2006, Edinburgh, Scotland.

Stuerme, M., Spaeth, S., & von Krogh, G. 2009. Extending Private-Collective Innovation: A Case Study. *R&D Management*, 39(2): 170-191.

Sullivan, B. N. 2010. Competition and Beyond: Problems and Attention Allocation in the Organizational Rulemaking Process. *Organization Science*, 21(2): 432-450.

Tiwana, A., Konsynski, B., and Bush, A. A. 2010. "Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics," *Information Systems Research* 21(4): 675- 687.

von Hippel, E., & von Krogh, G. 2003. Open Source Software and the 'Private-Collective'

Innovation Model: Issues for Organization Science. *Organization Science*, 14(2): 209-223.

von Hippel, E., & von Krogh, G. 2016. CROSSROADS—Identifying Viable "Need–Solution Pairs": Problem Solving Without Problem Formulation. Organization Science, 27(1): 207-221.

von Krogh, G., Haefliger, S., Spaeth, S., & Wallin M. W. 2012. Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development. *MIS Quarterly*, 36(2): 649-676.

von Krogh, G., Spaeth, S., & Lakhani, K. R. 2003. Community, joining, and specialization in open source software innovation: A case study. *Research Policy*, 32(7): 1217-1241.

von Krogh, G., & von Hippel, E. 2006. The Promise of Research on Open Source Software. *Management Science*, 52(7): 975-983.

Wasko, M., & Faraj, S. 2000. "It is what one does": Why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, 9(2-3): 155-173.

Wasko, M., & Faraj, S. 2005. Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, 29(1): 35-57.

West, J., & O'Mahony, S. 2008. The Role of Participation Architecture in Growing Sponsored Open Source Communities. *Industry & Innovation*, 15(2): 145-168.

Yoo, Y., Henfridsson, O., & Lyytinen, K. 2010. "The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research," *Information Systems Research* 21(4): 724-735.

Zbaracki, M. J. 2014. The garbage can model of organizational choice: Looking forward at forty. Organization Studies, 35: 454-459.