Barcelona School of Economics

WORKING PAPERS

Managerial Leadership, Truth-Telling, and Efficient Coordination

BSE Working Paper 1211
October 2020 (Revised: June 2024)

Jordi Brandts, David J. Cooper

bse.eu/research

Managerial Leadership, Truth-Telling, and Efficient Coordination

by Jordi Brandts* and David J. Cooper**

* Instituto de Análisis Económico (CSIC) and Barcelona School of Economics

** University of Iowa and University of East Anglia

Corresponding author: david-j-cooper@uiowa.edu

June 28, 2024

Abstract: We study the manager-agent (MA) game, a novel coordination game played between a manager and two agents. Unlike commonly studied coordination games, the MA game stresses asymmetric information (agents know the state of the world but managers don't) *and* asymmetric payoffs (for all states of the world, agents have opposing preferences over outcomes). *Efficient* coordination requires coordinating agents' actions *and* utilizing their private information. We vary how agents' actions are chosen (managerial control versus delegation), the mode of communication (none, structured communication, or free-form chat), and the channels of communication (i.e. who can communicate with each other). Achieving coordination *per se* is not challenging, but, averaging across all states of the world, total surplus only surpasses the safe outcome when managerial control is combined with three-way free-form chat. Unlike weak-link games, advice from managers to agents does not increase total surplus. The combination of managerial control and free-form chat works because under these conditions agents rarely lie about their private information. Our results suggest that common findings from the experimental literature on lying are not robust to changes in the mode of communication.

Keywords: Coordination, Experiments, Organizations, Communication, Truth-Telling

JEL Classification Codes: C92, D23, L20

Acknowledgements: The authors thank the National Science Foundation (SES-0214310), the Spanish Ministry of Economy and Innovation (Grant: ECO2020-114251GB-I00), the Severo Ochoa Program for Centers of Excellence in R&D (CEX2019-000915-S), and the *Generalitat de Catalunya* (Grant: 2017 SGR 1136). We thank Zeina Aboushaar, Caleb Bray, Adrià Bronchal, Joe Ballard, Ellis Magee, Luke Ortolani, and Lavinia Piemontese for valuable help as research assistants, and seminar participants at Alaska – Anchorage, Case Western Reserve, Copenhagen Business School, Durham, East Anglia, Edinburgh, Florida State, Indiana, Lausanne, Maryland, Paderborn, Paris School of Economics, Stavanger, UCSB, Virginia, the Arne Ryde Workshop, the London Experimental Workshop, and the NBER Organizational Economics Working Group for useful feedback.

1. Introduction: There is a large experimental literature studying whether managers can use various instruments, notably communication, to coordinate agents' actions on an efficient outcome. This research focuses on cases like the weak-link game where agents share common and known objectives with each other and their manager. In settings with aligned interests, efficient symmetric outcomes, and symmetric information, it is well-established that communication among agents and advice from a manager to agents both increase the probability of efficient coordination. Allowing managers to directly control agents' actions makes efficient coordination trivial. The manager's task is far more difficult when the underlying problem is asymmetric. Specifically, we are interested in settings that have the following four properties: (1) All parties would be better off if the agents coordinate on a common course of action. (2) Although all parties gain from coordination, they have differing preferences over which common course of action should be chosen. (3) The state of the world varies over time, changing which outcome is efficient in the sense of maximizing total surplus. (4) Agents know the state of the world, but the manager does not. The combination of Properties 2 and 4 implies that agents have an incentive to lie to their manager. Ideally, a manager achieves efficient coordination, inducing her agents to coordinate on a common course of action *and* doing so in a way that uses their private information to reach the efficient outcome. Achieving this is non-trivial given the agents' opposing objectives and the manager's lack of critical information.

Many problems like this arise within organizations. Managers and their agents must make choices about what inputs to use, what people to hire, what products to produce, and what strategies to pursue. There typically exist gains from coordination when these decisions are made. For example, costs are lower due to purchasing power if employees all use the same equipment or software, there are beneficial synergies if middle managers hire workers with complementary skills, and there are economies of scale if a firm's stores all sell the same products. But agents often have differing preferences over the available options that are independent from what is good for the organization as a whole. Workers like to use equipment and software they are already familiar with, middle managers prefer to hire people for the group they manage (e.g. the people in charge of product design always want more people for the product design group even if the firm needs more help in marketing), and store managers want product types that conform to tastes in their specific location. It is natural that agents who are "in the field" will be better informed than their manager. Workers who actually use a piece of software are probably better informed about its merits than their boss. Store managers who interact with customers are more likely to know the latest trends in consumer demand than an upper-level manager sitting in a glass tower. The problem is that these agents have little incentive to be truthful with their manager rather than trying to influence her into picking their preferred option.[1]

[1] A real-world example is the adoption of standards for new technologies such as HDTV and wireless in the 90s. Europe used a more centralized process than the US. This was especially true for wireless, where the US largely

Can coordination on an efficient outcome be achieved in the face of these difficulties? If the manager imposes a decision on her agents ("managerial control"), can she get her agents to reveal their information even though it is not in their interest? If so, does this depend on what type of communication is possible and, more broadly, what is the mechanism behind agents telling the truth against their interests? Is it better to delegate decision making to agents ("delegation"), giving up control in exchange for eliminating the need to extract information? Can the manager still play a useful leadership role while delegating authority by suggesting a course of action to her agents ("managerial advice")? The purpose of this paper is to address these questions.

We examine these issues using a novel coordination game, the manager-agent game ("MA game"). This is a three-player game with one player in the role of manager and two acting as her agents. The four properties listed above are all present: coordinating on a common action benefits all agents, agents have divergent preferences over possible outcomes, managers lack the necessary information to simply impose efficient coordination on their agents, and agents possess the relevant information but have little reason to truthfully reveal it. The MA game confronts subjects with a challenging coordination problem that accentuates the contrast between managerial control and delegation.

The experimental design features repeated play of the MA game in fixed groups. We vary how actions are chosen (delegation vs. managerial control) and what type of communication is available (none, structured communication with pre-specified messages, or free-form chat). The baseline treatment features delegation and no communication; this is expected to yield poor performance since there is no mechanism to address the challenging coordination problem faced by agents. Within each type of communication (structured communication or free-form chat) we consider three possible mechanisms for choosing actions: delegation with pre-play communication between agents, delegation with managerial advice, and managerial control with messages from agents to the manager.

Our analysis focuses on outperforming the "safe" outcome. This is the unique pure-strategy equilibrium outcome under managerial control for both the one-shot and finitely repeated MA games; the manager receives no information from her agents and imposes coordination. The resulting babbling equilibrium is generally not efficient in terms of maximizing total surplus; it provides a baseline for what successful coordination can achieve *without using agents' information*. Importantly, the safe outcome is the secure equilibrium under delegation (agents use their maxmin actions) and equalizes agents' payoffs on a round-

---

delegated the problem of choosing a standard to firms. This made some sense, as firms are better informed about the merits of various technologies than a central regulator, but also created a thorny coordination problem. Both network externalities and IP issues gave firms reasons to prefer differing technologies. Indeed, the US market suffered through an extended period where firms failed to coordinate on a single technology and was widely seen as lagging behind Europe (HDTV: Farrell and Shapiro, 1992; wireless: Gandal, Salant, and Waverman, 2003). Other related examples from the field include software adoption (Brynjolfsson and Kemerer, 1996) and product line selection (Thomas, 2011).

by-round basis. As such, the safe outcome provides an easy, if inefficient, way of solving the coordination problem under delegation. Unlike the safe outcome, achieving efficient coordination requires use of agents' information.

Turning to the data, total surplus is always higher with managerial control than with delegation. With no communication or structured communication, delegation invariably leads to coordination failure. Managerial control solves this problem but rarely achieves efficient coordination due to poor information transmission from agents to managers. With free-form chat, delegation with no managerial role achieves coordination, but agents opt for the safe outcome more often than efficient coordination. The safe outcome offers an easy way to coordinate and agents take advantage of this. Managerial advice helps little, as an increase in efficient coordination is offset by an increase in coordination failure. Managerial control solves the coordination problem *and* makes good use of agents' information. Only with the combination of managerial control and free-form chat is the total surplus significantly greater than the safe outcome, achieving roughly half of the possible efficiency gains.

Why does the combination of free-form chat and managerial control outperform the safe outcome when the theory predicts that agents should transmit no information to managers? Contrary to both the theory and the data with structured communication and managerial control, lying by agents is almost non-existent when free-form chat and managerial control are combined. This yields unambiguous transmission of information, making efficient coordination possible.

The preceding observation raises an important question: why is there so little lying with free-form chat? To address this question, we add a follow-up treatment that uses structured communication, but includes a number of features that mimic important aspects of free-form chat. Enriching structured communication has no effect on performance and less than half the effect of free-form chat on the frequency of lying. These observations strongly suggest that free-form communication is necessary per se for the sharp decrease in lying observed with the combination of managerial control and free-form chat.

Our work contributes to existing research in multiple ways. There is a large experimental literature about the effects of communication, advice, and leadership on efficient coordination. (See Section 2 for a full summary of related research.) This research focuses on settings like the weak-link game where the interests of all individuals are aligned and there is no dispute about the most desirable outcome. Communication from a leader, such as advice from the manager to agents, is effective in these settings, but it is relatively easy to act as a coordination device when everyone has the same information and aligned interests. Achieving efficiency via managerial control is trivial in a weak-link game, presumably explaining why this has not been studied.[2] Like a weak-link game (and unlike a battle-of-the-sexes game), the MA game has an inefficient symmetric equilibrium that makes it easy to coordinate. The challenge facing a leader is *not* to achieve coordination *per se*, but rather to achieve *efficient* coordination. Compared to weak-link games, efficient coordination is difficult in the MA game. Managers need to do more than just overcome strategic risk. They must either induce turn-taking through managerial advice or impose turn-taking via managerial control. Achieving efficient coordination via managerial control is non-trivial in the MA game due to asymmetric information; the manager has to acquire information that their agents have no incentive to provide. Managerial advice proves insufficient to outperform the safe outcome, unlike weak-link games, but the combination of managerial control and free-form chat performs surprisingly well.

Our work also relates to research, both theoretical and experimental, comparing centralization and decentralization. The recent literature sees this tradeoff in terms of a comparison between the benefits of coordination (meant in a somewhat different sense than here) and the costs of distorted information that accompany centralization. These tradeoffs play an important role in our work as well, but there are major differences. Unlike the games studied in papers like Alonso, Dessein, and Matouschek (2008a) and Rantakari (2008), the MA game with delegation is a true coordination game where the critical issue is selecting among multiple equilibria. We stress the role of active leaders using free-form communication to achieve efficient coordination. The effects of free-form communication are particularly relevant for the literature comparing centralization and decentralization. Performance with managerial control and free-form chat surpass the safe outcome because agents revealed more information than was consistent with their financial incentives. Even though the games are significantly different, the same insight presumably applies to comparisons of centralization and decentralization – existing theories may systematically underestimate the benefits of centralization due to overestimating the willingness of agents using free-form communication to distort reports about their information for strategic purposes.

Our work contributes to the large and growing literature on whether and when individuals are willing to lie. The typical finding is that individuals lie less than is payoff maximizing, adjust the frequency of lies in response to changing incentives (including both their own and other's payoffs), and frequently use partial lies (neither telling the truth nor lying to the full extent that would maximize profits). When agents are limited to sending a bare message about the state of the world, our data exhibits all of these standard patterns. It is striking that none of these patterns are present in the treatment with managerial control and free-form chat. Agents almost never lie, the frequency of lying does not respond to incentives, and partial

[2] There is little work on leadership in asymmetric games. The most relevant is the one-way communication treatment in Cooper, DeJong, Forsythe, and Ross's (1989) study of communication in the battle-of-the-sexes, but this is equivalent to communication between agents rather than managerial advice. See Section 2 for a detailed discussion. Managerial control is trivial in the battle-of-the-sexes due to the lack of asymmetric information.

lies are rare. The follow-up treatment eliminates many possible reasons for the sharp reduction in lying, such as agents fact-checking each other (i.e. calling out when the other agent lies), managers requesting truthful messages, and asynchronous communication. This suggests that some effect of free-form communication per se causes the reduction in lying. We speculate in the conclusion on what this effect might be, but it remains an open question why free-form communication has such a strong effect on lying.

Finally, the unwillingness of agents to lie plays a critical role in generating efficient coordination with managerial control and free-form chat. Previous work has found that pre-play communications in the form of free-form chat is more effective than restricted communication at fostering efficiency and cooperation (Brandts, Cooper, and Rott, 2019), but our work identifies a new channel by which this occurs.

2. Related Literature: There are a number of experiments showing that leaders can increase the likelihood of efficient coordination either by leading by example (e.g. Weber, Camerer, and Knez, 2004; Cartwright, Gillet, and van Vugt, 2013; Sahin, Eckel, and Komai, 2015) or by sending messages (Weber, Camerer, Rottenstreich, and Knez, 2001; Cooper, 2006; Brandts and Cooper, 2007; Brandts, Cooper, and Weber, 2015; Sahin et al., 2015; Cooper, Hamman, and Weber, 2020).[3] Unlike our work, these papers study symmetric games, mainly weak-link games. In a weak-link game, there is no dispute over what equilibrium should be chosen. The primary role of a leader involves overcoming strategic uncertainty. Choosing the efficient equilibrium is risky, and leaders help by establishing common beliefs that everyone will choose the efficient outcome. With the exception of Cooper et al. (2020), asymmetric information does not play an important role in the existing literature. Asymmetric payoffs and the resulting disputes are at the heart of the problem managers face in our experiment, and asymmetric information exacerbates the difficulty.[4]

Closely related, several experiments study the effect of advice on efficient coordination. This includes papers that study advice from either the experimenter (e.g. Van Huyck, Gillette, and Battalio, 1992; Brandts and MacLeod, 1995; Chaudhuri and Paichayontjivit, 2010) or from another subject who has previously played the game (Chaudhuri, Schotter, and Sopher, 2009). Advice can be effective, particularly if it is common knowledge and the interests of advisors and advisees are aligned. Once again, these papers about advice focus on symmetric games where players have aligned interests. We confront managers with the more challenging problem of resolving the conflicting interests of their agents.

[3] It has also been shown that leaders can increase contributions in public goods games, either leading by example or by transmitting their superior information about the state of the world. See Cooper and Hamman (2021) for a survey of this literature.

[4] In Cooper et al. (2020), the leader is better informed than the followers, and the primary problem created by asymmetric information is that the leader has an incentive to make the state of the world appear better than it is, but risks losing her credibility in the long run. In our paper, the problem is that the leader does not know which equilibrium maximizes total surplus and the well-informed followers have strong incentives to deceive her.

Many papers show that pre-play communication among players (as opposed to an external leader) leads to greater efficiency in social dilemmas (e.g. Dawes, MacTavish, and Shaklee, 1977; Isaac and Walker, 1988; Ostrom, Walker and Gardner, 1992; Charness and Dufwenberg, 2006; Cason and Mui, 2007; Cooper and Kühn, 2014) and symmetric coordination games with Pareto-ranked equilibria (e.g. Cooper, DeJong, Forsythe, and Ross, 1992; Blume and Ortmann, 2007; Kriss, Blume, and Weber, 2016; Blume, Kriss, and Weber, 2017). Especially relevant for our work, Cooper DeJong, Forsythe, and Ross (1989) study the effect of pre-play communication on coordination in a battle-of-the-sexes game, the best-known example of an asymmetric coordination game. Communication is limited to pre-play announcements of intended play. Without communication, coordination is difficult due to the lack of a focal equilibrium. With one-way communication, coordination rates are high as the sender can call for her preferred equilibrium and the receiver generally follows. This can be seen as an example of successful leadership in an asymmetric game. There are many differences between our set-up and the experiments of Cooper et al., but perhaps the most important is our focus on *efficient* coordination. Coordination is achieved in a number of our treatments, and we have little doubt that one-way communication would promote coordination as well. *Efficient* coordination is an entirely different matter. For Cooper et al., efficient coordination is a non-issue as the two pure-strategy equilibria in the battle-of-the-sexes game are not Pareto ranked.

Cooper et al. finds that two-way communication is less effective, although coordination rates improve somewhat with multiple rounds of two-way communication. Our treatment combining structured communication between agents with delegation resembles the Cooper et al. treatment with multiple rounds of two-way communication, but, due to two important differences, we anticipated that pre-play communication would be more effective in the MA game. Unlike the battle-of-the-sexes game, the safe outcome in the MA game provides a simple, safe way of coordinating without asymmetric payoffs. Second, we use partners matching while Cooper et al. uses strangers matching. This provides more opportunities for agents to reach an agreement, and also makes it possible to equalize (expected) payoffs while using asymmetric choices. In spite of our optimism, we also observed little effect relative to no communication.

We find large differences between our structured communication treatments and the parallel chat treatments. These findings parallel existing evidence that the pro-efficiency effects of communication are greater with free-form chat than structured communication (e.g. Lundquist, Ellingsen, Gribbe, and Johannesson, 2009; Ben-Ner and Putterman, 2009; Charness and Dufwenberg, 2010; Cooper and Kühn, 2014; Brandts, Charness, and Ellman, 2016). The mechanism underlying our result, specifically the shift to truth-telling by agents with free-form chat, differs from these previous studies.

An extremely active experimental literature on subjects' willingness to lie has developed over the past fifteen years (e.g. Gneezy, 2005; Erat and Gneezy, 2012; Fischbacher and Föllmi-Heusi, 2013; Gneezy, Kajackaite, and Sobel, 2018; Abeler, Nosenzo, and Raymond, 2019). Several striking regularities have emerged: (1) Subjects often tell the truth even when lying would pay more,[5] (2) the likelihood of lying is sensitive to incentives, and (3) partial lies (failing to either tell the truth or the payoff-maximizing lie) are common. In the treatment with structured communication and managerial control, where agents can only send bare messages about the state of the world, all of these regularities are present in agents' messages. However, lying is *not* sensitive to incentives and partial lies are rare in the treatment with managerial control and chat. Messages are observable in both cases and the message space in both cases is sufficient to communicate the full state of the world, yet the nature of truth-telling is quite different. Both theorists and experimenters have made a great deal of progress in understanding why individuals tell partial lies, but we are unaware of any results that explain the differing results with structured communication and chat. The conclusion discusses some possibilities, but a full explanation is beyond the scope of this paper.

Beyond the general literature on lie aversion, our treatment with structured communication and managerial control (**SC – MC**) is related to work by Lai, Lim, and Wang (2015) and Vespa and Wilson (2016). Both papers study information transmission in variations of the multidimensional cheap-talk model with multiple senders proposed by Battaglini (2002). Lai et al. find that information transmission is better with two senders than one and is particularly good when the receiver's interests are *in all states* aligned with one of the sender's interests along one of the two dimensions of the state space. Vespa and Wilson find that receivers perform poorly at extracting information in games where it is relatively difficult (but possible) to infer the state of the world from senders' messages. The MA game with structured messages and managerial control is also a cheap talk game with multiple senders but differs along a critical dimension from the games studied by Lai et al. (2015) and Vespa and Wilson (2016). In keeping with Battaglini (2002), they study games where the state space is multidimensional and messages fully reveal the state of the world in equilibrium. This contrasts with the MA game where the state space is unidimensional and messages are *not* fully revealing (or even informative) in equilibrium.[6] We intentionally made the informational problems under managerial control as severe as possible. That said, we also observe receivers (managers) struggling to extract information from senders' (agents') messages. If anything, the problem is even more severe as managers make systematic errors even when information extraction is trivial.

Our work has a clear relationship to the extensive literature comparing centralized and decentralized firm management. See Mookherjee (2006) for a survey of the older theoretical literature. Prominent recent examples in the theory literature include Hart and Moore (2008), Alonso, Dessein, and Matouschek (2008a, 2008b and 2015), Rantakari (2008), Hart and Holmstrom (2010), Dessein, Garicano, and Gertner (2010).

[5] Studies of cheap talk games find a similar bias towards telling the truth (Cai and Wang, 2006; Sanchez-Pages and Vorsatz, 2007), which could stem from either an aversion to lying or a failure to grasp the strategic benefits of lying. See Blume, Lai and Lim (2023) for contrary evidence.

[6] See Section 3 on Battaglini for a discussion of the unidimensional case.

Recent empirical studies using observational data include Thomas (2011) and McElheran (2014), and experiments by Evdokimov and Garfagnini (2019) and Hamman and Martínez-Carrasco (2020) compare centralization and decentralization.

Our work is not intended to test the predictions of any existing theory comparing centralization and decentralization, and our focus differs from the recent theoretical literature. Papers like Alonso et al. (2008a) and Rantikari (2008) concentrate on the relationship between the parameters of the game and the quality of information flowing between the various agents in equilibrium. The MA game with delegation is a true coordination game with multiple equilibria. Rather than focusing on how the equilibrium changes between organizational structures, we study the effect on the likelihood of efficient coordination with a stress on the roles of active leadership and free-form communication. As noted above, our results have some relevance for the theoretical literature comparing centralization and decentralization, but interested readers should see Evdokimov and Garfagnini (2019) for experiments directly testing some of the recent theoretical models. They find support for theoretical predictions from the models of Alonso et al. (2008a) and Rantakari (2008).

In our environment, the manager does not decide whether to delegate decision-making rights or keep control. Nevertheless, our work bears some relation to the experimental work on control, power, and delegation. The seminal work here is Fehr et al. (2013) and Bartling et al. (2014). For some more recent work see Pikulina and Tergiman (2020), Ferreira et al. (2020), and Neri and Rommeswinkel (2017).

3. The Manager-Agent Game: The MA game confronts subjects with a challenging environment that accentuates the tradeoffs between having the manager make decisions for her agents and delegating decisions to the agents.

$$\pi_{Ai} = k_1 - k_2 * coordination\ loss - k_3 * adaptation\ loss - k_4 * state\ loss \qquad \textbf{(Eq. 1)}$$

The MA game is played by two agents (A1 and A2) and a manager (M). Equation 1 gives the basic structure of Agent i's payoffs. They face three types of losses: (1) "Coordination losses" are losses from not choosing the same option as the other agent. In our simple example, it would be difficult to co-author a document if the two engineers used different software packages. Our model assumes that coordination is paramount, so the worst outcome is to have the agents fail to agree on an option. (2) "Adaptation losses" are losses due to deviations from an agent's most desired outcome (the agent has to "adapt" to the wants and needs of others). Adaptation losses do ***not*** depend on the state of the world. To maximize conflict, agents have diametrically opposed tastes in the MA game, with A1's most preferred option being A2's least preferred option (and vice versa). (3) "State losses" ***are*** state dependent, capturing that some options are inherently more or less attractive depending on the state of the world. If coordination is the foremost concern *and* agents care more about getting their most preferred option than the option that is best for the task at hand, it follows that $k_2 > k_3 > k_4 > 0$. Imposing these inequalities makes the game induced by any state of the world into a coordination game where the two agents have diametrically opposed interests.

The manager's payoff is the sum of the agents' payoffs, implying that management seeks to minimize total costs. This represents a setting where the manager is rewarded for how her unit does as a whole, and should not be interpreted as benevolence on the part of the manager. Because the agents have directly opposed interests, adaptation costs play no role in the manager's decisions under managerial control. Anything that makes one agent happier will necessarily make the other agent less happy. The misaligned incentives that play a central role in most principal-agent problems are also present in the MA game since agents care about whether or not coordination occurs at their preferred option but the manager does not.

3.1. Stage Game Payoff Functions: This sub-section formally describes the MA game. There are three players in the game, a manager (M) and two agents (A1 and A2). Let γ denote the state of the world: $\gamma \in \{1,2,3,4,5\}$. As standard nomenclature, we refer to the states of the world by the games they induce (e.g., Game 1 for $\gamma = 1$). The value of γ is randomly determined before players take any actions. Draws of γ are i.i.d. with each game equally likely. To ease comparisons across treatments, we used the same draw of games for all sessions (although different groups in a session faced different draws). Both agents know the draw of G, but the manager only knows the *ex-ante* distribution over games. The agents A1 and A2 choose (under delegation) or are assigned (under managerial control) actions α_1 and α_2: $\alpha_i \in \{1,2,3,4,5\}$. For simplicity, we use the term "outcome" to refer to the pair of actions (α_1, α_2) chosen by or assigned to agents A1 and A2.[7] Equations 2a, 2b, and 2c give the payoff functions for A1, A2, and M respectively.

$$\pi_{A1} = k_1 - k_2|\alpha_1 - \alpha_2| - k_3|\alpha_1 - 5| - k_4|\alpha_1 - \gamma| \quad \textbf{(Eq. 2a)}$$
$$\pi_{A2} = k_1 - k_2|\alpha_1 - \alpha_2| - k_3|\alpha_2 - 1| - k_4|\alpha_2 - \gamma| \quad \textbf{(Eq. 2b)}$$
$$\pi_M = \pi_{A1} + \pi_{A2} \quad \textbf{(Eq. 2c)}$$

A number of the MA game's features were chosen to accentuate specific aspects of the coordination problems facing managers and agents: (1) Without asymmetric information, the problem facing managers under managerial control is trivial as information transmission is a non-issue. (2) Having a single common shock rather than two independent shocks accentuates the difference between managerial control and delegation. Under delegation, asymmetric information plays no role but agents' conflicting preferences make coordination difficult. With managerial control, coordination is trivial but achieving efficiency

[7] From a technical point of view, the MA game is a game of imperfect information. The outcome of the game depends on the strategies chosen by players and the state of the world. We abuse terminology to simplify the exposition.

requires the manager to overcome the asymmetric information between her and her agents.[8] (3) The functional forms in Equations 1a and 1b use absolute values of differences, rather than squared differences as used by Alonso et al. (2008a) and Rantakari (2008). Because of this choice, there are multiple equilibria in the MA game with delegation rather than a single equilibrium.[9] Multiplicity plays a central role in our paper, as the main problem facing managers is trying to achieve efficient coordination rather than defaulting to the safe outcome. (4) We use five possible actions and five states of the world rather than two (as in a battle-of-the-sexes game) or three. Going from two to three possible actions adds the safe outcome (defined below) as an equilibrium that plays a critical role in subjects' choices. Going from three to five actions makes it easier to distinguish whether play is consistent with the efficient or safe outcome, since the two are equivalent less frequently, and easier to detect partial lies. (5) To accentuate their differing preferences, agents are paid based solely on their own payoffs rather than a weighted average over the two agents' payoffs. Implicitly, this eliminates incentive schemes that include revenue or profit-sharing components. (6) Finally, the interaction between the manager and agents under managerial control is modeled as a cheap talk game rather than a problem of mechanism design. Implicitly, we assume that the manager cannot commit to a mechanism for eliciting information.

For all treatments, $k_1 = 54$, $k_2 = 14$, $k_3 = 7$, and $k_4 = 4$. Table 1 displays the payoff tables for $\gamma = 1$, 3, and 5. Copies of all five payoff tables can be found in Appendix A. The three numbers in each cell of Table 1 correspond to the payoffs, denominated in ECUs, of A1 (π_{A1}), A2 (π_{A2}), and M (π_M). The row and column are the actions chosen by A1 and A2 respectively (or chosen for them by M). The row (R) and column (C) numbers correspond to the actions chosen by the agents (e.g. R3 $\equiv \alpha_1 = 3$; C4 $\equiv \alpha_2 = 4$).

Table 1: Stage Game Payoffs (k_1 = 54, k_2 = 14, k_3 = 7, and k_4 = 4)

Note: Each cell contains the payoffs for A1 (π_{A1}), A2 (π_{A2}), and M (π_M).

Game 1

	C1	C2	C3	C4	C5
R1	26, 54, 80	12, 29, 41	-2, 4, 2	-16, -21, -37	-30, -46, -76
R2	15, 40, 55	29, 43, 72	15, 18, 33	1, -7, -6	-13, -32, -45
R3	4, 26, 30	18, 29, 47	32,32, 64	18, 7, 25	4, -18, -14
R4	-7, 12, 5	7, 15, 22	21, 18, 39	35, 21, 56	21, -4, 17
R5	-18, -2, -20	-4, 1, -3	10, 4, 14	24, 7, 31	38, 10, 48

[8] Models like Alonso et al. (2008a) and Rantikari (2008) use two independent shocks. Using the terminology of our paper, this leads to a comparison of information flows between agents under delegation versus information flows from agents to managers under managerial control. Our paper focuses on the coordination problem with delegation and eliminates any issues due to asymmetric information.

[9] With absolute values, an agent maximizes his payoff by picking exactly the same action as the other agent. With squared differences, agents want to shade their choice away from their own preferred outcome and towards the other agent's action rather than matching the other agent's action.

Game 3

	C1	C2	C3	C4	C5
R1	18, 46, 64	4, 29, 33	-10,12, 2	-24, -13, -37	-38, -38, -76
R2	15, 32, 47	29, 43, 72	15, 26, 41	1, 1, 2	-13, -24, -37
R3	12, 18, 30	26, 29, 55	40, 40, 80	26, 15, 41	12, -10,2
R4	1, 4, 5	15, 15, 30	29,26, 55	43,29, 72	29, 4, 33
R5	-10, -10, -20	4,1, 5	18,12, 30	32, 15, 47	46, 18, 64

Game 5

	C1	C2	C3	C4	C5
R1	10, 38, 48	-4, 21, 17	-18, 4, -14	-32, -13, -45	-46, -30, -76
R2	7, 24, 31	21, 35, 56	7, 18, 25	-7, 1, -6	-21, -16, -37
R3	4, 10, 14	18, 21, 39	32, 32, 64	18, 15, 33	4, -2, 2
R4	1, -4, -3	15, 7, 22	29, 18, 47	43, 29, 72	29, 12, 41
R5	-2, -18, -20	12, -7, 5	26,4, 30	40, 15, 55	54, 26, 80

3.2: Equilibrium, Delegation: With delegation, each agent chooses an action and the manager is a passive bystander. Ignoring the payoff for M, all five games are coordination games with five pure-strategy Nash equilibria where the two agents choose the same action: ($\alpha_1 = \alpha_2 = 1$), ($\alpha_1 = \alpha_2 = 2$), ($\alpha_1 = \alpha_2 = 3$), ($\alpha_1 = \alpha_2 = 4$), and ($\alpha_1 = \alpha_2 = 5$). We refer to these outcomes as Outcome 1, Outcome 2, etc.

In all five games, there is a tension similar to the battle-of-the-sexes game since A1 most prefers Outcome 5 as an equilibrium and least prefers Outcome 1, the reverse is true for A2, and M prefers the equilibrium that maximizes total surplus. This implies that M always wants a different equilibrium than at least one of her agents and wants a different equilibrium than either A1 or A2 in Games 2, 3, and 4. Alternative principles for equilibrium selection, such as safety and efficiency, suggest different ways of resolving the tension stemming from agents' differing interests.

Unlike a battle-of-the-sexes game, but similar to a weak-link game, the MA game with delegation offers an equilibrium that is safe, simple, and fair, but is not efficient in the sense of maximizing total surplus. Outcome 3 (the "safe" outcome) is safe because $\alpha_i = 3$ is the maximin strategy for both agents in all five games, is simple because agents use the same action in all states of the world, and fair because it yields the same payoff to both agents. Except in Game 3, the safe outcome does *not* maximize total surplus. Despite this, the attractive features of the safe outcome give it drawing power in our data.

All five games have an equilibrium that maximizes total surplus. This is always equivalent to the game number (i.e., Outcome 1 in Game 1, Outcome 2 in Game 2, etc.). Efficient coordination, where the agents play the surplus-maximizing equilibrium in all states of the world, is procedurally fair (i.e., equalizes expected payoffs under the veil of ignorance about the state of the world; Bolton, Brandts, and Ockenfels,

2005) but yields asymmetric payoffs for all games except Game 3. Efficient coordination is also relatively complex because the agents must change their actions as the state of the world changes.

3.3: Equilibrium, Managerial Control: The following discussion is based on structured communication, but extends in a straightforward manner to free-form chat. With managerial control, agents do not choose rows and columns directly. After being informed about the state of the world (i.e., Game 1, Game 2, etc.) each agent independently sends a message to the manager indicating which state of the world has been selected ($\mu_i \in \{1,2,3,4,5\}$). After receipt of the two messages (μ_1 and μ_2), the manager chooses both a row and a column (α_1 and α_2). She has no knowledge about which game has been selected beyond the initial distribution over states of the world and whatever information she gleans from the agents' messages.

Conditional on enforcing coordination, Equations 1a and 1b imply that the manager does not care about the adaptation losses, but the agents do. Given their opposing interests, the agents have no incentive to be truthful with the manager. If both agents always report the game where the efficient outcome is best for them (Game 5 for A1, Game 1 for A2), the best the manager can do is to choose the safe outcome ($\alpha_1 = \alpha_2 = 3$).[10] Any benefits from the agents' private information are lost and the manager generally will not choose the efficient outcome.

More formally, we can prove the following theorem which implies that the only pure-strategy Perfect Bayesian equilibrium (PBE) are babbling equilibria where the safe outcome ($\alpha_1 = \alpha_2 = 3$) is always chosen. In Appendix B we show that the manager must pick the same row and column ($\alpha_1 = \alpha_2$) in any PBE. Given this, we henceforth refer to the manager as choosing a single action in response to the agents' messages.

Theorem: There does not exist a pure-strategy PBE for the MA game with managerial control where the manager chooses different actions for two different states of the world.

Proof: See Appendix B.

Finite repetition of the MA game with managerial control does not expand the set of possible pure-strategy equilibria to include informative equilibria. There is a unique pure strategy equilibrium payoff vector in the stage game. Doing backward induction, the set of equilibrium payoffs only expands if players can take advantage of differing payoffs across stage game equilibria to prevent deviations. If there were payoff differences between the equilibria, a punishment scheme could be constructed to support equilibria that are not babbling equilibria in the repeated game. But given that payoffs do not vary across equilibria, this is not possible.

---

[10] Given that payoffs are linear, this isn't transparent. Define a manager's error as the difference between the action she chooses (assuming $\alpha_1 = \alpha_2$) and her payoff maximizing choice. Choosing the safe equilibrium limits the size of manager errors. If she chooses the safe equilibrium, her average error is 1.2. If she chooses action 2 or 4, the average error rises to 1.4. Choosing 1 or 5, the average error goes up to 2.0.

The absence of an informative equilibrium does *not* reflect a generic property of cheap talk games with multiple senders; such games generically have an informative equilibrium when the state space is multidimensional (Battaglini, 2002). The MA game intentionally gives the two agents diametrically opposed interests over a unidimensional state space. The resulting lack of an informative equilibrium makes information transmission theoretically impossible with managerial control. This is in keeping with our goal of confronting subjects with a challenging environment that accentuates differences between the manager retaining control or delegating choices to her agents. In the quest for efficient coordination, managerial control exchanges the problem of having multiple equilibria for the problem of needing to get agents to reveal information against their interests.

The theory assumes messages are cheap talk, with agents incurring no costs, pecuniary or psychological, for sending false messages. If we add a psychological cost for sending false messages, as in Kartik (2009), it is trivial to construct cases where truthful revelation is consistent with an equilibrium. For example, let $c_L*|\mu_i - \gamma|$ be agent i's psychological cost of lying. If $c_L > k_3 - k_4$, there exists an equilibrium in which both agents truthfully reveal their information.[11]

There does not exist a pure strategy equilibrium which is informative, but there do exist mixed strategy equilibria which are partially informative. These equilibria are delicate; in Appendix B we show that no such equilibria exist if we require that an equilibrium is robust to the introduction of trembles and manager strategies are monotonic. Empirically, we see no evidence of such equilibria occurring. As such, we regard the partially informative equilibria as theoretical oddities rather than predictive of subject behavior.

4. Experimental Design and Hypotheses:

4.1 Experimental Design and Procedures: The initial design included seven treatments, broken into three broad categories by type of communication (no communication, structured communication, or chat). The six treatments with communication cross two types of communication (structured or free-form chat) with three different interaction structures between manager and agents (delegation with communication between agents, delegation with advice from managers, and managerial control). Before getting into the details of the experimental design, we pause to discuss the rationale for including both of these dimensions.

Structured communication and chat can be seen as extreme types of communication that serve different roles in our design.[12] Structured communication allows for clean tests of theoretical predictions since the message space matches the theoretical model being tested. More generally, the simplicity of structured

[11] If both agents send the same message, the manager chooses the corresponding outcome. If $\mu_i = 1$ and $\mu_j = 2$, where $i,j \in \{1,2\}$ and $i \neq j$, the manager chooses Outcome 2. If $\mu_i = 4$ and $\mu_j = 5$, where $i,j \in \{1,2\}$ and $i \neq j$, the manager chooses Outcome 4. Otherwise, the manager chooses Outcome 3.

[12] Section 6 presents a new treatment where the mode of communication is intermediate between these two extremes.

communication makes it easier to pin down the mechanism by which communication affects outcomes. Free-form chat is richer and offers a more realistic form of communication. There exists ample evidence that free-form chat has differing effects from structured communication (Brandts et al., 2019), generally being more effective at promoting cooperation. There is no existing work that we know of comparing the effects of structured and free-form communication on either coordination or truth-telling.

Turning to the other dimension of the design, the first interaction structure, delegation with communication between agents, gives the manager no role. The latter two treatments, delegation with advice from managers and managerial control, explore different ways in which a manager might try to achieve efficient coordination, either persuading the agents to coordinate efficiently, essentially acting as a focal point, or imposing actions upon the agents. In the absence of asymmetric information, it is trivial to achieve efficient coordination by imposing actions. But it is no longer obvious which approach will be most effective with the addition of asymmetric information. Our treatments emphasize different approaches to the manager's problem, letting us see which will be more effective.[13]

The seven initial treatments are as follows:

No Communication – Delegation (NC – D): This was the baseline treatment where subjects played the MA game with delegation, as described in Section 3.2, without any additional communication.

Structured Communication between Agents – Delegation (SC/S – D): This treatment was identical to the **NC – D** treatment, except pre-play communication between agents was added. Prior to the agents' choices of actions, each game began with three rounds of messages. Within each round of messages, the agents simultaneously chose a pair of messages suggesting actions for themselves and the other agent. The message space was limited in structured communication treatments; in **SC/S – D**, for example, the agents chose messages by clicking on radio buttons labeled with the five available actions and could not send any other messages. Agents observed each other's messages at the end of each round of messages. The purpose of having three rounds of messaging (rather than one) was to make it easier for agents to agree upon a course of action.

Structured Communication with Advice – Delegation (SC/A – D): This treatment was identical to the **NC – D** treatment, except the manager sent a message to the agents prior to each round of play. This message suggested actions for both agents in each of the five possible games. In other words, the manager advised

[13] The treatments with structured communication separately identify the effects of pre-play communication between agents and managerial advice. Given that neither had much effect in isolation, it seems safe to assume that the combination, paralleling **CH/A – D**, would also have little effect.

a course of action contingent on the realized state of the world. The full message (a 5 x 2 matrix) was shown to both S1 and S2 prior to their choice of actions. The agents knew that both received identical messages.

Structured Communication – Managerial Control (SC – MC): In this treatment, subjects played the game with managerial control as described in Section 3.3. In each round, the two agents viewed the state of the world (i.e., Game 1, Game 2, etc.) and sent simultaneous messages to the manager reporting the state of the world ($\mu_i \in \{1,2,3,4,5\}$). There was no requirement that these messages be truthful, a point emphasized in the instructions. After receipt of the two messages, M chose both a row and a column. There was no requirement that the row and column match.

Chat between Agents – Delegation (CH/S – D): This treatment was identical to the **NC – D** treatment, except pre-play chat between agents was added. The agents had two minutes to engage in free-form discussions via chat before choosing actions. They could discuss whatever they chose. In practice, discussions largely focused on the obvious topic, how to play the game. The manager saw the discussion but could not participate.

Chat with Advice – Delegation (CH/A – D): This treatment was identical to the **CH/S – D** treatment, except the manager could participate in the chat prior to her agents choosing actions. As in **SC/A – D**, the manager had no control over the agents' actions but could advise them. Free-form chat opened many possibilities beyond simply advising a course of action, such as explaining the rationale for adopting efficient coordination or persuading agents of the mutual benefits of trusting each other.

Chat – Managerial Control (CH – MC): This treatment was identical to the **SC – MC** treatment, except the structured messages about the state of the world were eliminated and replaced by free-form chat between the agents and their manager. The agents were not specifically instructed to share information about the state of the world, but this was a natural and typical topic of conversation, making the structured messages redundant. Again, there was a two-minute time limit. Unlike **CH/A – D**, the manager had control over the outcome and the agents had reason to not truthfully reveal the state of the world.

Beyond the seven treatments reported in the main text, we ran an additional four treatments. These were modifications of the **NC – D** and **SC – MC** treatments to examine secondary issues. The main experimental design holds incentives fixed to focus on the effects of changing decision-making rights and the types of available communication. The **HSL – D** and **HSL – MC** treatments examine the effect of changing incentives by increasing state losses ($k_4 = 6$ vs. $k_4 = 4$). This reduces the tension between agents, making the efficient outcome more attractive relative to the safe outcome. Play shifts towards efficient coordination, consistent with the change in incentives, but our qualitative conclusions are unaffected. In particular, lowering state losses does not significantly increase efficiency gains relative to **SC – MC** and efficiency gains remain significantly lower than in **CH – MC**. In other words, a strong increase in incentives to play the efficient equilibrium has significantly less impact than allowing free-form communication. The **STR – D** and **STR – MC** treatments used strangers matching rather than partners matching. This made efficient coordination harder, as expected, but the effects are not significant and our qualitative conclusions are unaffected. Appendix D provides more description of these four additional treatments and the results.

We used a between-subjects design, so each subject participated in just one of the treatments. There were three sessions per treatment and nine three-person groups per session, giving 27 subjects per session, 27 independent groups per treatment, and a total of 567 subjects in 189 independent groups.

Subjects played 18 rounds in all treatments. They were assigned the role of M, S1, or S2 at the beginning of the session and kept these roles throughout the session. Partners matching was used (i.e. participants were matched with the same two subjects throughout the entire experiment). In treatments with delegation, the participants in the M role were pure observers. We did this to keep the possible influence of other-regarding preferences constant across treatments.

The state of the world (i.e. the game being played) was randomly and independently determined at the beginning of each round. Common seeds were used across treatments, limiting the possibility that treatment effects could be driven by differing draws. At the end of each round, subjects received feedback about the realized game and the chosen actions. In the treatments with managerial control, this made it possible for managers to know if an agent had lied about the game being played.

Sessions were run using z-tree (Fischbacher, 2007). Each session began with instructions (see Appendix C). Participants had printed copies of the payoff tables for all five games. Sessions were run at the LINEEX lab at the University of Valencia, with undergraduate students from the university as participants. The payoffs were denominated in Experimental Currency Units, with 1 ECU = 0.2€. Participants received their cumulative earnings for all rounds. Including a 5€ show-up fee, average pay was 19.90€. Sessions lasted approximately an hour.

4.2. Hypotheses: **H1** draws on the theory developed in Section 3 to compare **NC – D** and **SC – MC**. Efficient coordination, which is an equilibrium in **NC – D,** uses agents' information to achieve the maximum possible total surplus. In **SC – MC**, only inefficient pure-strategy babbling equilibria exist. **H1** follows. This hypothesis was a straw man. The MA game with delegation resembles a battle-of-the-sexes game, a setting where coordination is known to be difficult in the absence of communication (Cooper et al., 1989). Even though the presence of a safe equilibrium should make coordination easier, we still doubted that agents could coordinate, let alone coordinate efficiently, in the absence of communication.

H1: *Total surplus will be greater in **NC – D** than in **SC – MC**.*

The theory predicts play of a babbling equilibrium in **SC - MC**, implying that agents' messages will be uninformative. **H2** follows. Once again, there were good reasons to be skeptical. Our design differed from most existing experiments, especially since more than one subject sent messages, but the general finding that individuals are reluctant to lie seemed likely to apply.

H2: (a) *In **SC – MC**, agents' messages will contain no useful information about the state of the world.* **(b)** *Total surplus will not exceed the payoff from the safe outcome (the babbling equilibrium).*

Turning to the treatments with structured communication and delegation, **SC/S – D** and **SC/A – D**, neither type of pre-play communication changes the theoretical prediction. We nevertheless expected total surplus to increase relative to **NC – D** in both cases. The different types of communication (between agents vs. managerial advice) emphasized different mechanisms by which communication might yield efficient coordination. Communication between agents gave them an opportunity to directly coordinate their choices prior to picking actions. The agents did not face any asymmetric information in **SC/S – D**, but lacked an obvious mechanism for resolving conflicts due to their divergent interests. Cooper et al. (1989) observe modest improvements from adding three rounds of bilateral pre-play structure communications to the battle-of-the-sexes game. Based on this evidence, we expected a modest increase in total surplus between the **NC – D** and **SC/S – D** treatments.[14] With managerial advice, we expected managers to act as a coordination device promoting efficient coordination. Because the power to choose actions resides with the agents, asymmetric information should not be an issue in **SC/A – D**. It is always in the manager's interest to promote efficient coordination, and unlike **SC/S – D**, agents in the **SC/A – D** treatment have a single, common source of guidance on how to play. We hoped that the benefits of certain coordination would overcome reluctance to accept a less preferred outcome. Thus, we anticipated that efficient coordination would be more likely with managerial advice than communication between agents. The preceding conjectures are summarized in **H3**. This hypothesis is stated relative to **NC – D**, as the structured communication treatments with delegation modify **NC – D**, but combining **H1** and **H3** yields a prediction that both treatments will also yield higher total surplus than **SC - MC**.

H3: *Total surplus will be greater in **SC/A – D** than **SC/S – D**, and greater in **SC/S – D** than **NC – D**.*

The final hypothesis covers the chat treatments. Many papers have compared the effects of structured communication versus chat. The general finding is that communication has a greater impact on outcomes with chat rather than structured communication (Brandts et al., 2019). Particularly relevant to our current work, Cooper and Kühn (2014) find that free-form communication outperforms structured communication

[14] There are differences in the structure of our experiment and game, described in Section 2 that increased our optimism about the relative performance of the **SC/S – D** treatment.

in a two period Bertrand game, *largely by improving coordination on an efficient equilibrium*. While the games are different, we expected that the ability to make unlimited asynchronous proposals along with the ability to explain proposals would similarly increase efficient coordination under delegation. We therefore expected the chat treatments to yield higher total surplus than the parallel structured communication treatments. We also expected the comparison of total surplus across chat treatments to match the order across structured communication treatments. Consistent with the theory above, we anticipated that agents would not truthfully communicate the state of the world in **CH – MC** leading to play of the safe outcome. We expected that chat would solve the coordination problem in the other two treatments (**CH/S – D** and **CH/A – D**) and would improve efficiency beyond the safe equilibrium. **H4** summarizes our conjectures.

H4: (a) ***CH/S – D*** *will yield higher total surplus than* ***SC/S – D*** *and* ***CH/A – D*** *will yield higher total surplus than* ***SC/A – D****. No difference is predicted between* ***SC – MC*** *and* ***CH – MC****.* **(b)** *Comparing chat treatments, total surplus will be highest is* ***CH/A – D****, followed by* ***CH/S – D*** *and* ***CH – MC****.*

5. Results: Section 5.1 gives an overview of the main treatment effects, and Section 5.2 examines the process underlying these treatment effects.

5.1. Treatment Effects: The discussion of treatment effects is based on the second half of the experiment (Rounds 10 – 18) when play has settled down.[15] Versions of Tables 2 and 3 based on the first half of the experiment can be found in Appendix A. Unless otherwise noted, statistical tests comparing treatments are Wilcoxon rank-sum tests and comparisons with total surplus from play of the safe outcome are Wilcoxon matched-pairs signed-rank tests. The p-values come from exact tests. An observation is the average value of the variable in question for a single group over Rounds 10 – 18. Total surpluses from the safe outcome are adjusted for the random draw of games.

Table 2 summarizes outcomes by treatment from Rounds 10 – 18. To maximize total surplus, the choices of the two agents need to be coordinated ($\alpha_1 = \alpha_2$) *and* these choices have to take advantage of the agents' information ($\alpha_1 = \alpha_2 = \gamma$). Along these lines, the first column of Table 2 reports the percentage of games where the choices were coordinated, and the second shows the frequency of efficient coordination *subject to the agents' choices being coordinated*. The final two columns give measures of overall performance: the third column of Table 2 shows average total surplus and the fourth column reports the average "efficiency gain." Total surplus is defined as the sum of the payoffs for A1 and A2, equivalent to the manager's payoff. Efficiency gain is defined as the difference between a group's total surplus for the nine-round block and the total surplus it would have achieved by playing the babbling equilibrium ($\alpha_1 = \alpha_2$

[15] The average change in total surplus across Rounds 1 – 9 is more than ten times larger than the change across Rounds 10 – 18 (6.40 vs. 0.60).

= 3) throughout, divided by the difference between the total surplus from efficient coordination and the total surplus from the babbling equilibrium. This transformation of the total surplus makes it easier to see how well groups do relative to the babbling equilibrium. Playing the efficient equilibrium yields an efficiency gain of 100%, while the babbling equilibrium leads to an efficiency gain of 0%. Negative efficiency gains reflect failure to coordinate. For treatments where coordination is high, the efficiency game is a good measure of how well groups make use of the agents' information

Table 2: Summary of Outcomes, Rounds 10 – 18

Treatment	% Coordinate	% Efficient (s.t. Coordinate)	Total Surplus	Efficiency Gain
NC – D	69.5%	39.1%	61.4	-108.3%
SC/S – D	77.8%	57.1%	65.7	-57.5%
SC/A – D	69.5%	58.7%	64.4	-77.4%
SC – MC	99.6%	45.9%	71.7	6.5%
CH/S – D	97.5%	48.1%	72.2	14.0%
CH/A – D	90.9%	60.6%	71.6	2.8%
CH – MC	100.0%	61.7%	75.2	44.3%

H1 hypothesized that total surplus would be greater in **NC – D** than **SC – MC**. This was a straw man, relying on the best-case scenario of efficient coordination for **NC – D**, and indeed **H1** is strongly rejected as total surplus is significantly greater in **SC – MC** than **NC - D** ($p < .001$). It is not difficult to see the reason for this difference. Managers understand the importance of coordination, leading to a coordination rate of almost 100% in **SC – MC**. Lacking a coordination device, coordination is significantly lower in **NC – D** ($p < .001$).

Result 1: *Total surplus is significantly higher in **SC – MC** than **NC – D.** The data are not consistent with **H1**. Significantly lower coordination rates largely explain the lower total surplus in **NC – D**.*

While performance is far stronger in **SC – MC** than **NC – D**, it does not follow that much use is made of the agents' information. The efficiency gain is only 6.5% for **SC – MC**, indicating that little of the possible gain over the babbling equilibrium is achieved. The difference between total surplus in **SC – MC** and the babbling equilibrium is not statistically significant ($p = .674$).

To help us better understand why performance varies across treatments, Table 3 summarizes the frequency of specific outcomes in Games 1, 2, 4, and 5. As defined previously, the safe outcome refers to mutual choice of 3 ($\alpha_1 = \alpha_2 = 3$) and the efficient outcome indicates coordinated choices matching the state of the world ($\alpha_1 = \alpha_2 = \gamma$). "Other" refers to any other outcome where the agents' actions are coordinated ($\alpha_1 = \alpha_2$), but at neither the safe nor the efficient outcome. Table 3 does not use data from Game 3 because the efficient and safe outcomes coincide in this case (hence, the coordination rates in Table 3 do not add up to the figures shown in the first column of Table 2).

Table 3: Frequency of Types of Coordination if $\gamma \neq 3$, Rounds 10 – 18

Treatment	% Safe	% Efficient	% Other
NC – D	48.1%	12.6%	4.9%
SC/S – D	36.1%	30.6%	8.2%
SC/A – D	32.8%	30.1%	4.4%
SC – MC	31.1%	36.1%	32.2%
CH/S – D	41.0%	36.1%	20.8%
CH/A – D	29.5%	44.8%	15.8%
CH – MC	34.4%	50.8%	14.8%

Table 3 makes it clear that coordination failure is *not* the only problem in **NC – D**. When agents coordinate and the safe and efficient outcomes do not coincide ($\gamma \neq 3$), they usually coordinate at the safe outcome (69% subject to coordinating) rather than the efficient outcome (18% subject to coordinating). The safe outcome provides a relatively easy route to coordination in the challenging environment of **NC – D**, and agents take advantage of this even though it means *not* using their information.

Matters are a bit more complex in **SC – MC**. In some ways performance is better than the babbling equilibrium. For $\gamma \neq 3$, the efficient outcome is slightly more common than the safe outcome. Unfortunately, average total surplus for the 32% of outcomes in the "other" category of coordination is *lower* than could have been achieved via the babbling equilibrium (62.1 vs. 67.7). These outcomes often do *not* involve shading the difference between safety and efficiency. For Games 1 and 5, 56% of "other" outcomes use actions that are *farther* away from the efficient outcome than the safe outcome. Managers attempt to use their agents' information, but often do so poorly.

Result 2: *In **NC – D,** agents generally coordinate by playing the safe outcome, implying a failure to use their information. Total surplus in **SC – MC** is almost identical to the babbling equilibrium prediction, consistent with **H2(b)**, but play is consistent with neither the babbling equilibrium (repeated play of the safe outcome) nor efficient coordination. This implies a failure to use the agents' information.*

The results in Table 2 provide little support for **H3**. Total surplus in Rounds 10 – 18 is higher for **SC/S – D** and **SC/A – D** than **NC - D**, but the differences are small and not statistically significant (p = .156 and p =.422 respectively). Total surplus is slightly lower in **SC/A – D** than **SC/S – D**, rather than higher as predicted. Neither **SC/S – D** nor **SC/A – D** does as well as **SC – MC**, with both differences significant across Rounds 10 – 18 (p = .056 and p = .009 respectively).

SC/S – D and **SC/A – D** have little effect on total surplus in Rounds 10 - 18 because neither increases the coordination rate significantly relative to **NC – D** (p = .325 and p = .947 respectively). To the limited extent that these treatments do better than **NC – D**, it is by making better use of agents' information. Subject to coordinating in $\gamma \neq 3$, the frequency of the efficient outcome rises from 19% in **NC – D** to 41% and 45% in **SC/S – D** and **SC/A – D** respectively.

Result 3: **SC/S – D** and **SC/A – D** *do not yield significantly higher total surplus than* ***NC – D****, and do significantly worse than* ***SC - MC****. The data do not support* ***H3****.*

The three treatments with free-form chat all yield significantly higher average total surplus across Rounds 10 - 18 than **NC – D** (p < .001 if all three cases). **H4(a)** fares well with delegation, but not with managerial control; **CH/S – D** and **CH/A – D** yield significantly higher total surplus than the parallel structured communication treatments (p = .021 and p = .005 respectively), but total surplus is significantly lower in **SC – MC** than **CH – MC** (p = .003). **CH – MC** also yields significantly higher total surplus than **CH/S – D** (p = .014). Strong performance in the **CH – MC** treatment is not due to chat *or* managerial control, but rather the conjunction of the two.[16]

Oddly, total surplus is not significantly higher in **CH – MC** than **CH/A – D** (p = .162) even though total surplus across Rounds 10 – 18 is *lower* on average in **CH/A – D** than **CH/S – D**. This reflects the high variance of outcomes in the **CH/A – D** treatment. The standard deviation of total surplus is more than double in **CH/A – D** (10.5) than the other two chat treatments (4.7 and 4.0 for **CH/S – D** and **CH – MC** respectively). Looking at the nine groups in the chat treatments that achieve a perfect total average surplus of 80, **CH/A – D** ties with **CH – MC** for the most at four apiece. But, if we look at the nine *worst* groups, **CH/A – D** leads again with five while **CH – MC** has none. It would be a mistake to describe performance in **CH/A – D** as either good or bad; a more accurate adjective would be "erratic."

CH – MC has by far the highest efficiency gain of any treatment (44.3%), and is the only treatment which yields significantly higher total surplus than repeated play of the safe outcome (p < .001).[17] Efficient coordination has two components: agents' choices must be coordinated and must reflect their information. **CH – MC** does well on both accounts. It is the only treatment where groups achieve 100% coordination in Rounds 10 – 18. **CH/S – D** does almost as well at achieving coordination, but **CH/A – D** has a lower coordination rate which largely explains its relatively low total surplus.[18] Not only is coordination 100% perfect in **CH – MC**, but play of the efficient equilibrium is significantly increased relative to either **SC –**

[16] Neither **CH/S – D** (p = .664) nor **CH/A – D** (p = .267) improve performance significantly over **SC – MC**.

[17] Equivalent test statistics for **CH/S – D** and **CH/A – D**, are p = .111 and p = .105 respectively.

[18] The average coordination rate in **CH/A – D** hides a great deal of heterogeneity; 19 of 27 groups achieve 100% coordination, but the other eight groups only have an average coordination rate of 69%.

MC (p = .027) or **CH/S – D** (p = 0.064). **CH – MC** outperforms these two treatments because of superior use of the agents' information. This is *not* true for **CH/A – D** where the rate of efficient coordination is only slightly lower than in **CH – MC (**p = .449**)**.

Result 4: *The data are only partially consistent with* ***H4(a).*** *All three chat treatments produce significantly higher total surplus than the parallel structured communication treatments –* ***H4(a)*** *predicts identical surplus for* ***SC – MC*** *and* ***CH – MC****.*

Result 5: *Across all seven treatments, the combination of free-form chat with managerial control yields the highest total surplus. This reflects both high levels of coordination and improved usage of the agents' information in* ***CH – MC****. This is not consistent with* ***H4(b)*** *which predicts that* ***CH – MC*** *will have the lowest total surplus across the three chat treatments.*

To summarize, either managerial control or free-form communication improves performance, *but only the combination of both beats repeated play of the safe outcome.* The only treatment that **CH – MC** fails to significantly outperform is **CH/A – D**, but this reflects the high variance of outcomes in the latter treatment. Given that **CH – MC** offers higher average total surplus *and* significantly lower risk than **CH/A – D**, it is difficult to argue that **CH – MC** is not doing better. The high performance of **CH – MC** reflects both high coordination rates and a relatively strong ability to use the agents' information. The next section digs into why this treatment does well on both dimensions relative to the other treatments, focusing on how well information is transmitted from agents to managers.

5.2. Process: This subsection examines the processes underlying the treatment effects described in Section 5.1 with a focus on information transmission in the treatments with managerial control (**SC – MC** and **CH – MC**). In both cases managers achieve almost perfect coordination, but only in **CH – MC** do managers take advantage of their agents' information to outperform the babbling equilibrium. We show that this reflects what information is communicated to managers and how they utilize it.

Unless otherwise noted, data from all rounds is used in this section. The eventual outcomes in late rounds are strongly affected by the process in early rounds, plus we often compare how managers behave between early (Rounds 1 – 9) and late (Rounds 10 – 18) rounds.

5.2.a. Structured Communication: Contrary to **H3**, neither treatment combining delegation with structured communication (**SC/S – D** and **SC/A - D**) has a significant impact on total surplus relative to **NC – D**. The fundamental problem in both cases is a failure to significantly improve the coordination rate. In **SC/S – D**, agents only reach an agreement in 63% of the observations, and the coordination rate falls to 40% without an agreement. Turning to **SC/A – D**, managers often fail to give good advice and, even when they do, agents often fail to follow it. Coordination is *not* recommended in 15% of managers' messages, a figure that improves little with experience (12% in late rounds). Managers often take a conservative approach when they do advise coordination, calling for the safe outcome rather than efficient coordination (37% for both if $\gamma \neq 3$). Agents often fail to coordinate even when recommended to do so (65%), a figure which improves little even if the safe outcome is suggested (71%).

Coordination is not the problem in **SC – MC**; rather, the issue is a lack of *efficient* coordination. Two things have to happen in **SC – MC** for a group to take advantage of the agents' information. The agents have to send messages that are informative about the state of the world, and the manager has to correctly interpret the information contained in their messages. The theory presented in Section 3 focuses on the first issue and concludes that information transmission will fail since the agents have no incentive to send informative messages. Built into the theory is an assumption that the messages would be interpreted correctly if informative. In reality, the messages sent by A1 and A2 contain useful information, but managers make frequent errors in using messages. Total surplus is about the same as predicted by the babbling equilibrium (repeated play of the safe outcome) because the advantages from better-than-expected information transmission are wiped out by errors in using this information.

Table 4: Messages as a Function of Game

		Game (Mapped)				
		1	2	3	4	5
Message (Mapped)	1	44	1	3	4	4
	2	2	71	3	2	2
	3	19	27	155	5	8
	4	32	27	21	143	8
	5	80	69	46	41	155

Table 4 displays the messages sent in **SC – MC** as a function of the game. The data from A2 players have been remapped to be from an A1's point of view, allowing us to combine data for the two roles.[19] If messages are uninformative, as the theory predicts, there should be no correlation between messages and the game being played. Instead, there is strong positive correlation ($\rho = .34$). Play of a babbling equilibrium implies that agents only tell the truth in 20% of the observations, but the observed likelihood of truth-telling is 58%. Even when it is least beneficial to do so (Game 1 for A1 or Game 5 for A2), 25% of messages tell the truth. If truth-telling is solely due to a failure to grasp the strategic value of lying, agents should lie more as they learn that lying pays. This is not the case, with 58% truth-telling in both Rounds 1 – 9 and Rounds 10 - 18. Purely self-interested agents should always send a message corresponding to their most preferred

[19] Recall that Outcome 5 is the most desired outcome for A1 and Outcome 1 is the most desired outcome for A2. We remap games for A2: G' = 6 - G. Messages are remapped in an analogous fashion: $\mu 2' = 6 - \mu 2$.

outcome. This is the most common type of lie, but 35% of self-serving lies are partial lies (i.e. the message lies strictly between the true game and the agent's preferred outcome).[20]

Result 6: *The messages sent by agents are informative. The data are not consistent with* ***H2(a)****. Partial lies are common.*

On aggregate, managers respond to the information in their agents' messages. Table 5 shows the managers' average choices as a function of the messages sent by the two agents. Cells with five or fewer observations are left blank due to the small amount of data, and we delete the small number of observations (7/486) where the manager did not choose the same action for her two agents. When the two messages coincide, the manager follows the messages closely (but not perfectly). When the two messages differ, the manager's choices generally increase in each agent's message (holding the other's message fixed). The response of managers to messages is strong and statistically significant.[21]

Table 5: Manager Choices as a Function of Messages

		Agent 2 (μ_2)				
		1	2	3	4	5
Agent 1 (μ_1)	1	1.38	---	---	---	---
	2	1.57	2.25	---	---	---
	3	2.42	2.58	2.98	---	---
	4	2.72	3.00	3.67	3.71	4.00
	5	2.90	3.08	3.52	4.50	4.79

Given that agents send useful information and managers respond to their agents' messages, why is total surplus no better than in the babbling equilibrium? The problem is that managers often make choices that seem to be clear errors. For example, when the two agents' messages match ($\mu_1 = \mu_2$), they are almost certainly telling the truth (98%). Not surprisingly, it is an empirical best response to assign both agents the action that corresponds to their messages ($\alpha_1 = \alpha_2 = \mu_1 = \mu_2$), but managers fail to do so in 18% of these observations. Managers making this type of error earn an average payoff of 66.6 ECUs, compared with 79.6 ECUs for those who play the best response. Another common error occurs when A1 and A2 send diametrically opposed messages by choosing $\mu_1 = 5$ and $\mu_2 = 1$. Obviously at least one of the agents is lying. The safe outcome is the empirical best response to diametrically opposed messages, but only 35% of

[20] Self-serving lies are shaded away from the actual game towards the agents' preferred outcome. There are a small number of messages (3.7%) that are shaded in the direction of the *other* agent's preferred outcome.

[21] To establish statistical significance, we ran a regression where the dependent variable is the common action chosen by the manager for her two agents, and the independent variables are the two messages. The parameter estimates are .361 and .392 with standard errors, corrected for clustering at the group level, of .051 and .049.

managers follow this course of action. This type of error also reduces average total surplus (66.9 vs 63.3 ECUs). Making matters worse, managers do not learn to avoid these errors. The frequency of the first type of error falls a bit between the first and second halves of the experiment (19% vs. 16%), and the frequency of the second error type increases slightly from 59% to 70%.

One possible explanation for managerial errors is repeated game effects. Over time, agents have the opportunity to establish a reputation for truthfulness. Agents potentially might benefit from building reputations for being truthful and then abandoning them at an opportune moment, and choices that we identify as managerial mistakes might actually represent optimal strategic decisions where managers take advantage of what they have learned about agents' truthfulness to increase their payoffs. In line with this, managers' choices are sensitive to the previous truthfulness of the agents, responding more strongly to messages from agents who have a history of being truthful. However, there is little indication that agents strategically manipulate their reputations, and the data strongly suggests that the managerial decisions we have identified as errors do indeed reflect mistakes rather than strategic choices as part of a reputational equilibrium. See Appendix D for a detailed exploration of these issues.

Managers' errors explain why total surplus in **SC - MC** is no better than repeated play of the safe outcome. To see how well managers could do just by avoiding obvious errors, suppose they adopt the following simple rule: If the agents' messages agree, choose the action that matches their messages; otherwise play the safe outcome. This rule yields an average total surplus of 73.1 compared to 71.0 for the babbling equilibrium and 70.8 for the average total surplus actually achieved by managers. The efficiency gain from the simple rule is 25.0% compared to the 6.5% actually achieved, and it yields significantly higher total surplus than either repeated play of the safe outcome ($p < .001$) or the realized total surplus ($p = .002$). Managers could easily outdo the babbling equilibrium, but fail to effectively use the information transmitted by their agents.

Result 7: *Managers in **SC – MC** respond to agents' messages but make frequent errors using the information contained in the messages, causing their failure to beat repeated play of the safe outcome.*

There are three specific things to take away from the various structured communication treatments. First, in all three treatments there is room for improvement. Even in **SC – MC**, the one case where coordination is *not* a problem, little advantage is taken of agents' information. Second, managers are error prone. Whether giving poor advice, being excessively conservative, or failing to grasp obvious information from their agents' messages, managers consistently make mistakes that hold down total surplus. Finally, and most importantly, even though there is no incentive to reveal their information, agents frequently do so. The manner in which they do so would not surprise anyone familiar with the literature on lie aversion; some agents tell the truth, but lying is common including the frequent use of partial lies.

5.2.b: Free-form Chat: To evaluate the impact of specific message types in the three chat treatments, we developed a systematic scheme for coding message content. The goal was to quantify communication that might be relevant for the play of the game, avoiding prejudgments about which sorts of messages were important. We employed the methods developed by Cooper and Kagel (2005). After reading a random sample of conversations, we developed a coding scheme. Two research assistants then independently coded the content of all chat conversations. No effort was made to force agreement among coders. For several categories (marked with asterisks on Table 6), the initial two coders had a Cohen's kappa of less than .5, indicating relatively low agreement. These categories were recoded by a third coder who was given extensive training in an attempt to improve the quality of the coding. The research assistants were not informed about any hypotheses the co-authors had about the messages. They were told that their job was to simply capture what had been said without concern to the possible effects of what had been said. Coding was binary – a message line was coded as a 1 if it was deemed to contain the relevant category of content and 0 otherwise. We had no requirement on the number of codings for a message line – a coder could check as many or few categories as he or she deemed appropriate. A number of the categories also had sub-categories. For example, the coding scheme has a category for suggesting what actions should be chosen and sub-categories for specific suggestions (e.g. suggesting play of the efficient outcome). A coder was free to check whatever sub-categories they deemed appropriate when the corresponding category was checked off. Our analysis of the coding uses averages across coders unless otherwise noted.

Table 6: Frequency of Coding Categories

Coding Category	**CH/S – D**	**CH/A – D**	**CH – MC**
# Messages (Manager)	n/a	3.33	4.56
# Messages (Agent)	4.46	3.72	5.31
Any Suggestion	93.1%	73.3%	90.7%
Suggest Safe Outcome	54.1%	37.6%	60.6%
Suggest Efficient Outcome	48.4%	41.0%	57.7%
Agreement to Suggestion	78.9%	54.0%	67.9%
Discuss Need to Coordinate *	6.4%	3.5%	4.0%
Discuss Fairness *	31.8%	34.6%	43.9%
Discuss Efficiency	39.4%	16.0%	37.7%
Questions About Rules of the Experiment *	11.7%	8.8%	15.0%
Questions About How to Play *	10.9%	6.0%	14.2%
Explanation *	21.7%	39.3%	32.3%
Ask What Game Is Being Played (M)	n/a	14.9%	19.4%
Truthfully Reveal Game	n/a	28.8%	68.4%
Lie About Game	n/a	0.0%	3.4%
Contradict (One tells truth, other lies)	n/a	0.0%	2.5%

Table 6 reports the frequency of the coding categories, broken down by treatment. Some of the categories are not relevant in **CH/S – D** since the manager cannot send messages, and hence no figures are reported. "Contradict" is not a category per se, but instead is a combination of the preceding two categories that accounts for cases where one agent truthfully reported what game was being played and the other lied. Table E1 in Appendix E provides a fuller description of the categories. The unit of observation is the entire conversation prior to play in a single round rather than a single message line within that conversation or messages from only one individual in the conversation. So, for example, in 93.1% of the pre-play dialogues in **CH/S - D**, at least one agent suggested what actions should be chosen.

Before discussing the content of messages, the first two lines of Table 6 report the average number of messages sent per round, broken down by role. Managers send significantly more messages in **CH – MC** than **CH/A – D** (p = .016), and agents send significantly more messages in **CH – MC** than either **CH/S – D** (p < .061) or **CH/A – D** (p = .005).[22] Recall that total surplus has high variance in **CH/A – D**. Underlying this, managers' behavior also has high variance in **CH/A – D**. The three most *and* the three least talkative managers come from **CH/A – D**, and, more generally, the variance in the frequency of messages sent by managers is higher in **CH/A – D** than **CH – MC** (StDev = 3.17 vs. StDev = 1.99). Our prediction of relatively high performance in **CH/A – D** depended on leadership by managers, but a surprisingly large fraction of managers fails to provide *any* leadership.

Turning to message content, recall that **CH/S – D** significantly improves total surplus relative to **SC/S – D**, the parallel treatment with structured communication. Performance in **SC/S - D** is limited by failures to agree on what actions should be used as well as a tendency to not agree on efficient coordination. Agreements are more frequent in **CH/S – D** than **SC/S – D** (79% vs. 63%). Given that agents almost always coordinate their actions if an agreement is reached (95%), the higher agreement rate translates into improved coordination and, by extension, higher total surplus. **CH/S – D** does not solve the second problem that plagued **SC/S – D**. In cases where the safe and efficient outcomes do not coincide ($\gamma \neq 3$), only 34% of agreements in **CH/S – D** call for play of the efficient outcome. This differs little from the 32% figure for **SC/S – D**. When agents agree on efficient coordination, they usually follow through (96%), but **CH/S – D** does no better than the safe outcome because such agreements occur too rarely.

Result 8: *Total surplus is higher in* ***CH/S – D*** *than* ***SC/S – D*** *because agreements are much more common in* ***CH/S – D****. This promotes coordination, but does not improve the likelihood of efficient coordination.*

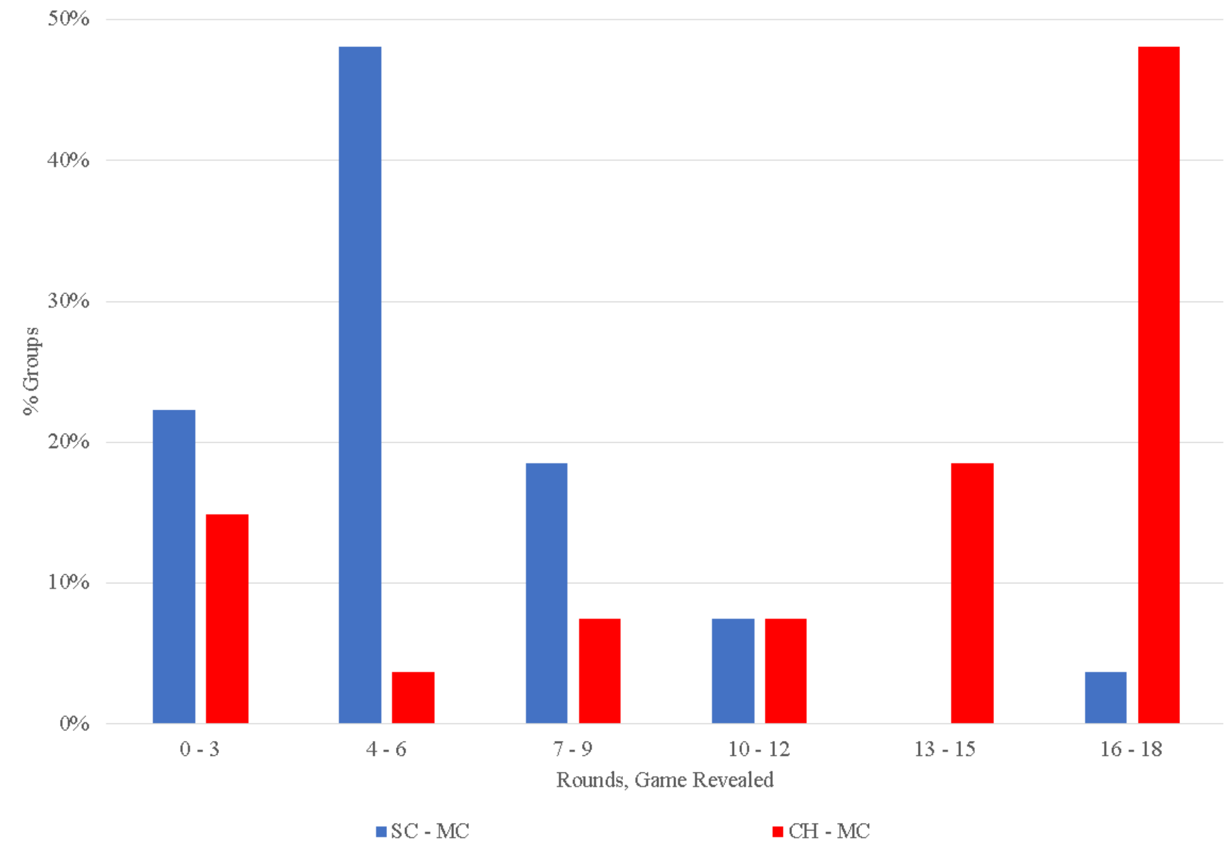
Total surplus is basically equal in **CH/A – D** and **CH/S – D**, but the factors driving performance differ. Subject to reaching an agreement when $\gamma \neq 3$, agreements on efficient coordination are more frequent in

[22] For agents, an observation for the statistical test is the average number of messages sent by the *pair* of agents in a group. The difference between **CH/S – D** and **CH/A – D** is not statistically significant (p = .308).

CH/A – D than **CH/S – D** (43% vs. 34%) and are usually followed (90%). The problem is that agreements of *any* kind are much less frequent in **CH/A – D** than **CH/S – D** (54% vs. 79%),[23] and failing to reach an agreement is associated with lower coordination rates (71% vs. 94%). This doesn't go away with experience, as the agreement rate *falls* slightly from 58% to 50% between the first and second halves of the experiment. The result is an odd combination of lower total surplus and more efficient coordination.

In **CH - MC**, coordination *per se* is trivial; the question is whether the manager can achieve *efficient* coordination given that she cannot observe the state of the world. A central finding of our work is that managers in **CH - MC** get remarkably good information about what game is being played, making efficient coordination possible. In most cases (68%), at least one agent truthfully reveals the game being played and only rarely (3%) does an agent lie about the game. Rather than falling, these figures improve slightly with experience from 65% and 5% in Rounds 1 – 9 to 72% and 2% in Rounds 10 – 18.

Figure 1: Distribution of Rounds that Game is Revealed



Information transmission is far cleaner in **CH – MC** than **SC – MC.** This can be seen in Figure 1. Define the game as being "revealed" to the manager in a round if at least one agent tells the truth without

[23] Consistent with the high variance of total surplus, there is more variance in groups' ability to reach agreements for **CH/A – D** than **CH/S – D**. Looking at the number of periods (out of 18) that a group reaches an agreement, the standard deviation is 4.53 for **CH/A – D** vs. 2.87 for **CH/S – D**.

contradiction from the other. An observation in Figure 1 is the number of rounds over the entire session where the game was revealed for a group. Figure 1 shows the distribution of how often agents in a group revealed the game. For example, in **SC – MC**, the game was revealed in 0 – 3 rounds for 22% of the groups. The distribution is shifted to the right for **CH – MC** relative to **SC – MC**. For 89% of groups in SC – MC, the game is revealed for *less* than half of the rounds. In **CH – MC**, the game is revealed for *more* than half of the rounds in 74% of the groups.

Managers get some useful information in **SC – MC** (see Section 5.2.a), but it often involves conflicting reports (69%) that are difficult to interpret. On top of this, managers in **SC – MC** often make mistakes extracting information from agents' messages. In **CH – MC**, managers receive some report about the game in 69% of observations. For 95% of these cases, they receive a truthful report without contradiction. Almost always, managers in **CH – MC** either have no information, and therefore do not face an information extraction problem, or have unambiguous information that makes information extraction trivial.

The high quality of information transmission in **CH – MC** is enormously important for efficiency. When the safe and efficient outcomes do not coincide ($\gamma \neq 3$), the frequency of efficient coordination rises from 18% when neither agent truthfully reveals the game to 52% if at least one tells the truth.[24] The frequency of efficient coordination changes little when one agent tells the truth and the other lies (53%), albeit based on very few observations. The truth wins in this environment.

Result 9: *Better transmission of the agents' information occurs in **CH – MC** than in **SC - MC**. This happens because agents frequently tell the truth, almost never lie, and rarely confront managers with conflicting reports. Truth-telling is strongly associated with efficient coordination.*

Accurate transmission also takes place in **CH/A – D**. Agents are far less likely to report what game is being played than in **CH – MC**, but always tell the truth when they do so. The lack of lies is less surprising for **CH/A - D** than **CH – MC**; there is little incentive to lie since the manager does not control what actions are chosen. Like **CH – MC**, accurate transmission promotes efficient coordination in **CH/A – D**. The frequency of efficient coordination is 51% when the game is truthfully reported (and $\gamma \neq 3$), compared with 33% otherwise.

The nature of truth-telling strongly differs between **CH - MC** and **SC - MC**. Most agents in **SC – MC** mix between telling the truth and lying; 69% both tell the truth in at least a third of the rounds and lie in at least a third of the rounds. There are only two agents that never lie and none that never tell the truth. Partial

[24] It may seem surprising that the rate of efficient coordination is not closer to zero when there is not a truthful report and $\gamma \neq 3$. In 87% of these cases, there is a suggestion that the efficient equilibrium should be played. These suggestions may serve as an indirect method of revealing the game, making a direct report unnecessary.

lying is common and agents are strategic about telling the truth, doing so more often when it is to their benefit to be believed. This can be seen in Table 7. As in Table 4, the games have been remapped for the A2 role so all observations are from the point of view of A1 (i.e. Outcome 1 is the worst outcome and Outcome 5 is the best). Agents are most likely to lie when the efficient outcome would be worst for them (γ = 1), and most likely to be truthful when it would be best for them (γ = 5).

Table 7: Frequency of Truth-Telling and Lying

Game (Remapped)	**SC – MC**		**CH – MC**	
	Truth	Lie	Truth	Lie
1	24.9%	75.1%	46.6%	2.5%
2	36.4%	63.6%	42.8%	3.1%
3	68.0%	32.0%	43.0%	2.2%
4	73.3%	26.7%	50.1%	1.9%
5	87.6%	12.4%	47.1%	0.3%

These patterns change in **CH – MC**. Mixing between truth-telling and lying is largely non-existent. There are 47 subjects in the agent role who send at least one message reporting what game is being played, averaging 9.8 reports over the course of 18 rounds. 35 of 47 reporting agents never lie and another 6 of 47 only lie once. None lie in more than 40% of their reports. Unlike **SC – MC**, there are *no* agents that both tell the truth in at least a third of the rounds and lie in at least a third of the rounds. Subjects mix, but it is almost entirely between telling the truth and not reporting. Subject to lying, partial lies are common (40% of lies), but in absolute terms partial lies are necessarily rare given the low overall rate of lying. Returning to Table 7, truth-telling is not sensitive to incentives.[25] Telling the truth is about as likely when it is most advantageous for an agent to lie (46.6% in Game 1) as when it is most advantageous to tell the truth (47.1% in Game 5).

Result 10: *The frequency of truth-telling, lying, and non-reports in* ***CH – MC*** *is not sensitive to what game is being played. Unlike* ***SC – MC****, the patterns of truth-telling in* ***CH – MC*** *do not parallel what is typically reported in the literature on lying.*

The different pattern of truth-telling in **CH – MC** suggests that the psychological mechanism underlying truth-telling is altered by the real-time, asynchronous communication available in this treatment. One possible reason for infrequent lying in **CH – MC** is that agents feel guiltier about lying to their manager

[25] Table 7 reports the frequency that an individual agent reports truthfully at some point during the pre-play communication. This differs from the figure reported under "Truthfully Reveal Game" in Table 6, which shows the frequency that at least one of the two agents reports truthfully at some point during the pre-play communication.

when they have been directly asked for a report. However, it is surprisingly rare for managers to request reports about what game is being played (19%), and the fraction of lies increases from 2% to 10% when a report is requested. Another possibility is that agents avoid lying because they are concerned about being "fact-checked." In both **SC – MC** and **CH – MC**, the manager knows *ex post* when an agent has lied, but in **CH – MC** it is possible for the other agent to call out a liar in real time. Indeed, in 40% of the observations where an agent lies, the other agent corrects them.[26] It may be more embarrassing to be actively called out as a liar than to merely be revealed as a liar.[27] We explore both of these potential explanations in the follow-up treatment reported in Section 6.

Failing to report what game is being played could be considered a "soft" lie. However, agents that don't report the game usually have little reason to do so. If one agent has truthfully revealed the game, there is little need for the other to reiterate this information. In line with this, the other agent has reported truthfully in 44% of the cases where an agent does not make a report. It is also pointless to report what game is being played if the safe outcome will be chosen regardless. Consider cases where the safe and efficient outcomes could be distinguished in the *previous* round ($\gamma \neq 3$). If the safe outcome was played in the previous round, *neither* agent makes a report for the current round in 44% of the observations; this makes sense if agents expect the manager to choose the safe outcome regardless of any new information. By contrast, if agents expect the efficient outcome to be played then they have an incentive to guide the manager's decision by reporting the current game. Indeed, when the efficient outcome was chosen in a previous round with $\gamma \neq 3$, *neither* agent reports for only 17% of the observations. Overall, 78% of non-reports occur in cases where either the other agent has told the truth or the safe outcome is used. Non-reports largely do not appear to be a form of deception.

5.2.c: The Effect of Chat Content: None of the preceding establishes a causal relationship between the content of pre-play communication and outcomes. Establishing causality is tricky because outcomes and the content of communication may both depend on lagged outcomes. Table 8 shows the results of probit regressions that control for lagged outcomes. Separate regressions are shown for each of the three treatments with chat. The dependent variable is either a dummy for coordination ($\alpha_1 = \alpha_2$) or efficient coordination ($\alpha_1 = \alpha_2 = G$). First round data are dropped to allow the use of lagged variables. There is no regression for coordination in **CH – MC** because there was 100% coordination following Round 1.

[26] This is different from the figure reported as "Contradict" in Table 6, which measures cases where one agent reported truthfully and the other lied. The 40% figure refers to "fact-checking" where one agent explicitly corrects a false report by the other (e.g. "It is Game 3." "No, it is really Game 2.").

[27] Fact-checking helps explain why the frequency of efficient coordination remains high when one agent tells the truth and the other lies, since fact-checking gives the manager guidance about which agent to believe.

As independent variables, all regressions include dummies for lagged outcomes (coordination failure, safe coordination, and efficient coordination with other coordination as the omitted category), game dummies, and a dummy for late rounds (Rounds 10 – 18). These are not reported to save space in the table. All regressions include the average coding for the categories reported in Table 6 with the following exceptions. The categories for "Lie About Game" and "Contradict" are highly collinear, so we only include the latter (we felt this was the more interesting of the two). There were no cases of contradicting reports in **CD/A – D**, so this variable is dropped. Including suggestions about what actions to play makes the regressions circular (subjects do what they say they should do), so these categories are omitted. We report marginal effects. Standard errors are corrected for clustering at the group level.

Table 8: Probit Regressions, Effects of Chat on Outcomes

Treatment	CH/S – D		CH/A – D		CH – MC
Dependent Variable	Coordination	Efficient Coordination	Coordination	Efficient Coordination	Efficient Coordination
Agreement	0.106*** (0.021)	0.217*** (0.060)	0.114*** (0.024)	0.184** (0.082)	0.051 (0.078)
Discuss Need to Coordinate	-0.031 (0.046)	0.231 (0.180)	0.032 (0.068)	0.080 (0.231)	-0.157 (0.199)
Discuss Fairness	0.010 (0.023)	-0.264*** (0.076)	-0.004 (0.024)	-0.088 (0.088)	-0.107 (0.105)
Discuss Efficiency	0.008 (0.026)	0.272*** (0.063)	0.050 (0.038)	0.262*** (0.098)	0.111 (0.090)
Questions About Rules	-0.053* (0.030)	-0.045 (0.148)	0.218*** (0.057)	0.112 (0.114)	-0.168 (0.104)
Questions About Play	0.036 (0.033)	-0.130 (0.148)	-0.061 (0.052)	0.159 (0.166)	-0.099 (0.149)
Explanation	-0.039 (0.031)	0.061 (0.125)	-0.076** (0.031)	-0.165* (0.086)	0.012 (0.098)
Ask What Game			-0.054** (0.024)	-0.121 (0.085)	0.127 (0.081)
Truthfully Reveal Game			0.022 (0.029)	0.202* (0.118)	0.294*** (0.094)
Contradict					-0.045 (0.234)

Notes: All models include 459 observations. Marginal effects are reported. Standard errors (in parentheses) are corrected for clustering at the group level. All regressions include controls for the game being played, a dummy for late rounds, and lagged outcomes. Coefficients for these variables are not reported to save space. Three (***), two (**), and one (*) star indicate significance at the 1%, 5%, and 10% levels using two-tailed tests.

We have stressed the importance of agreements for achieving coordination in **CH/S – D** and **CH/A – D**, and the regressions provide additional evidence of this. In both treatments, there is a strong positive relationship between reaching an agreement and either coordination or efficient coordination. Agreements play little role in **CH – MC**. The manager is a dictator in this treatment and does not need the agents to agree on a course of action. The regressions also support our observation that efficient coordination is likelier in both **CH/A – D** and **CH – MC** when at least one agent reports truthfully, with the effect being stronger in the latter case. Contradictions have little effect in **CH – MC**; as noted previously, the truth typically wins in this environment. It is interesting to note that discussing efficiency has a strong positive effect in **CH/S – D** and **CH/A – D**, but not in **CH – MC**. Once again, this illustrates the importance of control. Managers can impose efficient coordination in **CH – MC** without needing buy-in from their agents. Discussion of fairness plays an important role in **CH/S – D**, making the safe equilibrium more common, but plays surprisingly little role in the other two treatments. Perhaps it is difficult to argue persuasively that the safe outcome is fair when it harms the manager (and the manager has a voice).

Modifying the regressions in Table 8 lets us make an important point about the effectiveness of managers in **CH/A – D**. Studies of leadership in experimental economics typically focus on the *average* effect of having a leader, concluding that leadership does not matter if the average effect is zero. This ignores heterogeneity. The average effect of giving the manager an active role in **CH/A – D** is basically zero, but no group gets the average manager. Each group gets a specific individual who may be a better or worse leader than average. Outcomes in **CH/A – D** are highly variable, suggesting that some managers are better than others. Indeed, the data indicates that how the manager communicates in **CH/A – D** affects outcomes, and good managers are better communicators than their less successful peers.

To reach this conclusion, we first run a probit regression analogous to those reported in Table 8. This regression analyzes the effect of communication *by managers* in **CH/A – D** on the likelihood of efficient coordination. The independent variables of interest are the average coding of messages sent by the manager. The regressions control for a number of factors: lagged outcomes, the game being played, and a dummy for late rounds. We find that efficient coordination is less (more) likely if the manager suggests the safe (efficient) equilibrium. These effects are large (est. = -0.408 for suggesting the safe equilibrium, and est = 0.370 for suggesting the efficient equilibrium), and both effects are easily significant at the 1% level ($p = .004$ and $p < .001$ respectively). This is not a case where the result is circular, as the manager does not directly control the outcome in **CH/A – D**; the only thing a manager can do in this treatment is give advice.

The preceding results establish that the manager's messages affect outcomes even in the absence of direct control. This suggests that successful managers send better messages than other managers, which is indeed the case. To see this, divide managers into thirds by the average payoffs they achieve and label the top third as "good" managers. To create an index of suggestion quality, a manager gets a score of +1 if they suggest the efficient equilibrium only, -1 if they suggest the safe equilibrium only, and 0 if they suggest either both or neither. We then regress the suggestion quality on whether the group had a good manager, with controls for the game being played and a late period dummy. The effect of having a good manager is large (est. = 0.179), and statistically significant ($p = 0.049$). To summarize, good managers are more likely to make good suggestions, and groups that receive good suggestions from their manager are more likely to achieve efficient coordination. Two points can be taken from this: (1) The varying group outcomes in **CH/A – D** are not a matter of pure chance, but instead reflect differing performance by their managers. (2) The *average* effect of letting a manager give advice in **CH/A – D** is small, but this is *not* because advice given by managers does not affect outcomes. Rather, on aggregate, the positive effect of good managers is offset by the negative effect of bad managers.

6. Extended Communication: Performance in **CH – MC** is unambiguously better than in **SC – MC**. Underlying this, information transmission from agents to managers is unambiguously better in **CH – MC**, driven by a striking reduction in lying by agents. This raises an obvious question: what feature(s) of **CH – MC** leads to the low frequency of lies?

The free-form communication process in **CH – MC** allows for a number of possibilities that are not available in **SC – MC**: (1) Agents can be fact-checked in real time. (2) Manager**s** can request reports about the game as well as requesting that agents tell the truth. (3) Agents can tell managers that they reported truthfully. (4) Because communication is asynchronous, an agent can view the other agent's report before making their own and alter their report after seeing the other agent's report. (5) Agents have the option of not making a report about what game is being played. (6) The free-form nature of chat *per se* might affect behavior. Subjects must generate message content endogenously, can use any message rather than the limited set available with structured communication, and can frame messages in ways that subtly change their meaning from what is expressed by the pre-specified messages.

There is ample evidence from other papers that the sixth item listed above, free-form communication *per se*, changes behavior, generally leading to more prosocial outcomes.[28] However, *a priori*, the first five differences between **CH – MC** and **SC – MC** could also lead to less lying and can be implemented within structured communication. **CH – MC** does not provide conclusive evidence that any of these differences caused reduced lying. We therefore developed a follow-up treatment to explore whether any of the five differences other than free-form communication reduce lying. Like **SC – MC**, subjects only had a limited number of messages available. Unlike **SC – MC**, structured communication in the follow-up treatment captures features (1) to (5) listed above: fact-checking of lies was possible, managers could ask for (truthful) reports about the game being played, agents could assert that they told the truth, reports were asynchronous, and reports were not mandatory. We included all five of these differences, regardless of whether the **CH – MC** data suggested much impact, to give structured communication the best possible chance to reduce lying. If there is little impact with all of these differences present, it strongly suggests that free-form

[28] See Brandts et al. (2019) for a summary of the evidence.

communication per se is a necessary ingredient for reduced lying and improved information transmission in **CH – MC**.

Since the ability of subjects to communicate is extended relative to the **SC – MC** treatment, we refer to the new treatment as the *Extended Communication – Managerial Control* (**EC – MC**) treatment. Communication worked as follows. After the two agents observed which game was being played, there was a 45 second period in which all three players could communicate. This was done by pressing buttons containing prespecified messages. These messages could be sent as many times as a player desired. Once a message was sent, it was displayed in a message window on the screens of all three players along with the identity of the sender. The manager could send two possible messages: "What game are we playing?" and "Please tell me the truth." The agents could send four possible messages: "It is Game [agent entered a game number]", "I am telling the truth", "The other agent is telling the truth", and "The other agent is lying".

The **EC – MC** treatment allows for the first five aforementioned mechanisms. If one or more of these mechanism(s) leads to reduced lying and, by extension, improved information transmission, we should expect to see similar results in **EC – MC** as in **CH – MC**. However, if free-form communication itself is the critical ingredient, little difference should be observed between **SC – MC** and **EC – MC**.

Three sessions of **EC – MC** were conducted at LINEEX with the same subject pool and basic procedures as the main treatments. Each session contained 9 groups, giving 27 independent observations, like the other treatments.

Table 9: Comparison of MC Treatments, Rounds 10 – 18

	SC-MC	**EC-MC**	**CH-MC**
Total Surplus	71.7	71.3	75.2
Efficiency Gain	6.5%	1.7%	44.3%
% Coordinate	99.6%	96.7%	100.0%
% Efficient s.t. Coordinate	45.9%	55.7%	61.7%
Game Revealed (Group)	30.9%	49.8%	77.4%
Truth-Telling (Individual)	58.4%	63.6%	48.3%
Lie (Individual)	41.6%	28.6%	1.0%

Table 9 reports the main results for the **EC – MC** treatment, and compares them with results from the other two treatments with managerial control. Total surplus and efficiency gains are slightly lower for **EC**

– **MC** than **SC – MC**, rather than higher as expected, although the difference in total surplus is not statistically significant (p = .828). Like **SC – MC**, total surplus in **EC – MC** is not significantly better than the babbling equilibrium (p = .470) and is significantly lower than in **CH – MC** (p = .029). In terms of total surplus or efficiency, there is no evidence that the extra communication options afforded by **EC – MC** affects outcomes relative to **SC – MC**.

Messages are used frequently in **EC – MC**. Subjects almost always send at least one message in a round (95.2% in Rounds 10 – 18), and send an average of 4.5 messages per round.[29] Looking at the third line from the bottom of Table 9, **EC – MC** increases the probability that the game is revealed (at least one agent tells the truth without contradiction from the other) relative to **SC – MC**, albeit less than half as much as **CH – MC** (18.9% vs 46.5%). The increase over **SC – MC** is significant (p = .024), although the probability of revealing the game is still significantly lower than in **CH – MC** (p = .002). These differences in information transmission reflect decreased lying rather than increased truth-telling. The likelihood that an agent reveals the truth is roughly the same for Rounds 10 – 18 across the three treatments, but the likelihood of lying decreases as richer communication becomes possible (41.6% for **SC – MC**; 28.6% for **EC – MC**; 1.0% for **CH – MC**). When the true game is *not* revealed in **SC – MC** and **EC – MC**, it is usually because one agent tells the truth but is contradicted by the other agent telling a lie (79.8% for **SC – MC**; 83.6% for **EC – MC**). This rarely happens in **CH – MC** (9.1%), because agents who do not tell the truth typically do not report what game is being played rather than lying.[30]

Given that information transmission is better in **EC – MC** than **SC – MC**, why is the total surplus no higher? There is, as one would expect, a strong positive relationship in Rounds 10 – 18 between total surplus in **EC – MC** and whether the game is revealed by the agents or not (74.1 vs. 68.5). However, this relationship is stronger in **SC – MC** (78.4 vs. 68.7). To understand why, define a mistake by the manager as either a failure to coordinate or coordinating at an action that does not lie at or between efficient coordination and the safe equilibrium.[31] Subject to the game being revealed, managers make a mistake in 14.9% of observations for **EC – MC** vs. 1.3% for **SC – MC**. Thus, the benefits of better information are largely balanced out by more frequent managerial mistakes.[32]

[29] Agents sent almost twice as many messages per round as managers (5.9 vs. 3.0).

[30] For 89.1% cases where the game is *not* revealed in **CH – MC** for Rounds 10 – 18, neither agent reported what game was being played.

[31] For example, in Game 1, either not coordinating or coordinating at 4 or 5 are considered mistakes by the manager.

[32] **CH – MC** lies somewhere between with a mistake rate of 9.0% when the game is revealed. If managers *never* made mistakes when the game was revealed, instead playing the safe equilibrium, average total surplus across Rounds 10 – 18 would have been 71.7 in **SC – MC**, 73.0 in **EC – MC**, and 76.0 for **CH – MC**. Without mistakes, **EC – MC** would do a bit better than **SC – MC**, but not as well as **CH – MC**.

We previously outlined six differences between **CH – MC** and **SC – MC**, with five of these also being operational in **EC – MC**. Are any of these differences responsible for reduced lying in **EC – MC** relative to **SC – MC**? The first difference is the possibility of fact-checking. Fact-checking refers to cases where an agent who lies about what game is being played is accused of lying by the other agent. As in **CH – MC**, this is common in **EC – MC**. In Rounds 10 – 18, if an agent lies in their initial report, they are fact-checked 79.4% of the time. The problem is that being fact-checked has no immediate effect on truth-telling; no agent who is fact-checked responds by changing their report to the truth. Instead, it is very common for a fact-checked agent to double down by making additional reports that are also lies (90.7%).[33] Fact-checking does have a small delayed effect. If an individual lies and gets fact-checked, they are somewhat less likely to lie in the *next* game than if they are not called on their lie (37.1% vs. 45.2%). However, based on a probit controlling for the game being played and period, this effect is not statistically significant (p = .150). We previously speculated that fear of fact-checking led to reduced lying in **CH – MC**, but the **EC – MC** data provides little evidence in favor of this mechanism.

The second difference is the possibility that managers can request truthful reports. This difference explains much of the improved transmission of information in **EC – MC** relative to **SC – MC**. Consider whether the manager either asks for a report or asks agents to tell the truth *prior to any agent reporting what game is being played*.[34] Both types of messages are common in Rounds 10 – 18; managers ask for a report in 80.2% of rounds and ask for the truth in 32.9% of rounds. Asking for a report has no impact on the likelihood that the game being played is revealed (50.0% vs. 49.7%), but asking for the truth does have a substantial effect (44.2% vs. 61.3%). The latter effect could be biased downwards if requests for truth-telling are more likely when agents have lied in past periods. To control for this possibility, we run a probit regression where the dependent variable is whether the game was revealed truthfully. Independent variables are dummies for whether the manager had requested a report, requested the truth, and the lagged dependent variable. The regression also controls for the round and game being played. Standard errors are clustered at the group level. The estimated marginal effect of requesting the truth is large (est. = .188) and significant (p = .044).

Two things are worth noting at this point. First, the theory of guilt aversion (Charness and Dufwenberg, 2006) suggests a mechanism by which requests for truthful reports will reduce lying. Under guilt aversion, an agent's willingness to lie is sensitive to their (second-order) beliefs about whether the manager believes they will tell the truth. If requesting the truth implies that agents believe the manager believes the truth will

[33] It was generally very rare for agents to change an initial report. For Rounds 10 – 18, this only occurred for 0.9% of observations where an initial report was made.

[34] As documented below, agents rarely change their reports after an initial report. We therefore focus on what is said before reports are made, as requests after reports are made can have little effect.

be told, agents will feel guiltier about lying. If agents are guilt averse, requesting the truth will reduce lying as is observed in **EC – MC**. Second, the preceding cannot be the mechanism behind reduced lying in **CH – MC** because there are virtually no requests for truthful reporting. Our coding scheme for **CH – MC** did not even include a category for requesting the truth because it is so rare. To the extent that **EC – MC** reduces lying, it appears to do so via a different mechanism than **CH – MC**.

A third difference between **EC – MC** and **SC – MC** is that agents can tell managers that they have reported truthfully. This is a common message type; 61.8% of agents who send a report also send a claim that they are being truthful. The data gives no evidence that claiming to be truthful is associated with a reduced likelihood of lying. Instead, agents who claim to have told the truth in Rounds 10 – 18 are slightly *more* likely to lie (72.5% vs. 66.8%), subject to making a report about what game is being played.

The fourth difference between **EC – MC** and **SC – MC** is that the asynchronous nature of reporting makes it possible for an agent to view the other agent's report before reporting, and possibly alter their report. There appears to be an order effect in the data; limiting the data to agents who make a report, agents who report first are more likely to tell the truth than laggards (75.7% vs 61.7%) in Rounds 10 – 18.[35] However, there is a strong relationship between what game is being played and when agents report; an agent is almost twice as likely to report first when it is their most preferred game (67.8%) vs. their least preferred game (39.1%). In other words, they are more likely to report quickly in situations where they want to tell the truth. Running a probit that controls for the game being played and period, the effect of reporting first on truth-telling is not statistically significant (p = .395). Another possibility is that laggards respond to whether the first report is true. This appears weakly true in the raw data, but in an odd fashion; laggards are more likely to lie if the first agent to report was truthful rather than a lie (33.3% vs. 27.8%), but also more likely to not make any report (19.8% vs. 5.6%). This again is biased by the strong relationship between truth-telling and what game is being played. The results look more sensible in a probit where the dependent variable is whether a laggard lies and controls are included for the game being played and the period. A truthful first report significantly reduces the probability the laggard lies – the estimated marginal effect is .250 (p = .045). Digging further, we ran a multinomial logit with lying as the base category. A truthful first report makes laggards significantly more likely to not report (p = .052) and weakly (and not significantly) more likely to tell the truth (p = .120). In other words, laggards lie less when the first report was true, but this consists more of not reporting rather than telling the truth.

A final difference between **EC – MC** and **SC – MC** is that agents are required to make a report on what game was being played in **SC – MC**, but not in **EC – MC**. Agents usually still report what game is being

[35] There are 13 cases where the two agents were recorded as first reporting simultaneously (time is reported in second increments). In these cases, we include *both* agents as reporting first. If we limit the data to cases where one agent is unambiguously first, the percentage of truthful first reports increases slightly to 77.3%.

played (92% in Rounds 10 – 18). It is not random when non-reports occur. Almost always when one agent reports and the other does not, the agent that reports tells the truth (97%). In 66% of these cases, the agent who does not report sends a message saying that the other agent has told the truth. In other words, many subjects who do not explicitly report tell the truth implicitly. Non-reporting does not seem to matter much, except to the extent that Table 9 slightly under-reports the extent of truth-telling by individuals.[36]

To recapitulate, we identified six differences between **CH – MC** and **SC – MC**, any of which might explain why lying is less frequent and information transmission is better in **CH – MC**. **EC – MC** allows for five of these differences, but generates less than half of the effect of **CH – MC**, relative to **SC – MC**, on the likelihood that an agent lies or that the game is revealed. One factor that explains decreased lying in **EC – MC**, requests for truthful reporting, cannot explain decreased lying in **CH – MC**. Asynchronous communication explains some of the effect and could explain some of the decreased lying in **CH - MC**, but not enough to account for the near absence of lies is **CH – MC** (even when the first report is truthful, the rate of lying by laggards is well above the near absence of lies observed in **CH – MC**). These observations strongly suggest that the sixth difference plays a central role in the dramatic effect of **CH – MC** on truth-telling: free-form communication appears to be a necessary condition for the sharp decrease in lies observed in **CH – MC**.[37] In the conclusion we discuss why this might be the case.

Result 11: *The game is significantly more likely to be revealed in **EC – MC** than **SC – MC**, but significantly less likely than in **CH – MC**. The positive effect of **EC – MC** stems at least in part from requests by managers for truthful reports and asynchronous reporting. Given that less than half of the effects of **CH – MC** on information transmission and lying can be accounted for by the five factors incorporated into **EC – MC**, we conclude that free-form chat per se is a necessary ingredient for the dramatic decrease of lying in **CH – MC** relative to either **SC – MC** or **EC – MC**.*

7. Concluding Remarks: This paper studies coordination in a demanding experimental environment, the MA game. It combines four properties that characterize many organizational settings: coordinating on a common course of action benefits everyone, agents have divergent preferences over possible outcomes, managers lack the necessary information to simply impose efficient coordination on their agents, and agents have the necessary information but also have little reason to truthfully reveal it. Unlike the frequently studied weak-link game, the MA game stresses asymmetries; the manager doesn't know what game is being played, and the agents' interests are misaligned. Achieving coordination in the MA game is not difficult,

[36] The same caveat applies in **CH – MC**.

[37] This evidence is related to results reported by Charness and Dufwenberg (2006 and 2010). They find that promises made within free-form communication affect behavior differently from pre-formulated promises.

but achieving efficient coordination that uses agents' information is a challenge. It is well-established that either communication among players or external leadership (like managerial advice) increases efficient coordination in symmetric coordination games. Managerial control has not been previously studied as its likely effect in symmetric coordination games, unlike the MA game, is obvious. Our primary goal is to study the roles of communication and managerial control in achieving efficient coordination in the difficult environment of the MA game.

Achieving efficient coordination in the MA game requires two things: (1) The choices of the agents have to be coordinated and (2) the agents' information must be incorporated into the choice of action. Either free-form communication (chat) or managerial control are sufficient in isolation to solve the coordination problem, but a combination of chat *and* managerial control is necessary to use the agents' information sufficiently well to outperform the babbling equilibrium, gaining almost half of the possible gains over the babbling equilibrium. Free-form chat can seem like a magic bullet in experimental economics, solving all problems with coordination and/or cooperation. In the MA game, neither rich communication nor managerial advice is sufficient. Even though managers lack critical information, managerial control plays a valuable role in enabling groups to make effective use of agents' information.

The key feature that allows the combination of chat and managerial control to function so well is that information transmission is remarkably good. The MA game with managerial control gives agents strong incentives to lie, and with structured communication, agents often do so. Managers receive only limited information and struggle to use it effectively. When free-form communication is used, agents generally reveal what game is being played and rarely lie. Managers take advantage of their resulting good information to frequently impose efficient coordination.

The patterns of communication about the state of the world (i.e. what game is being played) in **CH – MC** are quite different from either what is observed in **SC – MC** or what is typically observed in experiments on truth-telling. That raises the question of why lying is so infrequent in **CH – MC**. Our initial conjecture was that a fear of being fact-checked (being called out on a lie in real time by the other agent) drives the low rate of lying. Indeed, fact-checking is common when lies are told in **CH – MC**. However, the results of the follow-up **EC – MC** argue against this explanation. Fact-checking is once again common in **EC – MC**, but there is little evidence that this leads to reduced lying. Apparently, the possibility of being called out on a lie, either in real-time or with a delay, is not sufficiently embarrassing to deter lying. Instead, the critical ingredient that leads to reduced lying appears to be free-form communication *per se*.[38]

[38] The use of free-form communication need not always lead to the almost complete elimination of lies we observe. For example, Lundquist et al. observe fewer deceptive lies with free-form messages, but a substantial fraction of lies still occurs (40%). This may reflect the differing structure of communication, as messages were in Lundquist et al. were restricted to a single one-way message as opposed to free-form bilateral (or trilateral) communication.

Our experiments cannot tell us why free-form communication *per se* decreases lying. One possibility is strictly mechanical; **EC – MC** omits some specific type of message that is necessary to prevent lying. If so, it is not obvious what such a message might be. The coding exercise reported in Section 5.2.b was intended to capture all content relevant to play of the game. The obvious candidates that might increase truth-telling are available in **EC – MC**. It is possible that there is some other type of message in the **CH – MC** chat, possibly not coded, that decreases lying. A second possibility is that free-form communication decreases lying via a mechanism that only functions if it emerges endogenously. Once again, it is not obvious what such a mechanism might be. Both fact-checking and asynchronous timing are endogenous in **EC – MC**, and managers must endogenously choose whether or not to ask for a report or request the truth. That leaves us with two possibilities that we view as more likely. Language can be subtle and nuanced, with a gap between what is literally said and what is meant. For example, in the American South, the phrase "bless your heart" often implies the opposite of what is being said.[39] Subtleties and nuances that induce greater truth-telling with free-form chat may be missing even with the relatively flexible structured communication of **EC – MC**.[40] Another possibility is that free-form communication changes preferences by reducing social distance between the managers and their agents. The older psychology literature on free-form communication and cooperation finds that free-form communication increases cooperation by promoting group identity (e.g. Orbell, Van de Kragt, and Dawes, 1988). Similar effects could account for reduced lying in **CH – MC**, with greater group identity increasing concerns about maintaining a positive social image by being truthful.[41] All of the preceding is obviously speculative. We establish that free-form chat is necessary to reduce lying in **CH – MC**, but exactly how that works is an open question that will require additional research to answer.

Free-form communication improves efficiency, but it is still necessary that players come up with the right thing to say. This point is made strongly by the **CH/A – D** treatment. Letting the manager give advice to the agents has a minimal effect *on average*, but this disguises a great deal of heterogeneity. Managers can successfully induce more efficient coordination, but this only works when they actually think to suggest efficient coordination to their agents.

---

[39] If said in a certain way, "He means well, bless his heart" is a gentle way of saying he is an idiot.

[40] Along similar lines, Charness and Dufwenberg (2010) find that "bare" promises do not have the same effect as promises expressed in the context of a free-form message. In both cases it is possible to promise trust-worthy behavior, but something about the richer language available with free-form messages makes subjects more likely to follow through on their promises.

[41] Cohn, Gesche, and Maréchal (2022) report that truthfulness is significantly higher when individuals interact with a human rather that a machine. They posit that an increased sense of "closeness" may be responsible for the result, increasing concerns about being judged.

We study an intentionally simple game designed to capture a set of features that are present in many organizations. A natural goal for follow-up work is abandoning some of that simplicity in exchange for greater verisimilitude. One possible approach is using subjects with real-world managerial experience as subjects in the manager role. Existing evidence suggests that using managers would *not* affect our results (for coordination games with leaders, see Cooper, 2006; for games in general, see Fréchette, 2015), but it would still be interesting to see how real-world managers approach the MA game. Another possibility is looking at decision making by groups. Many decisions within organizations are made by groups, and there is an extensive literature suggesting that groups and individuals do not make identical decisions either for games generally or coordination games specifically (Feri, Irlenbusch, and Sutter, 2010).

REFERENCES

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2019), "Preferences for Truth-Telling," *Econometrica*, 87(4), 1115-1153.

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

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2008b), "Centralization versus Decentralization: An Application to Price Setting by a Multi-Market Firm," *Journal of the European Economic Association*, 6 (2-3), 457-467.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2015), "Organizing to Adapt and Compete," *American Economic Journal: Microeconomics*, 7(2), 158-187.

Bartling, Björn, Ernst Fehr and Holger Herz (2014), "The Intrinsic Value of Decision Rights," *Econometrica*, 82(6), 2005-2039.

Battaglini, Marco (2002), "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70(4), 1379–1401.

Ben-Ner, Avner, and Louis Putterman (2009), "Trust, communication and contracts: An experiment," *Journal of Economic Behavior & Organization*, 70(1-2), 106-121.

Blume, Andreas and Andreas Ortmann (2007), "The Effects of Costless Pre-play Communication: Experimental Evidence from Games with Pareto-ranked Equilibria," *Journal of Economic Theory*, 132, 274–290.

Blume, Andreas, Peter H. Kriss, and Roberto A. Weber (2017), "Pre-play communication with forgone costly messages: experimental evidence on forward induction," *Experimental Economics*, 20(2), 368-395.

Blume, Andreas, Ernest K. Lai, and Wooyoung Lim (2023), "Mediated Talk: An Experiment," *Journal of Economic Theory*, 208 105593.

Bolton, Gary, Jordi Brandts, and Axel Ockenfels (2005), "Fair Procedures: Evidence from Games Involving Lotteries," *Economic Journal*, 115, 1054-1076.

Brandts, Jordi, Gary Charness, and Matthew Ellman (2016), "Let's Talk: How Communication Affects Contract Design," *Journal of the European Economic Association*, 14(4), 943-974.

Brandts, Jordi and David J. Cooper (2007), "It's What You Say Not What You Pay: An Experimental Study of Manager-employee Relationships in Overcoming Coordination Failure," *Journal of the European Economic Association*, 5(6), 1223-1268.

Brandts, Jordi, David J. Cooper and Christina Rott (2019), "Communication in Laboratory Experiments", in *Handbook of Research Methods and Applications in Experimental Economics*, edited by Arthur Schram and Aljaz Ule,, Edward Elgar Publishing.

Brandts, Jordi, David J. Cooper, and Roberto A. Weber (2015), "Legitimacy, Communication, and Leadership in the Turnaround Game," *Management Science*, 61(11), 2627-2645.

Brandts, Jordi and Bentley MacLeod (1995), "Equilibrium Selection in Experimental Games with Recommended Play." *Games and Economic Behavior*, 11(1), 36-63.

Brynjolfsson, Erik, Chris F. Kemerer (1996), "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market," *Management Science* 42(12), 1627-1647.

Cai, Hongbin, and Joseph Tao-Yi Wang (2006), "Overcommunication in strategic information transmission games," *Games and Economic Behavior*, 56(1), 7-36.

Cartwright, Edward, Joris Gillet and Mark van Vugt (2013), "Leadership by Example in the Weak-Link Game," *Economic Inquiry*, 51(4), 2028-2043.

Cason, Timothy N., and Vai-Lam Mui (2007), "Communication and coordination in the laboratory collective resistance game," *Experimental Economics*, 10(3), 251-267.

Charness, Gary and Martin Dufwenberg (2006), "Promises and Partnership," *Econometrica*, 74, 1579-1601.

Charness, Gary, and Martin Dufwenberg (2010), "Bare promises: An experiment," *Economics Letters*, 107(2), 281-283.

Chaudhuri, Ananish and Tirnud Paichayontvijit (2010), "Recommended play and performance bonuses in the minimum effort coordination game," *Experimental Economics*, 13, 346-363.

Chaudhuri, Ananish, Andrew Schotter and Barry Sopher (2009), "Talking Ourselves to Efficiency: Coordination in Inter-generational Minimum Effort Games with Private, Almost Common, and Common Knowledge of Advice," *The Economic Journal*, 119, 91-122.

Cohn, Alain and Michel Maréchal (2022), "Honesty in the Digital Age," *Management Science*, 68(2), pp. 827-845.

Cooper, David J. (2006), "Are Experienced Managers Experts at Overcoming Coordination Failure?" *Advances in Economic Analysis & Policy*, 6(2), Article 6.

Cooper, David J. and John R. Hamman (2020)"Leadership and Delegation of Authority," in Marie-Claire Villeval and Klaus Zimmermann (eds.), The Handbook of Labor, Human Resources and Population Economics, forthcoming, Berline, Springer Verlag.

Cooper, David J., John R. Hamman and Roberto Weber (2020), "Fool Me Once: An Experiment on Credibility and Leadership," *The Economic Journal*, 130, 631, pp. 2105 - 2133.

Cooper, David J., and John H. Kagel (2005), "Are two heads better than one? Team versus individual play in signaling games," *American Economic Review*, 95(3), 477-509.

Cooper, David and Kai-Uwe Kühn (2014), "Communication, Renegotiation, and the Scope for Collusion," *American Economic Journal: Microeconomics*, 6, 247-278.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1989), "Communication in the Battle of the Sexes Game: Some Experimental Results," *RAND Journal of Economics*, 20, 568-587.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1992), "Communication in Coordination Games," *The Quarterly Journal of Economics*, 107(2), 739-771.

Dawes, Robyn M., Jeanne McTavish, and Harriet Shaklee (1977), "Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation," *Journal of Personality and Social Psychology*, 35(1), 1.

Dessein, Wouter, Luis Garicano and Robert Gertner (2010), "Organizing for Synergies," *American Economic Journal: Microeconomics*, 2, 77-114.

Dryzek, John and Christian List (2003), "Social Choice Theory and Deliberative Democracy: A Reconciliation," *British Journal of Political Science*, 33(1), 1-28.

Erat, Sanjiv and Uri Gneezy (2012), "White Lies," *Management Science*, 58(4) 723-733.

Evdokimov, Piotr and Umberto Garfagnini (2019), "Communication and Behavior in Organizations: An Experiment," *Quantitative Economics*, 10(2), 775-801.

Farrell, Joseph and Carl Shapiro (1992), "Standard Setting in High Definition Television," *Brookings Papers on Economic Activity: Microeconomics*, 1992.

Fehr, Ernst, Holger Herz and Tom Wilkening (2013), "The Lure of Authority: Motivation and Incentive Effects of Power," *American Economic Review*, 103(4), 1325-59.

Feri, Francesco, Bernd Irlenbusch, and Matthias Sutter (2010), "Efficiency Gains from Team-based Coordination – Large-scale Experimental Evidence," *American Economic Review,* 100(4), 1892-1912.

Ferreira, Joao V., Nobuyuki Hanaki and Benoît Tarroux (2020), On the Roots of the Intrinsic Value of Decision Right: Experimental Evidence," *Games and Economic Behavior*, 119, 110-122.

Fischbacher, Urs (2007), "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2): 171-178.

Fischbacher, Urs and Franziska Föllmi-Heusi (2013), "Lies in Disguise – An Experimental Study on Cheating," *Journal of the European Economic Association*, 11(3):525–547.

Fréchette, Guillaume R. (2015), "Laboratory Experiments: Professionals versus Students", in G. Fréchette and A. Schotter (eds.), *Handbook of Experimental Economic Methodology,* Oxford: Oxford University Press.

Gandal, Neil, David Salant and Leonard Waverman (2003), "Standards in Wireless Telephone Networks," *Telecommunications Policy* 27, 325-332.

Gneezy, Uri (2005), "Deception: The Role of Consequences," *American Economic Review*, 95(1), 384-394.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel (2018), "Lying Aversion and the Size of the Lie," *American Economic Review*, 108(2), 419-453.

Hamman, John and Miguel A. Martínez-Carrasco (2020), "Delegation and Team Selection in Organizations: An Experimental Study," Working Paper.

Hart, Oliver and John Moore (2008), "Contracts as Reference Points," *Quarterly Journal of Economics*, 123, 1-48.

Hart, Oliver and Bengt Holmstrom (2010), "A Theory of Firm Scope," *Quarterly Journal of Economics*, *125*(2), 483-513.

Isaac, R. Mark, and James M. Walker (1988), "Communication and free-riding behavior: The voluntary

Karpowitz, Christopher and Tali Mendelberg (2011), "An Experimental Approach to Citizen Deliberation," in Druckman, J. N., D. P. Green, J. H. Kuklinski, and A. Lupia (eds), *Cambridge Handbook of Experimental Political Science*, Cambridge: Cambridge University Press.

Kartik, Navin (2009), "Strategic Communication with Lying Costs," *Review of Economic Studies,* 76, 1359-1395.

Kriss, Peter H., Andreas Blume, and Roberto A. Weber (2016), "Coordination with decentralized costly communication," *Journal of Economic Behavior & Organization*, 130, 225-241.

Lai, Ernest K, Wooyoung Lim and Joseph Tao-yi Wang (2015), "An experimental analysis of multidimensional cheap talk," *Games and Economic Behavior*, 91, 114-144.

Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson (2009), "The aversion to lying," *Journal of Economic Behavior & Organization,* 70(1-2), 81-92.

McElheran, Kristina (2014), "Delegation in Multi-Establishment Firms: The Organizational Structure of I.T. Purchasing Authority," *Journal of Economics and Management Strategy*, 23(2), 225-257.

Mookherjee, Dilip (2006), "Decentralization, Hierarchies and Incentives: A Mechanism Design Perspective," *Journal of Economic Literature*, 44(2), 367-390.

Myers, Daniel and Tali Mendelberg (2013), Oxford Handbook of Political Psychology. Oxford University Press, Ch. Political Deliberation, 699-734.

Neri, Claudia and Rommeswinkel, Hendrik (2017) Decision Rights: Freedom, Power, and Interference. Available at SSRN: https://ssrn.com/abstract=2485107 or http://dx.doi.org/10.2139/ssrn.2485107.

Orbell, John, Alphons J. C. van de Kragt, Robyn M. Dawes (1988), "Explaining Discussion-Induced Cooperation," *Journal of Personality and Social Psychology*, 54(5), 811-819.

Osborne, Martin and Ariel Rubinstein (1994), *A Course in Game Theory*, Cambridge: MIT Press.

Ostrom, Elinor, James Walker, and Roy Gardner (1992), "Covenants with and without a sword: Self-governance is possible," *American Political Science Review*, 86(2), 404-417.

Pikulina, Elena S. and Chloe Tergiman (2020), "Preferences for Power," *Journal of Public Economics*, 185, article 104173.

Rantakari, Heikki (2008), "Governing Adaptation," *The Review of Economic Studies*, 75, 1257-1285.

Sahin, Selhan Garip, Catherine Eckel and Mana Komai (2015), "An experimental study of leadership institutions in collective action games," *Journal of the Economic Science Association*, 1, 100-113.

Selten, R. (1975), "A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4(1), 25–55.

Thomas, Catherine (2011), "Too Many Products: Decentralized Decision Making in Multinational Firms," *American Economic Journal: Microeconomics*, 3, 280-306.

Van Huyck, John, Ann Gillette, and Raymond Battalio (1992), "Credible assignments in coordination games," *Games and Economic Behavior*, 4(4), 606-626.

Vespa, Emanuel and Alistair J. Wilson (2016), "Communication With Multiple Senders: An Experiment," *Quantitative Economics*, 7(1), 1-36.

Weber, Roberto A., Colin Camerer and Marc Knez (2004), "Timing and virtual observability in ultimatum bargaining and 'weak-link' coordination games." *Experimental Economics*, 7(1). 25-48.

Weber, Roberto A., Yuval Rottenstreich, Colin Camerer and Marc Knez (2001), The illusion of leadership: Misattribution of cause in coordination games," *Organization Science*, 12(5), 582-598.

Appendix A: Full Versions of Tables 1 – 3

Table A1: Stage Game Payoffs ($k_1 = 54$, $k_2 = 14$, $k_3 = 7$, and $k_4 = 4$)
Note: Each cell contains the payoffs for S1 (π_{S1}), S2 (π_{S2}), and M (π_M).

Game 1

	C1	C2	C3	C4	C5
R1	26, 54, 80	12, 29, 41	-2, 4, 2	-16, -21, -37	-30, -46, -76
R2	15, 40, 55	29, 43, 72	15, 18, 33	1, -7, -6	-13, -32, -45
R3	4, 26, 30	18, 29, 47	32, 32, 64	18, 7, 25	4, -18, -14
R4	-7, 12, 5	7, 15, 22	21, 18, 39	35, 21, 56	21, -4, 17
R5	-18, -2, -20	-4, 1, - 3	10, 4, 14	24, 7, 31	38, 10, 48

Game 2

	C1	C2	C3	C4	C5
R1	22, 50, 72	8, 33, 41	-6, 8, 2	-20, -17, -37	-34, -42, -76
R2	19, 36, 55	33, 47, 80	19, 22, 41	5, -3, 2	-9, -28, -37
R3	8, 22, 30	22,33, 55	36, 36, 72	22, 11, 33	8, -14, -6
R4	-3,8, 5	11,19,30	25, 22, 47	39, 25, 64	25, 0, 25
R5	-14, -6, -20	0,5,5	14, 8, 22	28, 11, 39	42, 14, 56

Game 3

	C1	C2	C3	C4	C5
R1	18, 46, 64	4, 29, 33	-10,12, 2	-24, -13, -37	-38, -38, -76
R2	15, 32, 47	29, 43, 72	15, 26, 41	1, 1, 2	-13, -24, -37
R3	12, 18, 30	26, 29, 55	40, 40, 80	26, 15, 41	12, -10,2
R4	1,4, 5	15, 15, 30	29,26, 55	43,29, 72	29, 4, 33
R5	-10, -10, -20	4,1, 5	18,12, 30	32, 15, 47	46, 18, 64

Game 4

	C1	C2	C3	C4	C5
R1	14, 42, 56	0, 25, 25	-14, 8, -6	-28, -9, -37	-42, -34, -76
R2	11, 28, 39	25, 39, 64	11, 22, 33	-3, 5, 2	-17, - 20, -37
R3	8, 14, 22	22, 25, 47	36, 36, 72	22, 19, 41	8, -6, 2
R4	5, 0, 5	19, 11, 30	33, 22, 55	47, 33, 80	33, 8, 41
R5	-6, -14, -20	8, -3, 5	22, 8, 30	36, 19, 55	50, 22, 72

Game 5

	C1	C2	C3	C4	C5
R1	10, 38, 48	-4, 21, 17	-18, 4, -14	-32, -13, -45	-46, -30, -76
R2	7, 24, 31	21, 35, 56	7, 18, 25	-7, 1, -6	-21, -16, -37
R3	4, 10, 14	18, 21, 39	32, 32, 64	18, 15, 33	4, -2, 2
R4	1, -4, -3	15, 7, 22	29, 18, 47	43, 29, 72	29, 12, 41
R5	-2, -18, -20	12, -7, 5	26,4, 30	40, 15, 55	54, 26, 80

Table A2: Summary of Outcomes

Rounds 1 – 9

Treatment	% Coordinate	% Efficient (s.t. Coordinate)	Total Surplus	Efficiency Gain
NC – D	46.1%	41.1%	53.5	-190.8%
SC/S – D	72.8%	46.3%	62.7	-89.0%
SC/A – D	55.1%	51.1%	57.9	-141.3%
SC – MC	97.5%	44.7%	70.0	-10.4%
CH/S – D	82.3%	43.0%	64.2	-70.2%
CH/A – D	75.7%	51.1%	64.0	-79.2%
CH – MC	98.4%	41.8%	71.3	5.2%

Rounds 10 – 18

Treatment	% Coordinate	% Efficient (s.t. Coordinate)	Total Surplus	Efficiency Gain
NC – D	69.5%	39.1%	61.4	-108.3%
SC/S – D	77.8%	57.1%	65.7	-57.5%
SC/A – D	69.5%	58.7%	64.4	-77.4%
SC – MC	99.6%	45.9%	71.7	6.5%
CH/S – D	97.5%	48.1%	72.2	14.0%
CH/A – D	90.9%	60.6%	71.6	2.8%
CH – MC	100.0%	61.7%	75.2	44.3%

Table A3: Types of Coordination

Rounds 1 – 9

Treatment	% Safe	% Efficient	% Other
NC – D	27.0%	9.0%	5.8%
SC/S – D	44.4%	20.1%	5.3%
SC/A – D	23.3%	23.8%	3.7%
SC – MC	28.6%	38.1%	31.7%
CH/S – D	36.5%	26.5%	16.9%
CH/A – D	28.0%	31.7%	15.3%
CH – MC	40.2%	33.3%	24.3%

Rounds 10 – 18

Treatment	% Safe	% Efficient	% Other
NC – D	48.1%	12.6%	4.9%
SC/S – D	36.1%	30.6%	8.2%
SC/A – D	32.8%	30.1%	4.4%
SC – MC	31.1%	36.1%	32.2%
CH/S – D	41.0%	36.1%	20.8%
CH/A – D	29.5%	44.8%	15.8%
CH – MC	34.4%	50.8%	14.8%

Appendix B: Equilibria in the MA Game

Proof of Theorem

This appendix discusses the equilibrium predictions for the MA game with managerial control. We begin by proving the theorem stated in the main text which shows that there do not exist informative pure strategy equilibria. We then explore whether there exist informative mixed strategy equilibria and whether such equilibria are plausible.

The MA game with managerial control is played by two agents (A1 and A2) and a manager (M). The game begins with both agents viewing the state of the world $\gamma \in \mathcal{G} = \{1,2,3,4,5\}$. The two agents then simultaneously choose messages to send to the manager. The manager observes the two messages, but not the state of the world, and chooses a pair of actions for the two agents. Let $\mu_i \in \mathcal{M}_i = \{1,2,3,4,5\}$ be the message chosen by Agent i and let $\alpha_i \in \mathcal{A}_i = \{1,2,3,4,5\}$ be the action chosen by the manager for Agent i. Payoffs for the two agents and manager (π_{A1}, π_{A2}, π_M) as a function of γ, α_1, and α_2 are given by Equations 2a, 2b, and 2c, reproduced below. Note that messages are cheap talk and do not appear in the payoff functions. As in the experiment, set $k_1 = 54$, $k_2 = 14$, $k_3 = 7$, and $k_4 = 4$.

$$\pi_{A1} = k_1 - k_2|\alpha_1 - \alpha_2| - k_3|\alpha_1 - 5| - k_4|\alpha_1 - \gamma| \quad \textbf{(Eq. 2a)}$$
$$\pi_{A2} = k_1 - k_2|\alpha_1 - \alpha_2| - k_3|\alpha_2 - 1| - k_4|\alpha_2 - \gamma| \quad \textbf{(Eq. 2b)}$$
$$\pi_M = \pi_{A1} + \pi_{A2} \quad \textbf{(Eq. 2c)}$$

For Agent i, define a strategy for the MA game with managerial control as a mapping from the set of games to the set of distributions over messages: $\sigma_i: \mathcal{G} \to \Delta(M_i)$. For the manager, a strategy is a mapping from the set of pairs of messages she could receive into the set of distributions over pairs of actions that she could impose: $\sigma_M: \mathcal{M}_1 \times \mathcal{M}_2 \to \Delta(\mathcal{A}_1 \times \mathcal{A}_2)$. Using these definitions and the payoff functions defined in Equations 2a, 2b, and 2c, we can define an agent's expected payoff as a function of the game, the message the agent chooses, and the strategies of the two other players: $\Pi_i(\gamma, \mu_i, \sigma_{-i}, \sigma_M) = \Pi_i\left(\gamma, \mu_i, \sigma_{-i}(\gamma), \sigma_M\left(\mu_i, \sigma_{-i}(\gamma)\right)\right)$. Define a ***belief map*** as a function from the set of pairs of messages into the set of distributions over games: $\beta: \mathcal{M}_1 \times \mathcal{M}_2 \to \Delta(\mathcal{G})$. Given the manager's beliefs, define $\Pi_M\left(\alpha_1, \alpha_2, \beta(\mu_1, \mu_2)\right)$ to be the manager's expected payoff as a function of the actions she chooses and her beliefs given the messages she receives from the agents.

In a Perfect Bayesian equilibrium, all players must maximize their payoffs subject to beliefs and the manager's beliefs must be consistent with the agents' strategies. Let $\widehat{\mathcal{G}_i}(\mu_i|\sigma_i)$ be the set of games in which Agent i could possibly send the message μ_i: $\widehat{\mathcal{G}_i}(\mu_i|\sigma_i) = \{\gamma \in \mathcal{G}: \sigma_i(\mu_i)(\gamma) > 0\}$. Let $\hat{\mathcal{G}}(\mu_1, \mu_2|\sigma_1, \sigma_2)$ be the set of games in which the pair of messages (μ_1, μ_2) could possibly be sent: $\hat{\mathcal{G}}(\mu_1, \mu_2|\sigma_1, \sigma_2) = \widehat{\mathcal{G}_1}(\mu_1|\sigma_1) \cap \widehat{\mathcal{G}_2}(\mu_2|\sigma_2)$.

Definition: A strategy profile $(\sigma_1^*, \sigma_2^*, \sigma_M^*)$ is a ***Perfect Bayesian equilibrium*** (PBE) if there exists a belief map $\beta: \mathcal{M} \to \Delta(\mathcal{G})$ such that the following hold:

i) If $\sigma_i^*(\mu_i)(\gamma) > 0$, then $\mu_i \in \underset{\mathcal{M}_i}{\operatorname{argmax}} \Pi_i(\gamma, \cdot, \sigma_{-i}, \sigma_M)$.

ii) If $\sigma_M^*(\alpha_1, \alpha_2)(\mu_1, \mu_2) > 0$, then $(\alpha_1, \alpha_2) \in \underset{\mathcal{A}_1 \times \mathcal{A}_2}{\operatorname{argmax}} \Pi_M\big(\alpha_1, \alpha_2, \beta(\mu_1, \mu_2)\big)$

iii) If $\gamma \in \hat{\mathcal{G}}(\mu_1, \mu_2|\sigma_1^*, \sigma_2^*)$, then $\beta(\gamma|\mu_1, \mu_2) = \frac{\sigma_1^*(\mu_1|\gamma)\sigma_2^*(\mu_2|\gamma)}{\sum_{\gamma' \in \hat{\mathcal{G}}(\mu_1, \mu_2|\sigma_1^*, \sigma_2^*)} \sigma_1^*(\mu_1|\gamma')\sigma_2^*(\mu_2|\gamma')}$

Before proving the theorem stated in the main text, it is useful first to prove the following lemma about the best response function of the manager.

Lemma 1: Following any $(\mu_1, \mu_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, given any belief mapping $\beta: \mathcal{M}_1 \times \mathcal{M}_2 \to \Delta(\mathcal{G})$, an optimizing manager must choose identical actions for the two agents: $\alpha_1 = \alpha_2$.

Proof: Suppose not. Suppose the manager chooses actions α_1^* and α_2^* such that $\alpha_1^* \neq \alpha_2^*$. This implies that the manager is choosing an outcome that is *not* a Nash equilibrium if the two agents are allowed to choose their own actions (i.e. not a pure strategy equilibrium in the MA game with delegation). Given that $k_2 > k_3 + k_4$, it follows from Equations 2a and 2b that the payoffs of both agents improve, regardless of beliefs, if the manager instead chooses $\alpha_1 = \alpha_2^*$. Since the manager's payoff equals the sum of the two agents' payoffs, as shown in Equation 2c, the manager's payoff also increases. This implies that the manager's initial choice could not have been optimal. **Q.E.D.**

Given the preceding lemma, we can refer to the manager's strategy as choosing a single outcome in response to the agents' messages:[42] $\sigma_M: \mathcal{M}_1 \times \mathcal{M}_2 \to \Delta(\mathcal{A})$ where $\mathcal{A} \in \{1,2,3,4,5\}$. Given some Perfect Bayesian equilibrium profile $(\sigma_1^*, \sigma_2^*, \sigma_M^*)$, define $\mathcal{A}^*: \mathcal{G} \to \Delta(\mathcal{A})$ as the implied mapping from games into distributions over outcomes. Let $\mathcal{A}_i \in \Delta\mathcal{A}$ be the vector where the i^{th} element equals one and all other elements equal zero.

A Perfect Bayesian equilibrium is ***fully informative*** if $\mathcal{A}^*(\gamma) = \mathcal{A}_\gamma$ for $\forall \gamma \in \mathcal{G}$; for Game i, the managers chooses Outcome i with certainty. For an equilibrium to be fully informative, the agents' messages must perfectly reveal the state of the world for $\forall \gamma \in \mathcal{G}$. A Perfect Bayesian equilibrium is ***partially informative***

[42] As elsewhere in the paper, we use the term "outcome" to refer to a pair of identical actions assigned by the manager to the agents.

if there exist $\gamma, \gamma' \in \mathcal{G}$ such that $\mathcal{A}^*(\gamma) \neq \mathcal{A}^*(\gamma')$. In a partially informative equilibrium, the outcomes selected by the manager respond to the agents' information. Note that any fully informative equilibrium is also partially informative. The following theorem shows that there does *not* exist a partially informative pure strategy Perfect Bayesian equilibrium.

Theorem 1: There does not exist a pure-strategy PBE for the MA game with managerial control where the manager chooses different actions for two different states of the world.

Proof: Suppose that such an equilibrium existed. Let γ and γ' ∈ $\mathcal{G}$ be two states where different actions are chosen by the manager in equilibrium: $\mathcal{A}^*(\gamma) \neq \mathcal{A}^*(\gamma')$. Without loss of generality, assume that A1 prefers $\mathcal{A}^*(\gamma)$ to $\mathcal{A}^*(\gamma')$ and A2 prefers $\mathcal{A}^*(\gamma')$ to $\mathcal{A}^*(\gamma)$. It follows from Equations 2a and 2b that $\mathcal{A}^*(\gamma) > \mathcal{A}^*(\gamma')$.[43]

It must be the case that $\sigma_1^*(\gamma) \neq \sigma_1^*(\gamma')$. In other words, A1 cannot send the same messages in γ and γ'. Proof is by contradiction. Suppose $\sigma_1^*(\gamma) = \sigma_1^*(\gamma')$. This implies that the manager's choice is determined solely by A2's message. Since A2 prefers $\mathcal{A}^*(\gamma')$ to $\mathcal{A}^*(\gamma)$, he should always send $\sigma_2^*(\gamma')$ whether the true state of the world is γ or γ'. But then the manager would choose $\mathcal{A}^*(\gamma')$ in both γ and γ'. A contradiction follows. By the same logic, $\sigma_2^*(\gamma) \neq \sigma_2^*(\gamma')$.

Suppose that A1 deviates by choosing $\sigma_1^*(\gamma)$ in γ'. In other words, A1 sends the message he is supposed to choose in γ' rather than the message he should send in γ. The resulting pair of messages ($\sigma_1^*(\gamma)$, $\sigma_2^*(\gamma')$) cannot make A1 better off or a profitable deviation from equilibrium would exist for A1. It follows that $\sigma_M^*\big(\sigma_1^*(\gamma), \sigma_2^*(\gamma')\big) \leq \sigma_M^*\big(\sigma_1^*(\gamma'), \sigma_2^*(\gamma')\big) = \mathcal{A}^*(\gamma') < \mathcal{A}^*(\gamma) = \sigma_M^*\big(\sigma_1^*(\gamma), \sigma_2^*(\gamma)\big)$. However, this implies that A2 can gain by choosing $\sigma_2^*(\gamma')$ in γ, rather than $\sigma_2^*(\gamma)$, since $\sigma_M^*\big(\sigma_1^*(\gamma), \sigma_2^*(\gamma')\big) < \sigma_M^*\big(\sigma_1^*(\gamma), \sigma_2^*(\gamma)\big)$; by Equation 2b and given that $k_3 > k_4$, decreasing the manager's choice increases A2's payoff. It follows that A2 has a profitable deviation from equilibrium. A contradiction follows. **Q.E.D.**

Mixed Strategy Equilibria

While there do not exist any pure strategy Perfect Bayesian equilibria which are partially informative, there do exist mixed strategy Perfect Bayesian equilibria that are partially informative and yield expected greater total surplus than a babbling equilibrium. We start our discussion of mixed strategy equilibria by describing two such equilibria. We then discuss the properties of these two equilibria that make it possible for them to be informative. Drawing on this discussion, we prove a result providing sufficient conditions for no partially

[43] We are somewhat abusing notation as $\mathcal{A}^*(\gamma)$ and $\mathcal{A}^*(\gamma')$ are vectors. Because the manager plays a pure strategy, both vectors must put full weight on a single outcome. The inequality is a comparison of the indices of said outcomes.

informative Perfect Bayesian equilibria to exist. We conclude Appendix B by discussing this result. We argue that partially informative equilibria, while theoretically possible, are not empirically plausible.

The following two equilibria are both mixed strategy Perfect Bayesian equilibria that are partially informative and yield greater expected total surplus than a babbling equilibrium.

Equilibrium 1: If $\gamma \in \{1,5\}$, $\mu_1 = 5$ and $\mu_2 = 1$. For $\gamma \in \{2,3,4\}$, both agents truthfully reveal what game is being played: $\mu_1 = \mu_2 = \gamma$. If $(\mu_1, \mu_2) \in \{(2,2), (3,3), (4,4)\}$, the manager chooses the actions that correspond to the agents' messages: $\alpha_1 = \alpha_2 = \mu_1 = \mu_2$. If $(\mu_1, \mu_2) \notin \{(2,2), (3,3), (4,4)\}$, the manager randomizes between the two extreme outcomes, $\alpha_1 = \alpha_2 = 1$ and $\alpha_1 = \alpha_2 = 5$, with equal weight on these two outcomes. If the manager observes a pair of messages that occur with positive probability in equilibrium, the resulting beliefs are determined by the application of Bayes Law. If $\hat{\mathcal{G}}(\mu_1, \mu_2 | \sigma_1, \sigma_2) = \emptyset$ then $\beta(1|\mu_1, \mu_2) = \beta(5|\mu_1, \mu_2) = ½$. In other words, following any pair of messages that occur with zero probability in equilibrium, the manager's beliefs put equal weight on $\gamma = 1$ and $\gamma = 5$.[44]

Equilibrium 2: If $\gamma \in \{1,2,3,5\}$, both agents randomize uniformly over $\mu_i \in \{1,2,3,5\}$. In other words, for all games except $\gamma = 4$, agents randomize uniformly over all messages except $\mu_i = 4$. If $\gamma = 4$, $\mu_1 = \mu_2 = 4$. If $\mu_1 \in \{1,2,3,5\}$ and $\mu_2 \in \{1,2,3,5\}$, the manager chooses the safe outcome: $\alpha_1 = \alpha_2 = 3$. If the manager observes $\mu_1 = \mu_2 = 4$, she chooses the corresponding outcome: $\alpha_1 = \alpha_2 = 4$. If the manager observes $\mu_i \in \{1,2,3,5\}$ and $\mu_j = 4$, she randomizes between the two extreme outcomes, $\alpha_1 = \alpha_2 = 1$ and $\alpha_1 = \alpha_2 = 5$, with equal weight on these two outcomes. If the manager observes a pair of messages that occur with positive probability in equilibrium, her beliefs are determined by the application of Bayes Law. If $\mu_i \in \{1,2,3,5\}$ and $\mu_j = 4$, implying $\hat{\mathcal{G}}(\mu_1, \mu_2 | \sigma_1, \sigma_2) = \emptyset$, then $\beta(1|\mu_1, \mu_2) = \beta(5|\mu_1, \mu_2) = ½$. In other words, following any pair of messages that occur with zero probability in equilibrium, the manager's beliefs put equal weight on $\gamma = 1$ and $\gamma = 5$.

The existence of partially informative mixed strategy equilibria like Equilibrium 1 and Equilibrium 2 implies that fully informative outcomes can be supported in the initial rounds of a finitely repeated version of the MA game with managerial control. These partially informative equilibria yield higher expected

[44] We thank the anonymous referee who suggested this equilibrium.

payoffs than the babbling equilibrium for all players, so the threat of reversion to the babbling equilibrium can be used as punishment to enforce truthful revelation of what game is being played.[45]

Not only do Equilibrium 1 and 2 provide examples of partially informative equilibria that yield greater expected total surplus than a babbling equilibrium, they also illustrate how any such equilibrium must be constructed. By definition, in a partially informative equilibrium there must exist $\gamma, \gamma' \in \mathcal{G}$ such that $\mathcal{A}^*(\gamma) \neq \mathcal{A}^*(\gamma')$. To yield greater expected total surplus than a babbling equilibrium, it is necessary that the two agents are not indifferent between the distribution of outcomes in γ and γ'. Without loss of generality, assume that A1 prefers $\mathcal{A}^*(\gamma)$ to $\mathcal{A}^*(\gamma')$ and A2 prefers $\mathcal{A}^*(\gamma')$ to $\mathcal{A}^*(\gamma)$. To make the agents' equilibrium strategies incentive compatible, it must be true that A1 cannot gain by following $\sigma_1^*(\gamma)$ when the state of the world is γ' and A2 cannot gain by following $\sigma_2^*(\gamma')$ when the state of the world is γ. The preceding constraints imply that $\sigma_M^*(\sigma_1^*(\gamma), \sigma_2^*(\gamma'))$ must serve as a punishment for possible deviations by *both* agents. The trick in building a partially informative equilibrium that yields greater expected total surplus than a babbling equilibrium is to find a strategy and belief mapping for the manager that serve this dual purpose.

Recall from Equation 1 that agents' payoffs are constructed from three types of losses: coordination losses, adaptation losses, and state losses. Punishing deviations by the agents must rely on using some combination of these three types of losses. By Lemma 1, any punishment that involves coordination losses (i.e. assigning different actions to the two agents) can be ruled out. Using just adaptation losses also won't work because the agents' preferences over outcomes are diametrically opposed. Any shift in the distribution over outcomes that increases adaptation losses for A1 must necessarily decrease adaptation losses for A2 and vice versa. Thus, any punishment must make use of state losses. An increase in state losses without a corresponding decrease in adaptation losses for one of the agents is accomplished by increasing the spread of outcomes. This can be seen in both Equilibrium 1 and Equilibrium 2 which use mixing between Outcome 1 and Outcome 5 as a punishment that harms *both* agents.

What sort of belief mapping is needed to support such a punishment? In either Equilibrium 1 or Equilibrium 2, mixing between Outcome 1 and Outcome 5 must be a best response to the manager's beliefs. Critically, choice of Outcomes 1 and 5 must be weakly better than choice of Outcomes 2, 3, and 4. This has important implications for the structure of the belief mapping. To understand why this is true, we now prove a second lemma about the structure of partially informative equilibria.

[45] Even allowing for mixed strategy equilibria, it is not possible to have an equilibrium in the finitely repeated MA game with managerial control which is fully informative for *all* rounds. Rather, the point is that fully informative outcomes can be supported in early rounds when the existence of multiple equilibria makes punishment feasible.

Definition: Let β be a belief mapping and let μ_1 and μ_2 be messages sent by A1 and A1. Game $\gamma \in \{1,2,3,4,5\}$ is ***pivotal*** subject to β, μ_1, and μ_2 if the following condition holds:

$$\sum_{\gamma'<\gamma} \beta(\gamma'|\mu_1,\mu_2) \leq \frac{1}{2} \leq \sum_{\gamma'\leq\gamma} \beta(\gamma'|\mu_1,\mu_2)$$

Intuitively, a pivotal game is the median of β subject to receiving messages μ_1 and μ_2. There are always either one or two pivotal games.

Definition: A ***pivotal outcome*** is the outcome corresponding to a pivotal game. For a belief mapping β and messages μ_1 and μ_2, if $\gamma \in \{1,2,3,4,5\}$ is a pivotal game then $\alpha_1 = \alpha_2 = \gamma$ is a pivotal outcome.

Lemma 2: In a Perfect Bayesian equilibrium, following any messages μ_1 and μ_2, only pivotal outcomes can receive positive weight in the manager's equilibrium strategy, $\sigma_M^*(\mu_1,\mu_2)$.

Proof: Suppose not. Let γ be the lowest pivotal game and suppose $\alpha < \gamma$ is an outcome that receives positive weight as part of the manager's strategy. Consider the consequences of moving all the weight placed on α to $\alpha + 1$. This does not affect the coordination losses and adaptation losses are irrelevant for the manager. The effect on the managers' payoff from this change must come solely through the state losses. Because $\alpha < \gamma$ and γ is pivotal, the manager believes that shifting from α to $\alpha + 1$ moves the action closer to the true game more often than not. It follows that expected state losses (as defined in Equations 2a and 2b) must decrease and, by extension, expected manager payoffs increase.[46] Therefore, $\sigma_M^*(\mu_1,\mu_2)$ could not be best response. A contradiction follows. Similar logic eliminates all cases where an action receiving positive weight is greater than the highest pivotal game. **Q.E.D.**

Given Lemma 2, very specific beliefs are needed to support mixing by the manager between Outcome 1 and Outcome 5. The belief mapping β must put equal weight on $\gamma = 1$ and $\gamma = 5$ following any pair of messages that occur with zero probability in equilibrium; these beliefs make the manager indifferent between choosing any of the five possible outcomes. No weight can be put on $\gamma = 2$, 3, or 4. By Lemma 2, if beliefs following a pair of messages that occur with zero probability did put weight on one or more of these games, σ_M^* could not put positive weight on both Outcome 1 and Outcome 5. One of the two can be a pivotal outcome given beliefs that put positive weight on some subset of $\gamma =$ 2, 3, or 4, but it is not possible for both Outcome 1 and Outcome 5 to be pivotal outcomes. This is a general property of the MA

[46] The expected gain in payoffs equals $k_4\left(1-2\sum_{\gamma''\leq\gamma'}\beta(\gamma''|\mu_1,\mu_2)\right)$. By definition of a pivotal game, $\sum_{\gamma''\leq\gamma'}\beta(\gamma''|\mu_1,\mu_2) < ½$. It follows that the manager's expected gain due to shifting from α to α' is positive.

game. Generating responses by the manager with high state losses, as is necessary to punish deviations, requires beliefs that only put weight on extreme outcomes.

While there exist partially informative equilibria, it is an open question whether said equilibria are empirically relevant: Is it plausible that the proposed equilibria could occur in an experimental setting? The preceding analysis indicates that beliefs in a partially informative equilibrium have three properties that make us question the empirical relevance of such equilibria. We first lay out these properties intuitively, illustrating them with Equilibrium 1 and Equilibrium 2, and then prove a formal result showing the relationship between these properties and the existence of partially informative equilibrium.

First, it is a given for any experimental data that subjects' decision making will be noisy. To be empirically relevant, an equilibrium ought to be robust to the presence of small trembles. Neither Equilibrium 1 nor Equilibrium 2 meet this standard. To generate the beliefs necessary to prevent agents from defecting, beliefs must put no weight on $\gamma = 2$, 3, or 4. The presence of trembles implies that there will be positive weight on these three states. Lemma 2 then implies that managers will *not* mix between the two extreme outcomes as required to punish deviations.

Second, even if the formalities of the model do not require any specific mapping between messages and meaning, agents' messages in the MA game with managerial control have a natural meaning, communicating a signal about the state of the world. If an agent increases their message (e.g. switches from $\mu_i = 2$ to $\mu_i = 3$), the obvious interpretation is that they are communicating a higher state of the world (e.g. $\gamma = 3$ rather than $\gamma = 2$). The manager might not believe the agent; after all, agents have an incentive to lie. But it would seem odd if the manager interprets a higher signal as meaning that the state of the world is lower. Equilibria in which beliefs are weakly monotonic are more likely to be empirically relevant.

Beliefs are *not* weakly monotonic in either Equilibrium 1 or Equilibrium 2. For Equilibrium 1, suppose $\gamma = 4$. The agents' messages should be $\mu_1 = \mu_2 = 4$ in equilibrium, and the manager's beliefs following receipt of this pair of messages put full weight on $\gamma = 4$. Suppose one of the agent deviates by sending $\mu_i = 5$. The manager's beliefs now put 50/50 weight on $\gamma = 1$ and $\gamma = 5$. After an increase in one agent's message, the manager goes from believing that $\gamma = 1$, the lowest possible state of the world, was not possible to believing that there is a 50% chance that it is the state of the world. The expected value of γ decreases from 4 to 3. An identical argument can be used to show that beliefs are not weakly monotonic in Equilibrium 2.

In both Equilibrium 1 and Equilibrium 2, the non-monotonicity of beliefs is directly related to how punishments must be constructed in equilibrium. Punishing deviations requires increasing state losses for both agents. Following a deviation, the manager's beliefs must concentrate weight on the extreme games, $\gamma = 1$ and $\gamma = 5$. This implies that a small increase in the message sent by one agent can lead to a large increase in the weight the manager's beliefs put on $\gamma = 1$, leading to a non-monotonicity.

Finally, in a Perfect Bayesian equilibrium players maximize subject to their beliefs. Technically, however, managers' strategies are defined as a function of the agents' messages rather than their beliefs. This can lead to the peculiar possibility that a change in an agent's message can change the manager's choice even though her beliefs are unaffected. We require that a manager's choices depend on her beliefs, not the messages that led to those beliefs. Specifically, if a change in the agents' messages does not affect the beliefs of the manager, it also does not affect the actions of the manager.

The preceding discussion suggests that partially informative equilibria depend on the absence of trembles, non-monotonic beliefs, and a direct relationship between beliefs and actions. Indeed, there is a direct link between these properties and the existence of partially informative equilibria: There exist no partially informative equilibrium which fulfill the following conditions: the equilibrium is trembling hand perfect, the manager has monotonic beliefs, and the manager's actions are determined solely by her beliefs.

Definition: A game is ***fully mixed*** if the strategy space for all players is restricted to put positive weight on all available actions. A ***fully mixed equilibrium*** is a Perfect Bayesian equilibrium of a fully mixed game.

Definition: A Perfect Bayesian equibrium is a ***trembling hand perfect*** equilibrium (Selten, 1975) if there exists a series of fully mixed games that converges to the unperturbed game and a series of fully mixed equilibria for this series of fully mixed games that converges to the Perfect Bayesian equilibrium.

For any trembling hand equilibrium of the MA game and any $(\mu_1, \mu_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, Lemma 2 implies that $\sigma_M^*(\mu_1, \mu_2)$ must either put 100% weight on either a single outcome or put 100% weight on two adjacent outcomes. This follows from the observation that if A1 and A2 use fully mixed strategies, the resulting beliefs must put strictly positive weight on all $\gamma \in \mathcal{G}$. Note that neither Equilibrium 1 nor Equilibrium 2 are trembling hand perfect equilibria.

Definition: Let $\mu_i, \mu'_i \in \mathcal{M}_i$ be two possible messages for Agent i such that $\mu'_i > \mu_i$. The manager's belief mapping is ***belief monotonic*** if the following relationship holds for any $\gamma \in \mathcal{G}$ and $\mu_j \in \mathcal{M}_j$ for Agent j ≠ i:

$$\sum_{\gamma' \geq \gamma} \beta\left(\gamma' | \mu'_i, \mu_j\right) \geq \sum_{\gamma' \geq \gamma} \beta\left(\gamma' | \mu_i, \mu_j\right)$$

In other words, holding the other agent's message fixed, the beliefs induced by μ'_i first order stochastic dominate the beliefs induced by μ_i. Assuming fully mixed strategies, it is straightforward to confirm that this condition on beliefs is equivalent to the following condition on Agent i's strategy σ_i. Notice that the following condition does not depend on μ_j.

$$\frac{\sum_{\gamma'\geq\gamma}\sigma_i(\mu'_i|\gamma')}{\sum_{\gamma'<\gamma}\sigma_i(\mu'_i|\gamma')} \geq \frac{\sum_{\gamma'\geq\gamma}\sigma_i(\mu_i|\gamma')}{\sum_{\gamma'<\gamma}\sigma_i(\mu_i|\gamma')}$$

Belief monotonicity guarantees that higher messages are (weakly) associated with beliefs that put more weight on higher states of the world. Neither Equilibrium 1 nor Equilibrium 2 is belief monotonic.

Definition: A manager's strategy is **belief dependent** if the following condition holds: if $\beta(\mu'_1,\mu'_2) = \beta(\mu_1,\mu_2)$, then $\sigma_M(\mu'_1,\mu'_2) = \sigma_M(\mu_1,\mu_2)$.

In conjunction with Lemma 2, belief monotonicity and belief dependence guarantee that the manager's strategy is a weakly increasing function of the messages sent by either agent.

Lemma 3: Hold $\gamma \in \mathcal{G}$ fixed. In a trembling hand perfect equilibrium with belief monotonicity and belief dependence, the same distribution over outcomes must follow any pair of messages (μ_1,μ_2) sent with positive probability.

Proof: Suppose there exist $\mu_1,\mu_1' \in \mathcal{M}_1$ and $\mu_2,\mu_2' \in \mathcal{M}_2$ such that all four messages are used with positive probability in equilibrium and $\sigma_M^*(\mu_1,\mu_2) \neq \sigma_M^*(\mu_1',\mu_2')$. Let $\overline{\mu}_i = max(\mu_i,\mu_i')$ and $\underline{\mu_i} = min(\mu_i,\mu_i')$. Monotonicity of the manager's strategy implies that $\sigma_M^*(\overline{\mu}_1,\overline{\mu}_2)$ strictly first order stochastic dominates $\sigma_M^*\left(\underline{\mu_1},\underline{\mu_2}\right)$. It must be the case that $\sigma_M^*\left(\overline{\mu}_1,\underline{\mu_2}\right) = \sigma_M^*\left(\underline{\mu_1},\underline{\mu_2}\right)$. If not, monotonicity of the manager's strategy, which follows from belief monotonicity and belief dependence, would imply that $\sigma_M^*\left(\overline{\mu}_1,\underline{\mu_2}\right)$ strictly first order stochastic dominates $\sigma_M^*\left(\underline{\mu_1},\underline{\mu_2}\right)$ and A1 can profitably deviation by sending $\overline{\mu}_1$ whenever he is supposed to send $\underline{\mu_1}$.[47] Given that $\sigma_M^*\left(\overline{\mu}_1,\underline{\mu_2}\right) = \sigma_M^*\left(\underline{\mu_1},\underline{\mu_2}\right)$ and $\sigma_M^*(\overline{\mu}_1,\overline{\mu}_2)$ strictly first order stochastic dominates $\sigma_M^*\left(\underline{\mu_1},\underline{\mu_2}\right)$, it follows that $\sigma_M^*(\overline{\mu}_1,\overline{\mu}_2)$ strictly first order stochastic dominates

[47] This statement relies on belief monotonicity, belief dependence, and Lemma 2. Together, the two imply that A1 cannot be worse off by sending $\overline{\mu}_1$ rather than $\underline{\mu}_1$ regardless of what message is chosen by A2 and must be strictly better off versus $\underline{\mu_2}$.

$\sigma_M^*\left(\overline{\mu}_1, \underline{\mu}_2\right)$. A2 therefore must have a profitable deviation by sending $\underline{\mu}_2$ whenever he is supposed to send $\overline{\mu}_2$. **Q.E.D.**

Theorem 2: For any trembling hand perfect equilibrium in which belief monotonicity and belief dependence hold, $\mathcal{A}^*(\gamma) = \mathcal{A}^*(\gamma')$ for any $\gamma, \gamma' \in \mathcal{G}$. It follows that total surplus in any trembling hand perfect equilibrium in which belief monotonicity and belief dependence hold cannot be higher than the babbling equilibrium.

Proof: Suppose not. There exists an equilibrium in which $\mathcal{A}^*(\gamma) \neq \mathcal{A}^*(\gamma')$ for some $\gamma, \gamma' \in \mathcal{G}$. By Lemma 3, the same distribution over outcomes must be taken following all possible pairs of messages sent with positive probability in equilibrium for γ. The same holds for γ'. Given this, it follows from Lemma 2 that the distribution of outcomes in one of these two games must strictly first order stochastic dominate the distribution of actions in the other. Without loss of generality, assume that $\mathcal{A}^*(\gamma')$ strictly first order stochastic dominates $\mathcal{A}^*(\gamma)$. Let (μ_1', μ_2') be a pair of messages sent with positive probability in equilibrium if the state of the world is γ' and let (μ_1, μ_2) be a pair of messages sent with positive probability in equilibrium if the state of the world is γ. Since, by Lemma 3, $\sigma_M^*(\mu_1', \mu_2') = \mathcal{A}^*(\gamma')$ and $\sigma_M^*(\mu_1, \mu_2) = \mathcal{A}^*(\gamma)$, it must be the case that $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$. Also note that either $\mu_1' \neq \mu_1$ and/or $\mu_2' \neq \mu_2$. If not, the same pair of messages could appear in both games, γ and γ', implying that $\mathcal{A}^*(\gamma) = \mathcal{A}^*(\gamma')$.

We now show that no such pair of messages can exist, leading to a contradiction.

It cannot be the case that $\mu_1' \leq \mu_1$ and $\mu_2' \leq \mu_2$ with strict inequality in at least one case. Proof is by contradiction. Suppose $\mu_1' \leq \mu_1$ and $\mu_2' \leq \mu_2$ with strict inequality in at least one case. By monotonicity of the manager's strategy, which follows from belief monotonicity and belief dependence, $\sigma_M^*(\mu_1, \mu_2')$ weakly first order stochastic dominates $\sigma_M^*(\mu_1', \mu_2')$ and $\sigma_M^*(\mu_1, \mu_2)$ weakly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2')$, but $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$. By transitivity, a contradiction follows.

It cannot be the case that $\mu_1' \geq \mu_1$ and $\mu_2' \geq \mu_2$ with strict inequality in at least one case. Proof is by contradiction. Suppose $\mu_1' \geq \mu_1$ and $\mu_2' \geq \mu_2$ with strict inequality in at least one case. It must be the case that $\sigma_M^*(\mu_1', \mu_2) = \sigma_M^*(\mu_1, \mu_2)$. If not, Agent 1 has a profitable deviation by sending μ_1' when it is supposed to send μ_1 in γ.[48] Since $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$, it follows that that $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1', \mu_2)$. This implies that Agent 2 must have a profitable deviation by sending μ_2 when it is supposed to send μ_2' in γ'. The argument is analogous to the one given for Agent 1 above. A contradiction follows.

Next, suppose $\mu_1' > \mu_1$ and $\mu_2' < \mu_2$. Again, proof is by contradiction. By monotonicity of the manager's strategy, $\sigma_M^*(\mu_1', \mu_2)$ weakly first order stochastic dominates $\sigma_M^*(\mu_1', \mu_2')$ which in turn strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$. It follows from transitivity that $\sigma_M^*(\mu_1', \mu_2)$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$. By monotonicity of the manager's strategy it follows that Agent 1 has a profitable deviation by sending μ_1' when it is supposed to send μ_1 in γ. A contradiction follows.

Finally, suppose $\mu_1' < \mu_1$ and $\mu_2' > \mu_2$. Again, proof is by contradiction. By monotonicity of the manager's strategy, $\sigma_M^*(\mu_1, \mu_2)$ weakly first order stochastic dominates $\sigma_M^*(\mu_1', \mu_2)$. Given that $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$, it follows from transitivity that $\sigma_M^*(\mu_1', \mu_2')$ strictly first order stochastic dominates $\sigma_M^*(\mu_1', \mu_2)$. Drawing on monotonicity once again, it follows that Agent 2 has a profitable deviation by sending μ_2 when it is supposed to send μ_2' in γ'. A contradiction follows. **Q.E.D.**

Our goal in stating and proving Theorem 2 is not to claim that informative equilibria are impossible. It can be seen from Equilibrium 1 and Equilibrium 2 that there exist partially informative Perfect Bayesian equilibria, and it is not terribly surprising that theoretical conditions can be found that eliminate such equilibria. Instead, our point of view is empirical in nature. For us, the central question is less whether an informative equilibrium exists but whether the informative equilibria are empirically plausible (i.e. is it plausible that such an equilibrium would spontaneously emerge in play by experimental subjects). Theorem 2 helps us answer that question.

The data, in combination with Theorem 2, suggests that partially informative equilibria are unlikely to play an important role in our experiment. The theorem leads us to not expect partially informative equilibra if subjects' decision making is noisy and manager strategies are a weakly monotonic function of agent's messages as implied by belief monotonicity and belief dependence. Both statements are consistent with the

[48] If $\mu_1' = \mu_1$, it trivially holds that $\sigma_M^*(\mu_1', \mu_2) = \sigma_M^*(\mu_1, \mu_2)$. Otherwise, if $\sigma_M^*(\mu_1', \mu_2) \neq \sigma_M^*(\mu_1, \mu_2)$, then, by belief monotonicity, $\sigma_M^*(\mu_1', \mu_2)$ strictly first order stochastic dominates $\sigma_M^*(\mu_1, \mu_2)$. It follows that Agent 1 strictly gains by sending μ_1' rather than μ_1 when Agent 2 sends μ_2. By belief monotonicity, Agent 1 cannot be worse off versus any other message that Agent 2 might send.

experimental data. It is almost a given that experimental data will be noisy. In **SC – MC**, the treatment where subjects directly play the MA game with managerial control, the distribution over games has full support following almost all possible messages.[49] We cannot directly observe beliefs, but beliefs that put zero weight on some games are not consistent with what actually occurs in the data. As for monotonicity, Table 5 and the associated statistical analysis in the main text provide evidence that managers' beliefs are weakly monotonic functions of agents' messages; the response of managers' outcome choices to an increase in the message sent by either agent is strong, positive, and statistically significant.

Moving away from the specific conditions of Theorem 2, Equilibrium 1 and Equilibrium 2 illustrate an important property that any partially informative equilibrium must have; the equilibrium must incorporate a response to deviations that punishes both agents by increasing state losses. This implies that outcomes will be concentrated at extreme outcomes when the messages of the two agents differ *even if neither message is extreme (i.e.* $\mu_i = 1$ *or* $\mu_i = 5$*)*. This is not true in the data. Using data from **SC – MC**, consider games where $\mu_1 \neq \mu_2$, $2 \leq \mu_1 \leq 4$, and $2 \leq \mu_2 \leq 4$. These are the cases where a partially informative equilibrium would lead us to expect use of Outcome 1 or Outcome 5. In reality, use of the extreme outcomes is very rare, only occurring for 3% of such observations. Not surprisingly, the most common outcome is the safe outcome (46%). When the agents' messages disagree, it is plausible that many managers believe that the messages are uninformative and any game is possible. The problem is such beliefs generally make the safe outcome the best response, not an extreme outcome. There is little evident of the type of punishment that underpins a partially informative equilibrium.

In **CH – MC**, unlike **SC – MC**, it is possible for the manager and agents to communicate freely, allowing them to coordinate on an equilibrium. A clever manager could possibly explain a partially informative equilibrium to the agents and then follow it, allowing them to do better than the babbling equilibrium. However, there is no empirical evidence of managers attempting this. Because truth telling is so prevelant in **CH – MC**, there is little reason for a manager to do so.

In our opinion, the partially informative equilibria are not empirically plausible. We see nothing in the data that suggests such equilibria are being played, and the nature of the equilibria makes it unlikely that such equilibria would spontaneously emerge. We see the partially informative equilibria of the MA game with managerial control as theoretical curiosities rather than realistic predictions for the data.

[49] The only case in **SC – MC** in which there is not full support is $\mu_1 = 1$, which was never sent for $\gamma = 2$. It is not a coincidence that $\mu_1 = 1$ is the rarest message in the dataset.

Appendix C: Translated Instructions

We include instructions for two of the treatments, **NC – D** and **SC – MC**. The rest of the instructions are available from the authors upon request. The Spanish words for row and column are "fila" and "columna". We have kept the original abbreviations, F and C, in the text and payoff tables.

INSTRUCTIONS (NC – D)

Thanks for coming to the experiment. You will receive 5 euros for participation in the experiment. Also, you will earn additional money during the experiment.

Participants have been randomly assigned to one of three roles: F, C and A. This role will be the same throughout the experiment.

There will be 18 separate periods. We will now present the instructions for the first block of nine periods. Later you will receive further instructions. In each period, you will be in a group of three participants, one in each role. The composition of each group is randomly determined at the beginning of the nine periods and stays constant during the nine periods. During the nine periods you will be with the same two persons. Also, at no time will you know the identity of who you are matched with.

Each period is independent from the others and develops in the following way. At the beginning of the period, the computer will randomly determine which of the following five games will be played.

In each of the cells the first number shown in yellow is the payoff that the person in the F role will receive, the second number shown in green is the payoff that the person in the C role will receive and the third number shown in red is the payoff for the person in the A role. As you can see all five games have five rows: f1, f2, f3, f4 and f5, and five columns; c1, c2, c3, c4 and c5. Observe also that the numbers in the different cells differ between the games.

Game 1

	c1	c2	c3	c4	c5
f1	26, 54, 80	12, 29, 41	-2, 4, 2	-16, -21,-37	-30,-46,-76
f2	15, 40, 55	29, 43, 72	15, 18, 33	1, -7, -6	-13, -32,-45
f3	4, 26, 30	18, 29, 47	32,32, 64	18, 7, 25	4, -18, -14
f4	-7, 12, 5	7, 15, 22	21, 18, 39	35, 21, 56	21, -4, 17
f5	-18, -2, -20	-4, 1,-3	10, 4, 14	24, 7, 31	38, 10, 48

Game 2

	c1	c2	c3	c4	c5
f1	22, 50, 72	8, 33, 41	-6, 8, 2	-20, -17, -37	-34,-42,-76
f2	19, 36, 55	33, 47, 80	19, 22, 41	5, -3, 2	-9, -28,-37
f3	8, 22, 30	22,33, 55	36, 36, 72	22, 11, 33	8, -14, -6
f4	-3,8, 5	11,19,30	25, 22, 47	39, 25, 64	25, 0, 25
f5	-14,-6,-20	0,5,5	14, 8, 22	28, 11, 39	42, 14, 56

Game 3

	c1	c2	c3	c4	c5
f1	18, 46, 64	4, 29, 33	-10,12, 2	-24, -13, -37	-38, -38, -76
f2	15, 32, 47	29, 43, 72	15, 26, 41	1, 1, 2	-13, -24,-37
f3	12, 18, 30	26, 29, 55	40, 40, 80	26, 15, 41	12, -10,2
f4	1,4, 5	15, 15, 30	29,26, 55	43,29, 72	29, 4, 33
f5	-10,-10,-20	4,1, 5	18,12, 30	32, 15, 47	46, 18, 64

Game 4

	c1	c2	c3	c4	c5
f1	14, 42, 56	0, 25, 25	-14, 8,-6	-28, -9, -37	-42, -34, -76
f2	11, 28, 39	25, 39, 64	11, 22, 33	-3, 5, 2	-17,-20,-37
f3	8, 14, 22	22, 25, 47	36, 36, 72	22, 19, 41	8, -6, 2
f4	5, 0, 5	19, 11, 30	33, 22, 55	47, 33, 80	33, 8, 41
f5	-6,-14,-20	8, -3, 5	22, 8, 30	36, 19, 55	50, 22, 72

Game 5

	c1	c2	c3	c4	c5
f1	10, 38, 48	-4, 21, 17	-18, 4,-14	-32,-13,-45	-46,-30, -76
f2	7, 24, 31	21, 35, 56	7, 18, 25	-7, 1,-6	-21,-16,-37
f3	4, 10, 14	18, 21, 39	32, 32, 64	18, 15, 33	4, -2, 2
f4	1,-4,-3	15, 7, 22	29, 18, 47	43, 29, 72	29, 12, 41
f5	-2,-18,-20	12,-7, 5	26,4, 30	40, 15, 55	54, 26, 80

Each of the five games has the same chance of being chosen in each period separately. That is in each period, each of the games will be chosen with 20% probability. Player F and player C will be informed of which game has been chosen, but player A will not be informed of which game has been chosen.

After having seen which game has been selected by the random draw, players F and player C will separately make decisions. Player F will choose between f1, f2, f3, f4 and f5 and player C will choose between columns c1, c2, c3, c4 and c5. Player A will not make any decisions.

The payoffs of players F, C and A will be the ones in the cell determined by the row chosen by F and the column chosen by C for the game selected by the random draw. Remember that players F and C will make their decisions independently from each other.

After each period everybody will be informed about what row was chosen by F and what column was chosen by C sent and about which game was randomly selected.

After this, a new period will start which will develop in the same way until reaching period 9. Remember that the persons you play with will not change from period to period.

Each ECU is worth 0,02 euros. At the end of the session you will receive 5 euros plus what you will have earned in all 18 rounds of the experiment.

You can ask questions at any time. If you have a question, please raise your hand and one of us will come to your place to answer it.

[Block 2] The rules will not change for the second block of nine periods. The persons you play with are the same as the first nine periods.

INSTRUCTIONS (SC – MC)

Thanks for coming to the experiment. You will receive 5 euros for participation in the experiment. Also, you will earn additional money during the experiment.

Participants have been randomly assigned to one of three roles: F, C and A. This role will be the same throughout the experiment.

There will be 18 separate periods. We will now present the instructions for the first block of nine periods. Later you will receive further instructions. In each period, you will be in a group of three participants, one in each role. The composition of each group is randomly determined at the beginning of the nine periods and stays constant during the nine periods. During the nine periods you will be with the same two persons. Also, at no time will you know the identity of who you are matched with.

Each period is independent from the others and develops in the following way. At the beginning of the period, the computer will randomly determine which of the following five games will be played.

In each of the cells the first number shown in yellow is the payoff that the person in the F role will receive, the second number shown in green is the payoff that the person in the C role will receive and the third number shown in red is the payoff for the person in the A role. As you can see all five games have five rows: f1, f2, f3, f4 and f5, and five columns; c1, c2, c3, c4 and c5. Observe also that the numbers in the different cells differ between the games.

Game 1

	c1	c2	c3	c4	c5
f1	26, 54, 80	12, 29, 41	-2, 4, 2	-16, -21,-37	-30,-46,-76
f2	15, 40, 55	29, 43, 72	15, 18, 33	1, -7, -6	-13, -32,-45
f3	4, 26, 30	18, 29, 47	32,32, 64	18, 7, 25	4, -18, -14
f4	-7, 12, 5	7, 15, 22	21, 18, 39	35, 21, 56	21, -4, 17
f5	-18, -2, -20	-4, 1,-3	10, 4, 14	24, 7, 31	38, 10, 48

Game 2

	c1	c2	c3	c4	c5
f1	22, 50, 72	8, 33, 41	-6, 8, 2	-20, -17, -37	-34,-42,-76
f2	19, 36, 55	33, 47, 80	19, 22, 41	5, -3, 2	-9, -28,-37
f3	8, 22, 30	22,33, 55	36, 36, 72	22, 11, 33	8, -14, -6
f4	-3,8, 5	11,19,30	25, 22, 47	39, 25, 64	25, 0, 25
f5	-14,-6,-20	0,5,5	14, 8, 22	28, 11, 39	42, 14, 56

Game 3

	c1	c2	c3	c4	c5
f1	18, 46, 64	4, 29, 33	-10,12, 2	-24, -13, -37	-38, -38, -76
f2	15, 32, 47	29, 43, 72	15, 26, 41	1, 1, 2	-13, -24,-37
f3	12, 18, 30	26, 29, 55	40, 40, 80	26, 15, 41	12, -10,2
f4	1,4, 5	15, 15, 30	29,26, 55	43,29, 72	29, 4, 33
f5	-10,-10,-20	4,1, 5	18,12, 30	32, 15, 47	46, 18, 64

Game 4

	c1	c2	c3	c4	c5
f1	14, 42, 56	0, 25, 25	-14, 8,-6	-28, -9, -37	-42, -34, -76
f2	11, 28, 39	25, 39, 64	11, 22, 33	-3, 5, 2	-17,-20,-37
f3	8, 14, 22	22, 25, 47	36, 36, 72	22, 19, 41	8, -6, 2
f4	5, 0, 5	19, 11, 30	33, 22, 55	47, 33, 80	33, 8, 41
f5	-6,-14,-20	8, -3, 5	22, 8, 30	36, 19, 55	50, 22, 72

Game 5

	c1	c2	c3	c4	c5
f1	10, 38, 48	-4, 21, 17	-18, 4,-14	-32,-13,-45	-46,-30, -76
f2	7, 24, 31	21, 35, 56	7, 18, 25	-7, 1,-6	-21,-16,-37
f3	4, 10, 14	18, 21, 39	32, 32, 64	18, 15, 33	4, -2, 2
f4	1,-4,-3	15, 7, 22	29, 18, 47	43, 29, 72	29, 12, 41
f5	-2,-18,-20	12,-7, 5	26,4, 30	40, 15, 55	54, 26, 80

Each of the five games has the same chance of being chosen in each period separately. That is in each period, each of the games will be chosen with 20% probability. Player F and player C will be informed of which game has been chosen, but player A will not be informed of which game has been chosen.

After having seen which game has been selected by the random draw, players F and player C will separately send messages to player A saying which game has been selected . This message can be truthful or not. Once player A has received the messages he will choose a row and column without knowing which game was selected.

The payoffs of players F, C and A will be the ones in the cell determined by the row and the column chosen by A for the game selected by the random draw. Remember that players F and C will send their messages independently from each other.

After each period everybody will be informed about what row and what column was chosen by A and about which game was randomly selected.

After this, a new period will start which will develop in the same way until reaching period 9. Remember that the persons you play with will not change from period to period.

Each ECU is worth 0,02 euros. At the end of the session you will receive 5 euros plus what you will have earned in all 18 rounds of the experiment.

You can ask questions at any time. If you have a question, please raise your hand and one of us will come to your place to answer it.

[Block 2] The rules will not change for the second block of nine periods. The persons you play with are the same as the first nine periods.

Appendix D: Information About Additional Treatments and Reputational Effects

Beyond the seven treatments reported in the main text, we ran an additional four treatments. These were modifications of the **NC – D** and **SC – MC** treatments, and involved exploring how behavior changes either with use of a strangers matching or an increase in the value of k_4, the parameter governing the state losses.

*Strangers (**STR – D** and **STR – MC**):* These treatments were identical to **NC – D** and **SC – MC** except for use of strangers matching; that is, groups changed from round to round. The matching was constructed so no participant met another person twice in a nine-round block (a point which the instructions stressed). At no time were participants informed about the identities of the other two people in their group. Because groups were not independent within a session, we conducted five sessions per treatment rather than three. There were 27 subjects in each session.

Our design focused on partners matching as the natural case since we are interested in the effect of organizational structure within long-lasting organizations. We conjectured that the repeated interactions helped groups, making it easier to coordinate on efficient coordination since taking turns was more direct and improving information transmission by strengthening the reputational reasons to tell the truth with managerial control (although the babbling equilibrium with play of the safe outcome is the unique pure strategy equilibrium for both types of matching). The Strangers treatments are a robustness check, testing whether the results of **NC – D** and **SC – MC** were sensitive to what type of matching was used, specifically whether surplus would be lower with strangers matching.

*High State Losses (**HSL – D** and **HSL – MC**):* These treatments were identical to **NC – D** and **SC – MC** except we increased state losses ($k_4 = 6$ vs. $k_4 = 4$). Increasing k_4 does not affect the theoretical predictions for the game under either delegation or managerial control since it remains true that adaptation losses are greater than state losses ($k_3 > k_4$), but the difference between adaptation and state losses is minimized. This reduces the tension between agents since the gain for moving from the efficient outcome to an agent's most preferred outcome is tiny. For example, moving from Outcome 1 to Outcome 5 in Game 1 gains S1 12 ECUs in **NC - D**, but only 4 ECUs in **HSL – D**.

The High State Losses treatments are a second robustness check. Previous experiments with asymmetric coordination games (e.g. the battle-of-the-sexes game) suggest that achieving coordination, let alone efficient coordination, will be challenging with delegation. Even if coordination occurs, the tension between agents makes the safe outcome attractive, sacrificing efficiency in order to achieve coordination. Under managerial control, achieving efficiency is difficult because the tension between agents provides a strong incentive to deceive the manager. The High State Losses treatments weaken tension between the agents, making the safe outcome less attractive. This lets us explore how the results change when the environment is less challenging.

Table D.1: Summary of Outcomes, Additional Treatments, Rounds 10 - 18

Treatment	% Coordinate	% Efficient (s.t. Coordinate)	Total Surplus	Efficiency Gain
NC – D	69.5%	39.1%	61.4	-108.3%
STR – D	46.7%	44.4%	55.0	-181.4%
HSL – D	59.3%	84.0%	60.0	-58.6%
SC – MC	99.6%	45.9%	71.7	6.5%
STR – MC	99.5%	35.5%	70.3	-9.1%
HSL – MC	97.5%	55.7%	68.2	13.5%

Table D.1 summarizes the results for the additional treatments, paralleling Table 2 in the main text. The results from **NC – D** and **SC – MC** are included as points of comparison. Performance is somewhat weaker in the two strangers treatments than the parallel partners treatments. The differences are not large and are not significant for either delegation (p = .147) or managerial control (p = .289). High State Losses should make matters easier, and efficiency gains are better with either delegation or managerial control.[50] This effect looks large with delegation, but the improvement is not significant for either delegation (p = .246) or managerial control (p = .191). Efficiency gains are significantly lower in **HSL – MC** than for **CH – MC** (p = .067). In other words, a strong increase in incentives to play the efficient equilibrium has significantly less impact than allowing free-form communication.

Reputation: In Section 5.2.a, we noted that a possible explanation for managerial errors in **SC – MC** is repeated game effects. Over time, agents have the opportunity to establish a reputation for truthfulness. Agents could benefit from building reputations for being truthful and then abandoning them at an opportune moment. Likewise, managerial mistakes could reflect attempts to take advantage of said reputations. Because **STR – MC** uses a strangers matching, making reputational effects unlikely, comparing **STR – MC** and **SC – MC** gives a clean measure of reputation effects.

There is strong evidence that managers in **SC – MC** respond to past information about agents' truthfulness. Define $T1_t$ as a dummy for whether S1 told the truth in Period t and $T2_t$ as the equivalent dummy for S2. Define the "Truth Index" for Period t as ($T1_{t-1} - T2_{t-1}$). This index is equal to zero if both agents told the truth in the previous period or both lied in the previous period. If only S1 told the truth, the Truth Index is equal to 1. Likewise, the Truth Index is equal to -1 if only S2 told the truth.

[50] The expected surplus from the babbling equilibrium is lower with High State Losses. Thus, lower surplus in the two HSL treatments doesn't imply weaker performance. Comparing efficiency gains puts the treatments on equal footing.

With this definition in hand, we ran two regressions (linear probability models) where the dependent variable is the outcome chosen by the manager.[51] The independent variables are the Truth Index as defined above, the messages sent by each of the agents, and dummies for the current period. Data from Round 1 is dropped as the Truth Index is based on lagged variables. Data is drawn from **SC – MC** and **STR – MC**.

The regression results are reported in Table D.2. Standard errors are corrected for clustering at the group level for **SC – MC** and the session level for **STR – MC**. Given the small number of clusters (five) for **STR – MC**, the results are biased in favor of significance. The critical estimate is the coefficient for the Truth Index. Recall that S1 always prefers higher outcomes, subject to coordination, and S2 always prefers lower outcomes. We would therefore expect a shift in favor of S1 if S1 has told the truth and S2 has not (Truth Index = 1). Likewise, a shift in favor of S2 is expected if S2 has told the truth and S1 has not (Truth Index = -1). Together, these observations imply that the coefficient for the Truth Index should be positive. For **SC – MC**, as shown by Model 1, this prediction is strongly confirmed. The effect isn't huge, with an estimated shift of about a third of an action in favor of a truthful agent, but is easily significant at any standard level (p = .001). This is *not* true in **STR – MC**, where reputation should not have any impact. For Model 2, the estimate for the Truth Index is small, nowhere close to statistical significance (p = .927), and has the wrong sign.

Table D.2: The Effect of Past Agent Truthfulness on Manager Choices

	Model 1	Model 2
Dataset	SC – MC	STR – MC
S1 Message	0.370***	0.369***
	(0.047)	(0.048)
S2 Message	0.399***	0.307***
	(0.051)	(0.062)
Truth Index	0.330***	-0.007
	(0.085)	(0.070)
Observations	454	757
R-squared	0.376	0.212

Notes: Standard errors (in parentheses) are corrected for clustering at the group level for Model 1 and for clustering at the session level for Model 2. Both regressions include controls for the current period, which are not reported to save space. Three (***), two (**), and one (*) stars indicate significance at the 1%, 5%, and 10% levels using two-tailed tests.

[51] The dependent variable relies on the manager choosing $A_1 = A_2$. We therefore drop all observations where the manager does not choose coordination ($A_1 = A_2$). Given that the coordination rate exceeds 98% for both SC – MC and STR – MC, the two treatments considered in these regressions, we doubt that the omitted data much matters.

While managers respond to the prior truthfulness of their agents, it does not appear that agents try to take advantage of this. If agents anticipate the response of managers to past truthfulness, they should tell the truth more often in **SC – MC** than **STR – MC**. Moreover, the difference should be most extreme for games where the cost of telling the truth is low. In other words, agents should build trust in low-importance games and take advantage of it when it suits them the most.

The top part of Table D3 reproduces Table 4 from the text, except, to ease comparisons between treatments, it reports percentages for each message as a function of the game.[52] The bottom part of Table D3 is the equivalent table for **STR – MC**.

Table D3: Messages as a Function of Game

SC - MC

		Game (Mapped)				
		1	2	3	4	5
Message (Mapped)	1	24.9%	0.5%	1.3%	2.1%	2.3%
	2	1.1%	36.4%	1.3%	1.0%	1.1%
	3	10.7%	13.9%	68.0%	2.6%	4.5%
	4	18.1%	13.9%	9.2%	73.3%	4.5%
	5	45.2%	35.4%	20.2%	21.0%	87.6%

STR - MC

		Game (Mapped)				
		1	2	3	4	5
Message (Mapped)	1	15.3%	2.5%	1.6%	1.2%	0.3%
	2	5.1%	21.2%	1.6%	1.2%	0.3%
	3	7.1%	12.3%	50.8%	2.5%	1.7%
	4	20.0%	16.6%	9.0%	55.4%	4.4%
	5	52.5%	47.4%	37.1%	39.7%	93.2%

The tables for the two treatments are similar, but not identical. Looking at the diagonals, the numbers are almost always higher for **SC – MC**. As predicted, agents are more truthful when reputation is an issue. However, the difference is not obviously larger for games that are low-

[52] Note that both messages and games are remapped to allow us to combine S1 and S2 data.

importance. Indeed, the only game where the percentage of truth-telling is *lower* for **SC – MC** is $\gamma = 5$, the case where telling the truth is least harmful for agents!

To make the preceding point on a more formal basis, we run a pair of probit regressions where the dependent variable is a dummy for whether a agent told the truth. Data is taken from all rounds of the **SC – MC** and **STR – MC** treatments. There are two observations per round, one for each of the agents. Independent variables include a dummy for the **SC – MC** treatment, an interaction between this treatment and the (mapped) game, dummies for the current game, and dummies for the current period. Standard errors are corrected for clustering at the group level for **SC – MC** and the session level for **STR – MC**. We report marginal effects, and the estimates for the (mapped) game and period dummies are suppressed to save space.

Model 1 only includes the treatment dummy for **SC – MC**. As expected, the estimate for the treatment dummy is large, positive, and statistically significant. Agents are more truthful when there is a benefit to maintaining a reputation. Model 2 adds the interaction term between the treatment dummy and the (mapped) game. If agents build trust in low-importance games and take advantage of it when it suits them the most, this parameter should be positive. From S1's point of view, they should be most truthful relative to **STR – MC** when $\gamma = 5$ and there is no benefit to lying and least truthful in $\gamma = 1$ when there is the most to gain by deceiving the manager. The estimate for the interaction term is weakly significant, but has the wrong sign. This does not imply that agents do not consider their reputations at all, but they do not seem to be sufficiently strategic that they carefully parse when a reputation is most worth maintaining.

Table D4: Truth-telling and Reputation

	Model 1	Model 2
SC - MC	0.141***	0.248***
	(0.048)	(0.076)
SC – MC *		-0.039*
Game (Mapped)		(0.022)
Observations	2,592	2,592
Log-likelihood	-1382.62	-1380.34

Notes: Marginal effects are reported. Standard errors (in parentheses) are corrected for clustering at the group level for Model 1 and for clustering at the session level for Model 2. Both regressions include controls for the current game (mapped) and period, which are not reported to save space. Three (***), two (**), and one (*) stars indicate significance at the 1%, 5%, and 10% levels using two-tailed tests.

While reputational effects play a role in **SC – MC**, we are dubious that the choices by managers which are identified as mistakes instead represent optimal strategic decisions where managers take advantage of what they have learned about agents' honesty to increase their payoffs. One reason for our doubts comes from comparisons of **SC – MC** and **STR – MC**. If manager choices we identify as mistakes are in fact largely due to reputational factors (or other factors due to repeated play), there should be far fewer mistakes in **STR – MC** where such factors cannot play a role. This is not the case. One common mistake is failing to follow the agents when they agree. In **SC – MC**, 18% of managers do not choose A1 = M1 and A2 = M2 if M1 = M2. This number is 17% is **STR – MC**, slightly lower but hardly a dramatic difference. Another commonly observed mistake occurs when M1 = 5 and M2 = 1. The optimal response to these uninformative messages is for the manager to choose the safe outcome, but only 35% of managers do so in **SC – MC**. This rises to 40% in **STR – MC**, only mildly better. An additional reason for our doubts is that these mistakes are costly on average (see p. 24). If mistakes actually represent managers using what they have learned to increase their payoffs, these choices should be associated with increased expected payoffs. Perhaps managers tried to take advantage of a history of honesty and failed, but the poor payoffs associated with mistakes make it unlikely that these choices reflected an unidentified optimal strategy.

The possibility of reputational effects exists for **CH – MC**, just as for **SC – MC**, but we see little evidence that the lack of lying in **CH – MC** reflects reputational concerns. Nothing in the chat indicates a role of reputation (i.e. we can imagine agents being warned to tell the truth lest they damage their reputation or being told that they are no longer believed because of a history of lying). Moreover, agents become more truthful with experience in **CH – MC**. A model in which agents build reputations for truthfulness in a strategic fashion would predict the opposite. The most striking empirical finding in our paper is the disappearance of lying in **CH – MC**. There is little reason to attribute this to reputational concerns.

Appendix E

Table E1: Detailed Description of Coding

1) Make a suggestion about what row/column should be chosen. (Coder always recorded the specific suggestion that was made.)
    a. Suggest safe outcome
    b. Suggest efficient outcome
2) Agree to proposal about what row/column should be chosen.
3) Discussion about what row/column should be chosen.
    a. Discuss need for coordination (pick same row & column). This requires more than making a suggestion that involves coordination. The message needs to indicate that the two players should be choosing the same thing (e.g. "We'll do better if we make the same choices" is coded. "Let's choose row 4 and column 4" is not coded.)
    b. Discuss fairness. This category includes any message that discusses the distribution of pay over the three players.
4) Discuss Efficiency: This includes discussion of maximizing total pay as well as explaining how and why rotation between players works.
5) Questions About Rules of the Experiment: This includes questions about either the rules of the experiment (e.g. "Do I choose a row or does [the manager] choose for me?") or the game (e.g. "Is the third number my payoff?").
6) Questions About How to Play: This was for conceptual questions rather than the frequent generic request that somebody suggest a row and column.
7) Explanation: This included explanations about the rules of the experiment or game, as well as explanations of a suggested way of playing the game.
8) (**CH/A – D** and **CH – MC** Only) M Asks What Game is Being Played.
9) (**CH/A – D** and **CH – MC** Only) Agents Report What Game is Being Played.
    a. Truthfully Reveal Game
    b. Lie About Game
    c. Conflict: This is used for cases where there was "fact-checking". (e.g. "S1: It is Game 3." "S2: No, it is Game 2.") This category is different from the "Contradict" category reported in Table 6, which is a combination of 9a and 9b. This category is the basis for the discussion of fact-checking in the text.