Attention in Organizations

Wouter Dessein and Andrea Prat
Columbia University

April 1, 2015

1 Introduction

Organizations – private firms, government agencies, and non-profit organizations – can be modeled as networks of agents who are working together toward a common set of goals. Arrow (1974) views organizations as ways to overcome the limits of individual agents. By bringing together multiple workers, organizations can perform tasks that are outside the reach of any individual. While this creates production opportunities it also poses a challenge. In order to be productive, workers must coordinate their actions. Often this requires communicating information that is dispersed throughout the organization.

However, we humans face cognitive limits: transmitting and absorbing information requires time and energy. Managers spend a considerable part of their work time communicating with other workers. Bandiera et al. (2011) report that over 80% of the work time of executive managers is spend in communication activities, such as meetings, phone conversations, events, conferences, etc. Mankins et al. (2014) find that senior executives devote more than two days every week to meetings involving three or more coworkers, and 15% of an organization's collective time is spent in meetings.

As Arrow (1974) noted, given the importance of communication both as an opportunity and as a cost, organizations will strive to optimize information flows between workers. This leads to two important predictions in organizational economics. First, communication patterns within an organization will not be random but they will, at least in part, be shaped by the goals of the organization. Second, the cost of communication will be an important factor in designing the organization. Different organizational charts imply different information flows and hence different costs. A number of scholars have developed and extended Arrow's (1974) insights into formal models. This literature forms a bridge between theories of rational inattention (Sims 2003), whereby agents pay a price to transmit or receive information from other agents, and network economics, where typically links between agents are not explicitly stated in terms of information transmission and agents' payoffs are not expressed in terms of actions to be taken with incomplete information. The term *attention network* appears appropriate for this type of models.

Attention networks are most closely related to the field of organizational economics (surveyed by Gibbons and Roberts 2013). Limits to attention play a crucial role in other theories in organizational economics such as Bolton and Dewatripont (1994) and Garicano (2000). We focus here on an explicit network-theoretic approach, leaving a more general discussion to the survey of organizational economics with cognitive costs by Garicano and Prat (2013). Attention networks can be used to discuss a central theme of organizational economics: coordination. Organizations exist to coordinate specialized workers. As emphasized by Adam Smith, breaking up a production process in specialized tasks allows to dramatically increase productivity. But while the division of labor has resulted in huge productivity gains in modern economies, it also creates a need for coordination of specialized activities. The main role of organizational networks, therefore, is to achieve coordination in the presence of the division of labor. The key feature of attention networks in organizations – the one highlighted by Arrow (1974) – is further that they are endogenous. Attention networks in organizations are designed, shaped and optimized for the goal of coordination. Communication is costly, and the decision to invest in communication is made consciously by agents, either individually or as a group.

The typical attention network model contains the following elements:

- A set of agents, each of which: observes some information from the environment, may choose to transmit (at a cost) information to other agents, and may choose to receive (at a cost) information from other agents.

- A set of tasks, which must be allocated to agents, who must make decisions on the basis of information that is available to them.

- A payoff function for individuals and for the organization, which depends on how well the decisions that are taken fit with the state of the world and with each other. Some models like Dessein and Santos (2006) and Dessein, Galeotti, Santos (2014) take a team-theoretic approach and assume that agents have a common objective. Other models, like Calvó, De Martí and Prat (2015) instead follow the game-theoretic tradition of assuming different objectives for the different agents.[1]

- An attention cost function that models the payoff implications of information transmission. Attention costs can relate to active communication when they are sustained by the sender (speaking, writing, providing product samples, etc.) or passive communication when they are incurred by the receiver (listening, reading, examining product samples, etc.) or they can be formulated at the group or organizational level (for example the time agents spend in meetings).

The equilibrium of an attention network describes how communication flows and how decisions are made within the organization. The models also predict who influences whom within an organization. When an agent receives new information, this affects not just his decision but also what signals he transmits to other agents and hence what actions they choose. Thus, agent i's influence has a natural meaning in this setting as how other agents' actions are affected by a change in i's information.[2]

While this literature is recent, there are already a number of results of relevance to organizational economics:

1. A decentralized attention network or a set of standardized operating procedures are two alternative ways of achieving coordination among members of an organization. The former is more likely to be optimal when local information is more important and communication costs are lower (Dessein and Santos 2006).[3]

2. Influence and communication patterns can be highly asymmetric even starting from a perfectly symmetric interaction function. When attention is scarce, it is optimal for an

[1] However, most of the results surveyed in this chapter hold whether one uses a team-theoretic or a game-theoretic approach. See Section 4 for a discussion.

[2] This brief survey focuses on the intensity of communication, assuming that the mode of communication – namely language – is given. As Arrow (1974) noted, we should expect organizations to affect the mode of communication as well, by developing a technical language, a *code*, that is suited to the type of problems they face. Cremer, Garicano, and Prat (2007) propose a model of codes and analyze its implications for the theory of the firm.

[3] In our model, decision-making is always decentralized. See Alonso, Dessein and Matouschek (2008, 2015) for two related models who study when centralized decision-making (with vertical communication) is preferred over decentralized decision-making (with horizontal communication).

organization to direct their members' attention to a small set of key agents (Dessein, Galeotti, Santos 2014).

3. Influence and communication patterns within an organization are highly interrelated. If we observe communication patterns – for instance through electronic records – we can use eigenvector centrality to rank the influence of the members of the organization (Calvó et al. 2015).

Models of attention networks are distinct from the rest of network economics in that they must include two elements: (i) Nodes represent Bayesian decision-makers; (ii) Links represent endogenous costly communication between nodes. Element (i) is shared with a number of economic network theories, including models of learning in networks (Chapter XX by Golub and Sadler). Regarding element (ii), a number of network models contain endogenous link formation (Chapter xx by Vannetelbosch and Mauleon), which often may be interpretable as a reduced form of communication: e.g two nodes connected by the link obtain higher payoffs by exchanging information. However, what characterizes attention networks model is that information transmission is modeled explicitly with a Bayesian set-up, as a costly, endogenous activity, whose benefit can be computed within the model in terms of better decision-making.

Attention networks are closely related to models of cheap talk in networks, like Koessler and Hagenbach (2010) and Galeotti, Ghiglino, and Squintani (2013). Those models include element (i). However, communication between nodes is endogenous but not costly. The focus is therefore on whether agents find it in their interest to reveal private information to other agents. Instead, in the models discussed in this chapter, agents would disclose everything if communication were free. The assumption that attention is costly is therefore crucial to all the results that we will soon discuss. Dewan and Myatt (2008) explore costly endogenous communication in a political economy setting.

Finally, networks have been central to sociology and have found a number of important application in the sociology of organizations, as discussed in Burt (2005).

The chapter is organized as follows. Section 2 first focuses on the building block of attention networks, endogenous communication between agents, and discusses the role of attention networks in achieving coordination among specialized agents (Dessein and Santos 2006). This section also compares the merits of attention networks relative to a more bureaucratic way of coordinating economic activity: coordination through centrally imposed standard operating procedures. Whereas Section 2 imposes attention networks to be symmetric in nature, Section 3 and 4 explore optimal attention networks and asymmetric communication patterns.

The shape of the information cost function is crucial in determining the properties of the endogenous communication network. If communication costs are convex, but not excessively so, even an ex ante symmetric set of agents will choose a corner solution that results in an asymmetric communication network (Section 3, based mainly on Dessein, Galeotti, Santos, 2014) If instead communication costs are sufficiently convex, the communication network will correspond to an interior solution. It will then be possible to characterize equilibrium networks in a general setting (which includes active and passive communication) and to establish a connection with eigenvector centrality (Section 4, based mainly on Calvó, De Martí and Prat, 2015).

Section 5 concludes by providing a short discussion of the empirical literature on attention networks, a fast-growing field thanks to the increasing availability of data on behavior within organizations.

2 The Role of Attention Networks within Organizations

We discuss attention networks in the context of the Dessein and Santos (2006; hereafter DS) model of an organization, in which multiple specialized agents work together and must coordinate their individual tasks. Coordination is made difficult by the need to adapt those tasks to a changing environment. We will use the DS model to study *endogenous* attention networks in organizations. Attention networks facilitate coordinated adaptation to a changing environment. Since organizational attention is scarce, attention networks are optimized to make optimal use of this scarce resource. Since attention is naturally modelled as a communication process, we will use the terms attention and communication interchangeably.

2.1 The Dessein-Santos Model

Production in DS requires the combination of n tasks, each performed by one agent $i \in \mathcal{N} = \{1, 2, ..., n\}$. The profits of the organization depend on (i) how well each task is *adapted* to the organizational environment and (ii) how well each task is *coordinated* with the other tasks. For this purpose, agent i must take a primary action, $a_{ii} \in \mathbb{R}$, and a coordinating action, $a_{ij} \in \mathbb{R}$, for each task $j \in \mathcal{N} \setminus \{i\}$.

Pay-offs.– Ideally, agent $i \in \mathcal{N}$ should set his primary action a_{ii} as close as possible to local information θ_i, a random variable with variance σ_θ^2 and mean $\hat{\theta}_i$. One can interpret $\hat{\theta}_i$ as the status quo or the 'standard operation procedure' (SOP) for task i, where we assume that $\hat{\theta}_i$ is known to all agents.[4] In contrast, only agent i observes θ_i. We refer to θ_i as the local information pertaining to task i and assume its realization is independent across tasks. Agent $j \neq i$, in turn, should set the coordinating action a_{ji} as close as possible to action a_{ii}. The expected misadapation and miscoordination losses to the organization then amount to $\Gamma = \sum_i \Gamma_i$ where

$$\Gamma_i = E\left[\phi(a_{ii} - \theta_i)^2\right] + \beta \sum_{j \neq i} E\left[(a_{ii} - a_{ji})^2\right] \tag{1}$$

where ϕ is the weight given to misadaptation and β the weight given to miscoordination. The parameter $\beta > 0$ can be interpreted as measuring task-interdependence. We take a team-theoretic perspective so that all agents choose their primary and coordinating actions in order to minimize expected coordination and adaptation losses, as captured by Γ.

Communication and Timing.– Agents send a message to each other about their primary action. Communication is assumed to be imperfect: Agent i's message is received and understood by agent j with probability p_i, while agent j learns nothing with probability $1 - p_i$. The timing of the game is assumed as follows. In stage 1, agent $i \in \mathcal{N}$ observes θ_i, chooses a_{ii} and communicates it to all agents $j \neq i$. In stage 2, agent j receives agent $i's$ message with probability p_i and sets $a_{ji} = a_{ii}$ when he learns the value of a_{ii} and sets $a_{ji} = E(a_{ii}) = \hat{\theta}_i$ otherwise. It follows that agent i chooses a_{ii} to minimize $E[\phi(a_{ii} - \theta_i)^2 + \beta(1 - p_i)(n-1)(\hat{\theta}_i - a_{ii})^2]$:

$$a_{ii} = \hat{\theta}_i + \frac{\phi}{\phi + \beta(n-1)(1-p_i)}(\theta_i - \hat{\theta}_i) \tag{2}$$

where we verify that, indeed, $E(a_{ii}) = \hat{\theta}_i$

Communication frictions and coordination-adaptation trade-offs. Note that if agent i can perfectly communicate his primary action to agent j, there is no trade-off between adaptation and coordination: agent i then optimally sets $a_{ii} = \theta_i$ and agent $j \neq i$ ensures coordination by setting $a_{ji} = a_{ii}$. In the presence of communication frictions, however, adaptation-coordination trade-offs arise. Indeed, assume $p_i = 0$ so that agent j receives no information about a_{ii} and therefore sets $a_{ji} = \hat{\theta}_i$. Then the more agent i adapts a_{ii} to θ_i the larger are the coordination costs with tasks $j \neq i$. By ignoring his local information θ_i and sticking to the standard operating procedure or the status quo $a_{ii} = \hat{\theta}_i$, however, agent i can ensure perfect coordination

[4] We endogenize the quality and precision of standard operating procedures in section 2.2.

with tasks $j \neq i$. The ratio

$$\alpha_i \equiv \phi/[\phi + \beta(n-1)(1-p_i)]$$

can be interpreted as the optimal degree of adaptiveness or discretion for agent i given the need for coordination β and the quality of communication p_i.

Substituting the equilibrium choices for primary and coordinating actions given communication frictions p_i for tasks $i \in \mathcal{N}$ yields expected losses

$$\Gamma = \sum_{i=1}^{n} \frac{\phi\beta(n-1)(1-p_i)}{\phi + \beta(n-1)(1-p_i)} \sigma_\theta^2 \tag{3}$$

Note that equilibrium costs are increasing in the variance σ_θ^2 (unexpected contingencies, change in the environment), the importance of adaptation as measured by ϕ, the importance of coordination β, as well as the division of labor n. In contrast, equilibrium losses are decreasing in the quality of communication p_i between agents.

As pointed out by DS, extensive specialization results in organizations that are increasingly inflexible and ignore local knowledge: from (2), a_{ii} is less correlated with θ_i as n increases. As DS show, the division of labor within organizations is therefore limited by the need for adaptation. We refer to DS for a study and analysis of the optimal degree of specialization in organizations. In particular, DS allows each agent to undertake several tasks in the production process, where a broader task-allocation improves coordination but reduces the gains from specialization. In the remainder of this chapter, we will take both the number of tasks n and the task-specialization of agents as given.

2.2 Attention Networks versus Standard Operating Procedures.

We now extend the DS model to highlight two very distinctive ways of coordinating economic activity in organizations:

- Coordination through (horizontal) communication networks.
- Coordination through (centrally imposed) Standard Operating Procedures (SOPs).

On the one hand, the firm can improve coordination by fostering bilateral communication between agents. For example, the organization can carve out plenty of time for meetings and information exchange between agents. The firm can further invest in communication networks and intra-nets, and improve horizontal communication by training its employees in interpersonal communication skills, by instituting a collaborative culture, or recruit employees with similar technical backgrounds and so on. Formally, we will assume that at a cost $F_N > 0$, the organization can create a high-functioning communication network where $p_i = p_H > 0$ for all $i \in \mathcal{N}$. If the organization does not invest in a communication network, then $p_i = p_L < p_H$ for all $i \in \mathcal{N}$. Wlog, we will set $p_L = 0$. We denote $d_C = 1$ when the firm invests in a horizontal communication network with $p = p_H$, and $d_C = 0$ otherwise.

Alternatively, the firm can improve coordination by investing in commonly understood standard operating procedures (SOPs). This creates a role for a center or headquarter manager to clearly communicate task guidelines and procedures to agents, and to update those procedures as the environment changes. If agents largely stick to such commonly understood task instructions, coordination can be achieved without any need for communication. How adaptive the organization is then depends on how well operating procedures capture task-specific information and how quickly the organizational environment changes, which may result in outdated SOPs. To capture the role of standard operating procedures, we extend the DS model by adding two ingredients:

1. The organization lives for two periods. The local information in the first period, denoted by $\theta_{i,1}$, is normally distributed with mean $\hat{\theta}_i$ and variance σ_θ^2. The local information in the second period, $\theta_{i,2}$, has a mean $\hat{\theta}_i + \varepsilon$ and variance σ_θ^2, where ε is normally distributed with mean 0 and variance σ_ε^2. The variance σ_ε^2 then reflects the amount of environmental change or turbulence there is in the environment.

2. Agent j does not observe the mean $\hat{\theta}_i$ directly, but knows that the mean $\hat{\theta}_i$ itself is a random variable with mean 0 and variance $\hat{\sigma}_\theta^2 > \sigma_\varepsilon^2$. At at a cost F_S per task, a headquarter manager can learn $\hat{\theta}_i$ in the first period and (perfectly) communicate it to the organization. The role of headquarters is thus to establish or improve standard operating procedures for each task and communicate those to the organization.[5]

We denote $d_S = 1$ when the firm establishes standard operating procedures for each task so that all agents observe $\hat{\theta}_i$ for $i \in \mathcal{N}$, and $d_S = 0$ otherwise. Abusing notation, we denote expected losses over two periods given investment choices (d_C, d_S) by $\Gamma(d_C, d_S)$.

[5] To simplify the analysis, we assume that headquarters only can establish operating procedures in the first period. If σ_ε^2 is large and F_S is small, however, it may be optimal to update operating procedures in the second period.

We distinguish between four cases.

Case 1: $(d_C, d_S) = (1, 1)$: The organization invests both in establishing a high-functioning communication network and establishing standard operating procedures. This case is almost identical to the benchmark DS model, except that the variance of the local information in the second period is given by $\sigma_\theta^2 + \sigma_\varepsilon^2$ rather than σ_θ^2. Expected losses to the organization over the two periods are given by

$$\Gamma(1,1) = n\frac{\phi\beta(n-1)(1-p_H)}{\phi+\beta(n-1)(1-p_H)}\sigma_\theta^2 + n\frac{\phi\beta(n-1)(1-p_H)}{\phi+\beta(n-1)(1-p_H)}(\sigma_\theta^2+\sigma_\varepsilon^2) + F_S + F_C$$

Case 2: $(d_C, d_S) = (0, 0)$. There is no communication between agents and no standard operating procedures are established. From the perspective of agent j, the random variable $\theta_{i,1}$ has a mean 0 and variance $\hat{\sigma}_\theta^2 + \sigma_\theta^2$; the random variable $\theta_{i,1}$ has a mean 0 and variance $\hat{\sigma}_\theta^2 + \sigma_\theta^2 + \sigma_\varepsilon^2$. Expected adaptation and coordination losses are given by

$$\Gamma(0,0) = n\frac{\phi\beta(n-1)}{\phi+\beta(n-1)}(\sigma_\theta^2+\hat{\sigma}_\theta^2) + n\frac{\phi\beta(n-1)}{\phi+\beta(n-1)}(\sigma_\theta^2+\sigma_\varepsilon^2+\hat{\sigma}_\theta^2)$$

Case 3: $(d_C, d_S) = (0, 1)$. There is no communication between agents, but there are standard operating procedures. From the perspective of agent j, the random variable $\theta_{i,1}$ has a mean $\hat{\theta}_i$ and variance σ_θ^2; the random variable $\theta_{i,2}$ has a mean $\hat{\theta}_i$ and variance $\sigma_\theta^2 + \sigma_\varepsilon^2$. Expected losses to the organization are given by

$$\Gamma(0,1) = n\frac{\phi\beta(n-1)}{\phi+\beta(n-1)}\sigma_\theta^2 + n\frac{\phi\beta(n-1)}{\phi+\beta(n-1)}(\sigma_\theta^2+\sigma_\varepsilon^2) + F_S$$

Case 4: $(d_C, d_S) = (1, 0)$. The firm invests in a communication network but not in standard operating procedures. From the perspective of agent j, the random variable $\theta_{i,1}$ has a mean 0 and variance σ_θ^2. If communication was not successful in the first period, the random variable $\theta_{i,2}$ has a mean 0 and variance $\sigma_\theta^2+\sigma_\varepsilon^2$ in the second period. If communication was successful in the first period, we assume wlog that also the mean $\hat{\theta}_i$ is known to agent j, so that information about $\theta_{i,1}$ is also informative about $\theta_{i,2}$. Expected losses to the organization are given by

$$\Gamma(1,0) = n\frac{\phi\beta(n-1)(1-p_H)}{\phi+\beta(n-1)(1-p_H)}(\sigma_\theta^2+\hat{\sigma}_\theta^2) + n\frac{\phi(1-p_H)(n-1)\beta}{\phi+(1-p_H)(n-1)\beta}(\sigma_\theta^2+(1-p_H)\hat{\sigma}_\theta^2+\sigma_\varepsilon^2) + F_C$$

Fixing $d_S \in \{0,1\}$, the benefits of *investing in a communication network* $(d_C = 1)$ equal

$$\Gamma(0,d_S) - \Gamma(1,d_S) \equiv \Delta_C(d_S)$$

Fixing $d_C \in \{0,1\}$, the benefits of *investing in standard operating procedures* ($d_S = 1$) equal

$$\Gamma(0, d_C) - \Gamma(1, d_C) \equiv \Delta_S(d_C)$$

It now easy to show that

(i) Communication networks and standard operating procedures are substitutes: $\Delta_C(0) > \Delta_C(1)$ and, similarly, $\Delta_S(0) > \Delta_S(1)$. Thus, investing in communication networks is less attractive if one also invests in standard operating procedures, and vice versa. Intuitively, better operating procedures reduce the value of horizontal communication as agent j can better predict the local information of agent i which, in turn, allows agent i to be more adaptive to his local information even in the absence of communication.

(ii) $\Delta_C(d_S)$ is increasing in σ_θ^2, σ_ε^2, $-F_C$, and p_H whereas $\Delta_S(d_C)$ is not affected by changes in σ_θ^2, σ_ε^2 and F_C and is decreasing in p_H. Hence, an increase in σ_θ^2, σ_ε^2, $-F_C$ or p_H makes investing in communication networks more attractive, whereas it does not affect or decreases the benefits of investing in standard operating procedures. Intuitively, σ_θ^2 reflects local information held by agent i which is not captured in high-quality operating procedures, whereas σ_ε^2 reflects how quickly operating procedures become obsolete. Both therefore make communication networks more valuable.[6] Standard operating procedures are further less valuable as the communication quality p_H improves, as SOPs are only useful in case communication fails.

(iii) $\Delta_S(d_C)$ is increasing in $-F_S$ whereas $\Delta_C(d_S)$ is not affected by F_S.

The following proposition follows directly from the above observations:[7]

[6] Note, however, that an increase in $\hat{\sigma}_\theta^2$, that is the variance of the optimal standard operating procedures both makes standard operating procedures and communication networks more attractive. Since attention networks and SOPs are substitutes, the comparative statics with respect to $\hat{\sigma}_\theta^2$ are ambiguous.

[7] Our results are similar to those obtained in Aoki (1986) in a different team-theoretic set-up. Building on Cremer (1980), Aoki compares the efficiency of vertical and horizontal information structures in coordinating operational decisions among interrelated units (shops) whose cost conditions are uncertain. Aoki then uses his model to compare stylized differences in the internal organization of large Japanese and US manufacturing firms. A horizontal information structure, similar to 'coordinating through communication networks' in our model, is said to be more represenative of how Japanese firms coordinate production in the 1970's and early 1980s. In contast, it is observered how US manufacturing firm tend to rely more on the use of a vertical information structure or what we refer to as 'standard operating procedures'.

Proposition 1 *$\Gamma(d_C, d_S)$ is supermodular in $d_C, -d_S, \sigma^2_\theta$, $\sigma^2_\varepsilon, p_H, -F_C$ and F_S. Hence:*

1. *Coordination through communication networks (standard operating procedures) is more (less) likely when*

   (i) local information is more important, that is σ^2_θ is larger, and/or there is more environmental change, that is σ^2_ε is larger.

   (ii) communication quality p_H is higher or the cost of communication networks, F_C, is lower.

   (iii) the cost of implementing high-quality standard operating procedures, F_S, is higher.

2. *Communication networks ($d_C = 1$) and standard operating procedures ($d_S = 1$) are substitutes:*

   (i) A decrease in the cost of communication networks F_C can result in a change from $d^*_S = 1$ to $d^*_S = 0$, but never the other way around.

   (ii) A decrease in the cost of establishing standard operating procedures F_S can result in a change from $d^*_C = 1$ to $d^*_C = 0$, but never the other way around.

3 Organizational Focus: Convexities in Attention Networks.

In the previous section, it was assumed that communication networks are symmetric – all agents observe or learn each other's actions with the same probability p. Drawing upon Dessein, Galeotti and Santos (2014, DGS hereafter), we now relax the assumption that communication networks are symmetric, that is we allow for $p_i \neq p_j$ and let an organization designer optimize over $[p_i, ..., p_n]$. Our starting point is that organizational attention is scarce, and communication networks optimally distribute this attention among the agents of the organization. We show that even in a symmetric environment where all agents are ex ante identical, optimal attention networks and information flows are often asymmetric ex post because of the complementarity between attention and decision-making. In particular, when organizational attention is scarce, a hybrid approach to coordinating economic activity is optimal where attention networks coordinate the tasks of a select number of agents and the remaining tasks are coordinated using standard operating procedures. Scarce attention thus creates convexities in the optimal allocation of attention, where all attention is (optimally) monopolized by a few agents. We first discuss this result in our baseline model and then discuss the robustness of our results to alternative communication technologies.

3.1 Baseline Model

Our starting point is that p_i - the quality of the communication about agent $i's$ action - is increasing in the organizational attention t_i devoted to agent i. Organizational attention is scarce, however, in that there is a fixed attention budget: $\sum_i t_i \leq T$. We can think of t_i as the "air-time" or "attention" agent i receives and T could be the length of time agents spend in meetings as opposed to production. For our analysis, we revert to the original DS model where there is only one period and where the mean $\hat{\theta}_i$ of the random variable θ_i is common knowledge. Relative to DS, however, we add an additional Stage 0 where the organizational designer optimally chooses the qualities $[p_1, ..., p_n]$ of the communication links. We assume that p_i follows a Poisson process, that is $p_i = 1 - e^{-\lambda t_i}$ where λ is the constant hazard rate that any agent $j \in \mathcal{N} \setminus \{i\}$ correctly learns the primary action taken by agent i. We can interpret $1/\lambda$ as a measure of the complexity of tasks. Note that the communication cost or attention t_i required to achieve a given communication quality p_i is increasing and convex in p_i. This reflects decreasing marginal returns to attention. Denoting by $P \equiv 1 - e^{-\lambda T}$ the maximum communication quality that can be achieved by focussing all organizational attention on one agent, the organizational attention constraint $\sum_i t_i \leq T$ can be rewritten as

$$\sum_{i \in \mathcal{N}} \log(1 - p_i) \geq \log(1 - P) \tag{4}$$

At stage 0, an organization designer then optimally chooses $[p_1, ..., p_n]$ subject to constraint (4), which will be binding at the optimum.

Two tasks, two agents. Assume now first that $n = 2$, so that the organization consists of two agents and (4) is equivalent to $(1-p_1)(1-p_2) \geq 1-P$. In order to increase the communication quality p_1, the organization then needs to reduce the communication quality p_2. The larger is the maximal communication probability P, the less there is a trade-off between good communication on task 1 versus task 2. Note that P will be large when attention is not scarce and/or tasks are not very complex.

Given that all agents are ex ante symmetric, all tasks are equally important, and interdependencies are symmetric, one may conjecture that the optimal communication network will be symmetric as well. Moreover, it is easy to verify that given constraint (4), $p_1 + p_2$ is uniquely maximized when $p_1 = p_2$. DSG show, however, that when organizational attention is scarce, that is whenever $P < \bar{P}$, it is optimal to focus all attention on one of the two agents that is $(p_1^*, p_2^*) \in \{(P, 0), (0, P\}$. Intuitively, having a high-quality communication link to agent i and letting agent i be very adaptive to his local information are complementary choices. The more agent i is adaptive to his local information, the less valuable are standard operating procedures in coordinating tasks, and the more important is communication to achieve coordination. More attention should therefore be focused on agent i. By the same token, it agent i is not responsive to his local information, then it is a waste of time to devote attention to agent i, as coordination is achieved appropriately by the common knowledge and adherence to standard operating procedures.

Assume that equilibrium primary actions are linear in θ_i and $\hat{\theta}_i$, that is

$$a_{ii} = \hat{\theta}_i + \alpha_i(\theta_i - \hat{\theta}_i),$$

where α_i can be interpreted as the adaptiveness (or discretion) of agent i. Substituting a_{ii} in (1) and taking expectations, expected losses equal

$$\Gamma = \phi(1-\alpha_1)^2\sigma_\theta^2 + \phi(1-\alpha_2)^2\sigma_\theta^2 + \beta(1-p_1)\alpha_1^2\sigma_\theta^2 + \beta(1-p_2)\alpha_2^2\sigma_\theta^2, \tag{5}$$

Inspecting (5), it is immediate that α_1 and p_1 are complementary choices. The more adaptive is agent 1, the larger are the benefits of improving communication about agent 1 in order to minimize Γ. If we had chosen an attention constraint with a constant rate of substitution between p_1 and p_2, for example $p_1 + p_2 \leq P$, then it would always be optimal to focus all attention on one agent so that either $p_1 = 0$ or $p_2 = 0$. More naturally, however, there are decreasing marginal returns to attention, as captured by the constraint (4). Indeed, from (4), the higher is p_1, the more one needs to reduce p_2 for any additional increase in p_1. Such decreasing marginal returns create a countervailing force against focussing all attention on the same task. The following proposition, taken directly from DSG, shows that an asymmetric attention networks is optimal if and only if attention is scarce:

Proposition 2 *Suppose $\beta > 1$. There exists a $\bar{P}(\beta)$ such that:*
(i) An asymmetric (focused) attention network is optimal, $(p_1^*, p_2^*) \in \{(P,0),(0,P)\}$ if and only if $P \leq \bar{P}(\beta)$
(ii) A symmetric (balanced) attention network is optimal, $(p_1^*, p_2^*) = (\tilde{p},\tilde{p})$ if and only if $P > \bar{P}(\beta)$, where $2\log(1-\tilde{p}) = \log(1-P)$
(iii) $\bar{P}(\beta)$ is increasing in the importance of coordination, β.

To summarize the above proposition, if organizational attention is scarce (T is small) or the environment are very complex (λ is small), then $P = 1 - e^{-\lambda T}$ is small as well, and it optimal to focus all attention on one agent, say agent 1. Agent 1 is then allowed to be very adaptive to his task, and coordination with agent 1 will be achieved through the attention network. In contrast, agent 2 will be forced to largely ignore his local information and coordination with this agent's tasks will be achieved through adherence to the commonly known standard operating procedure, $\hat{\theta}_2$.

If, on the other hand, organizational attention is abundant (T is large) or the environment is not very complex (λ is large), then P will be large, and it will be optimal to have a symmetric attention network where both agents divide attention equally. Intuitively, it is then feasible for each agent to communicate his primary action almost perfectly. Both agents can then be responsive to their local information and coordination will be achieved through the attention network for both agents. Standard operating procedures play a limited role in coordinating activity.

Large organizations. DSG extend their set-up to incorporate $n > 2$ tasks, in which case they show that the optimal communication network consists of ℓ leaders and $n - \ell$ followers. All attention is equally split among the ℓ leaders, whereas no attention is devoted to the $n - \ell$ followers. The ℓ leaders are very responsive to their local information and coordination with their task is achieved through the attention network. In contrast, coordination with the tasks of the $n - \ell$ followers is achieved by letting those agents stick closely the commonly known standard operating procedures. In other words, when attention is scarce, a hybrid approach to coordinating economic activity is optimal. Attention networks coordinate the tasks of a select number of agents and the remaining tasks are coordinated using standard operating procedures. The better the communication technology (the larger is λ), the larger is ℓ and the more the organization relies on attention networks rather than standard operating procedures. In contrast, the more interdependent are tasks, and the more important is the avoidance of coordination losses, the smaller is ℓ and the less the organization relies on networks.[8]

[8] In DGS, the number of leaders and who is a leader is determined in equilibrium. A number of other papers, such as Bolton, Brunnermeier and Veldkamp (2012), Dessein and Santos (2014) and Van den Steen (2014), also build on DS in order to study how a leader can achieve coordination among members of an organization. Communication networks are not endogenized, however, as communication is always between the exogenously appointed leader and the remainder of the organization.

3.2 Alternative communication technologies.

We now discuss the robustness of our results to alternative communication technologies. Let us denote by m_i be the information received by agent j regarding θ_i and define the *residual variance* about θ_i as $\mathrm{Var}(\theta_i|m_i) \equiv E(\theta_i - E(\theta_i|m_i))^2$. In our *baseline model*, we have assumed that agent $j \neq i$ observes θ_i with a probability $p_i = 1 - e^{-\lambda t_i}$ where t_i is the attention devoted to task i. Given this communication technology

$$E\left[\mathrm{Var}(\theta_i|m_i)\right] = \sigma_\theta^2(1-p_i) = \sigma_\theta^2 e^{-\lambda t_i} \tag{6}$$

Substituting (6) into (3), expected organizational losses can be written as :

$$\Gamma(t) = \sum_{i=1}^{i=n} \frac{\phi\beta(n-1)E(\mathrm{Var}(\theta_i|m_i))}{\phi\sigma_\theta^2 + \beta(n-1)E(\mathrm{Var}(\theta_i|m_i))}\sigma_\theta^2 \tag{7}$$

One can verify that $-\Gamma(t)$ is convex in attention t_i when t_i is small. As discussed above and shown by DGS, whenever organizational attention is scarce ($T < \bar{T}$), it follows that $-\Gamma(t)$ is maximized by setting $t_i = T/\ell$ for ℓ agents $i \in \mathcal{L} \subset \mathcal{N}$ and set $t_j = 0$ for $n - \ell$ agents $j \in \mathcal{N} \backslash \mathcal{L}$. Put differently, there are convexities in the optimal allocation of attention: tasks should either receive a lot of attention or no attention at all.

Instead of the above binary communication technology, assume now that θ_i is independently normally distributed for $i \in \mathcal{N}$ and that agents $j \neq i$ observe a noisy message $m_i = \theta_i + \epsilon_i$ about θ_i, where ϵ_i is (independently) normally distributed. Given linear decision rules $a_{ii} = \hat{\theta}_i + \alpha_i(\theta_i - \hat{\theta}_j)$ and $a_{ji} = E(\theta_i|m_i)$, one can show (see DGS) that expected organizational losses given an attention network t are again given by (7). We now consider *two alternative communication technologies* who only differ in how fast $E(\mathrm{Var}(\theta_i|m_i))$ decreases as a function of the attention t_i devoted to θ_i.

Rational Inattention and Entropy Information Costs Information Theory and the literature on Rational Inattention (Sims 2003) posit that the communication costs or "communication capacity" $C(m)$ required to send a message $m = (m_1, ..., m_n)$ about $\theta = (\theta_1, .., \theta_n)$ is equal to the reduction in entropy of θ following the observation of m. Following this literature, we posit that communication is optimized under the constraint

$$C(m) = H(\theta) - H(\theta|m) \leq T, \tag{8}$$

where $H(\theta)$ is the (differential) entropy of θ and $H(\theta|m)$ the entropy of θ conditional upon observing m. In other words, the attention capacity T of the organization puts a constraint on the total reduction in entropy following communication.[9] Given that m_i and the conditional distributions $F(\theta_i|m_i)$ are independently normally distributed for $i \in \mathcal{N}$, attention constraint (8) can be rewritten as

$$\sum_{i \in \mathcal{N}} \left(2 \ln \sigma_\theta^2 - 2 \ln Var(\theta_i|m_i)\right) \leq T$$

or still: $\sum_{i \in \mathcal{N}} t_i \leq T$, with

$$\text{Var}(\theta_i|m_i) = \sigma_\theta^2 e^{-2t_i} \tag{9}$$

Since (9) and (6) are equivalent up to a rescaling of the attention capacity, we obtain identical results as in our baseline model. Hence, whenever organizational attention T is scarce, normally distributed information and entropy information costs imply that the optimal attention network is asymmetric where a few agents monopolize all attention.

Sampling from a Normal Distribution An alternative way of modeling noisy communication is to assume that the number of i.i.d. signals agent j receives about θ_i is linear in the attention t_i devoted to task i. Let m_i be the average realization of t_i signals $s_{ik} = \theta_i + \varepsilon_{ik}$, with $k \in \{1, .., t_i\}$, where ϵ_{ik} is i.i.d. normally distributed with variance $\sigma_\varepsilon^2 : m_i = \frac{1}{t_i} \sum s_{ik}$. Then $m_i = \theta_i + \epsilon_i$ where $\sigma_\epsilon^2 = \sigma_\varepsilon^2 / t_i$ so that

$$\text{Var}(\theta_i|m_i) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + t_i \sigma_\theta^2} \sigma_\theta^2 \tag{10}$$

Whereas for t_i small, we have that $-\Gamma(t)$ is convex in t_i for communication technologies (9) and (6), it is now easy to verify that given communication technology (10), $-\Gamma(t)$ is always concave in t_i. Hence, given (10), the optimal allocation of attention is symmetric, that is $t_i^* = T/n$ for all $i \in \mathcal{N}$. Intuitively, regardless of the communication technology, the complementarity between attention and decision-making results in a convexity in the value of information. Indeed, from (7), $-\Gamma(t)$ is convex in $-E(\text{Var}(\theta_i|m_i))$. Convexities in the cost of communication, however, provide a countervailing force to focus all attention on a few tasks. Indeed, technologies (9), (6) and (10) all exhibit convex communication costs to reduce $\text{Var}(\theta_i|m_i)$, that is $\partial^2 \text{Var}(\theta_i|m_i)/\partial t_i^2 > 0$. It is only for communication technology

[9] Formally, this is equivalent with assuming that T is the (Shannon) capacity of the Guassian communication channel. The capacity of a channel is a measure of the maximum data rate that can be reliably transmitted over the channel. Shannon capacity has proven to be an appropriate concept for studying information flows in a variety of disciplines: probability theory, communication theory, computer science, mathematics, statistics, as well as in both portfolio theory and macroeconomics.

(10), however, that this convexity in the cost of communication dominates the convexity in the value of information for any value of t_i. In contrast, for technologies (9) and (6), the convexity in the cost of communication only dominates when t_i is sufficiently large.

In the next section, we study optimal attention networks in more complex environments: (i) tasks and interdependencies between tasks are asymmetric, (ii) communication has both an active and a passive component, and (iii) agents do not necessarily maximize a common objective function. To simplify the analysis, we will assume that communication costs are sufficiently convex, as in technology (10), allowing us to focus on interior solutions.

4 Communication and Influence in Attention Networks: Interior Solutions.

DSG showed that an ex-ante symmetric environment can give rise to a highly skewed communication patterns. This section considers asymmetric environments and studies how the ex ante asymmetry determines an ex post asymmetry in terms of communication and influence. The analysis will rely on Calvó de Martí, and Prat (2015), henceforth CDP. The main outcome will be a set of predictions on communication and influence flows, which can be seen as a formalization of Arrow's (1974) theory of organizational communication discussed in the Introduction.

Consider a set of N agents. Each agent i faces a local state of the world:

$$\theta_i \sim \mathcal{N}\left(0, 1/s_i\right),$$

where s_i denotes the precision of θ_i, i.e. $s_i = 1/Var\left(\theta_i\right)$. If the local states were correlated, agents' actions may be correlated in equilibrium even if agents do not communicate. We prefer to abstract from this form of direct correlation in order to focus on the role of communication. Therefore, we assume that θ_i is independent across agents.

Each agent i observes only θ_i and can then engage in communication with all other agents. Information transmission requires effort from both the sender and the receiver. The signal is more precise if the sender invests more in *active communication* (e.g. speaking or writing) and the receiver invests more in *passive communication* (e.g. listening or reading). Namely, agent i receives message y_{ij} from agent j, such that

$$y_{ij} = \theta_j + \varepsilon_{ij} + \eta_{ij},$$

where ε_{ij} and η_{ij} are two normally distributed noise terms

$$\varepsilon_{ij} \sim \mathcal{N}(0, 1/r_{ij}), \tag{11}$$

$$\eta_{ij} \sim \mathcal{N}(0, 1/p_{ij}), \tag{12}$$

and r_{ij} (resp. p_{ij}) is the precision of ε_{ij} (resp. η_{ij}). We assume that all stochastic terms are mutually independent (and independent of the θ's).

In the first stage of the game every agent chooses how much to invest in communication. Namely, agent i selects the values of two vectors: (i) The precision of the active communication part of all the signals he sends: $(r_{ji})_{j\neq i}$, for which he incurs cost $k_r^2 \sum_{j\neq i} r_{ji}$, where $k_r \geq 0$ is a parameter; (ii) The precision of the passive communication part of all the signals he receives, $(p_{ij})_{j\neq i}$, for which he incurs cost $k_p^2 \sum_{j\neq i} p_{ij}$, where $k_p \geq 0$ is a parameter (p is mnemonic for passive).

In the second stage of the game, every agent observes the signals he has received from the other agents and chooses the value of action $a_i \in (-\infty, \infty)$.

CDP can be formulated in three ways: as a non-cooperative game where each agent maximizes his own expected payoff, as a team-theoretical problem where each agent maximizes the sum of expected payoffs of all agents, or as a hybrid problem where communication investments are made cooperatively while actions are chosen selfishly. An earlier version of CDP (Calvó de Martí, and Prat 2009) considered all three versions and show they produce qualitatively similar solutions. The present chapter focuses exclusively on the first version.

The payoff of agent i is a classic quadratic objective function

$$u_i = -\left(d_{ii}(a_i - \theta_i)^2 + \sum_{j\neq i} d_{ij}(a_i - a_j)^2 + k_r^2 \sum_{j\neq i} r_{ji} + k_p^2 \sum_{j\neq i} p_{ij} \right), \tag{13}$$

where the term d_{ii} measures the adaptation motive, i.e. the importance of tailoring i's action to the local state, and the term d_{ij} represents the coordination motive, namely the interaction between the action taken by agent i and the action taken by agent j. For the rest of the paper we assume that the interaction terms are positive ($d_{ij} \geq 0$ for all i and all j).

The game has two versions according to whether investment in communication occurs before or after the agent observes his local state θ_i. The "before" version captures the idea that investment has a long-term component (e.g. two firms appoint liaison officers). The "after" version represents a shorter-term investment, like the direct cost of writing and reading a report. As Calvó. et al (2015) discuss, both versions have the same linear pure-strategy equilibrium. For concreteness, in this chapter, we focus on the "before" version. We refer to this game as $\Gamma(\boldsymbol{D}, \boldsymbol{k}, \boldsymbol{s})$, where $\boldsymbol{D} = (d_{ij})_{i,j}$, $\boldsymbol{k} = (k_r, k_p)$ and $\boldsymbol{s} = (s_i)_i$.

This can be seen as a game of communication and influence. In equilibrium, agents communicate with each other and they influence each other's decisions through the signals they communicate. The analysis of this game is divided in two parts. First, we provide a closed-form characterization of equilibrium. Second, we show that influence in equilibrium is approximated with an appropriately defined notion of eigenvector centrality, which can be computed directly on the interaction matrix $\boldsymbol{D}$.

Let us begin by characterizing equilibrium play. To do this, consider first a game with just two players. First, normalize the interaction matrix of agent i by dividing it by the sum of all interaction terms

$$\omega_{ij} = \frac{d_{ij}}{d_{i1} + d_{i2}}$$

The payoff – net of communication costs – of, say, agent 1 can now be written as

$$-\omega_{11}\overbrace{(a_1 - \theta_1)^2}^{\text{adaptation}} - \omega_{12}\overbrace{(a_1 - a_2)^2}^{\text{coordination}}$$

Focus on the second stage of the game. Given investments in communication, how do agents select their actions as functions of signals they receive. One can check this stage has a linear equilibrium of the following form:

$$\begin{aligned} a_1^* &= b_{11}\theta_1 + b_{12}y_2 \\ a_2^* &= b_{21}y_1 + b_{22}\theta_2 \end{aligned}$$

where the b-coefficients solve

$$\begin{array}{ll} b_{11} = \omega_{11} + \omega_{12}b_{21} & b_{22} = \omega_{22} + \omega_{21}b_{12} \\ b_{12} = \omega_{12}b_{22}\frac{r_{12}p_{12}}{s_2r_{12}+s_2p_{12}+r_{12}p_{12}} & b_{21} = \omega_{21}b_{11}\frac{r_{21}p_{21}}{s_1r_{21}+s_1p_{21}+r_{21}p_{21}} \end{array}$$

Note that b_{12} and b_{21} represent the influence of agents on each other. The influence of i's signal on j decision depends on how informative i's signal as well as how much j cares about coordinating with i.

Once we know what happens in the second stage, we can use backward induction to solve for equilibrium communication investments in the first stage. We find that active communication and passive communication are, respectively:

$$r_{21} = b_{21}\frac{\sqrt{d_{12}}}{k_r} \tag{14}$$

$$p_{12} = b_{12}\frac{\sqrt{d_{11}+d_{12}}}{k_p} \tag{15}$$

Investment in active and passive communication is not equal. The numerator of the right-hand-side of equation (14) contains only d_{12} while its counterpart in (15) contains both d_{11} and d_{12}. Passive communication offers a more direct return: the listener can make use of the signal he receives. Active communication is instead more indirect: the speaker makes an investment in order for the listener to use the resulting signal. Therefore, in this type of models passive communication has an intrinsic advantage. Of course this advantage can be undone by other considerations, like a lower cost of investing in active communication or the presence of economies of scope. However, everything else equal, agents will invest relatively more in listening than in speaking.

Once the two agent case is solved, one can more easily understand the general n-agent case. The logic is similar but we must add one important element: the possibility of indirect effects among agents. For example, if there are three agents, i may want to learn about j's state because he cares about m's action and he knows that m cares about j's state.

Two additional pieces of notation are needed. Let $\mathbf{\Omega}$ be the matrix of normalized interactions with typical element ω_{ij}. Let the matrix of normalized benefits be given by

$$h_{ij} = \begin{cases} \omega_{jj} & \text{if } i = j \\ -s_j\left(\frac{k_p}{\sqrt{D_i}} + \frac{k_r}{\sqrt{d_{ji}}}\right), & \text{otherwise.} \end{cases}$$

Provided that the cost of communication parameters k_r and k_p are sufficiently low (to avoid corner solutions), we have:

Theorem 3 *The game $\Gamma(\boldsymbol{D}, \boldsymbol{k}, \boldsymbol{s})$ has a linear equilibrium where:*

(i) Decisions are given by

$$\boldsymbol{b}_{\cdot j} = (\boldsymbol{I} - \mathbf{\Omega})^{-1} \cdot \boldsymbol{h}_{\cdot j} \quad \textit{for all } j;$$

(ii) Active communication is

$$r_{ij} = \frac{\sqrt{d_{ji}}b_{ij}}{k_r} \qquad \textit{for all } i \neq j;$$

(iii) Passive communication is

$$p_{ij} = \frac{\sqrt{D_i} b_{ij}}{k_p} \qquad \textit{for all } i \neq j$$

Theorem 3 is the generalization of the two-agent case. The inverse matrix $(\boldsymbol{I} - \boldsymbol{\Omega})^{-1}$, which captures the direct and indirect interactions of agents' actions on one another. It can be understood as an infinite series of higher-order normalized effects:

$$(\boldsymbol{I} - \boldsymbol{\Omega})^{-1} = \boldsymbol{I} + \boldsymbol{\Omega} + \boldsymbol{\Omega}^2 + \boldsymbol{\Omega}^3 + \cdots = \sum_{l \geq 0} \boldsymbol{\Omega}^l.$$

Theorem 3 can be seen as one way of formalizing Arrow's (1974) idea that communication and decisions pattern are shaped by the objectives of the members of the organization. Given underlying parameters that describe complementarities, information cost, and uncertainty, we can predict how much each agent will communicate in equilibrium and who much he will be influenced by other agents.

The second part of the analysis focuses on influence. Theorem 3 characterizes influence as a game-theoretic phenomenon. It turns out that this strategic approach is approximately equal to a much simpler network centrality concept.

To see, we need to additional definitions. First, define a sequence of games as follows. Fix D, s, k_r and k_p, and define the payoff function:

$$u_i = -\left(d_{ii} \left(a_i - \theta_i\right)^2 + \frac{1}{t} \sum_{j \neq i} d_{ij} \left(a_i - a_j\right)^2 + t^{\lambda} k_r^2 \sum_{j \neq i} r_{ji} + t^{\lambda} k_p^2 \sum_{j \neq i} p_{ij} \right),$$

where $t \in (0, \infty)$ and $\lambda > 1$. For every value of t we define a different game, which we can call $\Gamma\left(D, s, k_r, k_p, t\right)$. As t goes to zero, coordination becomes relatively more important than adaptation (and communication costs go down in order to guarantee that the solutions does not run into non-negativity constraints on communication intensities).

For every value of the parameter t, we have a natural definition of an agent's influence as the effect on all agent's actions (including his own) of an increase in his own state. Namely, the global influence of agent i, that we denote by $\mathcal{I}_i$, is

$$\mathcal{I}_i\left(t\right) = \sum_{j=1}^{n} b_{ji} \qquad k = 1, \ldots, n$$

Second, let us introduce an axiomatic network centrality concept, which seeks to assign an "importance" index to every node of a network purely on the basis of the strength of links between agents. The concept, referred to as eigenvector centrality, has been known since the 50's and has found a larger number of applications in a number of fields. Palacios-Huerta and Volij (2004) provided an axiomatization of the index and Golub and Jackson (2010) used it in network economics.

Let $\tilde{G}$ be the matrix with entries $\gamma_{ii} = 0$ for all i, and $\gamma_{ij} = \frac{d_{ij}}{\sum_{k \neq i} d_{ik}}$. The eigenvector index of agent i is ι_i, defined as the i-th component of the vector that solves:

$$\iota = \tilde{G}' \iota$$

and that satisfies $\sum_j \iota_j = 1$.

We can show that the game-theoretic influence index tends to the axiomatic influence index when t goes to zero:

Theorem 4 *As $t \to 0$, the relative global influence of agents converges to the ratio of eigenvector centrality indices weighted by an adaptation vs coordination ratio. Namely, for any i and j,*

$$\lim_{t \to 0} \frac{\mathcal{I}_i(t)}{\mathcal{I}_j(t)} = \frac{\iota_i}{\iota_j} \frac{\frac{d_{ii}}{D_{-i}}}{\frac{d_{jj}}{D_{-j}}}$$

In particular, if $d_{ii} = d_{jj}$ and $D_{-i} = D_{-j}$ for all $i, j \in N$, then we obtain that

$$\lim_{t \to 0} \frac{\mathcal{I}_i(t)}{\mathcal{I}_j(t)} = \frac{\iota_i}{\iota_j}$$

This result implies that, when t is sufficiently small, namely when coordination is more important than adaptation, eigenvector centrality is a good approximation of game-theoretic influence. This result is useful in practice because eigenvector centrality is easier to compute. It also creates a conceptual link between equilibrium influence in organizations and influence as defined in other contexts where eigenvector centrality is often used, such as search engines and bibliometrics.

5 Conclusions

The three previous sections discussed, respectively, three key findings of the theory of attention networks. First, a decentralized attention network or a set of standardized operating procedures are two alternative ways of achieving coordination among members of an organization. The former is more likely to be optimal when local information is more important and communication costs are lower. Second, influence and communication patterns can be highly asymmetric even starting from a perfectly symmetric interaction function. When attention is scarce, it is optimal for an organization to direct their members' attention to a small set of key agents. Third, influence and communication patterns within an organization are highly interrelated. If we observe communication patterns – for instance through electronic records – we can use eigenvector centrality to rank the influence of the members of the organization.

The rest of this section discusses two promising applications of the endogenous communication network framework. We first discuss empirical analyses and we then move to models that combine endogenous communication and behavioral biases.

The theories discussed in this chapter yield testable predictions on communication patterns within organizations. Both Dessein et al. (2014) and Calvó et al. (2015) characterize information flows as a function of the underlying primitives. These predictions can be tested, or used as a basis for estimation, provided one has data on communication patterns.

Until the Nineties, communication within organizations could only be measured through ethnographic studies. Mintzberg (1973) used personal observation to study how five top executives allocated their time to various activities. However, the IT revolution has created a wealth of data on communication patterns within organizations, such as email records and calendar information.

Palacios-Huerta and Prat (2011) analyze two datasets containing email communication traffic between all the executives of the same company (a European retailer). As suggested by Theorem 4 of Calvó et al. (2015), they compute the predicted influence of agents on the basis of their eigenvector centrality in the email traffic network. The email-based influence index of an agent turns out to be strongly correlated to the agent's influence as proxied by standard organizational variables, such as income and rank. Moreover, the discrepancies between the current email-based index of an agent and his or her current income and rank predict future promotions and dismissals.

Bandiera et al. (2011, 2014) analyze communication patterns within organizations from a different angle. They collect information on how chief executives officers of hundreds of companies around the world utilize their work time. In particular, they observe who the CEO spends time with. This includes internal constituencies such as the finance division or the marketing division, or external constituencies, such as customers or investors. The theories reviewed here predict that the allocation of this very scarce resource, CEO attention, should reflect the priorities of the company and the CEO.[10]

Let us finally turn our attention to behavioral theories that incorporate elements of endogenous communication within organizations. The models reviewed in this chapter build on Bayesian agents. However, the idea of organizations with endogenous information transmission can be extended to settings where agents have cognitive biases. In fact, an important question is how organizations will structure themselves in order to minimize the potentially detrimental effects of biased information processing. Bénabou (2013) considers a network of agents with anticipatory bias, which affects how they process and recall the information they observe. In such a setting, information avoidance may be beneficial or detrimental to welfare. If it makes bad news even worse for other agents (as in the case of risk spillovers), then it is detrimental. If it dampens the effect of bad news (as in the case of group morale), then it is beneficial.

Sethi and Yildiz (2013) consider endogenous communication networks where individual agents have subjective prior beliefs. In each period an agent receives a private signal and chooses to observe the opinion of another agent in the network. Observing an opinion provides information both about the state of the world and the prior of the agent whose opinion is observed. Long-run behavior may be history-dependent and inefficient, with some agents emerging as opinion leaders.

Makarov (2011), finally, studies an organization where employees display present-bias preferences and communication (e.g. email) can be high-priority or low-priority. In equilibrium, the organization suffers from social procrastination as agents spend excessive time on low-priority communication. In this setting, the organization may benefit from policies that restrict communication.

References

[1] Alonso, Ricardo, Wouter Dessein and Niko Matouschek (2008). "When does Coordination Require Centralization?" *American Economic Review*, 98 (1), 145-179.

[10] Halac and Prat (2014) study a principal-agent model, where the principal may invest in a non-dicrectly observable attention technology that provides recognition and reward to a hard-working agent. The model is analyzed in a dynamic setting.

[2] Alonso, Ricardo, Wouter Dessein and Niko Matouschek (2015). “Organizing to Adapt and Compete.” *American Economic Journal: Microeconomics*. Forthcoming.

[3] Aoki, Masahiko (1986). Horizontal vs. Vertical Information Structure of the Firm. *American Economic Review*, 76: 971-983.

[4] Arrow, Kenneth J. (1974). *The Limits of Organization*. New York, NY: Norton.

[5] Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun (2011). What Do CEOs Do? CEPR Discussion Paper 8235, February 2011.

[6] Bandiera, Oriana, Andrea Prat, and Raffaella Sadun (2014). Managerial Capital at the Top: Evidence from the Time Use of CEOs, Working paper. Columbia University. December 2014.

[7] Bénabou, Roland . Groupthink: Collective Delusions in Organizations and Markets, *Review of Economic Studies* 80, 429-462, 2013.

[8] Bolton, Patrick, Markus Brunnermeier and Laura Veldkamp (2013). “Leadership, Coordination and Corporate Culture.” *Review of Economic Studies*, 80(2), 512-537.

[9] Bolton, Patrick and Matthias Dewatripont (1994). The Firm as a communication network. *Quarterly Journal of Economics*. Vol. 109,4, pp. 809-839.

[10] Burt, Ronald S. (2005). *Brokerage & Closure*. New York: Oxford University Press.

[11] Calvó-Armengol, Antoni, Joan de Martí, and Andrea Prat. Endogenous Communication in Complex Organizations. Working paper, London School of Economics, 2009.

[12] Calvó-Armengol, Antoni, Joan de Martí, and Andrea Prat. Communication and Influence. *Theoretical Economics*, forthcoming, 2015

[13] Cremer, Jacques (1980). "A Partial Theory of the Optimal Organization of a Bureaucracy", *The Bell Journal of Economics*, 11(2), 683-693

[14] Cremer, Jacques, Luis Garicano, Andrea Prat. Language and the Theory of the Firm, *Quarterly Journal of Economics* 122(1): 373-407, February 2007.

[15] Dessein,Wouter, Andrea Galeotti, and Tano Santos (2014), “Rational inattention and organizational focus.” Columbia Business School, Mimeo.

[16] Dessein, Wouter and Tano Santos (2006). "Adaptive Organizations." *Journal of Political Economy*, 114 (50), 956-995.

[17] Dessein, Wouter and Tano Santos (2014). "Managerial Style and Attention." Columbia Business School, mimeo.

[18] Dewan, Torun and David Myatt. "The Qualities of Leadership: Direction, Communication, and Obfuscation." *American Political Science Review*, 102(3): 351–368, August 2008.

[19] Galeotti, Andrea, Christian Ghiglino, and Francesco Squintani. Strategic Information Transmission in Networks. *Journal of Economic Theory*, 148(5): 1751-1769, 2013.

[20] Garicano, Luis. (2000). Hierarchy and the Organization of Knowledge in Production. *Journal of Political Economy*, 108(5), 874–904.

[21] Garicano, Luis. and Andrea Prat (2013). "Organizational Economics with Cognitive Costs," In *Advances in Economics and Econometrics: Theory and Applications, Proceedings of the Tenth World Congress of the Econometric Society*, Cambridge University Press, 2013.

[22] Gibbons, Robert and John Roberts (eds.). *The Handbook of Organizational Economics.* Princeton University Press. 2013.

[23] Halac, Marina and Andrea Prat. Managerial Attention and Worker Engagement. CEPR Discussion Paper 10035. June 2014.

[24] Hagenbach, Jeanne and Frédéric Koessler. Strategic Communication Networks. *Review of Economic Studies* 77(3): 1072-1099, July 2010.

[25] Golub, Benjamin and Matthew O. Jackson (2010). Naive Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, 2(1):112-149.

[26] Mankins, Michael C, Chris Brahm, and Gregory Caimi. Your Scarcest Resource. *Harvard Business Review*, April 2015.

[27] Makarov, Uliana. Networking or not working: A model of social procrastination from communication (2011). *Journal of Economic Behavior and Organization* 80(3): 574–585.

[28] Palacios-Huerta, Ignacio and Andrea Prat (2010). Measuring the Impact Factor of Agents within an Organization Using Communication Patterns. CEPR Discussion Paper 8040.

[29] Palacios-Huerta, Ignacio and Oscar Volij. The Measurement of Intellectual Influence (2004). *Econometrica* 72(3), 963-977.

[30] Christopher A. Sims (2003). Implications of rational inattention. *Journal of Monetary Economics* 50 (3), 665-690.

[31] Sethi, Rajiv and Muhammet Yildiz. Perspectives, Opinions, and Information Flows (2013). *MIT Department of Economics Working Paper* 13-23.

[32] Van den Steen, Eric (2014). A Formal Theory of Strategy. Harvard Business School Working Paper.