A microstructural approach to self-organizing: The emergence of attention networks

Marco Tonellato
University of Trento and LMU Munich
marco.tonellato@unitn.it

Stefano Tasselli
University of Exeter, Erasmus University and ESC Rennes School of Business.
s.tasselli@exeter.ac.uk

Guido Conaldi
University of Greenwich
g.conaldi@gre.ac.uk

Jürgen Lerner
University of Konstanz
juergen.lerner@uni-konstanz.de

Alessandro Lomi
University of Italian Switzerland
alessandro.lomi@usi.ch

In press in ***Organization Science***

ABSTRACT

A recent stream of theoretical development investigates new forms of organizing as bundles of novel solutions to the universal problems of organizing: how to allocate organizational problems to organizational participants, and how to integrate participants' resulting efforts. In this paper, we add to this debate by reframing organizational attention allocation as a concatenation of self-organizing, micro-structural mechanisms linking multiple participants to multiple problems, thereby forming an attention network. In particular, we argue that, when managerial hierarchies are absent and authority is decentralized, visible acts of attention allocation become a fundamental coordination mechanism by which participants provide information to each other on how to integrate their efforts. We theorize that the observed structure of an organizational attention network is generated by the concatenation of four interdependent micro-mechanisms: Focusing, Reinforcing, Mixing, and Clustering. In a statistical analysis of attention networks in a large open-source software project, we found support for the four hypotheses about the self-organizing dynamics of the observed network associating organizational problems (software bugs) and organizational participants (volunteer contributors). We discuss the implications of attention networks for theory and practice by emphasizing the self-organizing character of organizational problem solving. We discuss the generalizability of our theory to a wide set of organizations where participants can choose which problems to allocate attention to and where the outcomes of these choices are publicly visible.

INTRODUCTION

A distinctive line of theoretical development examines organizations as systems of coordinated activities performed by multiple participants to solve multiple problems (Cohen et al. 1972, Galbraith 1973, Thompson 1967, Puranam 2018, Raveendran et al. 2020). Typically, this line of research starts with the assumption that individual goals are often inconsistent, ambiguous, or simply untraceable, links between actions and consequences are unclear, and participation in problem solving is fluid and evolving (March and Olsen, 1976). Consequently, organizations are viewed as social arenas that strive to design coordination mechanisms that reduce these sources of uncertainty to stabilize problem-solving and achieve organizational functioning (Cohen et al. 1972, Nadler and Tushman 1997, Padgett 1980, Padgett, Lee and Collier, 2003, Ocasio, 2012).

Available research on traditional organizations suggests a variety of such stabilizing mechanisms of coordination, including hierarchy-based structures (Mintzberg, 1979), exogenous distribution of tasks and activities (Galbraith 1973), organizational routines (Becker and Knudsen 2005) systems of formal incentives (Kretschmer and Puranam 2008) and decision-making architectures (Christensen and Knudsen 2010). All these examples assume that organizations are purposefully designed to exert direct managerial control over basic organizing principles such as division of labor, integration of effort, and the exercise of authority.

Despite these efforts, we know significantly less about how problem-solving activities self-organize in absence of direct authority exerted through managerial hierarchy and predefined task-allocation decisions (Lee and Edmondson 2017, Puranam et al. 2014). In recent developments of organizational theory and design, the term "self-organizing" describes the emergence of organizational structure as the result of spontaneous, evolving and interdependent interactions among participants and problems in a condition of decentralized, loose or even absent authority (Lee and Edmonson 2017, Massa and O'Mahony 2021). Examples of self-organizing dynamics are frequent in today's work arrangements, ranging from crowdsourcing (Majchrzak et al. 2021) to digital encyclopedias (Arazy et al., 2016, 2020), corporate experiments with "holacracy" (Robertson, 2015) and open-source online software platforms - the specific setting that we examine in this study (e.g., von Krogh and von Hippel 2006).

How do organizations control the process of problem allocation to participants in absence of a formal hierarchy providing an exogenous source of authority? Furthermore, how do organizations provide crucial information to participants to guide them in coordinating their work in absence of a formal workflow system? We seek to answer these questions by examining directly how coordination emerges endogenously to stabilize the relation between fluid streams of participants and problems (Cohen et al., 1972). In particular we argue that, in new forms of organizing – which are often driven by distributed communities (Gulati et al. 2012), built on architectures for decentralized collaborations (Fjeldstad et al. 2012) and populated by problems seeking the attention of problem-solvers (Haas et al. 2015, Piezunka and Dahlander, 2015) – activities are coordinated due to the visible, public and transparent acts by which active participants allocate attention to available problems.

We conceptualize organizational attention allocation as a concatenation of self-organizing, microstructural mechanisms based on the mutual interdependence between participants and problems. While traditional views of organizational design treat interdependence as a condition involving participants (or agents) and problems (or tasks) as separate entities (Puranam et al. 2012), our work draws from a more general view of participants and problems in organizations as standing in a dual relation of mutual dependence (Breiger 1974, Breiger and Mohr 2004, Tasselli and Kilduff 2021). This mutual dependence arises because, at any given moment, acts of attention allocation between participants and problems generate informational cues and create contingencies that influence how other actors might allocate their attention further. Rather than the design of explicit mechanisms for integration of effort between participants, attention allocation itself is the stimulus triggering further behavioral responses from other participants, such that organizations are coordinated by the very same processes that their members perform (Padgett and Powell 2012).

Conceptually, this process of allocation of attention may be seen as embedded in a relational structure involving two categories of interdependent entities: 'participants' and 'problems' (Carley 1991). We represent this system as a rectangular matrix where rows are participants, columns are problems and individual cells contain information about participants allocating attention to problems. We call this system an *attention network*. Unlike prior attempts to understand organizational networks as the informal representation of social or workflow relationships between participants (e.g., Kilduff and Brass 2010), the focus of our argument is that attention networks capture instead the emergence of self-organizing, interdependent mechanisms that stabilize decisions of attention allocation to problems in contexts characterized by lack of managerial control. We propose that the emergence of attention networks is generated by the concatenation of four basic micro-structural mechanisms, nested into each other in increasing order of complexity: attention focusing, attention reinforcing, attention mixing, and attention clustering. *Attention focusing* captures the tendency of participants to repeatedly allocate attention to the same issues, following a logic of familiarity. *Attention reinforcing* captures the tendency of participants to allocate attention to problems that already attracted a high degree of attention, following a logic of popularity. *Attention mixing* captures the interaction between the activity of participants and the popularity of problems, so that a disassortative mixing effect would indicate that participants that are more active are less attracted by popular problems, following a logic of devaluation. Finally, *attention clustering* captures the tendency of participants to extend their attention to future problems located in the vicinity of current problems, following a logic of proximity. We frame these mechanisms as theoretically grounded hypotheses, nested in each other in order of increasing complexity to extend our understanding of how attention represents a fundamental coordination mechanism in organizations.

We test these arguments through the empirical analysis of the evolutionary dynamics of organizational attention networks within a large open-source software project actively developed and maintained by a global community of participants. Given the "anarchistic" character of open-source productions (Lerner and Tirole 2001, p. 821), this empirical setting is appropriate and uniquely useful for our current purpose. Because development platforms that support open-source productions are explicitly designed to shape the allocation of collective attention, information on problems (called "software bugs") is reported (in a page called "bug report") and stored in virtual spaces (called "bug repositories" or "bug trackers") that participants (volunteer "software developers") may access freely at any time. An attention allocation event is recorded whenever a participant allocates attention to a problem by intervening on (or "touching") the corresponding bug report.

Our work contributes to contemporary research on organizations in at least three ways. First, we contribute to the literature on the microstructural perspective to organization design (Barney and Felin, 2013, Puranam 2018) by identifying organizational attention allocation as a self-organizing microstructure that replaces classical coordination mechanisms in traditional bureaucracies. A growing body of research investigates how participants follow self-organizing principles to collaborate (Deichmann et al., 2021; Fjeldstad et al. 2012), increase creativity (Majchrzak et al. 2021) and mobilize collective resources (Massa and O'Mahony 2021). However, no available research specifies the structural micro-mechanisms through which visible acts of attention allocation allow participants to solve the organizing problems of division of labor and integration of effort to achieve work coordination. Our empirical analysis – conducted, to the best of our knowledge, at an unprecedented level of detail – sheds light on the self-organizing dynamics by which visible acts of attention allocation produced by problem-solving agents provide necessary information to other agents that are self-selecting into problems to solve, revealing crucial complementarities between the solutions to the problems of task allocation and provision of information (Puranam 2018).

Second, we provide further insight to the scholarly debate on new forms of organizing (Lee and Edmondson 2017, Puranam et al. 2014) by focusing on an emerging empirical setting - open-source software communities - as an exemplary and theoretically-relevant case study that reveals how the universal problem of task allocation is solved – in absence of direct managerial control – by the information provided by participants' visible acts of attention allocation. Our study shows that the mechanisms of attention allocation that our theory postulates remain stable and clearly detectable after controlling for formal structures imposed by software modules, which help participants to search for problems that are better aligned with their interests and skills (Von Krogh et al. 2003). Furthermore, the results of our study address the known problem of under-provision of effort for "mundane but necessary" tasks that may easily escape collective attention in systems where a few popular problems attract a disproportionate share of attention (Puranam 2018, Lakhani and von Hippel 2003). We argue that the tendency of very active participants to divert their attention away from popular problems is fueled by a logic of devaluation, whereby some problems become less interesting to the highly engaged participants once a threshold of popularity is crossed.

Third, we contribute to the attention-based view (Ocasio 1997, 2011, Ocasio et al. 2022) by extending the original behavioral intuition that individual choices in organizations are influenced by the focus of attention of visible and relevant others (March and Olsen 1976, Ocasio 2011). We show that attention allocation is the product of "socially endogenous inferences" (Zuckerman 2012, p. 227) that give rise to an attention network, connecting and giving structural form to the relation between participants and problems (Cohen et al. 2012, Ocasio 2012) - an intuition originally discussed in seminal work on the Garbage Can Model (Cohen et al. 1972)[1]. When organizations are characterized by fluid participation that involves a constant churning of people and problems (March and Olsen 1984, p. 746), we show that the cumulative structural dynamics that we hypothesize leave "traces" that stabilize the link between participants and problems to form the attention network that we observe.

THE MICROSTRUCTURAL MECHANISMS OF SELF-ORGANIZING

The notion of organizations as systems of interdependent activities connecting participants (or agents) to problems (or tasks) was explicit in classic theories of organizations (e.g., March and Simon, 1958, Cohen et al., 1972), and is still central in contemporary theories of organization design (e.g., Puranam et al., 2014, Puranam, 2018, Raveendran et al. 2020). As explained by March and Simon (1993, p. 2), "Organizations are systems of coordinated action among individuals and groups whose preferences, information, interests, or knowledge differ." Classic theories of organizations emphasize the idea that boundedly-rational participants make decisions

[1] The self-organizing nature of attention networks that we introduce is deeply rooted in the basic behavioral premises of the Garbage Can Model (Cohen et al., 1972; March and Olsen, 1976). First, a key element of our theory hinges on the constant churning ("fluid participation" in Cohen et al.'s work) of organizational participants and problems - and hence opportunities for attention allocation. Second, participants in self-organizing contexts are likely to face a high degree of goal ambiguity ("problematic preferences" in Cohen et al.'s work), given the marginal role that managerial hierarchies and formal incentive structures play in the model. When these two conditions prevail, "attention focus, rather than utility, seems to explain much of the behavior" (March and Olsen, 1976: 15). In line with our hypotheses, decision-making becomes therefore less dependent on traditional rational choice logics, and more dependent on context-specific and situated attention allocation dynamics (Ocasio, 2012). As March and Olsen (1976:26) state: "What happens depends on how the situation fits into a mosaic of simultaneous performances involving other individuals, other places, other concerns, and the phasing of other events." Our theorizing of attention networks builds on this seminal work to provide a precise operational specification of the network principles and mechanisms of situated attention according to which attention allocation decisions are sensitive to the situational and structural context (Ocasio, 1997).

based on information characterized by a high degree of ambiguity (Galbraith, 1973, March and Olsen, 1976). For example, goals are often ill-defined and inconsistent across different organizational levels, causal connections between actions and their consequences are difficult to fully ascertain, and participation in problem-solving activities is characterized by frequent ebbs and flows that may affect the stability of outputs (Cohen et al 1972). Within this perspective, the "universal problem of organizing" – composed by division of labor and integration of effort – constitutes a continual process by which information is made less ambiguous as a result of coordinated actions of individual organizational participants (Puranam et al., 2012).

Traditionally, organization design scholars have focused on hierarchy as a critical mechanism for solving the problem of organizing (Keren and Levhari 1979, Mintzberg 1979, Padgett 1980). The provision of authority embedded in a hierarchical structure provides an effective mechanism for dividing work, allocating tasks, and giving necessary information for workflow coordination. However, the nature of work has significantly shifted in recent years. Organizational structures have become increasingly flatter, managerial authority has become significantly decentralized, and digital technology now allows for a great variety of ways to coordinate work (Lee and Edmondson 2017, Raveendran et al. 2020, Reitzig 2022). As a result, organizations strive to find new ways to tackle the problem of organizing while taking into account the recent shift in the nature of work (Puranam et al 2014). We contribute to this debate by bringing to bear the concept of organizational attention in contexts in which participants self-select into available problems to solve (Ocasio 1997, 2011, Ocasio and Joseph 2005).

In organization theory, there is widespread acknowledgment that attention is one of the bases for structuring individual behavior (March and Olsen 1976, p. 15; see also recent developments in attention-based views, e.g., Ocasio 2011). This research claims that what organizational participants do depends at least partly on the problems to which they devote their attention (Ocasio 1997, p. 188). However, this literature typically emphasized characteristics of problems competing for attention, such as length and breadth (Haas et al. 2015), urgency (Sullivan 2010), and degree of supplied information (Hansen and Haas 2001). Alternatively, the emphasis of previous work has been on the costs of (mis)coordination that (sub-)optimal communication between participants involves (for a recent review of this literature, see Prat and Dessein 2016). The self-organizing mechanisms of attention allocation that we introduce emphasize, instead, the fundamental dependence among individual acts of attention allocation connecting multiple participants to multiple problems. This interdependence arises because, at any given moment, the set of potential attention allocation opportunities is too large to be searched exhaustively by any individual participant (Simon 1978, p. 502). Consequently, under conditions of information transparency, distributed acts of attention allocation generate visible traces that reveal implicit participants' preferences and may thus be used as informational cues (Podolny 2001) to guide further attention allocation decisions by other participants, thereby providing vital information to their peers who strive to integrate and coordinate their efforts (Puranam 2018, Puranam et al. 2014).

What are – specifically – the structural mechanisms that drive and regulate these interdependent acts of attention allocation in organizations? Building on Schelling (1998, pp. 32-33), we define a micro-structural mechanism as a plausible hypothesis that explains the effect of a collective phenomenon (attention allocation, in our case) in terms of interaction between elementary agents (organizational participants and problems, in our case). Each specific mechanism that we hypothesize gives rise to a distinct, albeit interdependent attention structure – i.e., a determinant of the process by which participants choose to allocate their attention to problems (Ren and Guo 2011). Taken together, these mechanisms concatenate to form a dynamic attention network, in which edges connecting the nodes of an attention network are generated by individual, time-stamped events connecting participants to problems[2]. Overall, our set of hypotheses connects micro-level interdependencies to the aggregate structure of organizational attention[3].

---

[2] Of note, our notion of attention network differs from previous attempts at conceptualizing attention using a network perspective. For example Rhee and Leonardi (2018) looked at communication networks (i.e., social networks where nodes are people and edges are communication instances between people) and conceived attention as an actor-based attribute (i.e., a concentration index capturing the degree to which social actors pay uniform attention to all communication edges they are tied to, or concentrate their attention on a subset of these edges). On the contrary, our conceptualization of attention network entails a bipartite relation connecting participants to problems, and our structural micro-mechanisms define *exactly* in what ways participants allocate attention to problems.

[3] To facilitate exposition, the four hypotheses are introduced and discussed sequentially. In the empirical models, the four mechanisms concatenate to generate the organizational attention network that is actually observed. A useful way to interpret the hypotheses, thus, is as an interdependent set of interdependent claims about the relational micro-structures that regulate observed patterns of association between multiple participants and multiple problems.

Attention focusing

In an organizational world increasingly characterized by constant change in the sets of participants, problems, and decision opportunities, attention allocation cannot be reduced to a stable set of individual preferences (March 1991), because the choice set of organizational decision-makers is not defined *ex-ante* (Knudsen and Levinthal 2007)[4]. Allocating attention involves investing "energy, effort and mindfulness" (Ocasio 1997, p. 189) on a limited sample of problems selected from a larger population of problems. As an investment decision, the allocation of attention is only partially reversible because of the inertia determined by cognitive costs related to search, information acquisition, and learning (Conlisk 1996, Gabaix et al. 2006, Simon 1978). These problem-specific costs increase the likelihood that organizational participants will allocate their attention to more familiar problems. Consequently, what individuals focus their attention on is also likely to depend on what they become more familiar with. According to this "logic of familiarity," attention tends to become more focused over time as organizational participants distribute attention over a progressively narrower and relatively well-known set of issues (Levinthal and March 1993, p. 97). This expectation is consistent with available evidence on the effects of familiarity and experience on ease of recall of instances produced by classic experimental studies of choice under uncertainty (Tversky and Kahneman 1974).

In the presence of problem-specific (cognitive and learning) costs, the higher the participants' focus on familiar problems, the lower the cognitive cost of processing problem-specific information and recognizing problem-specific solutions (Reagans et al. 2005). Repeated attention granted to familiar problems gives rise to behavioral routines and related forms of recurrent behavior and habitual problem-solving that stabilize the relation between organizational problems and solutions to narrow the focus of attention (Cohen 2012). According to this view, our first hypothesis is consistent with classic predictions of behavioral theories of the firm about the reinforcing effects of familiarity on problem-solving behavior (Cyert and March 1963), and with more recent studies inspired

[4] As March usefully explains (1991, p. 109): "Individuals attend to some things and thus do not attend to others. The attention devoted to a particular decision by a particular potential participant depends on alternative claims on attention. Since those alternative claims are not homogenous across participants and change over time, the attention every particular decision receives can be both quite unstable and remarkably independent of the properties of the decision."

by this theoretical tradition (Christensen and Knudsen 2020) that investigate endogenous specialization as sustaining the division of labor in organizations. In the empirical context of our study, this prediction would be exemplified by a situation in which software developers decide to focus repeatedly on contributing to solve the same bugs – thereby increasing their task-based specialization – rather than seeking to engage a wider set of bugs. Building directly on these considerations, Hypothesis 1 provides the baseline expectation – based on a logic of *familiarity* – for the hypotheses that we develop next:

H1 (*Attention focusing hypothesis*)*: Participants are more likely to allocate attention to organizational problems that have attracted their attention in the past.*

Attention reinforcing

Attention focusing explains *why* participants are more likely to allocate their attention to familiar problems, but it does not explain *what* problems are more likely to attract attention in the first place, or how attention as a sampling mechanism actually works. Theories of selective attention have long recognized that salience affects the likelihood that a problem will receive attention (DiPrete and Eirich 2006, Taylor and Fiske 1978). Salience is partly an outcome of frequency-dependent social dynamics activated by the number of others who are visibly paying attention to the same problem, or issue (Salganik et al. 2006). In situations where acts of attention allocation performed by others are directly observable, attention allocation is subject to the tendency of popularity to breed popularity, i.e., to the law of accumulative advantage (Powell et al. 2005).

Attention-based accumulative advantage arises because popularity is interpreted as a signal of intrinsic interest, worth, attractiveness, and appropriateness. As Smith observes (2011, p. 64): "Increasing instances of a particular and, importantly, observable outcome signal the appropriateness of that outcome." The consequence of cumulative advantage of this kind is that attention allocation becomes progressively more concentrated on a limited number of issues or cases. The structural signature of this attention-reinforcing process is a skewed distribution of attention: few problems attract the attention of many, and most problems attract little or no attention at all. Similar forms of preferential attachment have been investigated in a variety of settings (Barabasi and Albert 1999, Dahlander and McFarland 2013, Powell et al. 2005). Studies of preferential attachment in social networks demonstrate that popular individuals are more likely to attract ties from other individuals – and hence become more popular (Merton 1968, Rivera et al. 2010).

In the more specific context of organizational attention, this general insight helps us to identify two meaningful sources of attention reinforcing, or cumulative advantage of popular problems. The first relates to the fact that the number of organizational participants allocating their attention to a problem increases the visibility of the problem. Participants may decide to allocate their attention to already popular problems to demonstrate their expertise, increase their reputation, or claim legitimate membership in their reference community (Dahlander and O'Mahony 2011, Shah 2006, Tajfel and Turner 1986). For example, popular problems may become public arenas where software developers become recognized as trustworthy and willing and able to contribute to problems of general interest for the project. A second source of attention reinforcing derives from the fact that participants engaging with popular problems have access to a larger and more diverse pool of knowledge. For example, allocating attention to popular problems allows software developers to access the experience of a larger sample of peer developers and learn from their experience and problem-solving practices (Boh et al. 2007, Reagans et al. 2005). Hypothesis 2 – based on a logic of problem *popularity* – summarizes this argument:

H2 (<u>*Attention reinforcing hypothesis*</u>): *Participants are more likely to allocate attention to popular organizational problems.*

Attention mixing

However reasonable as a general tendency describing the allocation of collective attention, reinforcing does have mitigating factors. There is empirical evidence, for example, that certain people tend to avoid choosing items of increasing popularity because they believe that their faddish character will make them transitory and short lived (Berger and LeMens 2009). Similarly, empirical studies have documented the systematic tendency of individuals to reduce their interest in an issue when the interest expressed by others in that same issue peaks (Berger and Heath 2008).

We argue that attention reinforcing is mitigated by a logic of devaluation of popular problems, and that the extent to which this happens depends on a specific characteristic of organizational participants, i.e., their level of activity. This particular argument hinges on assumptions about the mixing properties of attention networks – i.e., about the correlation between the relational characteristics of problems and participants. In our case, we expect attention allocation patterns to exhibit a disassortative mixing dynamic[5], as the attention of active participants spreads over less popular problems compensating, in part, for the tendency toward concentration induced by attention reinforcing.

Evidence of disassortativity in bipartite networks (like, for example, networks connecting individuals and issues) is widespread. Shang et al. (2010) have found evidence of disassortative mixing in the bipartite network affiliating participants in an internet-based recommender system website to music they play: very active consumers of music tended to listen to music tracks that few other users would select. Along similar lines, Grujić and Tadić (2009) report evidence of disassortative mixing in the bipartite network affiliating users and movies in a popular internet-based movie database: users who recommend many movies, recommend movies that are not recommended by many other users. In a recent study of Wikipedia – the open online encyclopedia – Lerner and Lomi (2020a) find that particularly active contributors (i.e., individuals with a high volume of editing experience) are less likely to modify articles that receive many edits from other contributors. In consequence, more active users seem to dedicate a larger share of their attention to less popular articles than less active users. These empirical studies report patterns of disassortative mixing in large bipartite systems composed of individuals and items competing for their attention, such as music tracks, movies, and webpages. In organizational attention networks, we expect disassortative mixing to be the outcome of a tendency whereby active people (i.e., people who pay attention to many items) are likely devalue popular items that become less interesting, at least to some participants, once a threshold of popularity is crossed (Berger and Heath 2008, Kovács and Sharkey 2014). In the context of organizational problem solving, the decline of interest in popular problems following devaluation is

[5] Reference to well-established network-analytic concepts might help to describe efficiently the mixing properties of attention networks (Newman 2002, 2003; Newman and Park, 2013; Pastor-Satorras et al. 2001). An attention network is assortative if active participants (participants who pay attention to many problems) are attracted by popular problems (problems that attract the attention of many participants). On the contrary, an attention network is disassortative if active participants allocate their attention to less popular problems. We note that the concept of assortativity we adopt is specific to 2-mode networks – networks defined *only between* distinct classes of objects (Lerner and Lomi, 2020b) - and differs from the more common concept of assortativity – or assortative mixing as used in game theory (Bergstrom, 2003) where networks are typically social, i.e., they connect objects *within* the same class. In the empirical context of our research disassortativity characterizes the consequence of the structural mechanism of 'mixing' that we introduce.

likely to be more pronounced for active participants whose available attention is more severely constrained by the intensity of their engagement (Kahneman 1973). While popularity of problems may be attractive to the average participant due to the potential increase in learning opportunities – as postulated in our *Attention reinforcing* hypothesis (H2) – it may also trigger effort-reducing cognitive heuristics (Tversky 1972) in more active participants. Popularity of problems may also strengthen social psychological effects such as diffusion of responsibility that depends directly on the number of participants already attracted to an issue (Darley and Latané 1968). Active participants with limited attention may then divert away from engaging with problems that: (i) require more intense computation and sense-making efforts to integrate a multitude of contributions from different participants (Castellaneta and Zollo 2014, Criscuolo et al. 2017), and (ii) seem to garner sufficient attention to guarantee their eventual resolution. Participants with extensive experience and limited attention might thus be more likely to focus on issues where their marginal contribution has higher impact. Consequently, active participants are likely to divert their attention away from popular problems.

More generally, our third hypothesis identifies a specific class of mechanisms that may be responsible for the endogenous emergence of division or labor and roles in organizations – a core concern in contemporary behavioral theories of organizations (Christensen and Knudsen, 2020; Knudsen and Srikanth, 2014). Hypothesis 3 – following a logic of *devaluation* – summarizes this argument:

H3 (*Attention mixing hypothesis*). *Active organizational participants are less likely to allocate attention to popular organizational problems.*

Attention clustering

Thus far, we have concentrated on behavioral tendencies of individuals (focusing, H1), characteristics of problems (reinforcing, H2), and on the *direct* interaction between individuals and problems (mixing, H3). But organizational participants are linked to each other also indirectly through their joint affiliation to problems that attract their attention (Carley 1991, Conaldi and Lomi 2013) – a necessary outcome of what Breiger (1974) identified as the duality of participants and problems in organizations. It is possible, therefore, that the *indirect* connection between problems and participants will make other proximate problems more likely to attract attention in the future. Our reasoning on this issue is firmly rooted in the classic behavioral insight that decision makers actively construct their choice set – and search is indeed a central aspect of organizational problem solving (Knudsen and Levinthal 2007). Organizational participants are more likely to allocate their attention to issues located in the neighborhood of issues that have attracted their attention in the past (Cyert and March 1963, Simon 1959). Building on this view, more recent research has identified a number of powerful factors that contribute to constrain the range of problemistic search in the neighborhood of current problems and solutions (Jung and Lee 2016). This view of search requires the definition of some notion of "neighborhood," so that the concept of "local search" may be specified (Stuart and Podolny 1996). To this end, suppose that two participants i_1 and i_2 are attracted by (i.e., allocate their attention to) the same problem m_2 at time t_1. Suppose, further, that i_2 (but not i_1) is also attracted by a second problem m_1. In this situation (i.e., at time t_1) problem m_1 is in the *neighborhood* of participant i_1 because i_1 can reach m_1 *indirectly* through the path $\{i_1 — m_2 — i_2 — m_1\}$[6]. If, at time t_2 ($>t_1$) participant i_1 decides to allocate attention to problem m_1 located in her neighborhood thus defined, then she generates a closed structure, or a "cluster" of two participants connected to the same two problems[7]. Unlike the more common forms of clustering, or "closure," in social networks involving three nodes (Bearman et al. 2014, Faraj and Johnson 2011, Shore et al. 2015), attention clustering involves *pairs* of participants allocating their attention jointly to the same *pairs* of problems forming a localized, and thus proximate, attention cluster.

We have established that the extent to which a problem may be considered "proximate" for a participant depends on the reachability of that problem through indirect connections with other participants. Attention clustering captures precisely the preferential tendency of participants to allocate their attention to more proximate problems – problems present in their neighborhood. This argument is consistent with the definition of "local search" that is central in behavioral theories of organizations (Cyert and March 1963, Levinthal and March 1993, Knudsen and Levinthal 2007). In the context of our study this prediction would translate into a

[6] This path indirectly connecting a participant to a problem through another participant is called a *three-path* in the analysis of bipartite networks (Wang et al. 2013). A three-path is the shortest possible indirect path linking a participant and a problem in a bipartite network. As such, the notion of three-path provides the analytical basis for an unambiguous definition of "neighborhood" (see also Pattison and Robins (2002) for a similar discussion in the context of social networks).

[7] The closed structure connecting two participants to the same two problems is called a *four-cycle* in the analysis of bipartite networks (Wang et al. 2013). The four-cycle is the analytical analogue of triadic closure in social networks.

situation in which (pairs of) developers are more likely to extend prior collaborations to future bugs. In our context, once two developers have allocated attention the same bug, they develop a shared understanding of that problem and a common ground for allocating shared attention to further problems (Lin et al. 2014). As a consequence, they will tend to find new bugs attracting their shared attention, thereby extending their prior problem-solving collaborations (Feld 1981: 1019). Hypothesis 4 – following a logic of problem *proximity* – summarizes our discussion:

H4 *(Attention Clustering): Participants tend to form attention clusters by allocating their attention to future problems in the neighborhood of their current problems.*

Tables 1 and 2 summarize the substantive implications of our mechanisms in the modern organizational environment, providing contextual examples in contemporary business and society, as well as underlying behavioral logics and resulting structural configurations. These examples suggest the possibility to generalize the interpretation of our mechanisms and hypotheses to multiple types of organizations beyond the specific empirical context. Table 3 instead summarizes from a formal perspective the four constructive mechanisms underlying our hypotheses. For each hypothesis, the table reports the antecedent configuration (at time t_1) in the left column, the change ("next event") prediction consistent with each mechanism in the central column, and the resulting (predicted) final configuration (at time t_2) in the right column.

--- Insert Tables 1-3 about here ---

RESEARCH DESIGN AND METHODS

Empirical setting

We study the dynamics of attention allocation in the *Apache HTTP Server* - a large and successful Free/Open Source Software (F/OSS) project. The term F/OSS typically refers to software products released under a license that allows inspection, use, modification and redistribution of the original software source code (Crowston et al. 2012). Developing teams (Lee and Cole 2003, Crowston and Scozzi 2008) contribute private effort to the production of what is effectively a public good (Spaeth et al. 2008, Von Krogh et al. 2012). Developers can be unpaid volunteers (Hars and Ou 2002, Lakhani and Wolf 2005), or paid by third parties (Henkel 2006, Stam

2009, Rolandsson et al. 2011). Their motivations vary widely and change over time (Von Krogh et al. 2012), ranging from ideological belief (Haruvy et al. 2003, Stewart and Gosain 2006) to pure enjoyment-based, intrinsic motivation (Lakhani and Wolf 2005), and to labor market signaling (Bitzer et al. 2017). Coordination happens mostly – if not only – online (Raymond 1999, Crowston and Scozzi 2008) and the distribution of attention is typically affected by the modular nature of the software (Baldwin and Clark 2006, Lerner and Tirole 2002), with developers concentrating their attention and activities on specific modules better aligned with their interests and skills, when not overviewing the overall code structure.

Open-source projects use instant messaging and mailing lists for technical discussions and support, code repositories for storing shared versions of the source code, and bug tracking systems for monitoring and tackling problems with the software (commonly referred to as software "bugs"). We consider the tackling of bugs – i.e., the actions aimed at resolving software problems that cause computer programs to behave in unintended and undesirable ways – as one of the most important activities affecting the quality of the software development process (Zhang and Kim 2010).

We are interested in understanding the mechanisms regulating the allocation of attention to organizational problems – an essential precursor to problem solving. Bug fixing absorbs a considerable amount of participants' attention (Crowston and Scozzi 2008). Indeed, the collective attention allocated to software bugs represents an important quality and reliability signal of the project (Crowston et al. 2003). Given the absence of centralized control and direct access to bug reports, developers typically self-manage the allocation of attention. Bug repositories provide an occasion for coordinating the highly decentralized activities of developers and for channeling collective attention.

Apache HTTP server is one of the most popular web server software, serving approximately 33% of all existing websites[8]. The development and maintenance of *Apache HTTP server* is overviewed by a Project Management Committee chaired by a Vice President. Members of the committee are appointed among the developers who

[8] Based on information reported by W3Techs (last access on January 30th 2023): https://w3techs.com/technologies/overview/web_server

have acquired significant merit in the project with their contributions. More generally, within *Apache HTTP server* – and all other projects overviewed by the Apache Software Foundation (ASF) – the right of contributors to modify the software is assigned by the community of developers and earned by showing commitment and active engagement. The ASF clarifies that all developers contribute to the project in their personal capacity – regardless of work affiliation (ASF 2022a, 2022b). Newcomers looking for ways to contribute to a project are explicitly encouraged by the ASF (2022a) to tackle a problem reported in the bug repository that stimulates their own interest. Given the freedom that developers experience when contributing to solve tasks in a very modular structure, the extensive reliance on volunteer participants and the complete transparency of their attentional processes, we think Apache represents a particularly apt environment for testing our hypotheses.

Data

We extracted the complete sequence of attention allocation events connecting project participants to software bugs stored in the official bug repository of the *Apache HTTP server*. Our sample includes all the bugs ever reported on releases 2.X of the software from the first bug report (March 2001) until the end of the observation period (March 2013). During the observation period a total of 13,526 actions by 4,338 unique developers on 6,000 distinct bug reports were recorded. Information on each individual event is exact to the second. Thus, the data set we constructed contain information on real-time attention allocation events observed throughout the life history of the project[9].

We adopt the definition of developer provided by ASF as individuals who "contribute to a project in the form of code or documentation. They take extra steps to participate in a project, are active on the developer mailing list, participate in discussions, provide patches, documentation, suggestions, and criticism" (ASF 2022b). In the rest of the paper, we use the generic term organizational "participants" to refer to 'developers' and the term organizational "problems" to refer to 'software bugs'. We collected the raw data by parsing the individual, publicly available, web pages of all bug reports included in the repository with the web-crawler software *Bicho*

[9] A dynamic visualization of the data we analyze in the empirical part of the study may be accessed by following the link: https://zenodo.org/record/7564503. The actual data we collected and used in the analysis are publicly available and may be found at the following address: https://github.com/juergenlerner/eventnet/tree/master/data/apache.

(Robles et al. 2009). In detail, a participant encountering a problem while running or developing the software usually starts the bug-reporting process and opens a bug report that is included as a new page in the bug repository. The repository provides the tracking infrastructure for describing, triaging, and resolving software bugs. A bug report is organized into a collection of pre-defined fields supplying information such as the name of the module, operating system, and release number that the bug affects. Bug reports also make the history of past actions visible and accessible to all participants. Fields in a bug report are populated when the report is first generated, and then updated by individual developers.

Dependent Variable

An attention allocation event linking participant *i* to problem *m* at time *t* is recorded whenever participant *i* modifies a field in ("touches") the bug report associated with problem *m*. The set of all observed attention allocation events represent the organizational attention network. What makes the organizational attention network a "network" is the aggregate structure of dependences that links individual acts of attention allocation. Our analysis focuses on the modeling of the time to the next observed attention allocation event conditional on the sequence of past events. More precisely, the dependent variable is the instantaneous probability of observing an attention allocation event linking a participant to a problem conditional on: (i) characteristics of the participant; (ii) characteristics of the problem, and (iii) the history of interaction between participants and problems within the project.

Independent Variables

It is possible to identify three broad sets of covariates included in our model. The first set includes covariates specifying the effect of direct theoretical interest that take the form of network statistics directly linked to our hypotheses. These covariates are defined exclusively in terms of sequences of attention allocation events linking participants and problems. Attention allocation sequences are inherently dynamic, thus the covariates of theoretical interest vary over time. The second set includes control covariates that account for a variety of other factors that may affect the probability of observing attention allocation events. Control covariates may refer to characteristics of participants or problems and may be time-constant or change over time in a way that does not depend on the history of attention allocation events. The third set of covariates includes interaction effects between control covariates and network effects.

Attention allocation mechanisms. We provide a verbal description of the statistics connected to the attentional mechanisms specified in the hypotheses. Then we describe the covariates included in the model to control for additional factors affecting the attention allocation events. Following the formal definition of "attention network," in **Appendix A** we develop the mathematical notation necessary to provide the formal definition of all the effects included in our empirical model specifications, which we now describe informally.

We start by defining *Cumulative attention* as a default, and therefore not directly hypothesized, attentional mechanism that serves as the basis for identifying the more specific mechanisms linked to our hypotheses. *Cumulative attention* simply involves a positive feedback mechanism regulating individual patterns of attention allocation. This generic baseline mechanism is commonly found in systems of social interaction characterized by increasing inequality in activities, outcomes and attainment (Merton 1973, Powell et al. 2005). *Cumulative attention* provides a useful baseline mechanism for modeling attention networks: it captures participants' average propensity to change their current level of attention to the project as a function of their history of participation (see Equation A.1 in Appendix A). A positive effect of *Cumulative attention* would provide evidence of self-reinforcing motivation to contribute attention: participants currently contributing a high level of attention to the project will be more likely to contribute an even higher level of attention in the future[10].

According to the *Attention focusing* hypothesis (H1), participants are more likely to allocate their attention to a problem if they have allocated their attention on the same problem in the past. A positive effect of *Attention focusing* would indicate that the greater is the number of attention allocation events connecting a participant to a specific problem, the higher is the likelihood of observing additional events reinforcing this connection (see equation A.3 in Appendix A).

[10] All the statistics are time-weighted according to a time decay parameter defined in equation A.2 in Appendix A. In this way, more recent events have a heavier weight on the prediction of the next event – a weight that progressively decreases for events in the more distant past.

According to the *Attention reinforcing* hypothesis (H2), popular problems are more likely to attract further attention, and hence become even more popular. To test this hypothesis, we define *Attention reinforcing* as a positive feedback effect driving the tendency of high levels of attention to generate further attention (see equation A.4 in Appendix A). A positive effect of *Attention reinforcing* would indicate that higher levels of attention received by a problem in the past, will lead to a higher likelihood of receiving an even more attention in the future – a sort of "popular for being popular" effect.

Attention mixing is defined as an interaction between cumulative attention and attention reinforcing (see equation A.5 in appendix A). A negative effect of *Attention mixing* would imply disassortativity, according to which active participants are less likely to allocate attention to more popular problems. At the project level, the consequence of disassortativity would be to divert (or disassociate) the attention of active participants away from popular problems to less popular ones. This is the specific prediction summarized in Hypothesis 3.

The fourth hypothesis involves *Attention clustering* (H4), a mechanism that stabilizes the association between participants and problems by "locking" the flow of attention within local clusters. We represent *Attention clustering* as the number of bipartite four-cycles (see Equation A.6 in Appendix A). A positive effect of *Attention clustering* would imply that participants tend to cluster their attention around shared problems over time. We call this effect attention "clustering" because bipartite systems cannot have closed cycles of odd length like, for example, the closed triangles that are commonly used to measure clustering tendencies in social networks (Newman and Park 2003)[11]. As we explained, attention clustering captures the local character of search by predicting that participants are more likely to extend their attention to problems located in the neighborhood of problems that are currently attracting their attention.

Control factors. The probability of observing an attention allocation event may be affected by a number of additional factors related to: (i) characteristics of participants; (ii) characteristics of problems; (iii) organization

[11] Like clustering defined for 1-mode (social) networks, clustering for bipartite networks involves path shortening behavior (or "closure"): the difference is that in bipartite networks, path shortening closes an open three-path connecting a participant indirectly to a problem through another problem and another participant (see equations A.6b and A.6c in Appendix A for the formal definition of 3-path. See also Table 1, bottom left panel for an intuitive graphical representation). Four-cycles are the most basic form of closure in bipartite systems (Wang et al. 2013).

structure, and (iv) interaction between these various factors and the observed sequences of attention allocation events. In our models, we include control variables at each of these levels. We define *Experienced participants* as an indicator variable taking value 1 for participants that were already active during previous release cycles of the project, and zero otherwise. Other conditions being equal, a positive effect would provide evidence that experienced participants are more likely to allocate their attention to problems. The variable *Institutional participants* is introduced to capture formal role differentiation. Participants are considered "institutional" if they use an email address ending with the official Apache domain (i.e., apache.org): in such case, the *Institutional participant* indicator variable equals 1, otherwise it is 0. There were 86 institutional participants out of 6,193 participants present during the observation period. A positive effect associated with this indicator variable implies that institutional participants are more likely to allocate their attention to problems.

The community of participants structured around the project has several ways to channel attention on problems. Examples include the assignment of problem priority and severity levels. Participants willing to allocate their attention to a specific problem assign a priority level ranging from 1 (highest priority problems) to 5 (lowest priority problem) to it. Assigning a priority level is an attempt to communicate how urgently participants would like to fix specific problems, and hence to direct their attention to them. We use this information to construct an ordinal problem priority indicator that we use to control for the potential differential attractiveness of problems assigned to different levels of priority. A negative effect associated to *Problem priority* would indicate that problems with high priority tend to attract additional attention (because problems with priority = 1 are high priority problems). Participants also assign a severity level to each software bug on a seven-point scale ranging from "enhancement" (a request for new features) to "blocker" - a bug that effectively prevents further development of the software. Assigning a severity level is an attempt to communicate how urgently bugs need to be fixed, and hence to direct attention to crucial problems. For example, if a problem has "blocker" status, then it must absolutely be fixed before the next release or project milestone. We use this information to construct a *Problem severity* indicator variable taking the value 1 if the problem is classified as a "blocker" and 0 otherwise. A positive estimate of the parameter associated with problem severity would suggest that severe problems tend to receive additional attention.

Problem latency is the age of a problem measured in days, i.e., the age of the problem computed for each active problem as the difference between the current time and the last time in which the problem was first reported - or reopened (Cohen et al. 1972). *Problem latency* is defined formally in Appendix A (Equation A.7). *Problem resolved* takes the value 1 when a problem status is changed to "resolved" and 0 otherwise. When this happens, the problem remains visible but is no longer "active." This time-varying indicator variable allows us to control for the decrease in attention over problems after they are resolved. A negative effect would indicate that attention allocation events are less likely to be performed on problems that are collectively considered resolved. Note that our design admits that resolved problems may be reopened in the future. For example, a software bug considered resolved might reappear in a subsequent release of the software. For this reason, attention allocated to resolved problems is likely to decrease, but does not necessarily or permanently drop to zero.

Problem recognition counts the number of comments associated with a problem. We include this variable to control for the differential propensity of problems to attract attention as a function of the discussion that problems generate. Many comments left by software users on a problem could signal that a problem is of general interest for the community at large. In turn, that could attract differential attention of the participants. *Problem recognition* is defined formally in Appendix A (Equation A.8)

F/OSS projects vary greatly in the extent to which they contain formal elements of organizational structure related to modularity. The necessity to adopt a modular structure increases as a function of the size and complexity of a project. *Apache HTTP server*, for example, is sufficiently large to sustain a formalized modular organization of its code. Examples of such modules include the "core" code, software extensions, and the corresponding software documentation. The software code is devoted to the basic processing of HTTP requests and responses. Each of the modules extends this basic functionality with a specific feature. For example, the module *mod ssl* provides *Apache HTTP server* with cryptographic capabilities now almost universally used by web servers. Alongside the software code composing core and modules, the code repository of the project contains software documentation that developers write and maintain as the project evolves. A software bug can affect the software as a whole, i.e., in multiple areas of the code base, across core and multiple modules. A software bug can affect only the core, one of the modules, or the documentation. Furthermore, a software bug may be reported as affecting the "build" - i.e., the compilation of the source needed before execution - as well as the installation procedure of the software. To assess the extent to which the modular structure of the project affects our results we assigned each software problem to one of five categories representing the different modules. This is the variable *Module*. The first category (omitted) includes all software bugs identified as affecting the software globally and serves as a baseline category. The second category includes all software bugs affecting the core module of *Apache HTTP server*. The third category includes all software bugs affecting build and installation of the software. The fourth category includes all software bugs affecting software documentation. The fifth category includes all software bugs affecting one of the other non-core modules of *Apache HTTP server*.

We incorporate information on the modular structure of the project by including *Preferential modularity* as a covariate controlling for the tendency of participants to work preferentially on problems within the same module. The interpretation of organizational structure provided by Cohen et al. (1972) is perfectly consistent with our quasi-experimental representation of *Preferential modularity* as an exogenous constraint on the access that solutions carried by participants have to problems that the project generates. A positive *Preferential modularity* effect would indicate that attention allocation events are more likely to occur within the same module where they occurred in the past. In other words, a positive estimate would provide evidence that participants tend to allocate their attention within individual modules. *Preferential modularity* is defined formally in Appendix A (Equation A.9). Participant-specific and problem-specific covariates may affect the probability of observing attention allocation events both directly, as well as indirectly through their interaction with the attention network. For this reason, in the empirical analysis we report in the next section we control for interaction effects that may reveal specific ways in which attributes of problems and participants interact with one another and with problem solving sequences. We define *Attention clustering within modules* as an interaction effect between *Module* and *Attention clustering*, because we want to investigate how formal organizational structure affects attention clustering between participants. A negative estimate of the coefficient associated with *Attention clustering within modules* would reveal a tendency of participants to cluster their attention on problems across modules. *Attention clustering within modules* is defined formally in Appendix A (Equation A.10).

Finally, *Time inactive* records for each participant *i* the time difference between the current time and the last time *i* was active in addressing any problem. We include time inactive to control for participants who are formally in the risk set, but do not actively contribute to the collective problem-solving process. Time inactive will make participants with long inactivity time hardly influence the estimation of other model effects. Time inactive is formally defined in equation A.11, Appendix A.

Table 4 summarizes the control factors included in the empirical model specifications, the class of objects to which factors pertain, the reason for inclusion, and the units of measurement. Mathematical definitions are reported in Appendix A.

--- Insert Table 4 about here ---

Relational Event Models

The relational event model that we implement in the empirical part of the paper exploits the full information contained in the sequence of time-stamped attention allocation events, and in their exact time ordering (Butts 2008, Perry and Wolfe 2013). The specific relational event model we adopt is described in detail and tested extensively in Lerner and Lomi (2020b).

The generating mechanisms represented in the model are defined in terms of event sequences that preserve the temporal information of individual problem-solving attempts. We use a Cox proportional intensity model incorporating both static and history-dependent covariates (Andersen and Gill 1982). We adopt well-established partial likelihood inference procedures to estimate the parameters of interest (Perry and Wolfe 2013). **Appendix B** provides additional background and information on the specification, estimation, interpretation, and evaluation of the relational event models. The model is estimated using the *eventnet* software (https://github.com/juergenlerner/eventnet) and the R package "Survival" Version 3.4-0 (Therneau 2022).

The complete information contained in the history of the attention network during the observation period was used to construct the vector of time-varying statistics, the "effects" described in the prior section. More specifically, a risk set including all participants and all problems recorded in our dataset is built and used to draw a sample of non-realized events needed for estimation (see Appendix B for details). When determining the event sequence used to construct our statistics, we note that some problems enter the risk set as they are reported, but no action is ever taken to solve them. Furthermore, to be defined the *Time inactive* control factor requires a second action to be taken by a participant at any point during the observation period. Thus, the event sequence used for estimation ends up consisting of 11,599 attention allocation events by 1,890 participants on 5,543 problems. The effects we estimate specify how the next attention allocation event (the dependent variable of the model) depends on specific configurations of time-structured sequences of past events. Parameter estimates tell the direction, magnitude, and significance of the theory-based mechanisms, and the effect of the control factors.

RESULTS

Analysis

The results of the analysis are reported in Table 5. Model 0 *(Null Model – M0)* is the benchmark model defined only in term of cumulative activity for each individual participant. According to Model 0, the next attention allocation event depends only on the history of past events. As a null model, Model 0 is considerably more challenging for alternative models than a model with no parameters. Model 1 *(Attribute control model- M1)* is the baseline model controlling only for attributes of the problems and the participants. According to M1 individual attention allocation events are independent and affected only by attributes of the participants and the problems, but do not dependent in any specific way from prior events. Model 2 *(Attention network model - M2)* introduces the effects of theoretical interest and is the focus of our discussion. Model 3 *(Organization structure model - M3)* examines the robustness of the estimates of theoretical interest when we consider element of formal organizational structure present in the project. Heuristically, the goodness of fit diagnostics reported at the bottom of the table indicate that the models estimated are significant, and that the full model (Model 3) improves significantly on the null model (Model 0), and on intermediate models – after accounting for differences in degrees of freedom.

-- Insert Table 5 about here --

The effects of control factors included in Model 1 are numerically stable across specifications and are generally consistent with our expectations. We discuss them briefly. *Institutional participants* and *Experienced participants* are significantly more likely to engage in attention allocation activities. For example, experienced participants are more than 148% more likely than non-experienced participants to allocate attention on problems within the project (exp(0.91) = 2.484). However, after one month of inactivity the probability than an inactive participant (*Time* inactive) will contribute again drops by approximately 39% (because exp(-0.0163*30) = 0.61).

Problems labelled "high priority" (*Problem priority*) do not have a significantly higher chance of attracting attention from participants. Severe problems (*Problem severity*) are approximately 15% more likely to attract attention. The longer a problem remains in the system (*Problem latency*), the less likely it is to attract attention. Problem resolved are predictably less likely to attract further attention. Because resolved problems may be eventually reopened, however, the odds of attracting further attention decline by almost 80% (because exp(-1.4596) = 0.23), but do not drop to zero immediately – a result echoing a "garbage can" view of organizational problems as recursive (Cohen et al. 1972, p. 10), and never quite resolved once and for all (Fioretti and Lomi 2010). Finally, a long list of comments attached to a bug report may be interpreted as a signal that a bug is being recognized by the community as particularly interesting, complex or worthy of discussion (Hooimeijer and Weimer 2007, Arya et al. 2019). The positive and significant effect of *Problem recognition* confirms this expectation: an increase in one standard deviation in the number of comments increases the hazard of generating additional attention by 41% (because exp(0.3430) = 1.41).

We now concentrate the focus of our discussion on Model 2 which incorporates the effects of theoretical interest. As an aid to interpretation, we note that significantly positive (negative) effects increase (decrease) the rate of events connecting participants to problems according to the mechanism generating the event sequence associated with the effect[12].

Attention focusing (H1). We hypothesized that participants working on specific problems will be more likely to

[12] Because the network-dependent effects are standardized, the interpretation of their magnitude is carried out in terms of standard deviations from mean. Standardization is useful in this case because the sample size and the way the covariates are constructed makes statistical significance alone unhelpful to evaluate the magnitude and strength of the effects of theoretical interest.

focus their attention on those problems in the future. We argued that this form of inertia in the attachment of participants to problems is important because it induces specialization, and because it stabilizes patterns of organizational attention. Because *Attention focusing* also sustains faster learning, attention focusing helps participants to develop skills that are, at least in part, problem specific. Estimates support the focusing hypothesis: An increase by one standard deviation in the level of attention that a participant has allocated to one specific problem prior to his current decision point, increases the odds that they will decide to pay attention to the same problem again by 23% (exp(0.2150) = 1.234).

Attention Reinforcing (H2). Hypothesis 2 summarizes our prediction that popular problems – problems attracting the attention of many participants – are more likely to attract additional attention in the future. As we have discussed, *Attention reinforcin*g may be due to uncertainty about the quality of a problem such as, for example, its level of difficulty, or it may be the consequence of participants' attempts to gain access to a broader pool of knowledge by joining conversations attended by many other participants. We found strong support for the *Attention reinforcing* hypothesis: An increase by one standard deviation in the level of attention that a problem has received within the community increases the odds that the same problem will attract further attention by approximately 15% (because exp(0.1354) = 1.145).

Attention mixing (H3). The results support our hypothesis for *Attention mixing*: the negative and significant coefficient shows that the attention that active participants allocate to popular problem is reduced, indicating a tendency for disassortativity. A participant that is by one standard deviation more active than the average participant, will experience a decrease in the attention reinforcing effect by approximately 3% (exp(-0.0320) = 0.97). That is – for such a participant – a problem that has received one standard deviation more attention in the past has the odds of receiving attention by that participant increased only by approximately (15-3)% = 12%. A participant that is more active than average by approximately 42 events will exhibit no preferential tendency to allocate attention to more popular problems (exp(-0.0320*6) = 0.83, will offset the 15% increase in the attention that a more popular problem receives from a participant with average activity). Other conditions equal, this results show that very active participants will exhibit a preferential tendency *against* popular problems. Figure 1 provides a graphical examination of this mitigation effect, by showing to what extent the marginal effect of *Attention reinforcing* decreases the more *Cumulative attention* increases.

Attention clustering (H4). We hypothesized that participants are more likely to allocate attention to problems in the neighborhood of their current problems. Attention clustering captures the tendency of participants to extend their current collaborative experience to future problems located in their neighborhood. Results provide solid support for the *Attention clustering* hypothesis (H4). An increase by one standard deviation in the number of three-paths *indirectly* connecting a participant to a problem will increase by slightly less than 4% the odds that this participant will address this problem *directly* in the future (exp(0.0349 = 1.355). This effect may seem small, but – considering that a participant-problem pair can be indirectly connected by a high number of three-paths – even one extra event closing a three-path could have a large impact on the number of 4-cycles, or attention clusters, that are generated.

Table 6 summarizes the qualitative implications of our findings. Hazard ratios in Table 6 are calculated based on the estimates of Model 2 in Table 5 under *ceteris paribus* assumptions. Consequently, the figures discussed in the text and summarized in the table are provided as an aid to a heuristic interpretation of the estimated parameters in the light of our hypotheses.

-- Insert Table 6 about here --

Supplementary sources of structuring in organizational attention networks

Attention focusing, reinforcing, mixing and clustering may be viewed as intended consequences of attention structures shaped, at least in part, by formal organization design (Ocasio and Joseph, 2005). We entertain this possibility, in a supplementary analysis that accounts for the formal elements inherent in the organizational structure of the project. Doing so allows us to assess the role of formal structure in matching interests and skills of participants to characteristics of problems as an alternative explanation for the allocation of attention – and hence for the self-organizing nature of attention networks. Such structure in open-source software projects takes the form of modularity (Baldwin and Clark 2006).

Modularity allows different functional areas to be developed almost independently, thus speeding the evolution of the global software system. In more general terms, modularity exemplifies a case of "specialized access - or an attempt to force solutions to specialize in the kinds of problems that can be associated to them" (Cohen et al. 1972, p. 6). In open-source software projects, the primary function of modularity is to help participants search efficiently for opportunities to contribute aligned with their interests and skills (MacCormack et al. 2006). For this reason, modularization is a common organizing principle in open-source software projects (Von Krogh et al. 2003), but it is particularly salient in *Apache HTTP server* which is one of the largest open-source software in existence. Because of its size and complexity, the overall project is parsed into modules that reflect major functional areas of the software.

Model 3 incorporates the effects of modularity. The modularity effects should be interpreted with respect to the omitted baseline category ("software bugs affecting the software as a whole"). The effect of *Preferential modularity* (ratio of the participant's past contributions that are allocated to problems in the same module as the focal problem) reveals a strong tendency of participants to concentrate their attention on problems within software modules. According to the estimates of Model 3 participants are approximately 51% more likely to allocate attention to problems within the same module (exp(0.4094) = 1.51). Interpreted in tandem with the effect of *Attention focusing*, the effect of *Preferential modularity* implies a tendency towards specialization within modules, in accordance with the general expectation that participants tend to allocate attention to problems that belong to their area of expertise.

Accounting for elements of formal structure organizing problems by skills and interests of participants does not affect the significance, strength, or direction of the mechanisms underlying our hypotheses. Therefore, our results seem to be extendable to contexts where the "rules of the game" (Ocasio 1997) allow for some degree of specialized access to problems – at least in its "upper limits" of who may attend to what (March and Olsen 1976, p. 40) – instead of the more general unsegmented access typical of organized anarchies (Cohen et al. 1972, Ocasio 2012). Attention focusing, reinforcing, mixing and clustering continue to operate once the effect of organizational structure on the allocation of attention is accounted for. Thus, empirical evidence is stacked against the intuitive explanation emphasizing individual interests and personal skills as the main drivers of attention. What our results reveal is that individual interests and skills as reflected in the modular structure of the project, do not explain away how software developers allocate their attention to software problems. This conclusion is strengthened further by the effect of *Attention clustering within modules* that we include to account for the tendency of participants with similar interest and skills to be attracted by the same problems with the same module. The effect is far from statistical significance – a noteworthy outcome given sample size. This means that patterns of extended collaboration captured by *Attention clustering* cannot be simply explained away by similarity of individual interests among participants.

Additional Robustness Checks

To further establish the validity of our results in a variety of situations where the allocation of attention is potentially driven by alternative (self-) organizing logics, we performed a series of additional robustness checks. Specifically, we tested whether the self-organizing patterns of attention structures are sensitive to different levels of *intensity* (or *effort*) in attention allocation. Furthermore, we checked if our results are affected by different empirical interpretations of the theoretical assumption of "attention scarcity" postulated in classic behavioral (Simon 1947), attentional (Ocasio 2012) and garbage-can theories (Cohen et al. 1972). In particular, we checked whether the attention network mechanisms operate differently in presence of higher problem *crowding*, which denotes higher scarcity of collective attention (Piezunka and Dahlander 2015). Finally, we checked the sensitivity of our results to some alternative definitions of our sampling strategy, by excluding a) problems that are reopened once solved and b) events produced by extremely active participants. In all cases, the main substantive patterns of our results remain robust. Further details about the specific tests and the respective tables of additional results can be consulted in **Appendix C** of our supplementary material.

DISCUSSION

Through the lens of attention networks, the space of possibilities for the allocation of attention in organizations appears as potentially very large, but almost completely empty. The attention allocation events that are actually observed cover only a very small sub-space of the events that may be possible. The subnetwork of observed events is small, and highly structured by the theoretical mechanisms of attention allocation that we have postulated. This general conclusion resonates clearly with classic behavioral theories of organization (Cyert and March, 1963; Levinthal and March, 1993; Ocasio, 2011; Simon, 1947), but also bears important implications for contemporary theories of organizational attention, and for future research on the allocation of attention within organizations. We conclude with a discussion of the theoretical contributions of the study, its practical implications, limitations, and future directions.

Contributions to theory and research.

We need theories of organizations that bring renewed focus on the investigation of contemporary organizational environments – including crowdsourced production, holacracies and bossless organizations – in which individuals have increasing independence in choosing the tasks and issues they contribute to, and in which their activities are often publicly visible to other organizational members (e.g., Puranam et al. 2014, Alexy et al. 2021). In these contexts, problem-solving activities take place under conditions of high decentralization of authority, fluid participation, and variable attention lent by volunteer participants. Under conditions in which managerial hierarchies and organizational systems of task allocation and control are weak or absent, visible acts of attention allocation become a fundamental stabilization mechanism by which participants provide information to each other about how to coordinate their efforts. Considering attention as a resource that connects participants and problems in a relation of mutual constitution brings to the fore the issue of its role in enabling coordination in self-organizing contexts (Hansen and Haas, 2001; Ocasio and Hoffman, 2001).

In this paper, we address these core theoretical concerns by reframing socially transparent acts of attention allocation as a microstructural coordination mechanism allowing organizational participants to provide information to each other about organizing work around available problems. This study contributes to the scholarly debate on the micro-foundations of organization design, i.e., "the micro-level processes, behaviors, and interactions that aggregate to yield the organization's overall structure" (Raveendran et al. 2020: 829; for further discussion of the micro-foundations of social networks, see Tasselli, Kilduff and

Menges 2015). These micro-foundational interactions between participants and problems are not based on exogenous task allocation decisions or pre-determined workflows that participants know ex ante. Instead, they emerge from participants' efforts to coordinate work as they face ambiguous task demands and fluid organizational boundaries and observe how their peers operate. In this sense, our analytical framework is consistent with a view of boundedly-rational participants who learn about their interests and uncertain future preferences by interacting with actual problems attracting their attention (March and Olsen 1976), in a context where "contributors do not know if their efforts will result in a suitable working product" (Von Krogh et al. 2003, p. 1219). By investigating the emergence of attention networks as the byproduct of interdependent acts of attention allocation, our work responds to recent calls for a better "understanding [of] the involvement of agents in the design process […] in scenarios in which the tasks are not clear-cut." (Raveendran et al 2020: 829) and resonates with the idea that "interdependence is endogenous to the organization design process […] because it arises during agent interactions" (Raveendran et al 2020: 831).

It's worth noting that our study hinges on the assumption that organizational participants have access to transparent information about how others allocate their attention. Observability underlies the formation of "socially endogenous inferences" (Zuckerman 2012) upon which we base our theorizing and facilitates the emergence of stigmergic coordination - i.e., "implicit coordination mediated by changes to a shared work product" (Rezgui and Crowston 2018, p. 1; Moffett et al. 2021). This feature of our study reveals an important implication for the debate on the micro-structural approach to organization design, insofar as "we would expect the process of search for new forms of organizing to stabilize around clusters of complementary solutions" (Puranam 2018: 155). In particular, we show that the transparency of attention constitutes a valuable complementary solution to task allocation via self-selection, because it provides the necessary informational cues for participants to coordinate their effort, and thus for attention networks to self-organize and dynamically emerge in organizations.

We introduced the notion of attention networks to emphasize the self-organizing character of organizational problem solving. Our emphasis, specifically, is not on the static characterization of this network, but on the interdependent mechanisms of its transformation and emergence (Gibson 2012, Padgett and Powell 2012). The operational concept of attention network that results from our set of hypotheses provides an illustration of how contemporary organizations are to some extent defined, and transformed, by the very activities that their members perform, resonating with the idea "auto-catalysis" (Padgett 2018, Padgett et al. 2003, Padgett and Powell 2012). Every visible act of attention allocation involves the creation of new connections between participants and problems, and hence implies change in the organizational structure via change in the activities and issues that organizational participants choose to focus on. Hence "structure" and "change" are constructs that can only be understood in reference to one another: they represent a duality that entails the concept of organizing itself. If traditional firms and societies have been described as "centralized, bureaucratic, and inflexible" (Thompson 1967, p. 108), self-organizing contexts allow individuals to broaden their "range of aspirations" (p. 114) by giving the freedom to allocate time and effort to issues that are prioritized by the individuals themselves. Our set of connected hypotheses shows that the mechanisms linking people to problems and tasks follow a structural logic that, accumulating over time, contributes to shape organizational problem solving.

Interestingly, as shown in the models that we have specified and estimated, the microstructural mechanisms associated with each individual hypothesis concatenate to generate the dynamic attention network that we actually observe. Not only, as predicted and tested by structuralist sociologists (e.g., Blau 1960, Freeman 1978), social interactions between individuals follow network patterning – being subjected to structural properties and regularities – but, as we discover, also the very relations connecting organizational participants and the tasks they perform are exposed to structural mechanisms that can be interpreted and measured from a network perspective. We have shown that this result is unaffected by observable individual characteristics of the participants (e.g., their experience and institutional affiliation), intrinsic features of the problems (e.g., their level of difficulty), and by the effect of formal (exogenous) organizational structures that regulate the matching between individual skills and problem characteristics (i.e., the project modules).

Implications for practice

Our study bears implications for managers and practitioners interested in the design of self-organizing and distributed productions. Far from being confined to the world of open-source software projects, self-organizing is becoming increasingly popular also in the context of traditional corporate organizations, where the pressure to give employees and managers freedom in self-selecting into tasks is often associated with the aim to stimulate collective creativity and empower idea generation (e.g., Cross et al. 2021; see also Table 2 for concrete examples on the generalizability of our hypotheses). However, attempts to give personnel more independence often find internal resistance from decision makers, who are afraid of losing control over their workforce's activities. How can organizations and managers find the right balance between giving freedom to organizational members to self-allocate time and attention to problems, and still managing to keep control over decision making and problem solving? As shown by our results, this tension is at least partly misplaced. Attention tends to self-organize following structural patterns, thus finding, in serendipitous ways, its own order. More specifically, our mechanism of *attention mixing* seems particularly important as it mitigates the effect of *attention reinforcing* by distributing the attention of active participants over a broader set of problems. Concretely, this mechanism helps the sustainability of crowdsourced production systems, which are characterized by the presence of a large number of diverse issues that may easily escape attention, while a limited number of popular issues tend to attract a disproportionate share of collective attention (Huberman et al. 2009). In these settings, survival depends crucially on the willingness of a community of volunteers to allocate attention to "mundane but necessary tasks" (Lakhani and von Hippel, 2003: 923). Yet these tasks are essential to the survival of the project, although they are neither popular, nor capable of motivating repeated engagement (Shah 2006). Moreover, the disassortative patterns of *attention mixing* distributes collective attention over a larger set of problems, tackling well-known problems of under or over-provision of effort for certain tasks – an outcome that students of organizations have struggled to find a solution for despite very detailed and extensive research (Faraj and Johnson 2011; Puranam 2018, Stewart 2005, von Krogh et al. 2012).

Limitations and conclusions

This study has several limitations, and two deserve special attention as they invite future research to follow clear directions. The first limitation relates to the empirical scope of the study, which assumes that acts of attention allocation are transparent to other peer participants. This scope condition applies to most arguments based on "social transparency" (Stuart et al. 2012) and "ambient awareness" (Leonardi 2015), or free access to information on the behavior, opinion, orientation or evaluation that might reveal the preferences of others. The scope of our study is thus relevant to a wide range of contexts designed precisely to support information sharing through social media functionalities. Examples of self-organizing forms that rely on transparency as a condition for information provision (Puranam et al. 2014) include external (Piezunka and Dahlander 2015) and internal (Deichmann et al. 2021) crowdsourcing platforms for idea development; enterprise-based systems of collaboration, such as internal communities of practice (Haas et al. 2015), platform-based systems of decentralized problem solving such as open-source software (Von Krogh and Von Hippel 2006) and distributed innovation (Kogut and Metiu 2001). Clearly, the realism of the assumptions underlying our hypotheses decreases as the opacity of information about what others do increases. However, the confidence in the generalizability of our hypotheses derives not only by the growing diffusion of new forms of organizing, but also by the general tendency to make traditional organizations increasingly more open and transparent. These tendencies are diffusing from emergent online communities to established corporate entities (e.g., Dahlander and Magnusson 2005, Fosfuri et al. 2008), including the cases of the Zappos "holacracy" (Robertson 2015), examples of 'agile' organizational networks (Tasselli and Caimo 2019), or the cases of Valve and Morningstar documented by Lee and Edmondson (2017). Our arguments extend to a wide set of contexts in which participants decide to allocate attention to "many problems seeking solutions" (Haas et al. 2015, p. 681) and where peer participants can observe their actions and decisions.

A second limitation is inherent in the almost exclusive focus on individual acts of attention allocation, which precluded analysis of possible outcomes of such acts. Whenever participants allocate attention on a bug report, their act contributes an event edge in our attention network. The dynamics of these events are the focus of our analytical interest in this paper. This observation scheme is not inspired by a focus on the real effectiveness of the collective resolution of problems, on how long problems remained unresolved within the project, or on how durable (or stable) were the solutions that contributors implemented (Fioretti and Lomi 2008). Addressing these issues would require a different research design, one oriented toward the consequences of problem-solving behavior rather than on attention as one of its main antecedents. The study says little about attentional selection - or the consequences of attention allocation decisions (Ocasio 2011). We hope our study will encourage to address issues of organizational effectiveness that we could not pursue here.

Despite its limitations, we believe that our study contributes new elements to the understanding of the self-organizing mechanisms underlying attention allocation in organizations. The world of self-organizing is a world of variable attention, fluid participation, and evolving problems with little or often no centralized control. The rapid diffusion of such new organizational forms force us to broaden our view of what "organizations" are, and under what conditions they can exist and operate effectively. Contemporary open-source software projects are probably the organizational archetype of self-organizing contexts that capture the properties of the ideal formulation of "organized anarchy" prophesized by Cohen, March and Olsen more than half a century ago. We have presented an empirical example of how coordination of efforts may emerge and self-organized out of ordered time-sequences of individual acts of attention allocation connecting organizational participants (software contributors) to organizational problems (software bugs). In the case we presented here, the linkages between participants and problems were "less consequential than temporal" because "attention to problems seems to be determined as much by the time of their arrival as by assessment of their importance" (March and Olsen, 1984:743). But the theoretical vision that inspires the current study was articulated decades before innovation in information technology made open-source software projects possible, or even only conceivable. In this sense, the current study celebrates the success of organization theories – and theorists whose insight, vision and imagination time proved prescient.

REFERENCES

Alexy O, Poetz K, Puranam P, Reitzig M (2021) Adaptation or persistence? Emergence and revision of organization designs in new ventures. *Organ. Sci.* 32(6):1439–1472

Andersen PK, Gill RD (1982) Cox's regression model for counting processes: A large sample study. *Ann. Stat.* 10:1100–1120.

Apache Software Foundation (2022a) Apache Community Development Website. Retrieved August 25, 2022, https://community.apache.org/.

Apache Software Foundation (2022b) How it works. Retrieved August 25, 2022, http://www.apache.org/foundation/how-it-works.html.

Arazy O, Lindberg A, Rezaei, Mostafa M, Samorani M (2020) The evolutionary trajectories of peer-produced artifacts: Group composition, the trajectories' exploration, and the quality of artifacts. *MIS Q*. 44(4):2013-2053.

Raveendran M, Luciana S, Ranjay G (2020) The role of interdependence in the micro-foundations of organization design: Task, goal, and knowledge interdependence. *Acad. Manag. Ann*. 14:828-868.

Arazy O, Daxenberger J, Lifshitz-Assaf H, Nov O, Gurevych I (2016) Turbulent stability of emergent roles: The dualistic nature of self-organizing knowledge coproduction. *ISR*. 27(4):792-812.

Arya D, Wang W, Guo JLC, Cheng J (2019) Analysis and detection of information types of open-source software issue discussions. *2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*, 454–464.

Baldwin CY, Clark KB (2006) The architecture of participation: Does code architecture mitigate free riding in the open-source development model? *Manage. Sci.* 52(7):1116–1127.

Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science*. 286(5439):509–512.

Barney JAY, Felin T (2013) What are microfoundations? *Acad Manage Perspect*. 27(2):138–155.

Bearman PS, Moody J, Stovel K (2014) Chains of affection: The structure of adolescent romantic and sexual networks. *Am. J. Sociol.* 110(1):44–91.

Becker MC, Knudsen T (2005) The role of routines in reducing pervasive uncertainty. *J. Bus. Res*. 58(6):746-757.

Berger J, Heath C (2008) Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes. *J. Pers. Soc. Psychol.* 95(3):593–607.

Berger J, Le Mens G (2009) How adoption speed affects the abandonment of cultural tastes. *P. Natl. Acad. Sci.* 106(20):8146–8150.

Bergstrom, T. C. (2003). The algebra of assortative encounters and the evolution of cooperation. *Int. Game Theory Review*, *5*(03), 211-228.

Bitzer J, Geishecker I, Schröder PJ (2017) Is there a wage premium for volunteer OSS engagement? – signalling, learning and noise. *Appl. Econ.* 49(14):1379–1394.

Blau PM (1960) Structural effects. *Am. Sociol. Rev.* 25(2):178–193.

Boh W, Slaughter SA, Espinosa JA (2007) Learning from experience in software development: A multilevel analysis. *Manage. Sci*. 53(8):1315–1331.

Breiger R (1974) The duality of persons and groups. *Soc. Forces.* 53(2):181–190.

Breiger RL, Mohr, JW (2004) Institutional logics from the aggregation of organizational networks: Operational procedures for the analysis of counted data. *Computational & Mathematical Organization Theory*. 10(1):17–43.

Butts CT (2008) A relational event framework for social action. *Sociol. Methodol.* 38(1):155–200.

Carley K (1991) A theory of group stability. *Am. Sociol. Rev.* 56(3):331–354.

Castellaneta F, Zollo M (2015) The dimensions of experiential learning in the management of activity load. *Organ. Sci.* 26(1):140–157.

Christensen M, Knudsen T (2010) Design of decision-making organizations. *Management Science*. 56(1):71–89.

Christensen M, Knudsen T (2020). Division of roles and endogenous specialization. *Ind. Corp. Change*. 29(1):105–124.

Cohen MD (2012) Perceiving and remembering routine action: Fundamental micro-level origins. *J. Manage. Stud.* 49(8):1383–1388.

Cohen MD, March JG, Olsen JP (1972) A Garbage Can Model of organizational choice. *Admin. Sci. Quart.* 17(1):1–25.

Cohen, MD, March JG, Olsen JP (2012) "A Garbage Can Model" at 40: A solution that still attracts attention. Lomi A, Harrison JR, eds. *The Garbage Can Model of Organizational Choice: Looking Forward qt Forty* (*Research in the Sociology of Organizations, Vol. 36)* (Emerald Group Publishing Limited, Bingley, UK), 19–30.

Conaldi G, Lomi A (2013) The dual network structure of organizational problem solving: A case study on open-source software development. *Soc. Networks.* 35(2):237–250.

Conlisk J (1996) Why bounded rationality? *J. Econ. Lit.* 34(2):669–700.

Criscuolo P, Dahlander L, Grohsjean T, Salter A (2017) Evaluating novelty: The role of panels in the selection of R&D projects. *Acad. Manage. J.* 60(2):433–460.

Cross R, Gardner HK, Crocker A.(2021) For agile transformation, choose the right people. *Harv. Bus. Rev.* 99(2):60-71.

Crowston K, Annabi H, Howison J (2003) Defining open-source software project success. *Proc. 24th Int. Conf. Inform. Syst.* (ICIS 2003), 28.

Crowston K, Scozzi B (2008) Bug fixing practices within Free/Libre Open Source Software development teams. *J. Database Manage.* 19(2):1–30.

Crowston K, Wei K, Howison J, Wiggins A (2012) Free/Libre Open Source Software development: What we know and what we do not know. *ACM Comput. Surv.* 44(2):1–35.

Cyert RM, March JG (1963) *A Behavioral Theory of The Firm* (Prentice-Hall, Englewood Cliffs, NJ).

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

Dahlander L, McFarland DA (2013) Ties that last: Tie formation and persistence in research collaborations over time. *Admin. Sci. Quart.* 58(1):69–110.

Dahlander L, O'Mahony S (2011) Progressing to the center: Coordinating project work. *Organ. Sci.* 22(4):961–979.

Darley JM, Latané B (1968) Bystander intervention in emergencies: Diffusion of responsibility. *J. Pers. Soc. Psychol.* 8(4pt1):377–383.

Deichmann D, Gillier T, Tonellato M (2021) Getting on board with new ideas: An analysis of idea commitments on a crowdsourcing platform. *Res. Policy.* 50(9):104320.

DiPrete TA, Eirich GM (2006) Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annu. Rev. Sociol.* 32:271–297.

Faraj S, Johnson, SL (2011) Network exchange patterns in online communities. *Organ. Sci.* 22:1464–1480.

Feld SL 1981 The focused organization of social ties. *Am J. Sociol.* 86:1015–1035.

Fioretti G, Lomi A (2008) The garbage can model of organizational choice: An agent-based reconstruction. *Simul. Model. Pract. Th.* 16(2):192–217.

Fioretti G, Lomi A (2010) Passing the buck in the garbage can model of organizational choice. *Comput. Math. Organ. Th.* 16(2):113–143.

Fjeldstad ØD, Snow CC, Miles RE, Lettl C (2012) The architecture of collaboration. *Strateg. Manag. J.* 33(6):734–750.

Fosfuri A, Giarratana MS, Luzzi A (2008) The penguin has entered the building: The commercialization of open-source software products. *Organ. Sci.* 19(2):292–305.

Freeman LC (1978) Centrality in social networks conceptual clarification. *Soc. Netw.* 1(3):215–239.

Gabaix X, Laibson D, Moloche G, Weinberg S (2006) Costly information acquisition: Experimental analysis of a boundedly rational model. *Am. Econ. Rev.* 96(4):1043–1068.

Galbraith J (1973) *Designing Complex Organizations* (Reading, MA).

Gibson DR (2012) Taking turns and talking ties: Networks and conversational interaction. *Am. J. Soc.* 110(6):1561–1597.

Grujić J, Mitrović M, Tadić B (2009) Mixing patterns and communities on bipartite graphs on web-based social interactions. *16th Int. Conf. on Digital Signal Processing*. (IEEE, Santorini), 1–8.

Gulati R, Puranam P, Tushman M (2012) Meta-organization design: rethinking design in interorganizational and community contexts. *Strat. Mgmt. J.* 33:571–586.

Haas MR, Criscuolo P, George G (2015) Which problems to solve? Attention allocation and online knowledge sharing in organizations. *Acad. Manage. J.* 58(3):680–711.

Hansen MT, Haas MR (2001) Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company. *Admin. Sci. Quart.* 46(1):1–28.

Hars A, Ou S (2002) Working for free? Motivations for participating in open-source projects. *Intern. J. Electron. Comm.* 6(3):25–39.

Haruvy E, Prasad A, Sethi SP (2003) Harvesting altruism in open-source software development. *J. Optim. Theory Appl.* 118(2):381–416.

Henkel J (2006) Selective revealing in open innovation processes: The case of embedded Linux. *Res. Policy.* 35(7):953–969.

Hoffman AJ, Ocasio W (2001) Not all events are attended equally: Toward a middle-range theory of industry attention to external events. *Organ. Sci.* 12(4):414–434.

Hooimeijer P, Weimer W (2007) Modeling bug report quality. *Proc. 22nd IEEE/ACM Int. Conf. on Automated Software Engineering (ASE '07)* (Association for Computing Machinery, New York, NY, USA), 34–43.

Huberman BA, Romero DM, Wu F (2009) Crowdsourcing, attention and productivity. *J. Inf. Sci.* 35(6):758–765.

Jung HJ, Lee JJ (2016) The qUEST for originality: A new typology of knowledge search and breakthrough inventions. *Acad. Manage. J.* 59(5):1725–1753.

Kahneman D (1973) *Attention and Effort* (Prentice Hall, New Jersey).

Keren M, Levhari D (1979) The optimum span of control in a pure hierarchy. *J. Manag. Sci.* 25(11):1162–1172.

Kilduff M, Brass DJ (2010) Organizational social network research: Core ideas and key debates. *Acad Manag Ann.* 4(1): 317–357.

Knudsen T, Levinthal DA (2007) Two faces of search: Alternative generation and alternative evaluation. *Organ. Sci.* 18(1):39–54.

Knudsen T, Srikanth K (2014) Coordinated Exploration: Organizing joint search by multiple specialists to overcome mutual confusion and joint myopia. *Adm. Sci. Q.* 59(3):409–441.

Kogut B, Metiu A (2001) Open-source software development and distributed innovation. *Oxford Rev. Econ. Pol.* 17(2):248–264.

Kovács B, Sharkey AJ (2014) The paradox of publicity: How awards can negatively affect the evaluation of quality. *Admin. Sci. Quart.* 59(1):1–33.

Kretschmer T, Puranam P (2008) Integration through incentives within differentiated organizations. *Organization Organ. Science*. 19(6): 860–875.

Lakhani, von Hippel (2003) How open-source software works: "free" user-to-user Assistance. *Res. Policy*. 32(7): 923–943.

Lakhani KR, Wolf RG (2005) Why hackers do what they do: Understanding motivation and effort in Free/Open Source Software projects. Feller J, Fitzgerald B, Hissam S, Lakhani KR, eds. *Perspectives on Free and Open Source Software* (MIT Press, Cambridge, MA), 3–23.

Lee GK, Cole RE (2003) From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organ. Sci.* 14(6):633–649.

Lee MY, Edmondson AC (2017) Self-managing organizations: Exploring the limits of less-hierarchical organizing. *Res. Organ. Behav.* 37:35–58.

Leonardi PM 2015 Ambient awareness and knowledge acquisition. *MIS Quart.* 39(4):747–762.

Lerner J, Lomi A (2020a) Reliability of relational event model estimates under sampling: How to fit a relational event model to 360 million dyadic events. *Network Sci.* 8(1):97–135.

Lerner J, Lomi A (2020b) The free encyclopedia that anyone can dispute: An analysis of the micro-structural dynamics of positive and negative relations in the production of contentious Wikipedia articles. *Soc. Networks* 60:11–25.

Lerner J, Tirole J (2001) The open-source movement: Key research questions. *Eur. Econ. Rev.* 45(4-6):819–826.

Lerner J, Tirole J (2002) Some simple economics of open-source. *J. Ind. Econ.* 50(2):197–234.

Levinthal DA, March JG (1993) The myopia of learning. *Strateg. Manag. J.* 14(S2): 95–112.

Lin YR, Keegan B, Margolin D, Lazer D 2014. Rising tides or rising stars? Dynamics of shared attention on Twitter during media events. *PloS oneOne*. 9(5):e94093.

MacCormack A, Rusnak J, Baldwin CY (2006) Exploring the structure of complex software designs: An empirical study of open-source and proprietary code. *Manage. Sci.* 52(7):1015–1030.

Majchrzak A, Malhotra A, Zaggl MA (2021) How open crowds self-organize. *Acad. Manag: Discov.* 7(1):104–129.

March JG (1991) How decisions happen in organizations. *Hum-Comput. Interact.* 6(2):95–117.

March JG, Olsen JP (1984) The new institutionalism: Organizational factors in political life. *Am. Polit. Sci. Rev.* 78(3):734–749.

March JG, Olsen JP 1976 Attention and the ambiguity of self-interest. March JG, Olsen JP, eds. *Ambiguity and Choice in Organizations* (Universitetsforlaget, Bergen, Norway), 38–53.

March JG, Simon HA (1958) *Organizations* (Wiley, New York).

March JG, Simon HA (1993*)* Introduction to the Second Edition. *Organizations, 2nd ed.* (Blackwell Publishers, Cambridge, MA).

Massa FG, O'Mahony S (2021) Order from chaos: How networked activists self-organize by creating a participation Architecture. *Adm. Sci. Q.* 66(4):1037–1083.

Merton RK (1968) The Matthew effect in science: The reward and communication systems of science are considered. *Science*. 159(3810):56–63.

Merton RK (1973) *The Sociology of Science: Theoretical and Empirical Investigations* (University of Chicago press, Chicago).

Mintzberg H (1979) *The Structuring of Organizations*. (Englewood Cliffs, New Jersey).

Moffett MW, Garnier S, Eisenhardt KM, Furr NR, Warglien M, Sartoris C, Ocasio W, Knudsen T, Bach LA, Offenberg J (2021) Ant colonies: Building complex organizations with minuscule brains and no leaders. *J. Organ. Des.* 10(1):55–74.

Nadler D, Tushman M (1997) *Competing by Design: The Power of Organizational Architecture*. (Oxford University Press, Oxford).

Newman ME (2002) Assortative mixing in networks. *Phys. Rev. Lett.* 89(20):208701.

Newman ME (2003) Mixing patterns in networks. *Phys. Rev. E.* 67(2):026126.

Newman ME, Park J (2003) Why social networks are different from other types of networks. *Phys. Rev. E.* 68(3):036122.

Newman, M. E., Park, J. (2003). Why social networks are different from other types of networks. *Physical review E*, *68*(3), 036122.

Ocasio W (1997) Towards an attention-based view of the firm. *Strateg. Manag. J.* 18(S1):187–206.

Ocasio W (2011) Attention to attention. *Organ. Sci.* 22(5):1286–1296.

Ocasio W (2012) Situated attention, loose and tight coupling, and the Garbage Can Model. Lomi A, Harrison JR, eds. *The Garbage Can Model of Organizational Choice: Looking Forward at Forty (Research in the Sociology of Organizations, Vol. 36)* (Emerald Group Publishing Limited, Bingley, UK), 293–317.

Ocasio W, Joseph J (2005) An attention-based theory of strategy formulation: Linking micro-and macroperspectives in strategy processes. *Strateg. Process*. 22:39–61.

Ocasio, W., Yakis-Douglas, B., Boynton, D., Laamanen, T., Rerup, C., Vaara, E., Whittington, R. (2022). It's a Different World: A Dialog on the Attention-Based View in a Post-Chandlerian World. *J. Manag. Inq.*

Padgett JF, Lee D, Collier N (2003) Economic production as chemistry. *Ind. Corp. Industrial and Corporate Change*. 12(4):843–877.

Padgett JF (2018) Faulkner's assembly of memories into history: Narrative networks in multiple times. *Am J. SociolAmerican Journal of Sociology*. 124(2): 406–478.

Padgett JF, Lee D, Collier N (2003) Economic production as chemistry. *Ind. Corp. Change.* 12(4):843–877.

Padgett JF, Powell WW (2012) The problem of emergence. Padgett JF, Powell WW, eds. *The Emergence of Organizations and Markets* (Princeton University Press, Princeton, NJ), 1–30.

Padgett, JF (1980) Managing garbage can hierarchies. *Adm. Sci. Q*. 583–604.

Pastor-Satorras R, Vázquez A, Vespignani A (2001) Dynamical and correlation properties of the Internet. *Phys. Rev. Lett.* 87(25):258701.

Pattison P, Robins G (2002) Neighborhood–based models for social networks. *Sociol. Methodol.* 32(1):301–337.

Perry PO, Wolfe PJ (2013) Point process modelling for directed interaction networks. *J. R. Stat. Soc. B* 75(5):821–849.

Piezunka H, Dahlander L (2015) Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Acad. Manag. J.* 58(3):856–880.

Podolny JM (2001) Networks as the pipes and prisms of the market. *Am. J. Sociol.* 107(1):33–60.

Powell WW, White DR, Koput KW, Owen-Smith J (2005) Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *Am. J. Sociol.* 110(4):1132–1205.

Prat A, Dessein W (2016) Attention in organizations. Bramoullé Y, Galeotti A, Rogers B, eds. *The Oxford Handbook of the Economics of Networks* (Oxford University Press, Oxford, UK), 675–697.

Puranam P (2018) *The Microstructure of Organizations*. (Oxford University Press, Oxford, UK).

Puranam P, Alexy O, Reitzig M (2014) What's "new" about new forms of organizing? *Acad. Manag. Rev.* 39(2):162–180.

Puranam P, Raveendran M, Knudsen T (2012) Organization design: The epistemic interdependence perspective. *Academy of Management Review*. 37(3):419–440.

Raymond E (1999) The cathedral and the bazaar. *Knowledge, Technology & Policy.* 12(3):23–49.

Reagans R, Argote L, Brooks D (2005) Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Manag. Sci.* 51(6):869–881.

Reitzig M (2022) How to get better at flatter designs: considerations for shaping and leading organizations with less hierarchy. *J. Organ. Des. Journal of Organization Design*. 11(1):5–10.

Ren CR, Guo C (2011) Middle managers' strategic role in the corporate entrepreneurial process: Attention-based effects. *J. Manag.* 37(6):1586–1610.

Rezgui A, Crowston K (2018) Stigmergic coordination in Wikipedia. *Proc. 14th Int. Symposium on Open Collaboration* (Association for Computing Machinery, New York), 1–12.

Rhee L, Leonardi PM (2018) Which pathway to good ideas? An attention-based view of innovation in social networks. *Strat. Manag. J.* 39(4):1188–1215.

Rivera MT, Soderstrom SB, Uzzi B (2010) Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annu. Rev. Sociol.* 36:91–115.

Robertson BJ (2015) *Holacracy: The New Management System for A Rapidly Changing World* (Henry Holt and Company, New York).

Robles G, González-Barahona JM, Izquierdo-Cortazar D, Herraiz I (2009) Tools for the study of the usual data sources found in libre software projects. *Int. J. Open Source Softw. Process.* 1(1):24–45.

Rolandsson B, Bergquist M, Ljungberg J (2011) Open-source in the firm: Opening up professional practices of software development. *Res. Policy.* 40(4):576–587.

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science.* 311(5762):854–856.

Schelling TC (1998) Social mechanisms and social dynamics. Hedström P, Swedberg R, eds. *Social Mechanisms: An Analytical Approach to Social Theory.* (Cambridge University Press, New York), 32–44.

Shah SK (2006) Motivation, governance, and the viability of hybrid forms in open-source software development. *Manage. Sci.* 52(7):1000–1014.

Shang MS, Lu L, Zhang YC, Zhou T (2010) Empirical analysis of web-based user-object bipartite networks. *Europhys. Lett.* 90(4):48006.

Shore J, Bernstein E, Lazer D (2015) Facts and figuring: An experimental investigation of network structure and performance in information and solution spaces. *Organ. Sci.* 26(5):1432–1446.

Simon HA (1947) *Administrative Behavior: A Study of Decision Making Process in Administrative Organization, 2nd ed.* (Free Press, New York).

Simon HA (1959) Theories of decision-making in economics and behavioral science. *Am. Econ. Rev.* 49(3):253–283.

Simon HA (1978) Rationality as process and as product of thought. *Am. Econ. Rev.* 68(2):1–16.

Smith EB (2011) Identities as lenses: How organizational identity affects audiences' evaluation of organizational performance. *Admin. Sci. Quart.* 56(1):61–94.

Spaeth S, Haefliger S, von Krogh G, Renzl B (2008) Communal resources in open-source software development. *Inform. Res.* 13(1).

Stam W (2009) When does community participation enhance the performance of open-source software companies? *Res. Policy.* 38(8):1288–1299.

Stewart D (2005) Social status in an open-source community. *Amer. Sociol. Rev.* 70(5):823–842.

Stewart KJ, Gosain S (2006) The impact of ideology on effectiveness in open-source software development teams. *MIS Quart.* 30(2):291–314.

Stuart HC, Dabbish L, Kiesler S, Kinnaird P, Kang R (2012) Social transparency in networked information exchange: A theoretical framework. *Proc. ACM 2012 Conf. on Computer Supported Cooperative Work*, 451–460.

Stuart TE, Podolny, JM (1996) Local search and the evolution of technological capabilities. *Strateg. Manage. J.* 17(S1):21–38.

Sullivan BN (2010) Competition and beyond: Problems and attention allocation in organizational rulemaking processes. *Organ. Sci.* 21(2):432–450.

Tajfel H, Turner JC (1986) The social identity theory of intergroup behavior. Worchel S, Austin W, eds. *Psychology of Intergroup Relations* (Nelson-Hall, Chicago), 7–24.

Tasselli S, Caimo A (2019) Does it take three to dance the Tango? Organizational design, triadic structures and boundary spanning across subunits. *Soc. Networks*, 59, 10-22.

Tasselli S, Kilduff M (2021) Network agency. *Acad. Manag. Annals*, 15(1), 68-110.

Tasselli S, Kilduff M, Menges JI (2015) The microfoundations of organizational social networks: A review and an agenda for future research. *J. Manag.* 41(5), 1361-1387.

Taylor SE, Fiske ST (1978) Salience, attention, and attribution: Top of the head phenomena. *Adv. Exp. Soc. Psychol.* 11(C):249–288.

Therneau, TM (2022) Survival: A package for survival analysis in R (Version 3.4-0). Retrieved August 25, 2022, https://CRAN.R-project.org/package=survival.

Thompson JD (1967) *Organizations in Action: Social Science Bases of Administrative Theory* (McGraw-Hill, New York).

Tversky A (1972) Elimination by aspects: A theory of choice. *Psychol. Rev.* 79(4):281–299.

Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science.* 185(4157):1124–1131.

von Krogh G, Haefliger S, Spaeth S, Wallin MW (2012) Carrots and rainbows: Motivation and social practice in open-source software development. *MIS Quart.* 36(2):649–676.

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

von Krogh G, von Hippel E (2006) The Promise of research on open-source software. *Manage. Sci.* 52(7):975–983.

Wang P, Pattison P, Robins G (2013) Exponential random graph model specifications for bipartite networks: A dependence hierarchy. *Soc. Networks.* 35(2):211–222.

Zhang H, Kim S (2010) Monitoring software quality evolution for defects. *IEEE Softw.* 27(4):58–64.

Zuckerman EW (2012) Construction, concentration, and (dis)continuities in social valuations. *Annu. Rev. Sociol.* 38(1):223–245.

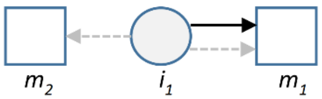
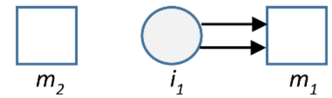
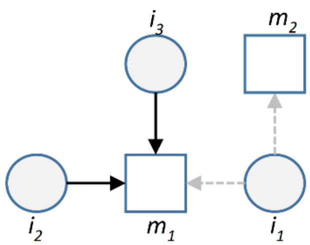
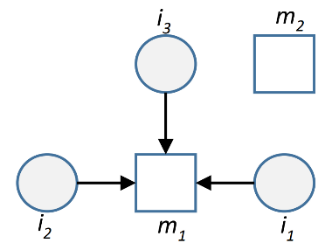
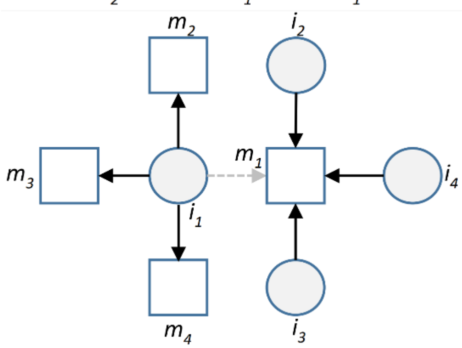
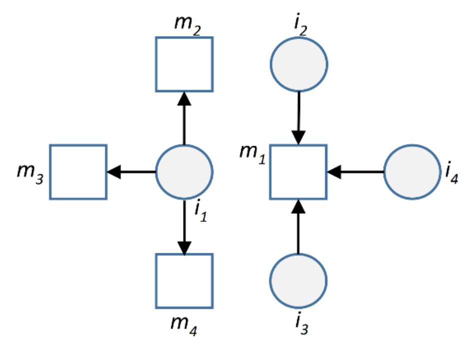
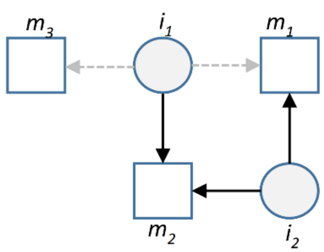
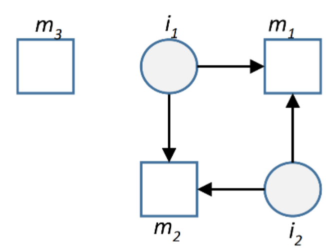
Table 1: summary of attention network mechanisms with examples

Attention network mechanism	Description	Inherent behavioral logic	Contextual example	Structural outcome
Attention focusing	Acts of attention allocation are dedicated to problems that are *familiar* (i.e. addressed in the past by the focal participant).	Logic of *familiarity*: people manifest a preference for problems that already captured their own attention due to dynamics of habituation.	A software developer tends to focus on contributing repeatedly to solve a specific bug rather than spreading attention to different bugs.	Attention is repeatedly allocated to the same problems.
Attention reinforcing	Acts of attention allocation are dedicated to problems that are *popular* (i.e. addressed by many others).	Logic of *popularity*: people manifest a preference for problems that already captured the attention of other people, due to an increased perception of interest and opportunities.	A software developer tends to attend to bugs that exhibit a wealth of prior activities by other developers rather than to less attended bugs.	Attention becomes progressively more concentrated on a limited number of problems.
Attention mixing	Acts of attention allocation by participants who are highly active are progressively diverted away from popular problems.	Logic of *devaluation*: people's preference for popular problems fades away as participants become more active, due to a decreased interest for popular items.	A software developer, while becoming more active in the community, increasingly tends to dedicate attention to bugs left unattended by the rest of the community.	Attention of active participants spreads over less popular problems compensating, in part, for the tendency toward concentration of attention to already popular problems.
Attention clustering	Acts of attention allocation are dedicated to problems that are situated in the *neighborhood* of already attended problems.	Logic of *proximity*: people manifest a tendency to allocate their attention to proximate problems in their neighborhood, due to a behavioral tendency for local search.	A software developer tends to extend previous collaborations with other developers to future bugs.	Attention develops around a *neighborhood* of participants who collaborate by giving attention to the same sets of problems.

Table 2: Generalizability of attention network mechanisms to other settings

Setting	Attention focusing	Attention reinforcing	Attention mixing	Attention clustering
Distributed software development (e.g., GitHub)	Software developers focus repeatedly their attention to specific projects rather than spreading it across many different projects.	Software developers contribute attention to projects that are well attended by other developers rather than to less popular projects.	While accumulating activity, software developers increasingly divert their attention away from popular to less popular projects.	Software developers extend previous collaborations by attending to the development of proximal projects in their attention network.
Crowdsourced innovation (e.g., idea generation platforms)	Platform participants concentrate their attention on contributing repeatedly to the same ideas	Platform participants allocate attention to ideas that received a high volume of contributions from other participants.	While accumulating activity on the platform, participants diversify their scope, diverting their attention away from popular ideas.	Participants extend their collaboration with other participants by allocating attention to proximal ideas in their attention network.
Bossless organizations (e.g., holacracy)	Employees in self-managing teams focus their attention on engaging repeatedly the same type of tasks	Employees allocate their attention to tasks that received a high volume of attention by peer employees	While accumulating activity, employees divert their attention away from popular tasks to less popular ones.	Employees extend their collaboration across teams by engaging in proximal tasks in their attention networks
Scientific research production	Researchers accumulate specialized knowledge by focusing their attention repeatedly on specific topics.	Researchers are attracted by topics that are well attended in their scientific community.	While accumulating activity, researchers broaden their expertise by engaging with topics that are less popular in the scientific community.	Researchers extend collaboration with other researchers to proximate topics in their attention network.

Table 3: Summary of hypotheses

Time = t1	Predictions	Time = t2 (>t1)
	**Attention focusing (H1)** Preferential tendency of a participant *(i1)* to allocate attention to a problem *(m1)* at time *t2* that has attracted his attention at time *t1*.	
	**Attention reinforcing (H2)** Preferential tendency of a participant *(i1)* to allocate attention at time *t2* to problems *(m1)* that attracted a higher level of collective attention at time *t1*.	
	**Attention mixing (H3)** Preferential tendency of a participant *(i1)* that is very active at time *t1* against allocating attention to problems that are very popular *(m1)* at time *t2*.	
	**Attention clustering (H4)** Preferential tendency of a participant *(i1)* to allocate attention to a problem (*m1*) at time *t2*, if that problem is in its neighborhood – i.e., it is reachable through a three-path *(i1– m2 – i2 – m1)* indirectly connecting *i1* to *m1*.	

Legend: Grey circles are participants (i1, i2). White squares are problems (m1, m2, m3, m4). Arrows represent the relation "allocate attention to." Dashed gray arrows denote potential opportunities of attention allocation. Solid black arrows denote observed attention allocation events. Therefore, potential opportunities can only be present at time t_1 (i.e., in the left column of the table).

Table 4: Variables, units and measures

Effect	Variable type	Unit of measurement	Measure	Included in the model to capture
Institutional participant	Binary	Dimensionless	Indicator variable = 1 if participant's email address ends in "apache.org," and = 0 otherwise	*Preferential tendency of institutional participants to become involved in problem solving activities*
Experienced participant	Binary	Dimensionless	Indicator variable = 1 if participant was active in prior release cycles, and = 0 otherwise	*Preferential tendency of experienced participants to become involved in problem solving activities*
Problem priority	Ordinal	Dimensionless	Priority level assigned to the problem (5 priority levels where 1 = highest priority)	*Differential tendency of problems to attract attention as a function of their assigned level of priority*
Problem severity	Binary	Dimensionless	Indicator variable = 1 if problem is classified as severe in the bug report, and = 0 otherwise	*Differential tendency of problems to attract attention as a function of their assigned level of severity*
Problem latency	Numerical	Units (days)	Problem age	*Differential tendency of problems to attract attention as a function of their age*
Problem resolved	Binary	Dimensionless	Indicator variable = 1 if problem is resolved, and = 0 otherwise	*Differential tendency of problems declared resolved to attract further attention*
Problem recognition	Numerical	Units (messages)	Number of comments generated by a problem	*Differential tendency of problems to attract attention as a function of the comments they have generated*
Time inactive	Numerical	Units (days)	Number of days elapsed since last contribution	*Time of inactivity of participants*
Cumulative attention	Numerical	Units (Events)	Overall number of events connecting participants to problems observed within the observation period	*Overall volume of problem solving activity within the project*
Module indicators	Binary	Dimensionless	Indicator variable = 1 for problems belonging to one of the five main modules of the project	*Tendency of formal organizational structure to channel attention of participants toward specific classes of problems*

Table 5: Cox Regression Model: Partial likelihood estimates of bipartite relational event models (standardized estimates)

	Null (M0)	Control (M1)	REM (M2)	Org Structure (M3)
Cumulative attention	2.5376 (0.0233) ***	0.8813 (0.0144) ***	0.7808 (0.0143) ***	0.7848 (0.0143) ***
Experienced participant		0.9060 (0.0476) ***	0.9767 (0.0557) ***	1.0444 (0.0568) ***
Institutional participant		1.0971 (0.0361) ***	0.9537 (0.0392) ***	0.9736 (0.0395) ***
Problem priority		-0.0334 (0.0256)	-0.0175 (0.0275)	-0.0341 (0.0279)
Problem severity		0.1409 (0.0323) ***	0.1519 (0.0347) ***	0.1531 (0.0356) ***
Problem latency		-0.0019 (0.0000) ***	-0.0017 (0.0001) ***	-0.0017 (0.0001) ***
Problem resolved		-1.4596 (0.0356) ***	-1.7946 (0.0418) ***	-1.7897 (0.0420) ***
Problem recognition		0.5253 (0.0108) ***	0.3430 (0.0143) ***	0.3440 (0.0145) ***
Time inactive		-0.0163 (0.0003) ***	-0.0135 (0.0003) ***	-0.0135 (0.0003) ***
Attention focusing (H1)			0.2150 (0.0066) ***	0.2029 (0.0065) ***
Attention reinforcing (H2)			0.1354 (0.0116) ***	0.1320 (0.0117) ***
Attention mixing (H3)			-0.0320 (0.0028) ***	-0.0308 (0.0027) ***
Attention clustering (H4)			0.0349 (0.0039) ***	0.0452 (0.0070) ***
Module 1				-0.1276 (0.0634) *
Module 2				-0.2006 (0.0660) **
Module 3				0.1717 (0.0732) *
Module 4				-0.5103 (0.0557) ***
Preferential modularity				0.4094 (0.0243) ***
Attention clustering w/in modules				-0.0128 (0.0069)
Log-Likelihood	-33,021.85	-12,844.57	-11,434.08	-11,266.95
AIC	66,045.7081	25,707.1446	22,898.1567	22,575.9039
LR test	-	40,355	43,176	43,510
Num. events	11,599	11,599	11,599	11,599
Num. obs.	1,170,871	1,170,871	1,170,871	1,170,871

*** $p < 0.001$, **$p < 0.01$, *$p < 0.05$; *3-path* effects estimates not reported in table.

Table 6: Qualitative implications of estimates for parameters of theoretical interest.

Hypotheses	Hazard ratio (M2, Table 5)	Heuristic interpretation
(H1) Attention focusing	exp(0.2150) = 1.240	An increase by one standard deviation in the level of attention that a participant has allocated to one specific problem in the past increases the odds that he will decide to pay attention to the same problem again by approximately 24% (please refer to text for further discussion).
(H2) Attention reinforcing	exp(0.1354) = 1.145	An increase by one standard deviation in the level of attention that a problem has received within the community in the past increases the odds that the same problem will attract further attention by approximately 15% (please refer to text for further discussion).
(H3) Attention mixing	exp(-0.0320) = 0.969	A participant that is by approximately 7 events (one standard deviation) more active than the average participant, experiences a reduction in the strength of attention reinforcing effect only by approximately 3% (please refer to text for further discussion).
(H4) Attention clustering	exp(0.0349) = 1.036	An increase by approximately 10 (one standard deviation) in the number of three-paths indirectly connecting a participant to a problem will increase by slightly less than 4% the odds that this participant will allocate attention directly to this problem in the future (please refer to text for further discussion).

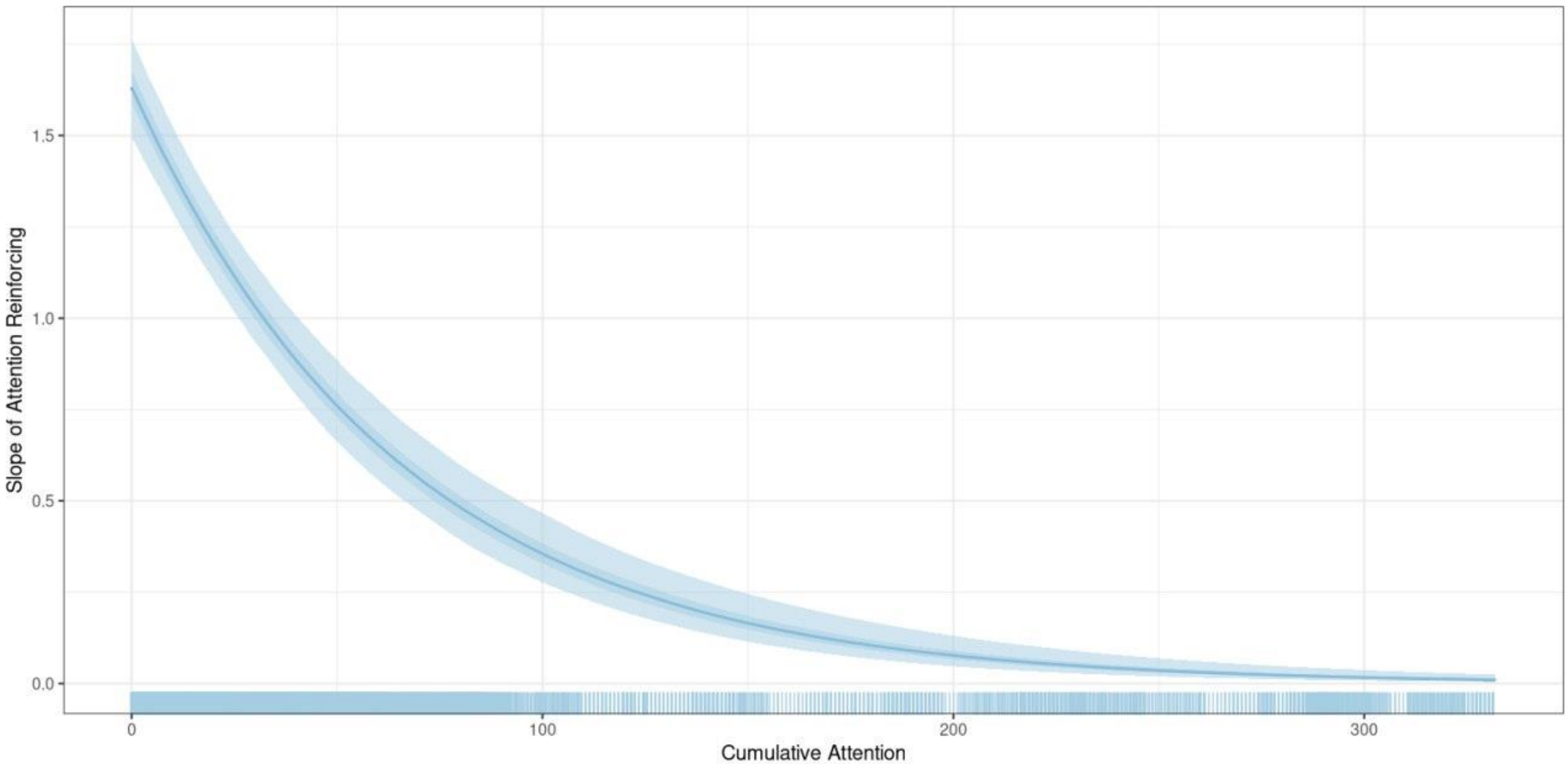


Figure 1: Johnson-Neyman plot (with 95% CI) of *H3: Attention Mixing* interaction term. The plot shows how the marginal effect (i.e., the slope) of *Attention reinforcing* decreases as the *Cumulative attention* of participants increases. The marginal effect is calculated on the hazard ratio, therefore values above 1 on the y axis denote a positive effect of *Attention reinforcing* (on the probability of an attention allocation event) and values below 1 a negative effect

Appendix A: Notation and definitions

Appendix A establishes the basic notation needed for describing with accuracy the effects of theoretical interests and the control factors included in the empirical model specification and discussed in the paper.

Organizational attention networks

Consider an organization as a set of problems $M = \{m_1, m_2, \ldots, m_k\}$, and a set of participants $I = \{i_1, i_2, \ldots, i_l\}$. The models estimated in the paper interpret an individual act of attention allocation as a relational event $e = (t_e, i_e, m_e)$ connecting participant i_e to problem m_e at time t_e. An attention network is the set of all relational events $E = \{e_1, e_2, \ldots, e_n\}$ connecting individuals to problems at the given event times. Membership in the sets M and I is updated at every event time t_e – which in the data we collected is precise to the second.

An attention network is empty (denoted E_0) if no organizational participant ever pays attention to any organizational problem. The attention network is complete (denoted as $E_{m,l}$) if every participant pays attention to every organizational problem at every given time. As Cohen et al. (1972) note, the allocation of attention in organizations typically generates relational patterns that fall within these two extreme cases: some organizational participants attend to some organizational problem some of the times.

Baseline attention allocation mechanisms

As explained in the main text, *Cumulative attention* provides the basic attention allocation mechanism that we use as a baseline. *Cumulative attention* of participant $i \in I$ is defined as:

$$\text{Cumulative attention}(t, i) = \sum_{m \in M} \sum_{e=1}^{N_{im}(t^-)} w\left(t, T_{im}^e\right), \qquad \textit{(A.1)}$$

where T_{im}^e is the timestamp associated with event e connecting participant i to problem m. The count $N_{im}(t^-)$ is the number of past events from participant i to problem m that have happened strictly before t. The function $w(t, T_{im}^e)$ assigns a temporal weight to events so that every event observed has an immediate effect when is recorded and then a delayed effect on subsequent events that decays over time. The precise definition of the temporal weight for the event e is given by:

$$w(t, t_e) = \exp\left(-(t - t_e) \cdot \frac{\ln(2)}{T_{1/2}}\right) \qquad \textit{(A.2)}$$

for a given half life $T_{1/2} > 0$, where t_e is the time of the event included in the computation of the network statistics, and t is the current time, i. e., the time in which the partial likelihood function is evaluated (see Appendix B). Following Brandes et al. (2009); Lerner et al. (2013), Equation (A.2) assigns a weight that progressively decreases toward zero to events that happened in the more distant past. In this way, current attention allocation events have a continuously decreasing effect on future events. In our models, we set the half-life equal to 10 days.

Hypotheses

Attention focusing is the cumulative number of attention allocation events connecting participant i to problem j before time t – downweighted by the elapsed time – and is defined as:

$$\text{Attention focusing}(t, i, j) = \sum_{e=1}^{N_{ij}(t^-)} w\left(t, T_{ij}^e\right), \tag{A.3}$$

where t^- indicates the complete history of the event network up to t, and the time-weighting function $w\left(t, T_{ij}^e\right)$ is defined in Equation (A.2).

Attention reinforcing is defined simply as the current number of attention allocation events performed on a problem – downweighed by the elapsed time:

$$\text{Attention reinforcing}(t, j) = \sum_{g \in I} \sum_{e=1}^{N_{gj}(t^-)} w\left(t, T_{gj}^e\right), \tag{A.4}$$

where T_{gj}^e is the timestamp of attention allocation event e from participant g to problem j, and the function $w\left(t, T_{gj}^e\right)$ defined in Equation (A.2) assigns a temporal weight for the event e.

To examine mixing patterns (or assortativity), *Attention mixing* is simply the interaction of cumulative attention and attention reinforcing:

$$\text{Attention mixing}(t, i, j) = \sum_{m \in M} \sum_{e=1}^{N_{im}(t^-)} w\left(t, T_{im}^e\right) \cdot \sum_{g \in I} \sum_{e=1}^{N_{gj}(t^-)} w\left(t, T_{gj}^e\right). \tag{A.5}$$

In a bipartite data structure, such as the one we analyze, closure cannot involve an odd number of links. Therefore, *Attention clustering* takes the form of a bipartite four-cycle and can be formally defined as follows:

$$\begin{aligned}\text{Attention. clustering}&(t, i, j) \\ &= \sum_{m \neq j} I\left[N_{im(t^-)} > 0\right] \cdot \sum_{g \neq i} I\left[N_{gj(t^-)} > 0\right] \cdot I\left[N_{gm(t^-)} > 0\right],\end{aligned} \tag{A.6}$$

where $I[x]$ is an indicator function that equals to 1 if statement x is true, otherwise $I[x] = 0$. In Equation (A.6), g and m index participants and problems, respectively. *Attention clustering* gives the number of three-paths indirectly connecting the participant i of the focal dyad with its problem j via another participant g different from i and another problem m different from j; these three-paths are closed to four-cycles by an event on (i, j). This statistic is contingent on the number of 3-paths attached to the participant i and the number of 3-paths attached to the problem j, defined in the following Equation (A.7) and Equation (A.8), respectively.

The number of three paths attached to the participant i of a given dyad (i, j) is defined as follows:

$$\text{3path.participant}(t,i,j) = \sum_{j'} \sum_{m \neq j'} I\left[N_{im(t^-)} > 0\right] \cdot \sum_{g \neq i} I\left[N_{gj'(t^-)} > 0\right] \cdot I\left[N_{gm(t^-)} > 0\right]. \tag{A.7}$$

In contrast to the four-cycle statistic, *3paths.participant* does not require that the three-paths starting at participant i end at the problem j of the focal dyad but instead allows that these three-paths can end at any problem j' that may, but does not have to, be different from j. The statistic *3paths.participant* is one of the two antecedents of the four-cycle statistic since it gives the three-paths that may be closed to a four-cycle in the next event initiated by i. The other antecedent of the four-cycle statistic is the number of three-paths attached to the problem j of the given dyad (i,j) defined by:

$$\text{3paths.problem}(t,i,j) = \sum_{i'} \sum_{m \neq j} I\left[N_{i'm(t^-)} > 0\right] \cdot \sum_{g \neq i'} I\left[N_{gj(t^-)} > 0\right] \cdot I\left[N_{gm(t^-)} > 0\right]. \tag{A.8}$$

Similar to *3paths.participant*, the statistic *3paths.problem* does not require that the three-paths ending at problem j start at the participant i of the focal dyad but instead allows that these three-paths can start at any participant i' that may, but does not have to, be different from i. The statistic *3paths.problem* is the second of the two antecedents of the four-cycle statistic since it gives the three-paths that may be closed to a four-cycle in the next event received by j.

Control factors

Problem latency records the age of each problem and is defined as:

$$\text{Problem latency}(t,j) = t - T_j^{\text{opened}}, \tag{A.9}$$

where T_j^{opened} is the time when problem j is fist reported – or reopened if it had already been declared resolved in the past.

Problem recognition is defined as:

$$\text{Problem recognition}(t,j) = \sum_{i=1}^{l} \sum_{e=1}^{C_{ij}(t^-)} w\left(t, T_{ij}^{e}\right), \tag{A.10}$$

where $C_{ij}(t^-)$ is the set of comment events from participant i to problem j before time t, T_{ij}^{e} is the time stamp of comment event e, and the time-weighting function $w\left(t, T_{ij}^{e}\right)$ is defined in Equation (A.2).

Preferential modularity is defined as:

$$\text{Preferential modularity}(t, i, j) = \frac{\sum_m I\,[z(m) = z(j)] \cdot I[N_{im}(t^-) > 0]}{\sum_m I\,[N_{im}(t^-) > 0]}\,, \qquad (A.11)$$

where $z(j)$ is the module index of problem j. The statistic is the proportion of problems in the same module as problem j out of all problems addressed by participant i.

Attention clustering within modules for participant i to problem j at time t is defined as follows:

$$\begin{aligned} &\text{Attention clustering w.m.}(t, i, j) \\ &= \sum_{m \neq j} I\,[z(m) = z(j)] \cdot I\big[N_{im(t^-)} > 0\big] \cdot \sum_{g \neq i} I\,\big[N_{gj(t^-)} > 0\big] \\ &\cdot I\big[N_{gm(t^-)} > 0\big]\,, \end{aligned} \qquad (A.12)$$

where the first indicator equals 1 if problems j and m are in the same module, otherwise $I[z(m) = z(j)]$ is zero. The effect described in Equation (A.12) is used to examine the preferential tendency of attention to cluster within software modules, i. e., to examine the extent to which formal organizational structure affects the joint allocation of attention to the same problems.

Finally, *Time inactive* records for each participant i the time difference between the current time and the last time i was active in addressing any problem. It is defined as

$$\text{Time inactive}(t, i) = t - T_i^{\text{last active}}(t)\,, \qquad (A.13)$$

where $T_i^{\text{last active}}(t)$ is the maximum time stamp, strictly before the current time t, when participant i addressed any problem. time inactive(t, i) is undefined for participants i that have never been active before t. Indeed, we define that participants enter the risk set right after their first event (first events of participant are not modeled). Once they have entered the risk set, participants never leave it again – but the time inactive provides an important control mechanism that let the probability that participants initiate any further event tend to zero as their time inactive increases. Participants with long time inactive will hardly influence the estimation of other model effects; they are “almost out of the risk set”. This is a more principled way to deal with participants who apparently have decided not to contribute again than defining an arbitrary crisp cut-off after which participants are removed from the risk set. Strictly speaking, even after a prolonged time of inactivity, participants *could* still decide to become active again – even though it is increasingly unlikely.

Appendix B: Model estimation, interpretation, and evaluation

Point process models for bipartite networks

The model we implement in the empirical part of the paper is based on a bipartite extension of the point process models for directed social interaction networks proposed by Perry and Wolfe (2013). Introduced by Butts (2008) as a strategy for the analysis of social networks, this class of models is also known as relational event models. The advantage of this approach over more conventional models for networks is its ability to analyze sequences relational events directly, rather than as aggregate network "ties". Our data require that the model be adapted to two-mode networks – networks containing two classes of nodes, with relations defined only between nodes in different classes (Everett and Borgatti 2013). The counting process framework developed in the analysis of repeated events within event history analysis provides the statistical foundation for the models that we develop (Aalen et al. 2008).

Modeling the evolutionary dynamics of attention network connecting participants to problems starts by defining a counting process $N_{ij}(t)$ on the dyad linking participant i and problem j. The counting process $N_{ij}(t)$ increases (or "jumps") by one unit whenever participant i allocates attention to problem j at time t. When we have l participants and k problems, the total number of counting processes is $l \times k$. Following Perry and Wolfe (2013), each process is modeled by a conditional intensity function $\lambda_{ij}(t)$ taking form of the Cox proportional intensity model (Cox 1972; Cook and Lawless 2007):

$$\lambda_{ij}(t \mid H_{t^-}) = R_{ij}(t)\lambda_0(t) exp[\theta^T s(t,i,j)]\,, \tag{B.1}$$

where H_{t^-} is the complete network history right before time t, $s(t,i,j)$ is the vector of time-varying statistics, and θ is a vector of coefficients to be estimated from data. $R_{ij}(t)$ is the "at-risk" indicator which equals 1 if participant i can perform actions on problem j at time t. This happens when the participant is active (that is, right after her first event) and the bug report is opened so they are both in the risk set at time t. Otherwise, $R_{ij}(t) = 0$. The "at risk" indicator function plays a central role in our models because it records the continuous change in the flow of problems and participants, thus controlling what actions are possible at any given moment – i. e., the "opportunity set." When all elements of $s(t,i,j)$ are set to zero, the intensity equals the baseline rate $\lambda_0(t)$. To account for potential baseline rate changes during the observation time, we assume a non-parametric form of $\lambda_0(t)$, a flexible and widely used approach in survival and event history analysis (Andersen and Keiding 2002; Perry and Wolfe 2013; Vu et al. 2011).

Model estimation

Thanks to the non-parametric choice of the baseline rate $\lambda_0(t)$, the effects associated with the network statistics discussed in the paper can be estimated by maximizing the partial likelihood (Andersen et al. 2012).

$$\mathrm{PL}(\theta) = \prod_{e \in E} \frac{\exp[\theta^T s(t_e, i_e, j_e)]}{\sum_{(i,j)\in R(t_e)} \exp\,[\theta^T s(t_e, i, j)]}\,, \tag{B.2}$$

where E is the set of attention allocation events; $R(t_e)$, contains every dyad (i,j) for which the indicator $R_{ij}(t_e)$ is equal to 1. Perry and Wolfe (2013, Appendix B) provide a proof of the consistency of inference based on maximum partial likelihood for this model.

Defining a counting process for events on each participant-problem pair makes computation unfeasible since the number of events may also be very large. To alleviate this computational constraint, we employ the nested case-control sampling approach (Borgan et al. 1995). Under this sampling method, for each event included in the sample, we randomly select a subset of non-events (case controls) from the current risk set $R(t)$ to compute the denominator sum in the partial likelihood (B.2). This results in the sampled partial likelihood of the form (Borgan et al. 1995):

$$\widetilde{\mathrm{PL}}(\theta) = \prod_{e \in E} \frac{\exp[\theta^T s(t_e, i_e, j_e)]}{\sum_{(i,j) \in \tilde{R}(t_e)} \exp\,[\theta^T s(t_e, i, j)]}, \tag{B.3}$$

where $\tilde{R}(t_e)$ includes the case and only the sampled controls at the event time t_e. For our current analysis, we sample up to 100 controls for each observed event (Lerner and Lomi 2020). This results in a final data set of 11,599 cases and 1,170,868 nested controls for the estimation. Most commercial or open-source statistical software can be used for parameter estimation based on this sampled partial likelihood. The results that we report are based on the *survival* package (Therneau and Grambsch 2013) in the R software for statistical computing.

Parameter interpretation

We interpret estimated network effects in terms of hazard ratios, a common concept in survival analysis (Aalen et al. 2008). The hazard ratio Ψ_p of a network statistic s_p is defined as the ratio of the intensity function for dyads with the statistic value $s_p(t, i, j) = v + 1$ to the intensity function of those with one unit smaller in that network statistic, i. e. $s_p(t, i, j) = v$, while holding all other statistics constant. It can also be thought of as the odds that attention allocation events will occur on dyads with $s_p(t, i, j) = v + 1$ over those with $s_p(t, i, j) = v$, all other statistics being equal. The hazard ratio can be estimated by the formula $\Pi_p = \exp(\beta_p)$, where β_p is the maximum likelihood estimate of the parameter corresponding to the network statistic $s_p(t, i, j)$.

Appendix C: Robustness Checks

Attention Intensity Levels. In constructing our dataset, we record an attention allocation event whenever a developer chooses to allocate attention on – and thus visibly "touch" – a bug report without distinguishing between the varying levels of effort required. All actions leave visible cues potentially catching the attention of other participants; thus, our modelling approach reflects the fact that the intensity of the effort behind each specific act of attention allocation does not directly affect the theoretical arguments that underlie our hypothesized mechanisms. However, could it be that attention allocation patterns do vary significantly depending on the level of effort put into the acts visible to participants? To answer this question, we estimated new models including interaction effects between the four variables capturing the attention mechanisms we hypothesized and a new variable called *High attention effort*. In our empirical setting, the intensity of attention acts can be inferred by considering the nature of the bug report modification that each act represents. We consider acts that involve the direct production or review of software code – intended as a patch for the focal bug – as a proxy for "high attention effort". Conversely, "low attention effort" acts are those addressed at more mundane tasks contributing to the description, general classification, and maintenance of software bugs (Lakhani & von Hippel, 2003). The results of these additional tests show that all four main effects are still significant and consistent with our hypotheses (see Table C.1). *Attention focusing* and *Attention mixing* show significant interactions going in the same direction of their respective main effect, thus representing a reinforced effect for high attention effort events. Our additional findings suggest that the dynamics that underlie the self-organizing properties of the attention structures that we investigate are substantively similar for high and low attention efforts.

Crowding. In our modelling approach the idea that attention is limited is just assumed – in line with our theoretical framework– and thus not directly tested. However, within that assumption, to which problem a participant allocates attention could depend on the amount of choice opportunities available, a concept referred to as "crowding" in related literature (e.g., Piezunka and Dahlander, 2015). According to this view, attention could become more limited as there are more choice opportunities available, and individuals could become more selective in allocating their limited attention to competing issues. Do crowding levels have a significant effect of the mechanisms of attention allocation we hypothesize? To answer this question, we coded a new variable *Crowding* – similarly to what done by Piezunka and Dahlander (2015) – by counting, for each attention allocation event recorded, all problems "at risk" of attracting attention acts. We also applied an exponential decay function to the count, with 60 days half-life, thus giving more emphasis to newer problems (we tested alternative specifications of 30 and 90 days with similar results). We then interacted *Crowding* with the four variables capturing the attention mechanisms we hypothesize. The results in Table C.2 show that all four main effects are still significant and consistent with our hypotheses. While all new interaction effects are statistically significant, our main effects maintain the same direction and significance once the moderator is included in the model, with crowding only affecting the relative magnitude of the effects. These results suggest that the mechanisms underlying our hypotheses are robust to this "ecological" conceptualization of limited attention.

Returning Problems. In our modelling approach, problems that are considered solved remain in the risk set as they could be re-opened at a later stage and still attract the attention of participants. It is however reasonable to expect a reduced attractiveness of problems marked as resolved and our control variable Problem resolved confirms this intuition consistently in our models. However, distinct attention mechanisms pertaining to problems resolved and re-opened could exist and potentially confound our results. To address the point, we re-estimated our models excluding problems from the risk set after they were marked as resolved once. The number of bugs excluded from the analysis is not high (approximately 200) and the estimates are stable and fully consistent with our previous results (see Table C.3).

Extreme Outliers. In our empirical setting the number of recorded attention allocation acts vary significantly across participants and is not normally distributed. Indeed, we expect the mechanisms underlying our hypotheses to produce an uneven concentration in attention allocation and our modelling approach is suited to handle this skew. Nonetheless, could it be that the results we find are only driven by the actions of the most active participants? To answer this question, we re-estimated our models excluding the more severe outliers amongst the participants. We computed the 90% quantile of cumulative attention over all events and removed all observations above that threshold from the dataset. The results in Table C.4 show that the new estimates are stable and consistent with our previous results.

Table C.1: Cox Regression Model: High vs. low attention effort

	Model 1	**Model 2**
Cumulative attention	0.9125 (0.0147) ***	0.8098 (0.0142) ***
Experienced participant	1.4348 (0.0554) ***	1.2900 (0.0556) ***
Institutional participant	1.1685 (0.0373) ***	1.1428 (0.0377) ***
Problem priority	-0.0250 (0.0270)	-0.0223 (0.0275)
Problem severity	0.1710 (0.0347) ***	0.1498 (0.0350) ***
Problem latency	-0.0020 (0.0000) ***	-0.0018 (0.0000) ***
Problem resolved	-1.9565 (0.0412) ***	-1.8208 (0.0411) ***
Problem recognition	0.3474 (0.0153) ***	0.3340 (0.0151) ***
Time inactive	-0.0170 (0.0003) ***	-0.0155 (0.0003) ***
Attention focusing (H1)	0.2357 (0.0067) ***	0.2096 (0.0066) ***
Attention reinforcing (H2)	0.1608 (0.0119) ***	0.1448 (0.0120) ***
Attention mixing (H3)	-0.0342 (0.0028) ***	-0.0322 (0.0027) ***
Attention clustering (H4)	0.0449 (0.0070) ***	0.0444 (0.0070) ***
Module 1	-0.1431 (0.0609) *	-0.1132 (0.0616)
Module 2	-0.1652 (0.0640) **	-0.1386 (0.0645) *
Module 3	0.1328 (0.0709)	0.1175 (0.0719)
Module 4	-0.5976 (0.0542) ***	-0.5334 (0.0547) ***
Preferential modularity	0.4003 (0.0232) ***	0.3830 (0.0234) ***
Attention clustering w/in modules	-0.0055 (0.0069)	-0.0045 (0.0069)
High attention effort		-3.0232 (0.0849) ***
Attention focusing * High attention effort		0.1266 (0.0208) ***
Cumulative attention * High attention effort		-0.4138 (0.0990) ***
Attention reinforcing * High attention effort		-0.0442 (0.0402)
Attention clustering * High attention effort		-0.0287 (0.0289)
Attention mixing * High attention effort		-0.2865 (0.0931) **
AIC	29,255.1392	24,287.3073
Num. events	11,599	11,599
Num. obs.	2,330,137	2,330,137

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table C.2: Cox Regression Model: Interacting attention crowding with effects of interest

	Model 1	**Model 2**
Cumulative attention	0.7706 (0.0139) ***	1.0765 (0.0240) ***
Experienced participant	1.2818 (0.0560) ***	1.2079 (0.0589) ***
Institutional participant	1.1832 (0.0377) ***	1.2699 (0.0389) ***
Problem priority	-0.0526 (0.0276)	-0.0446 (0.0287)
Problem severity	0.1494 (0.0349) ***	0.1579 (0.0357) ***
Problem latency	-0.0017 (0.0000) ***	-0.0018 (0.0001) ***
Problem resolved	-1.7850 (0.0407) ***	-1.7857 (0.0424) ***
Problem recognition	0.3223 (0.0154) ***	0.2941 (0.0154) ***
Time inactive	-0.0150 (0.0003) ***	-0.0139 (0.0003) ***
Attention focusing (H1)	0.2112 (0.0066) ***	0.2513 (0.0077) ***
Attention reinforcing (H2)	0.1298 (0.0117) ***	0.1815 (0.0126) ***
Attention mixing (H3)	-0.0285 (0.0027) ***	-0.1173 (0.0067) ***
Attention clustering (H4)	0.0509 (0.0067) ***	0.0416 (0.0081) ***
Module 1	-0.1247 (0.0614) *	-0.1120 (0.0631)
Module 2	-0.1498 (0.0641) *	-0.1098 (0.0653)
Module 3	0.1419 (0.0713) *	0.1515 (0.0730) *
Module 4	-0.5139 (0.0543) ***	-0.4888 (0.0554) ***
Preferential modularity	0.3780 (0.0234) ***	0.3843 (0.0239) ***
Attention clustering w/in modules	-0.0052 (0.0067)	-0.0040 (0.0070)
Attention focusing * Crowding		0.0636 (0.0063) ***
Cumulative attention * Crowding		-0.4220 (0.0244) ***
Attention reinforcing * Crowding		-0.0531 (0.0121) ***
Attention clustering * Crowding		0.0273 (0.0077) ***
Attention mixing * Crowding		0.0857 (0.0062) ***
AIC	23,256.4512	22,405.9749
Num. events	11,599	11,599
Num. obs.	1,170,870	1,170,870

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table C.3: Cox Regression Model: Exclusion of problems attracting attention of developers after resolution

	Model 1	**Model 2 (excluding resolved problems)**
Cumulative attention	0.8171 (0.0144) ***	0.7604 (0.0137) ***
Experienced participant	1.2219 (0.0558) ***	1.2506 (0.0559) ***
Institutional participant	1.1508 (0.0379) ***	1.1406 (0.0380) ***
Problem priority	-0.0447 (0.0274)	-0.0133 (0.0284)
Problem severity	0.1379 (0.0351) ***	0.1734 (0.0351) ***
Problem latency	-0.0018 (0.0000) ***	-0.0018 (0.0001) ***
Problem resolved	-1.8088 (0.0412) ***	-1.8164 (0.0416) ***
Problem recognition	0.3507 (0.0153) ***	0.3192 (0.0156) ***
Time inactive	-0.0148 (0.0003) ***	-0.0153 (0.0003) ***
Attention focusing (H1)	0.2054 (0.0065) ***	0.2090 (0.0067) ***
Attention reinforcing (H2)	0.1255 (0.0117) ***	0.1424 (0.0123) ***
Attention mixing (H3)	-0.0310 (0.0027) ***	-0.0284 (0.0029) ***
Attention clustering (H4)	0.0465 (0.0069) ***	0.0527 (0.0067) ***
Module 1	-0.1408 (0.0622) *	-0.0717 (0.0624)
Module 2	-0.1968 (0.0650) **	-0.1304 (0.0646) *
Module 3	0.1266 (0.0719)	0.1274 (0.0723)
Module 4	-0.5065 (0.0550) ***	-0.4660 (0.0548) ***
Preferential modularity	0.3562 (0.0237) ***	0.3556 (0.0236) ***
Attention clustering w/in modules	-0.0040 (0.0069)	-0.0059 (0.0066)
AIC	23,022.6114	22,901.9335
Num. events	11,599	11,369
Num. obs.	1,170,871	1,147,640

*** $p < 0.001$, **$p < 0.01$, *$p < 0.05$

Table C.4: Cox Regression Model: Observations above the 90% quantile of *Cumulative attention* removed

	Model 1	**Model 2 (excluding observation > 90% quantile of *Cumulative attention*)**
Cumulative attention	0.8171 (0.0144) ***	1.2232 (0.0204) ***
Experienced participant	1.2219 (0.0558) ***	1.0198 (0.0584) ***
Institutional participant	1.1508 (0.0379) ***	1.1627 (0.0407) ***
Problem priority	-0.0447 (0.0274)	-0.0536 (0.0291)
Problem severity	0.1379 (0.0351) ***	0.1519 (0.0378) ***
Problem latency	-0.0018 (0.0000) ***	-0.0018 (0.0001) ***
Problem resolved	-1.8088 (0.0412) ***	-2.0719 (0.0465) ***
Problem recognition	0.3507 (0.0153) ***	0.3597 (0.0163) ***
Time inactive	-0.0148 (0.0003) ***	-0.0129 (0.0003) ***
Attention focusing (H1)	0.2054 (0.0065) ***	0.2272 (0.0071) ***
Attention reinforcing (H2)	0.1255 (0.0117) ***	0.1746 (0.0121) ***
Attention mixing (H3)	-0.0310 (0.0027) ***	-0.1164 (0.0062) ***
Attention clustering (H4)	0.0465 (0.0069) ***	0.0442 (0.0082) ***
Module 1	-0.1408 (0.0622) *	-0.1033 (0.0670)
Module 2	-0.1968 (0.0650) **	-0.1749 (0.0699) *
Module 3	0.1266 (0.0719)	0.1733 (0.0771) *
Module 4	-0.5065 (0.0550) ***	-0.4707 (0.0595) ***
Preferential modularity	0.3562 (0.0237) ***	0.3598 (0.0253) ***
Attention clustering w/in modules	-0.0040 (0.0069)	-0.0087 (0.0082)
AIC	23,022.6114	20,078.9980
Num. events	11,599	10,439
Num. obs.	1,170,871	1,169,235

*** $p < 0.001$, **$p < 0.01$, *$p < 0.05$

Bibliography

Aalen, O., Borgan, O., and Gjessing, H. (2008), Survival and event history analysis: a process point of view, Springer Science & Business Media.

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (2012), Statistical models based on counting processes, Springer Science & Business Media.

Andersen, P. K. and Keiding, N. (2002), "Multi-state models for event history analysis," Statistical methods in medical research, 11, 91–115.

Borgan, Ø., Goldstein, L., and Langholz, B. (1995), "Methods for the analysis of sampled cohort data in the Cox proportional hazards model," The Annals of Statistics, 1749–1778.

Brandes, U., Lerner, J., and Snijders, T. A. (2009), "Networks Evolving Step by Step: Statistical Analysis of Dyadic Event Data," in Proc. 2009 Intl. Conf. Advances in Social Network Analysis and Mining (ASONAM), IEEE, pp. 200–205.

Butts, C. T. (2008), "A relational event framework for social action," Sociological Methodology, 38, 155–200.

Cohen, M. D., March, J. G., Olsen, J. P., et al. (1972), "A garbage can model of organizational choice," Administrative science quarterly, 17, 1–25.

Cook, R. J. and Lawless, J. (2007), The statistical analysis of recurrent events, Springer Science & Business Media.

Cox, D. (1972), "Regression Models and Life-Tables," Journal of the Royal Statistical Society. Series B (Methodological), 34, 87–22.

Everett, M. G. and Borgatti, S. P. (2013), "The dual-projection approach for two-mode networks," Social Networks,35, 204–210.

Lerner, J., Bussmann, M., Snijders, T. A., and Brandes, U. (2013), "Modeling Frequency and Type of Interaction in Event Networks," Corvinus Journal of Sociology and Social Policy, 4, 3–32.

Lerner, J. and Lomi, A. (2020), "Reliability of relational event model estimates under sampling: how to fit a relational event model to 360 million dyadic events," Network Science, 8, 97–135.

Perry, P. O. and Wolfe, P. J. (2013), "Point process modelling for directed interaction networks," Journal of the Royal Statistical Society: Series B (Statistical Methodology), 75, 821–849.

Therneau, T. M. and Grambsch, P. M. (2013), Modeling survival data: extending the Cox model, Springer Science & Business Media.

Vu, D. Q., Hunter, D., Smyth, P., and Asuncion, A. U. (2011), "Continuous-time regression models for longitudinal networks," in Advances in neural information processing systems, pp. 2492–2500.