11-10

September 6, 2011

ESMT Working Paper

DYNAMIC COORDINATION VIA ORGANIZATIONAL ROUTINES

ANDREAS BLUME, UNIVERSITY OF PITTSBURGH
APRIL M. FRANCO, ROTMAN SCHOOL OF MANAGEMENT, UNIVERSITY OF TORONTO
PAUL HEIDHUES, ESMT

ISSN 1866-3494

Abstract

Dynamic coordination via organizational routines[+]

Author(s):* Andreas Blume, University of Pittsburgh
April M. Franco, Rotman School of Management, University of Toronto
Paul Heidhues, ESMT

We investigate dynamic coordination among members of a problem-solving team who receive private signals about which of their actions are required for a (static) coordinated solution and who have repeated opportunities to explore different action combinations. In this environment ordinal equilibria, in which agents condition only on how their signals rank their actions and not on signal strength, lead to simple patterns of behavior that have a natural interpretation as routines. These routines partially solve the team's coordination problem by synchronizing the team's search efforts and prove to be resilient to changes in the environment by being ex post equilibria, to agents having only a coarse understanding of other agents' strategies by being fully cursed, and to natural forms of agents' overconfidence. The price of this resilience is that optimal routines are frequently suboptimal equilibria.

Keywords: coordination games, organizational routines, decentralized information, ex-post equilibria, cursed equilibria, multi-agent learning, rational learning

JEL Classification: C73, D23

* Contact: Paul Heidhues, ESMT, Schlossplatz 1, 10178 Berlin,
Phone: +49 (0) 30 21231-1536, paul.heidhues@esmt.org.

+ We thank Tim Grebe, Zhike Lei, Ariel Rubinstein, Yuuso Välimäki, Mary Zellmer-Bruhn, and especially Robert Gibbons for helpful comments. We are also grateful to seminar and conference audiences at the Berlin Behavioral Economics Workshop, University of Bonn, UNC/Duke Theory Workshop, Johns-Hopkins University, University of Pittsburgh, MIT/Sloan Seminar in Organizational Economics, Econometric Society Summer Meetings, SFB/TR GESY conference in Mannheim, conference on Strategic Information Acquisition and Transmission in Munich, and the Society for Economic Dynamics. Paul Heidhues acknowledges financial support from the German Science Foundation (SFB/TR GESY).

...the essence of an organizational routine is that individuals develop sequential patterns of interaction which permit the integration of their specialized knowledge without the need for communicating that knowledge.

R.M. Grant [*Organization Science*, 1996]

1 Introduction

Much of the knowledge in organizations is held by individuals and thus distributed within organizations. This creates opportunities for organizations to come up with mechanisms that generate value from and sustain competitive advantage through integrating this distributed knowledge, as noted by Grant [1996]. Importantly, some of the knowledge held by the members of an organization is not in declarative form and thus not easily communicated; e.g. Polanyi [1966] emphasizes that tacit knowledge is incommunicable, Hayek [1945] that "knowledge of the particular circumstances of time and place" is hard to centralize, and March and Simon [1958] that knowledge transfer within organizations is severely limited by "language incompatibility".[1] In all of these circumstance other mechanisms than communication have to be put in place to coordinate the use of such decentralized knowledge. Sometimes organizational routines, that is *persistent patterns of behavior among members of an organization with distributed knowledge*, can serve as the mechanism for knowledge integration.[2]

These patterns of behavior among multiple interacting agents may be more or less well adapted to the problem at hand and yet difficult to undo given their equilibrium nature. Routines that are robust within organizations and not easily transferred across organizations can explain the empirical puzzle of persistent performance differences among organizations that otherwise operate under similar conditions (see Gibbons [2010]).[3] In particular, well-adapted routines add

[1] In a similar vein, Dewatripont and Tirole [2005] highlight that "the acts of formulating and absorbing the content of a communication are privately costly, and so *communication is subject to moral hazard in teams...*". Thus even in environments in which there is common interest about the decision to be taken, successful communication cannot be taken for granted. Stasser and Titus [1985] show experimentally how communication fails to aggregate information when individuals have common interests over actions for every state of the world. In their case, this is a result of communication being consensus confirming; discussion focusses on commonly held information and supports choices that are optimal given individuals' prior information.

[2] Organizational routines have long been an object of study (e.g. Nelson and Winter [1982]) and continue to attract attention as a unit of analysis of organizational behavior (e.g. Cohen and Bacdayan [1994]). Much of that literature is reviewed in Becker [2004], who notes that the terminology surrounding routines is not entirely settled but mentions patterned behavior and distributed knowledge as frequently being associated with routines.

[3] Gibbons [2010] comprehensively surveys the evidence on persistent performance differences and argues that "it seems important for organizational economists to study" these. Classic case studies include Salter [1960] on the observed performance differences in the British pig-iron industry during the early 20th century, Argote, Beckman and Epple [1990] on the US war-ship production during World War II, and Chew, Bresnahan and Clark [1990] on the persistent performance differences between plants from a commercial food company. Syverson [2011] surveys the literature on the determinants of productivity, and also highlights the relationship between persistence of firm-

to the dynamic capabilities of organizations and may create persistent competitive advantages. In that spirit we propose a simple stylized model of an organization that faces a dynamic decision problem that can be addressed in a variety of ways—including organizational routines. We offer a natural sense in which routines are robust and therefore might be difficult to unseat, even if they are Pareto inferior among routines. Optimal routines exist, but are difficult to find given the huge multiplicity of routines and given that optimality depends on details of the environment, which may change over time. Even optimal organizational routines in our model are not optimal equilibria for the organization and do not solve the informationally constrained social planner's problem. We thus have a representation of organizational dynamics in which we can naturally distinguish between full-information optimality, in which all the distributed knowledge of the organization's members has been made public, informationally constrained optimality, in which the organization optimally utilizes the private information of its members, and robust optimality, where the organization adopts a routine that is best among routines. Hence in addition to generating persistent performance differences among organizations from multiple equilibria, we point out that these equilibria differ in terms of robustness and that even among robust equilibria optimal ones may not be easy to identify. Our model therefore formalizes the argument of Cohen and Bacdayan [1994] that organizational routines are typically suboptimal since they are not tailored to every specific situation. We furthermore emphasize the robustness of these organizational routines to a variety of behavioral biases, and show through an example that seemingly suboptimal routines may be optimally selected by a management realizing that the members of the organization are excessively overconfident.

Formally, we investigate dynamic coordination among members of a problem-solving team who receive private signals about which of their actions are required for a (static) coordinated solution and who have repeated opportunities to explore different action combinations. There is exactly one profile of actions that results in a positive payoff and the problem is solved once the team identifies that profile. This "success profile" remains the same during the entire course of the interaction. Team members can explore different action combinations by trial and error. They cannot communicate either directly, or indirectly through observing each others actions. At the beginning of time each team member gets a signal that indicates for each of her actions the probability of that action being part of the success profile. In later periods the only additional information of each team member is the history of her own actions up to that point in time.

We show that the set of equilibria of the game that we investigate can naturally be split into two classes, ordinal equilibria and their complement, cardinal equilibria. Ordinal equilibria, in

and plant-level productivity and the nature of intangible capital—that is know-how embodied in organization—as one of the big research questions to be addressed.

which by definition players condition only on how their signal ranks their actions and not the strength of their signal, are remarkably robust and have a natural interpretation as routines. They are *ex post* equilibria and therefore do not depend on the distributions of signals, players' beliefs about these distributions, or higher-order beliefs etc. They also are (fully) cursed—that is consistent with players having a coarse perception of how other players' information affects their play (Eyster and Rabin [2005])—, and robust to natural specifications of overconfidence by team members. In an ordinal equilibrium the only information a player needs to assess the optimality of her own strategy is the pattern of behavior of other players, regardless of how that behavior depends on other players' information.

We identify organizational routines as *patterns of behavior among multiple interacting agents with distributed knowledge.* Distributed knowledge is a characteristic of the environment we study. Patterns of behavior are attributes of a class of equilibria in this environment: In an ordinal equilibrium the members of the organization make only limited use of the private information that is available to them and conditional on a rank-ordering of their actions follow a fixed predetermined schema of action choices. One play according to an ordinal equilibrium is then to be thought of as one instantiation of the routine. The recurrence that is widely held to characterize routines is captured by the independence of ordinal equilibrium behavior from some of the details of the game; neither need agents know the exact generating process for their private information, nor need they know what other player believe this process to be. Thus the same behavior pattern remains an equilibrium across an entire array of possible situations. We can think of routines in our setting either as the result of learned behavior, e.g. if after each play of the game actions and payoffs become public, or as the result of infrequent managerial intervention. According to the latter interpretation, whenever the expected benefits of resetting a routine exceed the costs of information acquisition, management collects data to identify the true signal generating process and prescribes a routine that is optimal for that process. Routines in that case are the result of optimizing behavior subject to deliberation and informational constraints, akin to standard operating procedures.

To summarize our results, we find that routines partially solve the team coordination problem. They synchronize the team's search efforts and help avoid repetition inefficiencies where the same action profile is tried more than once. They are resilient to changes in the environment (signal distributions, agents' beliefs about these distributions, beliefs about these beliefs etc.) and therefore can serve as focal points across a range of search problems. Routines are fully cursed equilibria and thus robust to a lack of full strategic sophistication by team members. Furthermore, routines are robust to various forms of information-processing mistakes—such as overconfidence in the ability to predict one's correct action—of the team members. This resilience of routines, however, comes with a two-fold cost: First, routines may become outdated; a routine that was optimal (among routines) for a given set of conditions may not fit current conditions. Second, even optimal routines are generally suboptimal problem-solving strategies for the team; under a wide range of conditions the team would be better off to give more discretion to its members by letting their behavior be more sensitive to the quality of their information. We also, however, highlight through a simple example that the latter conclusion depends on the team members being fully rational: in the presence of information-processing mistakes such as overconfidence by team members, routines can be strictly optimal.

The paper is organized as follows. In the next section we discuss related literature. In Section 3 we provide an illustrative example in which we highlight our main findings; in Section 4 we set up the general model; in Section 5 we characterize the set of ordinal equilibria, discuss the robustness of these routines to distributional misspecifications and behavioral biases, prove that routines are typically suboptimal problem-solving approaches, and characterize the optimal problem-solving solution; and in Section 6 we discuss possible extensions of our framework that can formally address a variety of questions informally raised in the organizational routine literature.

2 Related Literature

We analyze how organizations coordinate their search efforts over time. Conceptually, our framing of coordination as a constrained maximization problem is reminiscent of the approach introduced by Crawford and Haller [1990]. In general, efforts to coordinate can be affected by a variety of constraints, including strategic uncertainty, lack of precedent, conflicting incentives, absence of communication, imperfect observability, and private information as well as behavioral biases of the team members such as lack of strategic sophistication and mistakes in information processing. Crawford and Haller [1990] study the question of how to achieve static coordination by way of repeated interaction in an environment where the constraint is that players lack a common language for their actions and roles in a game. They model such absence of a common language through requiring that players use symmetric strategies and treat all actions symmetrically that have not been distinguished by past play. Coordination in their setting is achieved via the common observation of precedents that are created by the history of play and that help desymmetrize actions and player roles.

In contrast, we principally focus on the constraint that is imposed by players having private information about payoffs, while ruling out communication and making actions unobservable.

As in Crawford and Haller [1990], incentives are perfectly aligned. Therefore we have a team problem and can frame the coordination question as one of maximizing the team's joint payoff subject to its informational, observational, rationality, and communication constraints.

Coordinating as quickly as possible is also at the heart of Alpern's [1976] *telephone problem*:[4] There is an equal number of telephones in two rooms. They are pairwise connected. In each period a person in each room picks up the receiver on one of the phones. The goal is to identify a working connection in minimum expected time. Unlike in Crawford and Haller's work, in the telephone problem there is uncertainty about which action combination leads to coordination (i.e. a working connection). Hence players face a two-fold constraint. In addition to lacking a common language that would permit them to implement an optimal search pattern from the outset, they also cannot use observations of past actions to create precedents for search patterns.

Blume and Franco [2007] study dynamic coordination in a search-for-success game in which players have an identical number of actions, some fraction of action profiles are successes and, as in the telephone problem, players cannot observe each others' actions. They show that in an optimal strategy that respects the symmetry constraints of Crawford and Haller, players will revisit action profiles by chance, and that this may occur even before all possibilities of guaranteeing the visit of a novel profile has been exhausted. Blume, Duffy and Franco [2009] find experimental evidence for such behavior in a simple version of the search-for-success game.

In contrast to this literature, where symmetry is the principal constraint, in the present setting coordination on an optimal search pattern is difficult because the problem-solving knowledge is distributed throughout the organization: Each player knows privately for each of her actions how likely it is that this action is required for a coordinated solution. Implementing the *ex post* optimal search pattern, however, requires knowing every team members' private information.

Our modeling of routines as equilibrium behavior is reminiscent of Chassang [2010]. He studies efficient cooperation between agents with conflicting interests and asymmetric information about what productive actions are available. Over time the common history helps reduce the asymmetric information and enables players to coordinate better. Optimal equilibria in his model are history-dependent, learning remains incomplete, and hence his model generates endogenous performance differences among otherwise similar organizations. Limiting behavior in Chassang's setting is routine in the sense that agents eventually settle on a fixed set of acceptable actions rather than exploring new ones. Our model shares the feature that routines are equilibria. In contrast to and complementing Chassang our emphasis is on the robustness of action patterns. Behavior is routine in this sense if it is not sensitive to the details of the

[4]For related problems, see Alpern and Gal [2003].

environment and hence agents make only limited use of the information available to them.[5] Our routines encompass a subclass of equilibria that are robust to misspecifications of the environment, a variety of behavioral biases and rationality constraints but whose robustness comes at the cost of suboptimality. While, unlike Chassang, we do not endogenously predict performance differences, the robustness of our routines can account for their persistence.

More broadly, the problem that interests us is related to other models of rational learning of payoffs in games, e.g. Wiseman's [2005] work on repeated games with unknown payoff distributions and Gossner and Vieille's [2003] work on games in which players learn their payoffs from experience. Another prominent example is the work on social learning, e.g. Banerjee [1992], Bikchandani, Hirshleifer and Welch [1992], and the recent book by Chamley [2004].

We have emphasized that certain forms of knowledge are not easily communicated (Polanyi [1966], Hayek [1945], March and Simon [1958]).[6] Imperfect communication in organizations has been examined by Crémer, Garicano and Prat [2007] who study optimal organizational codes subject to a coarseness constraint in a static common-interest environment with private information and Ellison and Holden [2010] who look at the emergence of coarse codes in a dynamic setting where they assume that it is difficult to communicate complete contingent plans.[7] In the extreme, communication is ruled out entirely. This is the case in the Condorcet-jury-theorem literature (e.g. Austen-Smith and Banks [1996] and McLennan [1998]), which studies how agents aggregate decentralized knowledge via voting but without communication.

[5] Miller [2011] uses an *ex post* incentive compatibility condition to express robustness and select equilibria in repeated games with private monitoring and applies this approach to understand price wars in cartels. He finds that under certain conditions robust collusion is inefficient and may require price wars. This parallels the finding in our model that optimal routines, and hence optimal *ex post* equilibria, are suboptimal.

[6] Coordination failures from a lack of communication have been documented for various organizations and events. For example, Amy C. Edmondson [2004] attributes the frequent lack of learning from failure in health care teams to inadequate communication. She finds in her empirical work that "process failures in hospitals have systemic causes, often originating in different groups or departments from where the failure is experienced, and so learning from them requires cross departmental communication and collaboration." Lack of communication also contributed to the failure of the rescue mission during the Iran hostage crisis. In the interest of maintaining secrecy and through it operational security the rescue team maintained complete radio silence. All communication between helicopters was through light signals and when helicopters became separated in a dust cloud vital information was not communicated. According to the Rescue Mission Report of the Department of the Navy [1980]: "The lead helicopter did not know that #8 had successfully recovered the crew from #6 and continued nor that #6 had been abandoned in the desert. More importantly, after he reversed course in the dust and landed, the lead could not logically deduce either that the other helicopters had continued or that they had turned back to return to the carrier. He did not know when the flight had disintegrated. He could have assumed that they had become separated before he reversed course and unknowingly proceeded. Alternatively, they could have lost sight of him after turning and, mistaking his intentions, continued back to the carrier. Lastly, #5 might have elected to continue had he known that his arrival at Desert One would have allowed the mission to continue and that VMC existed at the rendezvous." (VMC=visual meteorological conditions.)

[7] More distantly related is the literature studying how incentive problems limit the scope for communication of decentralized knowledge in organizations. See for example Alonso, Dessein and Matouschek [2008] and the references therein.

Here we also limit ourselves to the no-communication case, in part to avoid having to choose a particular coarse-communication regime but also to develop a benchmark for what different communication regimes can achieve.

3 An Illustrative Example

In this section, we introduce a simple example that illustrates our more general findings regarding organizational routines. There are two players $i = 1, 2$, each of whom has two actions. Of these four possible action combinations, one leads to successful coordination with a contemporaneous payoff that is normalized to one, while all other action profiles yield a payoff of zero. If players successfully coordinate in the first period, the game ends. Otherwise, they choose an action in Period 2, after which the game ends. Payoffs from the second period are discounted according to a common discount factor $\delta \in (0, 1)$.

Each player has some private knowledge regarding the likelihood of each of her own actions being part of the success profile. Formally, Player i has the action set $A_i = \{a_{i1}, a_{i2}\}$. Before choosing an action, Player i receives a private signal vector $\omega_i = (\omega_{i1}, \omega_{i2})$. The signal component ω_{ij} is the probability that a success requires action a_{ij} by Player i. We assume that conditional on the signals $\omega_1 = (\omega_{11}, \omega_{12})$ and $\omega_2 = (\omega_{21}, \omega_{22})$, the success probability of an action profile (a_{1j}, a_{2k}) equals the product of the individual signals $\omega_{1j} \cdot \omega_{2k}$. We refer to this property as *action independence* in the more general setup. Since $\omega_{i2} = 1 - \omega_{i1}$, the signal ω_i can be identified with ω_{i1} in our example. We furthermore assume in this example that ω_{i1} is uniformly distributed on $[0, 1]$ and that the players' signals are independently distributed. Formally, this *signal-independence* assumption requires that the probability that $\omega_{11} < x$ and $\omega_{21} < y$ equals $x \cdot y$ for all x and y with $0 \leq x, y \leq 1$.

To simplify notation, denote the higher of Player 1's two signals (the first order statistic of her signals) by α, i.e. $\alpha := \max\{\omega_{11}, \omega_{12}\}$. Similarly, for Player 2, define $\beta := \max\{\omega_{21}, \omega_{22}\}$. α and β are the first order statistics of the uniform distribution on the one-dimensional unit simplex. Note that α and β are independently and uniformly distributed on the interval $[\frac{1}{2}, 1]$. In the sequel, when talking about Player 1's action, it will be often convenient to refer to his α (or high-probability) action and her $1 - \alpha$ (or low-probability) action, and similarly for Player 2.

We now use this example to develop intuition for finding and comparing equilibria that carries over to the general class of games with an arbitrary (finite) number of players and actions per player, and with an arbitrary time horizon that we introduce in the next section. In our example one can identify classes of equilibria, characterize optimal behavior, and illustrate the difficulties arising in joint search more generally. We highlight that there are multiple Pareto-ranked equilibria and that in the search for optimal equilibria it suffices to investigate convex-partition equilibria in which a player's signal space is partitioned into convex subsets over which the player chooses the same action sequence. Furthermore, there are routine equilibria that avoid repetition inefficiencies, but they are suboptimal; the optimal equilibrium exhibits both repetition inefficiency, i.e. with positive probability players repeatedly try the same action profile, and search-order inefficiency, where less promising profiles are tried before more promising ones. In contrast to the optimal equilibrium, however, routines are robust to strategic naivete—they are cursed equilibria—and overconfidence in the sense that the payoff achieved when using these routines remains constant when introducing various degrees of the above biases, while the payoff of attempting to play the optimal strategy profile decrease in the presence of these biases. We also show that a manager who is aware that her agents are sufficiently overconfident, strictly prefers a routine to a more flexible (cardinal) problem-solving approach in this example.

Returning to our example, it is immediately clear that the full-information solution (or *ex post*-efficient search), which a social planner with access to both players' private information would implement, is not an equilibrium in the game with private information. The social planner would prescribe the α-action to Player 1 and the β-action to Player 2 in the first period, and in the second period would prescribe the profile $(\alpha, (1-\beta))$ if $\alpha(1-\beta) > (1-\alpha)\beta$, and the profile $((1-\alpha), \beta)$ otherwise. The players themselves, who only have access to their own information, are unable to carry out these calculations and cannot decide which of the two players should switch actions and who should stick to her first-period action. This raises a number of questions: What is the constrained planner's optimum, i.e. which strategy profile would a planner prescribe who does not have access to the players' private information? What are the equilibria of the game?

Two simple strategy profiles are easily seen to be equilibria. In one, Player 1 takes her α action in both periods and Player 2 takes her β action in the first and her $1-\beta$ action in the second period. In the second equilibrium, Player 2 stays with her β action throughout and Player 1 switches. In these equilibria, players condition only on the rank order of their actions according to their signal (which action is the α action) and not on signal strength (the specific value of α). They never examine the same cell twice. These equilibria are *ex post* equilibria; i.e., each Player i's behavior remains optimal even after learning the other Player j's signal. As long as we maintain action independence, these strategy profiles remain equilibria regardless of each player's signal distributions. In addition these equilibria are fully cursed: The non-switching player need not know that the other player switches from a high- to a low probability action. All she needs to know is that the other player switches. Similarly, all the switching player needs to know is that the other player does not switch. She need not know that the non-switching player sticks to her high-probability action. Thus, these equilibria are robust to changes in the environment and to player ignorance about the details of how the other player's private information affects behavior. If we imagine players facing similar problems (say with varying individual signal distributions) repeatedly over time, this robustness makes these equilibria natural candidates for being adopted as routines: One player is designated (perhaps by management) to always stay put and the other to always switch regardless of the new problem.

While these routine equilibria are robust and avoid repetitions, they make only the first-period decision sensitive to the players' information; the switching decision does not depend on the signal. One may wonder whether it would not be better to tie the switching probability to the signal as well. Intuitively, a player with a strong signal, α close to one, should be less inclined to switch than a player with a weak signal, α close to one half. In order to investigate the existence of equilibria in which signal strength matters in addition to the ranking of actions, we need to describe players' strategies more formally.

A strategy for Player i has three components: (1) $p_1^i(\omega_{i1})$, the probability of taking action a_{i1} in period 1 as a function of the signal; (2) $q_1^i(\omega_{i1})$, the probability of taking action a_{i1} in period 2 after having taken action a_{i1} in period 1 as a function of the signal; and (3), $q_2^i(\omega_{i1})$, the probability of taking action a_{i1} in period 2 after having taken action a_{i2} in period 1 as a function of the signal. We show in the appendix, using the fact that actions are unobservable, that for any behaviorally mixed strategy that conditions on Player i's signal ω_{i1} there is a payoff equivalent strategy that conditions only on her signal strength α and *vice versa*. Intuitively, because Player j does not observe which action i chooses, i's payoff depends only on the associated signal strength and not the name of the chosen action. More precisely, consider two different signals ω'_{i1} and ω''_{i1} that give rise to the same α. Hence, these signals differ only in that one identifies action 1 and the other action 2 as the high-probability action (H). Without loss of generality, suppose that ω'_{i1} identifies action 1 as the high-probability action so that $\alpha = \omega'_{i1} = 1 - \omega''_{i1}$. Define $p^i(\alpha) \equiv (1/2)\ p_1^i(\omega'_{i1}) + (1/2)\ (1 - p_1^i(\omega''_{i1}))$, which is the probability of taking the high-probability action in period 1 as a function of the signal strength α. Defining $q_h^i(\alpha)$ and $q_l^i(\alpha)$ similarly (again the intuitive obvious but tedious formal argument is in the Appendix), we can thus express Player i's strategy using the following reduced-form probabilities: (1) $p^i(\alpha)$, the probability of taking the high-probability action in period 1 as a function of the signal; (2) $q_h^i(\alpha)$, the probability of taking the high-probability action in period 2 after having taken the high-probability action in period 1 as a function of the signal; and (3), $q_l^i(\alpha)$, the probability of taking the high-probability action in period 2 after having taken the low-probability action in period 1 as a function of the signal.

We will also make use of the fact (verified in the Appendix for the general setup) that in our game Nash equilibria can be studied in terms of mappings from players' signals to distributions over sequences of actions. Intuitively, since j's first-period choice is unobservable, Player i cannot condition on Player j's past behavior. Hence, we can think of i as choosing the entire two-period action sequence upon observing her signal ω_i.

Now fix a strategy for Player 2. We are interested in the payoff of Player 1 for anyone of her possible signal-strength types α, for any possible action sequence she may adopt, and for any possible strategy of Player 2. In writing down payoffs, we will use the fact that in equilibrium Player 2 will never stick to her low-probability action in the second period after having used her low-probability action in the first period, i.e. $q_l^2(\beta) = 1$ for all $\beta \in [\frac{1}{2}, 1]$ in every equilibrium. Intuitively, by switching away from the low-probability action in period two, a player ensures that a new cell is explored for certain, and independent of the behavior of the other player the induced cell has a higher success probability than when sticking to the low-probability action.

Using the fact that β is distributed between 1/2 and 1 with density 2, the payoff of type α of Player 1 when choosing the high-probability action in both periods is:

$$
\begin{aligned}
(1) \qquad \mathrm{HH}(\alpha) \;=\; & \int_{\frac{1}{2}}^{1} 2 \;\Big[\underbrace{\alpha\beta}_{\substack{\text{success prob.}\\ \text{of HH cell}}} + \delta \underbrace{(1-q_h^2(\beta))}_{\substack{\text{prob. that}\\ \text{2 switches}}} \underbrace{\alpha(1-\beta)}_{\substack{\text{success prob.}\\ \text{of HL cell}}}\Big] \underbrace{p^2(\beta)}_{\substack{\text{prob. of 2}\\ \text{initially}\\ \text{playing H}}} \, d\beta \\
+ & \int_{\frac{1}{2}}^{1} 2 \;\Big[\underbrace{\alpha(1-\beta)}_{\substack{\text{success prob.}\\ \text{of HL cell}}} + \delta \underbrace{q_l^2(\beta)}_{\substack{\text{prob. that}\\ \text{2 switches}}} \underbrace{\alpha\beta}_{\substack{\text{success prob.}\\ \text{of HH cell}}}\Big] \underbrace{(1-p^2(\beta))}_{\substack{\text{prob. of 2}\\ \text{initially}\\ \text{playing L}}} \, d\beta
\end{aligned}
$$

Similarly, Player 1's payoff from taking the high-probability action in the first and the low-probability action in the second period, when his type is α, equals

$$
\begin{aligned}
(2) \quad \mathrm{HL}(\alpha) \;=\; & \int_{\frac{1}{2}}^{1} 2 \;\left[\alpha\beta + \delta\, q_h^2(\beta)\,(1-\alpha)\beta + \delta\,(1-q_h^2(\beta))\,(1-\alpha)(1-\beta)\right] p^2(\beta)\, d\beta \\
+ & \int_{\frac{1}{2}}^{1} 2 \;[\alpha(1-\beta) + \delta\,(1-\alpha)\beta]\,(1-p^2(\beta))\, d\beta
\end{aligned}
$$

Finally, Player 1's payoff from taking the low-probability action in the first and the high-probability action in the second period, when his type is α, equals

$$
\begin{aligned}
(3) \qquad \mathrm{LH}(\alpha) \;=\; & \int_{\frac{1}{2}}^{1} 2 \; \left[(1-\alpha)\beta + \; \delta\, q_h^2(\beta)\; \alpha\beta + \; \delta\; (1-q_h^2(\beta))\; \alpha(1-\beta)\right] \; p^2(\beta)\; d\beta \\
+ \; & \int_{\frac{1}{2}}^{1} 2 \; [\; \delta\; \alpha\beta + (1-\alpha)(1-\beta)] \; (1-p^2(\beta))\; d\beta
\end{aligned}
$$

As argued above, the sequence of actions LL is strictly dominated for all $\alpha > \frac{1}{2}$.

It follows by inspection that all three of these payoffs are linear in α and that $\mathrm{HH}(\cdot)$ is strictly increasing in α. Intuitively, the better the signal the higher the payoff from choosing the more promising action in both periods. Also, when being sure that a particular action is correct, it is always (weakly) better to select this action independent of how one's partner behaves, i.e. $\mathrm{HH}(1) \geq \mathrm{HL}(1)$ and $\mathrm{HH}(1) > \mathrm{LH}(1)$. At the other extreme, when both actions are equally likely to be correct, the first-period choice does not matter (i.e. $\mathrm{HL}\left(\frac{1}{2}\right) = \mathrm{LH}\left(\frac{1}{2}\right)$) and switching to ensure that a new cell is investigated in the second period is weakly dominant $\mathrm{HL}\left(\frac{1}{2}\right) \geq \mathrm{HH}(\frac{1}{2})$. These properties are illustrated in FIGURE 1.

Note also that $\mathrm{HL}\left(\frac{1}{2}\right) = \mathrm{HH}\left(\frac{1}{2}\right)$ is only possible if Player 2 switches with probability zero, i.e. if $q_h^2(\beta) = 0$ for almost all β. In that case, since the sequence LL is played with probability zero in equilibrium, Player 2 must either play HL or LH with probability one. But if Player 2 switches with probability one, HH is the unique best reply (up to changes on a set of measure zero), which in turn requires that Player 2 plays HL with probability one.

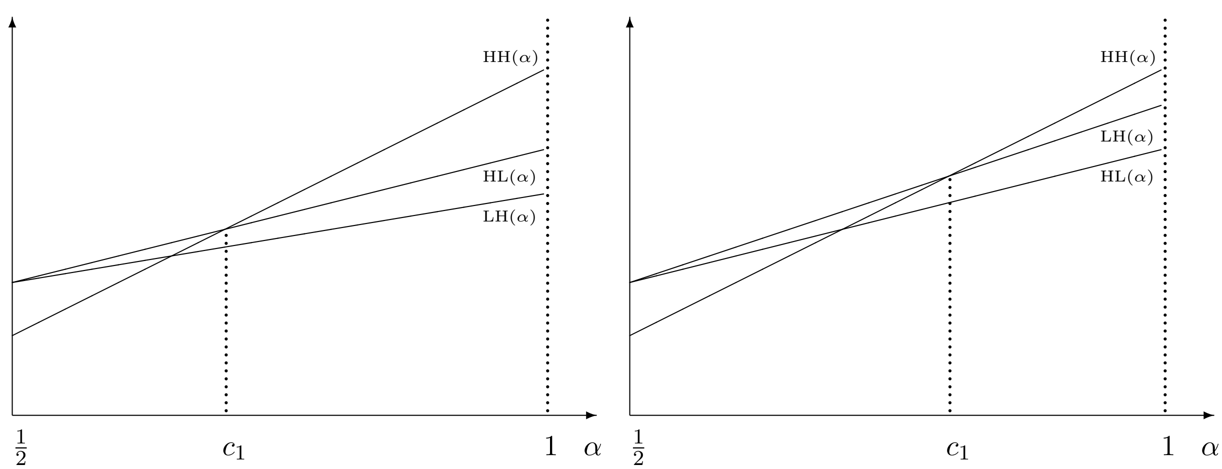

FIGURE 1

We begin by considering equilibria in which $\mathrm{HL}(1) \neq \mathrm{LH}(1)$, as depicted in FIGURE 1. This implies that in equilibrium Player 1 (similarly for Player 2) either plays HH for all α, or HL for all α, or LH for all α, or there exists a critical value c_1 such that she plays HL for $\alpha \leq c_1$ and HH for $\alpha > c_1$, or there exists a critical value c_1 such that she plays LH for $\alpha \leq c_1$ and HH for $\alpha > c_1$. In addition, against a player using only the action sequences HH and HL, the action sequence LH is never optimal, because in that case HL is a better response. This leaves only two possible types of equilibria for which HL(1) $\neq$ LH(1):

1. *HL-equilibria* in which Player i has a cutoff c_i such that she uses HL for α below this cutoff (and HH above the cutoff), and

2. *LH-equilibria* in which Player i has a cutoff c_i such that she uses LH for α below this cutoff (and HH above the cutoff).

FIGURE 1 illustrates the payoff structure for different action sequences as they would look in these two types of equilibria for interior cutoffs, i.e. $c_i \in (0,1)$. The left panel illustrates an HL-equilibrium and the right panel an LH-equilibrium.

Because a Player i with a cutoff signal c_i must be indifferent between playing HH and HL, cutoffs in any HL-equilibrium must satisfy the system of equations:

$$
\begin{aligned}
(4) \qquad & \int_{c_{3-i}}^{1} c_i \beta d\beta + \int_{\frac{1}{2}}^{c_{3-i}} \left[c_i\beta + c_i(1-\beta)\delta\right] d\beta \\
= & \int_{c_{3-i}}^{1} \left[c_i\beta + (1-c_i)\beta\delta\right] d\beta + \int_{\frac{1}{2}}^{c_{3-i}} \left[c_i\beta + (1-c_i)(1-\beta)\delta\right] d\beta \qquad i = 1, 2.
\end{aligned}
$$

Conversely, because LH is never an optimal response to the other player playing only HL and HH, any solution to this system of equations corresponds to an HL-equilibrium. There are exactly three solutions in the relevant range of $c_i \in [\frac{1}{2}, 1], i = 1, 2$. These are, $(c_1, c_2) = (.5, 1)$, $(c_1, c_2) = (1, .5)$, and $(c_1, c_2) \approx (0.760935, 0.760935)$. The cutoffs $(c_1, c_2) = (.5, 1)$ and $(c_1, c_2) = (1, .5)$ correspond to the two routine equilibria discussed above. In the third equilibrium, players play the high-probability action in both periods when being sufficiently confident that their high-probability action is correct and otherwise attempt the high-probability action first but switch to the low probability action following a failure in order to induce a new cell. All three HL-equilibria are unaffected by the players' level of impatience: in the first period players investigate the most promising cell, and thereafter both want to maximize the (myopic) probability of success since they are in the final period. These equilibria, therefore, are also robust to players having different discount factors. The property that routine equilibria are independent of the discount factor, and robust to players having different discount factors, extends more generally.

We now turn to LH-equilibria. A necessary condition for having an LH-equilibrium is that players do not have an incentive to deviate to HL for any α. Given the linearity of the payoff functions, this condition is satisfied if at each Player i's cutoff c_i we have $HH(c_i) = LH(c_i) \geq HL(c_i)$. As a result, we have an LH-equilibrium if the following conditions are satisfied:

$$\begin{aligned}(5) \qquad & \int_{c_{3-i}}^{1} c_i\beta d\beta + \int_{\frac{1}{2}}^{c_{3-i}} \left[c_i\beta\delta + c_i(1-\beta)\right] d\beta \\ &= \int_{c_{3-i}}^{1} \left[c_i\beta\delta + (1-c_i)\beta\right] d\beta + \int_{\frac{1}{2}}^{c_{3-i}} \left[c_i\beta\delta + (1-c_i)(1-\beta)\right] d\beta \qquad i = 1, 2.\end{aligned}$$

and

$$(6) \qquad \int_{\frac{1}{2}}^{c_{3-i}} c_i\beta d\beta \geq \int_{\frac{1}{2}}^{1} (1-c_i)\beta d\beta \qquad i = 1, 2.$$

The solutions of the system of equations (5) in the relevant range of $c_i \in [\frac{1}{2}, 1], i = 1, 2$, depend on δ. For $\delta = 1$, there are three solutions: $(c_1, c_2) = (.5, 1)$, $(c_1, c_2) = (1, .5)$, and a symmetric solution. We establish in the appendix that for $\delta < 1$ there is a unique solution to equation (5), which is symmetric ($c_1 = c_2 = c$) and increasing in the discount factor δ. This unique solution is an equilibrium provided that it satisfies condition (6), which is equivalent to

$$4c^3 + 2c - 3 \geq 0.$$

The smallest value, c^*, of c that satisfies the above inequality is $c^* \approx 0.728082$. The corresponding discount factor for which c^* is a symmetric solution to the system of equations (5) is $\delta^* \approx 0.861276$. Hence for $\delta \in (\delta^*, 1)$ there exists a unique solution with a common cutoff $c(\delta)$ that is strictly increasing in the discount factor δ. Intuitively, if players are very impatient, i.e. $\delta = 0$, then independent of the other player's behavior, each player wants to maximize the probability of a success in the first period and will therefore initially choose her high- probability action. Thus an LH-equilibrium does not exist when players are very impatient. When players are very patient, on the other hand, their primary concern is with finding a success in either period. In that case, against a player who only uses HH and LH, playing LH may be attractive because it ensures both that two different action profiles are examined and it takes advantage of a complementarity between action sequences that switch in the same order.[8] In the

[8]Given (almost) any realization of signal strengths α and β, for $\delta = 1$ conditional on both players switching, they receive higher payoffs if they switch in the same order. This can be seen as follows: For $\delta = 1$, the difference in payoffs between switching in the same order and in opposite orders equals $[\alpha\beta-(1-\alpha)(1-\beta)]-[(1-\alpha)\beta-\alpha(1-\beta)] = 1 - 2\alpha(1-\beta) - 2\beta(1-\alpha)$. The derivative of the right-hand side (RHS) of the equation with respect to β equals $2\alpha - 2$ and therefore is negative for almost all α and if we evaluate RHS at the lowest possible value of β, i.e. $\beta = \frac{1}{2}$, then RHS equals $1 - \alpha - (1-\alpha) = 0$. Hence, for almost all values of α and β the RHS, and thus the payoff difference between switching in the same and in opposite orders, is positive.

limit when players are perfectly patient ($\delta = 1$), the cutoff converges to that of the symmetric HL-equilibrium, since perfectly patient players care about which cells are investigated, but not in which order.

The next proposition summarizes our discussion thus far:

Proposition 1 *The entire set of equilibria in which neither player is indifferent between HL and LH for all signal realizations has the following form: For all $\delta \in (0, 1)$, there exists a symmetric HL-equilibrium with common cutoff $c \approx 0.760935$ and there exist two asymmetric HL-equilibria with cutoffs $(c_1, c_2) = (.5, 1)$ and $(c_1, c_2) = (1, .5)$, respectively. Furthermore, there is a critical discount factor $\delta^* \approx 0.861276$ such that for all $\delta \in (\delta^*, 1)$ there exists a symmetric LH-equilibrium with common cutoff $c(\delta)$, which is strictly increasing in δ, where $c(\delta^*) \approx 0.728082$ and $c(1) \approx 0.760935$. Conversely, no LH-equilibrium exists for $\delta < \delta^*$.*

Proposition 1 completely characterizes the set of equilibria that satisfy the condition HL(1) $\neq$ LH(1) for both players. Under some conditions, there also exist equilibria with HL(1) = LH(1) for at least one player. (We construct such equilibria in the appendix.) Since in these equilibria one or both of the players are indifferent between HL and LH over a range of signal strengths that has positive probability, we call these IN-equilibria. In an IN-equilibrium at least one of the players either randomizes between LH and HL over some range of signal strengths or one can partition a subset of the set of possible signal strengths into sets where she either plays LH or HL. In either case, IN-equilibria can be ignored in the search for optimal strategy profiles. Players would be better off if both players switched to playing HL over the relevant range: If a single player switches payoffs are not affected because of indifference; if then the other player switches as well payoffs strictly increase because HL is strictly better than LH against HL. To find the optimal equilibrium, we thus only have to compare the payoffs from the equilibria characterized in Proposition 1. For each player, all of these equilibria are simple in the sense that they assign a particular action sequence to a convex subset of her signal space (here the unit interval).[9]

Furthermore, when considering the equilibria of Proposition 1, we can immediately rule out that an LH-equilibrium is optimal: To see this, simply change both players' strategies to HL-strategies, without changing the cutoff. Under the original strategies, there are three possible events, each arising with strictly positive probability: Both players follow an HH-sequence; both follow an LH sequence; and, one follows an LH-sequence while the other follows an HH sequence. Clearly LH is not optimal against HH and therefore in this instance the new strategy yields a strict improvement. Also, both players following HL rather than LH yields a strict improvement

[9]Below we will illustrate that this feature of optimal equilibria generalizes to other distributions and an arbitrary number of players and periods.

for impatient players. Thus in two events there is a strict payoff improvement, in the remaining event payoffs are unaffected, and all three events have strictly positive probability.

It is, however, not immediately clear whether to prefer the symmetric HL-equilibrium or the asymmetric HL-equilibria. In either, there is positive probability that profiles are searched in the wrong order. The symmetric equilibrium makes the second-period switching probability sensitive to a player's signal, which seems sensible. At the same time, it introduces an additional possible source of inefficiency. Players may not succeed in the first round despite having signals so strong that they do not switch in the second round. In that case, they inefficiently search only one of the available profiles.

It would be a straightforward matter to calculate and compare payoffs from symmetric and asymmetric equilibria directly. We will follow a different line of reasoning, whose logic parallels the one we use in Lemma 1 and Proposition 12 for the general model. Start with the asymmetric HL-equilibrium in which $c_1 = \frac{1}{2}$ and $c_2 = 1$. Consider the (informationally-constrained) social planner who raises c_1 from $\frac{1}{2}$ and lowers c_2 from 1 by the same small amount γ. The social planner thus induces Player 1 to switch rather than to stick with her high-probability action whenever both of her actions are (approximately) equally promising and at the same time induces Player 2 to stick to her high-probability action whenever she is (approximately) certain that her high-probability action is correct. This does not change first-period actions or payoffs, and the second-period payoff as a function of γ is proportional to

$$\pi(\gamma) = \int_{\frac{1}{2}}^{1-\gamma} \int_{\frac{1}{2}}^{\frac{1}{2}+\gamma} (1-\alpha)(1-\beta) d\alpha d\beta + \int_{1-\gamma}^{1} \int_{\frac{1}{2}}^{\frac{1}{2}+\gamma} (1-\alpha)\beta d\alpha d\beta + \int_{\frac{1}{2}}^{1-\gamma} \int_{\frac{1}{2}+\gamma}^{1} \alpha(1-\beta) d\alpha d\beta.$$

It is straightforward to check that $\left.\frac{\partial \pi(\gamma)}{\partial \gamma}\right|_{\gamma=0} = 0$ and $\left.\frac{\partial^2 \pi(\gamma)}{\partial \gamma^2}\right|_{\gamma=0} > 0$. Hence, the social planer can improve on the two asymmetric equilibria. In common interest games an optimal strategy profile is a Nash equilibrium, and we prove that an optimal strategy exists for our general model below. This implies that for any arbitrary strategy profile σ, either σ is an equilibrium or there exists an equilibrium σ^* with $u_i(\sigma^*) > u_i(\sigma)$ for $i = 1, 2$. Thus the pair of cutoff strategy profiles with cutoffs $c_1 = 1/2 + \gamma$ and $c_2 = 1 - \gamma$ with an appropriately small value of γ either is an equilibrium or there exists an equilibrium that strictly dominates it. Furthermore, an optimal strategy profile must be one of the partition-equilibria characterized in Proposition 1. Therefore, we have the following observation:

Proposition 2 *For any $\delta \in (0, 1)$, in the two-player two-action two-period game with signals that are independently and uniformly distributed, the symmetric HL-equilibrium is the optimal* *equilibrium and at the same time the optimal strategy that an informationally-constrained social planner would implement.*

The example nicely illustrates that routines are suboptimal with fully rational players. This raises the question why players would select such a Pareto-dominated equilibrium. Furthermore, in the example there is no given routine that stands out, which leads to the further question of how players would select a particular routine. We informally think of routines as being selected by the management of the organization, which makes recommendation to the players of how to behave. This, in turn, raises the question under what circumstance management would want to select a problem-solving routine. The following example highlights that routines can be optimal if agents are not fully rational. We begin by arguing that routines can be optimal when agents are overconfident.

Suppose that a player interprets her signal as having a first-order statistic of $(1-x)\alpha + x$, where $x \in (0, 1)$. In this stylized example, x is a measure of a player's overconfidence. As x approaches 1, a player always believes with (almost certainty) to know what her correct action is while the true probability is still uniformly distributed. For the sake of the example, suppose both players are equally overconfident (have the same x) and consider the payoff of a symmetric HL-type equilibrium in which players are meant to play HH when very confident and HL otherwise. In particular, we suppose that an overconfident player correctly predicts for what signals her fellow team member switches and consider the true signal at which she is indifferent between switching and not switching. That is for any given true cutoff signal c_{3-i} of her fellow team member, a player with a perceived signal $\tilde{c}_i \equiv (1-x)c_i + x$ must be indifferent between switching and not switching. Now replacing c_i with the perceived signal $(1-x)c_i+x$ in Equation 4, shows that if Player i becomes extremely overconfident($x \to 1$), then her true cutoff signal approaches $(1/2)$ for any $c_{3-i} > 1/2$. This implies that in the symmetric equilibrium as both players become extremely overconfident ($x \to 1$), the true equilibrium cutoff signal approaches $1/2$.[10] Intuitively, as long as there is a small probability of the other player switching, an extremely overconfident player will be to reluctant to switch herself, and as her overconfidence gets extreme (x approaches 1) she will almost never do so.

Clearly, however, as the common cutoff $c \to 1/2$ the players' payoff is less than in a routine. Observe also that routines remain equilibria when players are overconfident. Even if Player i is extremely confident that she knows what action is correct, if Player j never switches it is

[10] Formally, take any sequence of equilibrium cutoffs $c(x)$ as $x \to 1$. This sequence must have a convergent subsequence. Suppose the convergent subsequence converges to some cutoff $\hat{c} > 1/2$. Then for any $\epsilon > 0$, there exists an $\bar{x}$ such that for all $x > \bar{x}$, $c(x) \in (\hat{c} - \epsilon, \hat{c} + \epsilon)$. This however contradicts the above established fact that for any $c > 1/2$, the cutoff signal her team member responds with goes to $1/2$.

optimal to switch for Player i. Hence, in our example, when players are sufficiently overconfident it becomes strictly optimal for the management to implement a routine, and the payoffs of the routine are fully robust to players' overconfidence. Indeed, straightforward calculations reveal that for all $x > .374$ routines perform better than the symmetric overconfident HL-equilibrium. We emphasize that overconfidence can justify the use of routines in

Proposition 3 *For any $\delta \in (0,1)$, in the two-player two-action two-period game with signals that are independently and uniformly distributed, if players are sufficiently overconfident, then the payoff of an ordinal equilibrium is higher than that of the overconfident symmetric HL-equilibrium.*

Routines are also robust to other types of biases documented and modeled in behavioral economics. For example, suppose agents are strategically naive in the sense that they play cursed equilibria. In a fully cursed equilibrium, each player best responds to the actual distribution of actions sequences by the other player but fails to take into account how this distribution of action sequences depends on the other player's type. In an ordinal equilibrium, one player—say 1—always switches. It is then clearly optimal for Player 2 to always select her high probability action even if not realizing that Player 1 switches from her high to her low-probability action. Similarly, given that Player 2 does not switch, it is clearly optimal for Player 1 to do so. Hence routines are fully cursed equilibria, and in this sense robust to strategic naivete of team members. In contrast, the optimal equilibrium is not robust to such strategic naivete.

To see this, consider a symmetric fully-cursed equilibrium in which agents play HH when having high signals and HL when having a low signal. In such a symmetric fully cursed equilibrium with cutoff first-order statistic c, Player 2 switches with probability $2(c-(1/2))$. Given this behavior, a fully cursed Player 1 is indifferent between switching and not switching when having a first-order statistic α if

$$(7) \qquad \alpha\left[\int_0^1 xdx\right] + \delta\alpha 2\left(c-\frac{1}{2}\right)\left[\int_0^1 xdx\right] = \alpha\left[\int_0^1 xdx\right] + \delta(1-\alpha)\left[\int_0^1 xdx\right].$$

Furthermore, using that in a symmetric equilibrium $\alpha = c$, the fully cursed equilibrium cutoff satisfies: $2c^2 - 1 = 0$, so that the common cursed cutoff is equal to $\sqrt{1/2}$. Now calculating the true payoff when players use the above cutoff for $\delta = 1$ shows that the expected payoff is 0.753 while the payoff of the ordinal equilibria is 0.75. In our example, thus, lack of strategic sophistication severely reduces the benefits of optimal equilibria over routines—although in this specific case not completely eliminating it. It is natural to also consider teams in which members exhibit some combination of cursedness and overconfidence, or are unsure about either the cursedness or overconfidence of fellow team members. We highlight in the next section that routines are *fully* robust to relaxing the rationality constraint simultaneously in these directions.

4 The General Model

One can think of our model as a formal representation of the following stylized "safe problem": A group of individuals wants to open a safe. Each of them has access to a separate dial in an isolated room. There is a single combination of dial settings that will open the safe. The group repeatedly tries out different combinations. It is impossible to communicate or to observe the actions of other group members. Initially, each individual privately and independently receives a signal that indicates for each of her dial settings the probability of it being correct, i.e. being part of the combination that will open the safe. The probability that any given combination is correct is the product of the corresponding signals.

Each Player i out of a finite number I of players has a finite set of actions A_i that has cardinality m^i; we will slightly abuse notation by using I to denote both the set of players and its cardinality. $A := \times_{i=1}^{I} A_i$ denotes the set of action profiles. A typical element of A_i is denoted a_i and we write $a = (a_i, a_{-i}) \in A$ for a typical action profile. There is a single *success profile* $a^* \in A$ with a common positive payoff $u(a^*) = 1$, and the common payoff from any profile $a \neq a^*$ equals $u(a) = 0$. The location of the success profile a^* is randomly chosen from a distribution $\omega \in \Omega := \Delta(A)$ over the set of all action profiles. The distribution ω itself is randomly drawn from a distribution $F \in \Delta(\Delta(A))$, the set of distributions over distributions of success profiles. This permits us to express the idea that players are not only uncertain about the location of the success profile, but also that each player has some information regarding the location that is unknown to others. Formally, after ω is chosen, each Player i learns ω_i, the marginal distribution over Player i's actions. Thus, if $\omega(a)$ denotes the probability that ω assigns to the profile a being the success profile, $\omega_i(a_{ij}) = \sum_{j_1,\ldots,j_{i-1},j_{i+1},\ldots,j_n} \omega(a_{1j_1}, \ldots a_{ij} \ldots a_{nj_n})$ is the probability that a success requires Player i to take action a_{ij}. Denote the set of Player i's marginal distributions ω_i by Ω_i.

We make two assumptions that limit how much players can infer about the signals of others from their own signals. We assume *action independence,* which requires that each ω in the support of F be the product of its marginals, i.e. $\omega = \Pi_{i=1}^{I} \omega_i$. Furthermore, we require *signal independence,* which requires that F is the product of its marginals, F_i, i.e. $F(\omega) = \Pi_{i=1}^{I} F_i(\omega_i)$.[11] Upon observing her signal a player therefore does not revise her belief about how confident other players are about which of their actions are required for a success profile.

[11] We discuss possible consequences of violations of these assumptions in the Appendix.

Players choose actions in each of $T < \infty$ periods, unless they find the success profile, at which point the game ends immediately. Players do not observe the actions of other players. Therefore a player's strategy conditions only on the history of her own actions. Denote the action taken by Player i in period t by a_i^t. Then Player i's action history at the beginning of period t is $h_{it} := (a^0, a_i^1, \ldots, a_i^{t-1})$, where a^0 is an auxiliary action that initializes the game. We let $h_t = (h_{1t}, \ldots, h_{It})$ denote the period-t action history of all players. The set of all period-t action histories of Player i is denoted H_{it}, where we adopt the convention that $H_{i1} = \{a^0\}$. The set of period-t action histories of all players is H_t and the set of all action histories of all players is $H := \cup_{t=1}^T H_t$. A (pure) strategy of Player i is a function $s_i : H_{it} \times \Omega_i \to A_i$ and we use s to denote a profile of pure strategies. For any pure strategy profile s and signal vector ω, let $a^t(s, \omega)$ denote the profile of actions that is induced in period t. Similarly, define $A^t(s, \omega) := \{a \in A | a^\tau(s, \omega) = a \text{ for some } \tau \leq t\}$ as the set of all profiles that the strategy s induces before period $t + 1$ when the signal realization is ω. A behaviorally mixed strategies σ_i for Player i is a (measurable) function $\sigma_i : H_{it} \times \Omega_i \to \Delta(A_i)$. We use Σ_i^T to refer to the set of such strategies in the T-period game. $\Sigma^T := \times_{i \in I} \Sigma_i^T$ is the set of mixed strategy profiles in the T-period game. Players discount future payoffs with a common factor $\delta \in (0, 1)$. Thus, if t^* is the first period in which the success profile a^* is played, the common payoff equals δ^{t^*-1}; if the success profile is never played the common payoff is zero.

We will now formally describe payoffs. For that purpose define $\{a \notin h_t\}$ as the event that action profile a has not occurred in history h_t. Furthermore let the probability of reaching the initial history $\text{Prob}(h_1 | \sigma, \omega, a) = 1$ and for $t > 1$, with $h_t = (h_{t-1}, a')$ denoting the action history h_{t-1} followed by the action profile a', recursively define the probability of reaching history h_t given σ, ω and given that a is the success profile through

$$\text{Prob}(h_t | \sigma, \omega, a) := 1_{\{a \notin h_{t-1}\}} \prod_{i \in I} \sigma_i(a'_i | h_{i,t-1}, \omega_i) \text{Prob}(h_{t-1} | \sigma, \omega, a).$$

Then expected payoffs from strategy profile σ are given by

$$\int_{\omega \in \Omega} \sum_{a \in A} \sum_{h_t \in H} \delta^{t-1} \prod_{i \in I} \sigma_i(a_i | h_{i,t}, \omega_i) \text{Prob}(h_t | \sigma, \omega, a) \omega(a) dF(\omega),$$

where $\prod_{i \in I} \sigma_i(a_i | h_{i,t}, \omega_i) \text{Prob}(h_t | \sigma, \omega, a)$ denotes the (unconditional) probability that action profile a is played following history h_t and $\omega(a)$ is the probability that a is the success profile. We will denote the expected payoff from strategy profile σ by $\pi(\sigma)$ and Player i's expected payoff from strategy profile σ conditional on having observed signal ω_i by $\pi_i(\sigma; \omega_i)$. Observe that expected payoffs are well-defined since $\Delta(A)$ is a finite-dimensional unit simplex, and F is a distribution over this simplex. For simplicity, we assume throughout the paper that F has full support on

$\Delta(A)$. The timing of the game is as follows: (1) Nature draws a distribution $\omega \in \Delta(A)$ from the distribution F. (2) Each player receives a signal ω_i. (3) The success profile is drawn from the realized distribution ω. (4) Players start choosing actions.

One of our objectives in this paper is to demonstrate that routines are often suboptimal, and hence we compare them to optimal strategies. Regarding optimal strategies, some facts are worth noting. First, since we are studying common interest games, i.e. the payoff functions of the players coincide, there is a simple relation between optimality and (Bayesian Nash) equilibrium. An *optimal strategy profile must be a Nash equilibrium* since all players have a common payoff and if there were a profitable deviation for one player, then a higher common payoff would be achievable, contradicting optimality.[12] Second, as long as an optimal strategy profile exists, this observation has the following useful corollary: *any equilibrium that is payoff-dominated by some strategy is also payoff-dominated by an equilibrium strategy.* We use this fact repeatedly throughout.

The third noteworthy fact is that *optimality implies sequential rationality in common interest games.* Specifically, any optimal outcome of a common interest game can be supported by a strategy profile σ that is an *essentially perfect Bayesian equilibrium (EPBE)* (see Blume and Heidhues [2006] for a the formal definition and detailed discussion of EPBE)[13], i.e. one can partition the set of all histories into relevant and irrelevant histories so that σ is optimal after all relevant histories regardless of play after irrelevant histories. In general games it is frequently the case that Nash equilibria are supported by specific behavior off the path of play, which may not be sequentially rational. In an optimal strategy profile of a common-interest game, however, following the prescribed behavior on the path of play is optimal independent of what players do off the path of play. This can be seen as follows. Classify any history off the path of play of an optimal profile σ as irrelevant and any other history as relevant. Now suppose that there is a partial profile $\hat{\sigma}_{-i}$ that agrees with σ_{-i} on the path of play and a deviation σ_i' of Player i from σ_i that is profitable against $\hat{\sigma}_{-i}$. Then, since we have a common interest game, the strategy profile $(\sigma_i', \hat{\sigma}_{-i})$ yields a higher payoff for all players than σ, which contradicts optimality of σ.

Below, after formally introducing and characterizing them, we also show that routines are sequentially rational by proving that any ordinal equilibrium outcome in our setting can be supported by an EPBE.

[12] This is also used in Alpern [2002], Crawford and Haller [1990], and McLennan [1998].

[13] In finite games the outcomes that are supported by perfect Bayesian equilibria coincide with those supported by EPBEa. The use of EPBE, however, allows one to focus on the economically relevant aspects of the sequential rationally requirement because it does not require one to specify behavior after irrelevant histories, which although economically irrelevant can be technically challenging. Furthermore, if following irrelevant histories continuation equilibria do not exist in infinite games, EPBE is a superior solution concept.

5 Organizational Routines

5.1 Characterization of Routines

In this section we identify and characterize a class of equilibria that have a natural interpretation as organizational routines. In these ordinal equilibria players use strategies that condition only on the rank order of signals not their value, which implies that independent of the concrete signal realization team members always switch actions in a pre-specified order, thereby inducing the common pattern of behavior that we interpret as a particular problem-solving routine.

	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
$a_{1,1}$	1	2	3
$a_{1,2}$	4	5	6

FIGURE 1

For an informal introduction of these routines consider, for example, the matrix of action profiles in the stage game in FIGURE 1. In the figure $a_{i,j}$ denotes the j-th action of Player i, wlog ranked in the order of the corresponding signals, i.e. if we denote by $\alpha_{i,j}$ the probability that the j-th action of Player i is part of a success profile, then $\alpha_{i,j} \geq \alpha_{i,j+1}$ for all i and j. For convenience, the six action profiles have been numbered. Then there is an ordinal equilibrium in which the profile labeled t is played in period $t = 1, ..., 6$. In this equilibrium, Player 1 plays her most probable action in the first three periods during which Player 2 begins with her most likely action, then tries the next most likely action, and finally attempts her least likely action. Thereafter Player 1 switches to her least likely action and Player 2 repeats the previous sequence of actions. Taking the action sequence of the other player as given, in each period both players select the action that is most likely to lead to a success. On the other hand, there is no ordinal equilibrium in which players play the sequence of profiles $1, 3, 2, 4, 5, 6$: given that Player 1 is playing her most probable action in the first three periods, Player 2 can deviate from such a candidate equilibrium and in the first three periods select her action in the order of likelihood of leading to a success, thereby inducing the profile $1, 2, 3, 4, 5, 6$, which yields a higher payoff. This discussion suggest that a defining characteristic of ordinal equilibria is that each player in every period selects the action that is most likely to lead to a success. Propositions 4 and 5 indeed show that all ordinal equilibria are characterized by players selecting such *maximal* actions—which are precisely defined below—in every period. This, however, gives rise to a rich class of equilibria including some counterintuitive incomplete search equilibria that nevertheless satisfy (the spirit of) trembling-hand perfection as we illustrate below. Proposition 6 establishes that the entire class of ordinal equilibria is sequentially rational, and Proposition 7 specifies the optimal ordinal equilibrium or optimal routine.

We now turn to formally introducing routines. For tie-breaking purposes, it is convenient to introduce a provisional ranking of Player i's actions, where all provisional rankings have equal probability and Player i learns the provisional ranking at the same time as she learns ω_i. Using this provisional ranking to break ties where necessary, for any signal ω_i, we can generate a vector $r(\omega_i)$ that ranks each of Player i's actions a_{ij}, from the highest to the lowest probability of that action being required for a success. A strategy is ordinal if it only conditions on whether an action is more likely then another—i.e. has a higher rank—and not on how much more likely a particular action is. More precisely, a strategy σ_i of Player i is ***ordinal*** if there exists a function $\tilde{\sigma}_i$ such that $\sigma_i(h_{it}, \omega_i) = \tilde{\sigma}_i(h_{it}, r(\omega_i))$ for all $h_{it} \in H_{it}$ and all $\omega_i \in \Omega_i$. A profile σ is ordinal if it is composed of ordinal strategies; otherwise, it is ***cardinal.***

For any action history h_t define $A(h_t) := \{a \in A | a \in h_t\}$ as the set of all action profiles that have occurred before time t in history h_t. Given a strategy profile σ and any private history (ω_i, h_{it}) that is consistent with that profile (i.e. for which h_{it} has positive probability given σ and ω_i), let $A^t_{-i}(\sigma_{-i}, h_{it}, \omega_i) = \{a_{-i} \in A_{-i} | \text{Prob}(a^t_{-i} = a_{-i} | h_{it}, \omega_i, \sigma_{-i}) > 0\}$ be the set of partial profiles that have positive probability in period t given Player i's information σ_{-i}, h_{it}, and ω_i. For a strategy profile σ and any private history (ω_i, h_{it}) that is consistent with that profile, we say that the action a_{ij} is ***promising*** for Player i provided that given her information $(\sigma_{-i}, h_{it}, \omega_i)$ there is positive probability that it leads to a success.[14] An action a_{ij} is ***rank-dominated*** for a strategy profile σ following history (ω_i, h_{it}) if there exists a promising action $a_{ij'}$ such that $\omega_{ij'} > \omega_{ij}$. An action a_{ij} is ***maximal*** for a strategy profile σ following history (ω_i, h_{it}) if it is promising and rank–undominated or if no promising action exists. Roughly speaking, given the behavior of all other players, a maximal action has the highest probability of finding a success in the current period. We begin by observing that players must choose maximal actions on the path of play of any ordinal equilibrium.

Proposition 4 *If a profile of ordinal strategies σ is an equilibrium, then for every ω_i and every history of actions h_{it} that has positive probability given σ and ω_i, Player i plays a maximal action.*

[14] Formally, thus, given a strategy profile σ and a private history (ω_i, h_{it}) that is consistent with that profile, the action a_{ij} is promising for Player i if $\text{Prob}\left\{\{(a_{ij}, a_{-i}) \notin A(h_t)\} \cap \{(a_{ij}, a_{-i}) | a_{-i} \in A^t_{-i}(\sigma_{-i}, h_{it}, \omega_i)\} | \omega_i, h_{it}, \sigma_{-i}\right\} > 0$.

The proof of the proposition proceeds by noting that whenever there is a period in which Player i plays a non-maximal action, there is a signal realization ω_i that puts zero probability of success on all actions below this maximal action and positive success probability on the maximal action. In this case, however, Player i can profitably deviate by playing the maximal action in that period and thereby increasing the probability of success in that period. This either moves success probability forward—if this cell was going to be investigated in a later period anyhow—or simply increases the probability of success and hence contradicts that playing a non-maximal action can be optimal for all signal realizations in a candidate ordinal equilibrium.

Conversely, we observe next that if players choose maximal actions along the path of play of an ordinal strategy profile, then this strategy profile is an equilibrium. To this end, we say that Player i's strategy σ_i is *maximal* against the partial profile σ_{-i} if it prescribes a maximal action for Player i for every signal ω_i and every action history h_{it} that has positive probability given σ and ω_i. A strategy profile σ is maximal if σ_i is maximal against σ_{-i} for all players i.

Proposition 5 *If a profile σ of ordinal strategies is maximal, then it is an equilibrium.*

The proof proceeds in four steps: (1) We show that if a profile of ordinal strategies σ is maximal, then for every Player j every pure strategy in the support of σ_j induces the same actions in periods in which there is a positive probability of a success. (2) We conclude from (1) that if σ_i is maximal against σ_{-i}, then it is maximal against all s_{-i} in the support of σ_{-i}. (3) We show that if σ_i is maximal against a pure strategy profile s_{-i} then it is a best reply against that profile. And finally, (4) we appeal to the fact that if σ_i is a best reply against every s_{-i} in the support of σ_{-i}, then it is a best reply against σ_{-i} itself.

Propositions 4 and 5 show that ordinal equilibria have a simple structure: Actions profiles that are higher (in a vector sense based on the players' signal) are tried before lower profiles. There is substantial multiplicity of such equilibria because the ordering is not complete and therefore coordination on an ordinal equilibria is difficult. If, however, coordination on an ordinal equilibrium is achieved by some mechanism this equilibrium will prove remarkably robust.

We now argue that every ordinal equilibrium outcome is sequentially rational by proving that it can be supported by an EPBE. For any ordinal equilibrium profile σ classify histories on the path of play as relevant and all other histories as irrelevant. Take any strategy profile $\tilde{\sigma}$ that coincides with σ on the path of play. We need to argue that playing according to σ_i remains a best response to $\tilde{\sigma}_{-i}$ for any history on the path of play. In an ordinal equilibrium there exists a commonly known first period τ with the property that either a success is achieved with probability one in a period $t \leq \tau$ or τ is the final period. Because a deviation of Player

i is not detected prior to period τ, it does not change the behavior of all other players in any period $t \leq \tau$. Since given the behavior of all other players, Player i plays a maximal action in every period $t \leq \tau$, a deviation by i cannot increase her expected payoff conditional on finding a success prior to τ, and it must lower it whenever a success is found after period τ. Hence, it remains optimal to play according to σ_i on the path of play. We thus have:

Proposition 6 *Any ordinal equilibrium outcome can be supported by an EPBE and thus is sequentially rational.*

Observe that the equilibria characterized in Propositions 4 and 5 include (i) equilibria in which all profiles are examined without repetition, (ii) equilibria in which search stops before all profiles have been examined, and (iii) infinitely many Pareto-ranked equilibria in which search is temporarily suspended and then resumed. Reconsider the example illustrated in in FIGURE 1, where $a_{i,j}$ denotes the j-th action of Player i, and wlog we ranked these actions in the order of the corresponding signals. Then, (i) there is an equilibrium in which the profile labeled t is played in period $t = 1, \ldots, 6$, (ii) another equilibrium in which the profile labeled t is played in period $t = 1, \ldots, 4$ after which profile 1 is played forever, and (iii), for any k with $k > 0$ and $k < T - 4$ there is an equilibrium in which the profile labeled t is played in period $t = 1, \ldots, 4$ after which profile 1 is played for k periods followed by play of profiles 5 and 6.

Somewhat counter-intuitively, such ordinal equilibria in which search ends prematurely, or is temporarily suspended, survive elimination of dominated strategies. To see this, return to our example with two players, two actions, a uniform signal distribution, but now with $T \geq 4$ periods. For the row player let H (L) denote taking the high (low) probability action, regardless of the value of the signal. For the column player, use lower case letters, i.e. h and l, to describe the same behavior. Then $G_1 G_2 \ldots G_T$ with $G_t \in \{H, L\}$ is the strategy of the row player that prescribes taking the action G_t in period t regardless of the value of the signal, and similarly for the column player.

Suppose that the row player believes that the column player uses the strategy $hlhhhhh \ldots h$ with probability $1 - \epsilon$ and the strategies $lllhl \ldots l$, $llllhl \ldots l$, $\ldots$, $llll \ldots lhl$ and $llll \ldots llh$ each with probability $\frac{\epsilon}{T-3}$. Then the strategy $HHLH \ldots H$ is a unique best reply for almost every realization of the row player's signal (and a best reply for every signal realization). This implies that $HHLH \ldots H$ is an undominated strategy for the row player. By an analogous argument it follows that $hlh \ldots h$ is an undominated strategy for the column player. Hence, we have an equilibrium in undominated strategies in which search terminates after the third period, even if until that point there has been no success and there are arbitrarily many future search periods left.

If we discretize the game by considering signal distributions with a finite support, we have a finite game and can check equilibria for trembling-hand perfection. It is well known that in finite two-player games the set of (normal form) perfect equilibria coincides with the set of equilibria in undominated strategies. As a consequence, in the discrete approximation of our game the equilibrium ($HHLH\dots H$, $hlh\dots h$) is (normal form) perfect. In this sense, the equilibrium is robust and similar constructions can be found in the case with more players and or actions.

Although even these clearly suboptimal equilibria satisfy the spirit of trembling-hand perfection, in our interpretation of routines as being selected by management, it is implausible that such routines would be selected. Management would clearly prefer to select an optimal routine to an obviously suboptimal one and hence we now show that an optimal routine exists and characterize it.

Denote the random variable that is Player i's signal by $\tilde{\omega}_i$, to distinguish it from the signal realization ω_i, and let $\tilde{\omega}_{i(n)}$ stand for the nth (highest) order statistic of the random vector $\tilde{\omega}_i$. Define $\overline{\omega}_{i(n)}$ as the expectation $\mathbb{E}[\tilde{\omega}_{i(n)}]$ of the nth order statistic of Player i's signal. For every realization ω_i, use $a_{i(n_i)}(\omega_i)$ to denote the action of player i with the n_ith highest signal according to ω_i. For any i and n, let $a_{i(n)}$ denote the rule of playing the nth highest action for any signal realization ω_i. Refer to the rule $a_{i(n)}$ as Player i's nth rank-labeled action and to every $(a_{1(n_1)}, a_{2(n_2)}, \dots, a_{I(n_I)})$ as a rank-labeled action profile.

Proposition 7 *An optimal ordinal equilibrium exists and in any optimal ordinal equilibrium agents play a sequence of rank-labeled action profiles in the order of their* ex ante *success probability without repetition.*

Conceptually, finding an optimal routine is simple. For each team member determine the expected values for the order statistics of that team member's signal vector; for each profile of such expectations multiply these expected values; and, play profiles in decreasing order of the magnitude of these products. Note however that while routines are easy to find and do not depend on detailed knowledge of signal distributions, finding an optimal routine may be computationally burdensome and optimality will generally depend on which distributions signals are drawn from.

5.2 Robustness of Routines

Having specified the class of ordinal equilibria and found the optimal equilibrium in this class, we turn to highlight the robustness of these routine equilibria. Loosely speaking, we begin by showing that the class of ordinal equilibria coincides with the class of *ex post* equilibria. An equilibrium is an *ex post* equilibrium if each player's strategy remains a best response even after learning the other players' private information. Since *ex post* equilibria induce best replies for every signal distribution, they do not depend on the distribution that generates players' signals or the beliefs that players have about how signals are generated. Another consequence in our setting is that an outsider without knowledge about how signals are generated or how players form their beliefs could step in and help coordination by suggesting an ordinal equilibrium profile. That ordinal equilibria are *ex post* is intuitive given the fact that our informal discussion at the beginning of this section made no references to the underlying signal distribution; once I know that one of my partners switches after the first period with probability one, it is a best response for everyone else to stick to their high probability action independent of the signal realizations; and once everyone else does not switch, is is clearly a best response for the designated player to switch independent of her and other players' private information. Similarly, in later periods—as long as they are promising—exactly one player will switch to a lower probability action and given the behavior of other players, this is optimal independent of the signal realization. But it is worth emphasizing that we also show that basically only ordinal equilibria are *ex post*. Hence a management that wants a robust solution with respect to the underlying distribution needs to select a routine. Finally, in Proposition 10 we show that routines are also hyper-cursed equilibria: they are robust to a wide variety of incorrect beliefs by the team members. A management having to deal with less than perfectly rational agents may thus benefit from selecting routines, which are robust to a variety of behavioral biases documented in the literature on behavioral economics.

5.2.1 Routines are *ex post* Equilibria

The following proposition, which focuses on pure strategies, establishes a simple and clean equivalence of the set of ordinal and the set of *ex post* equilibria. We discuss more general results for mixed strategies below (Proposition 9).

Proposition 8 *The set of pure-strategy* ex post *equilibria coincides with the set of pure-strategy ordinal equilibria.*

It is straightforward to see that every ordinal equilibrium σ is an *ex post* equilibrium, even if we allow for mixed strategies: According to Proposition 4, in an ordinal equilibrium in every period in which there is a positive probability of a success, a player plays a maximal action. The property of an action being maximal for Player i does not depend whether or not the entire signal vector ω is known; i.e., an action that is maximal for Player i when ω_{-i} is private information remains maximal when ω_{-i} is made public. Thus even when the signal vector ω is publicly known there is no instantaneous gain from switching to a different action in any period in which there is a positive probability of a success. Evidently, in periods in which there is no positive probability of success, even with knowledge of ω_{-i}, there is no instantaneous gain from deviating from σ. Since actions are not observed, a player can also not hope to affect future play of others by switching from a maximal to a non-maximal action. Therefore, it remains optimal to maximize the instantaneous probability of success by taking a maximal action and hence σ is an *ex post* equilibrium. The converse is established in the appendix.

Since ordinal equilibria are *ex post* even when we allow for mixed strategies, one may wonder whether mixed-strategy *ex post* equilibria also coincide with ordinal equilibria. To understand intuitively, why this is not the case reconsider the example illustrated in FIGURE 1. Then if $T = 6$ there exists an ordinal equilibrium in which in periods $t = 1, \cdots, 5$ the corresponding action profile is played, so that on the path of play all but the a priori-least likely action profile have been played prior to the final period. In the final period Player 1 plays her first action ($a_{1,1}$) while Player 2 randomizes (with any given probability) between the two more likely actions ($a_{2,1}, a_{2,2}$); in this incomplete-search equilibrium no player can deviate in the final period and induce a positive probability of success. Note also that this ordinal equilibrium is *ex post*, which follows from the above arguments for the first 5 periods and the fact that independent of the signal realization no player can induce a success in the final period. If this randomization by Player 2, however, conditions on more than her ordinal ranking of signals, the resulting equilibrium is not ordinal and yet *ex post*. What our next proposition establishes, is that *ex post* equilibria differ from ordinal equilibria only with regard to such inconsequential randomization in which players randomize between multiple maximal action—i.e. randomize in non-promising periods. Thus, subject to this minor qualification, the behavior in *ex post* equilibria coincides with that in ordinal equilibria: The set of (mixed-strategy) *ex post* equilibria share with ordinal equilibria the property of inducing maximal actions in every period.

Proposition 9 *In any* ex post *equilibrium every player plays a maximal action in every period.*

Note that the above Propositions 7 and 9 imply that an optimal routine is also an optimal *ex post* equilibrium, and that an optimal *ex post* equilibrium is a routine.

5.2.2 Robustness to Behavioral Biases

We now turn to illustrate the robustness of routines to behavioral biases, beginning with a lack of strategic sophistication by players that is formally incorporated in the concept of cursed equilibria. In a fully cursed equilibrium (Eyster and Rabin [2005]) every type of every player best responds to the correct probability distribution over the other players' actions that is induced by their equilibrium strategies, but does not properly attribute these actions to the other players' private information.[15] To state this condition formally in our setting, we need to introduce some notation: We denote a sequence of actions for Player i in the T-period game by λ_i. Given, a signal profile ω, a profile λ of such action sequences induces an expected payoff $u_i(\lambda, \omega)$. For any strategy profile σ (with a slight abuse of notation) denote by $\sigma_{-i}(\lambda_{-i}|\omega_{-i})$ the probability with which players other than i follow the partial profile of action sequences λ_{-i} if their signals are given by ω_{-i}. Since $\sigma_{-i}(\lambda_{-i}|\omega_{-i}) = \prod_{j \neq i} \sigma_j(\lambda_j|\omega_j)$, it is measurable and thus we can define the expected average play of others as

$$\overline{\sigma}_{-i}(\lambda_{-i}|\omega_i) := \int_{\Omega_{-i}} \sigma_{-i}(\lambda_{-i}|\omega_{-i}) dp_i(\omega_{-i}|\omega_i),$$

where $p_i(\cdot|\omega_i)$ is Player i's posterior distribution over the other team members' signals conditional on her own. Then σ is a fully cursed equilibrium if for every i, $\omega_i \in \Omega_i$ and every action sequence $\hat{\lambda}_i \in \text{supp}[\sigma_i(\cdot|\omega_i)]$,

$$\hat{\lambda}_i \in \arg\max_{\lambda_i} \int_{\Omega_{-i}} \sum_{\lambda_{-i} \in \Lambda_{-i}} u(\lambda, \omega) \overline{\sigma}_{-i}(\lambda_{-i}|\omega_i) dp_i(\omega_{-i}|\omega_i),$$

where $u(\lambda, \omega)$ denotes the (common) expected payoff if the signal realization is ω and players follow the profile of action sequences λ. Note that because of signal independence in our case $p_i(\omega_{-i}|\omega_i) = F_{-i}(\omega_{-i})$, and $\overline{\sigma}_{-i}(\lambda_{-i}|\omega_i)$ simplifies to

$$\overline{\sigma}_{-i}(\lambda_{-i}) = \prod_{j \neq i} \int_{\Omega_j} \sigma_j(\lambda_j|\omega_j) dF_j(\omega_j).$$

While cursedness captures the idea that players underestimate the extent to which other players' actions depend on their information, our ordinal equilibria are robust to many other biases in information processing. To illustrate this, we will considerably strengthen the cursedness requirement in several dimensions, and show that pure-strategy ordinal equilibria satisfy these conditions.

First, while cursedness requires a player to correctly predict the distribution of other players' action sequences, we can relax this assumption and ask instead that the player merely correctly predicts the support of the distribution. Formally, we strengthen the robustness requirement by asking that σ satisfies the condition that for every Player i, own signal $\omega_i \in \Omega_i$, profile of

[15] As is well known, a fully cursed equilibrium corresponds to an analogy-based equilibrium (Jehiel [2005]) with a "private information analogy partition" (Jehiel and Koessler [2008])—i.e. a partition that groups together all types of other players.

other players' action sequences $\lambda_{-i} \in \cup_{\omega'_{-i}} \operatorname{supp}[\sigma_{-i}(\lambda_{-i}, \omega'_{-i})]$, and every own action sequence $\hat{\lambda}_i \in \operatorname{supp}[\sigma_i(\cdot|\omega_i)]$:

$$\hat{\lambda}_i \in \arg\max_{\lambda_i} u(\lambda, \omega) \;\; \forall\, \omega_{-i}.$$

We refer to any σ that has this property as a ***strongly cursed equilibrium***. In a strongly cursed equilibrium, a player must best respond to any action sequence played by others on the path of play—independent of what the true type of other players is.[16] It is thus robust to *any misperception* of how other players' equilibrium behavior depends on their information. For example, the frequency with which a player thinks her partner's play a given sequence need not match this frequency in equilibrium. In our organizational interpretation in which a routine is selected by a management that cannot observe the team members' private information, this requirement can also be interpreted as a *weak-accountability condition* in the sense that every player believes that her fellow team members choose only action sequences that can be justified in front of the management as being consistent with the management's order for some possible private signal realization.

Second, a player may also put positive weight on some action sequences that are not played in equilibrium as long as she correctly predicts when other players switch between actions and when previously chosen actions are repeated. As an example, think of a player who has three actions of which the third action is always the least likely action. Consider a candidate equilibrium in a two-period game in which she is meant to always play the most likely action. Then, for example, we allow her partner to misperceive her behavior of not switching as always playing the third action even though this is never the most likely action. To capture this formally, for any set $\tilde{\Lambda}_{-i}$ of profiles of action sequences of other players use $\mathcal{L}(\tilde{\Lambda}_{-i})$ to denote the set that is obtained by replacing any action sequence $\lambda_k = (a_k^1, \ldots a_k^T)$ of any Player $k \neq i$ by $\ell(\lambda_k) = (\ell(a_k^1), \ldots \ell(a_k^T))$ where ℓ is any permutation of Player k's set of actions A_k. Then we ask that σ satisfy the condition that for every Player i, own signal $\omega_i \in \Omega_i$, $\lambda_{-i} \in \mathcal{L}\left(\cup_{\omega'_{-i}} \operatorname{supp}[\sigma_{-i}(\lambda_{-i}, \omega'_{-i})]\right)$, and every own action sequence $\hat{\lambda}_i \in \operatorname{supp}[\sigma_i(\cdot|\omega_i)]$:

$$\hat{\lambda}_i \in \arg\max_{\lambda_i} u(\lambda, \omega) \;\; \forall\, \omega_{-i}.$$

Third, in addition to the above misperceptions of other players' behaviors, we can allow for a player to misinterpret her own signal as longs as the ranking of her own signals remains

[16]Strongly cursedness is a far stronger requirement as cursedness; for example in the standard independent-private value auction environments every Nash equilibrium is (fully) cursed but not necessarily strongly cursed. In the second-price independent private-value auction, the dominant strategy equilibrium is strongly cursed. In general, however, *ex post* equilibria need not be cursed and hence also not strongly cursed.

correct.[17] For example, in a setting in which a player's signals are the result of her ability to understand and analyze the basic problem, overconfidence may result in her thinking that the most likely action is part of a success profile with a higher than appropriate probability. Similarly, if a player's signal comes from repeatedly drawing from her signal distribution, the belief-in-small-numbers bias may often lead to overconfidence. Whatever the exact driver of incorrect own beliefs, a player will want to stick to the prescribed play, and in addition the true expected payoffs of the ordinal equilibrium are unaffected by these biases.

Putting it all together, we say that an equilibrium σ is a ***hyper-cursed equilibrium*** if for any Player i, true signal ω_i and perceived own signal $\tilde{\omega}_i$ that satisfies $r(\tilde{\omega}_i) = r(\omega_i)$, $\lambda_{-i} \in \mathcal{L}\left(\cup_{\omega'_{-i}} \text{supp}[\sigma_{-i}(\lambda_{-i}, \omega'_{-i})]\right)$, and every own action sequence $\hat{\lambda}_i \in \text{supp}[\sigma_i(\cdot|\omega_i)]$:

$$\hat{\lambda}_i \in \arg\max_{\lambda_i} u(\lambda, (\tilde{\omega}_i, \omega_{-i})) \ \ \forall\ \omega_{-i}.$$

As our next result shows, in our setting pure-strategy ordinal equilibria are hyper-cursed equilibria, as well as conventional Bayesian equilibria. Intuitively, what matters for a given player is that the other players follow a particular pattern of play—i.e. of switching between their various actions—and not on how the realization of this pattern depends on players' signal realizations. In terms of our example in Section 3, even if the other team member incorrectly plays her low-probability action first, it is optimal to respond with playing one's high-probability action in the first period. Furthermore, if my fellow team member doesn't switch, it is optimal to switch in the second period and if my team member switches, it is optimal to keep playing the high-probability action. And if players play a routine they only condition on the rank of their signals and hence a misperception of their own signal strength is inconsequential as long the ranking of own signals is unaffected. Thus routines are hyper cursed.

Proposition 10 *Every pure-strategy ordinal equilibrium is hyper cursed.*

We have the following trivial consequence:

Corollary 1 *Every pure-strategy ordinal equilibrium is fully cursed.*

We have focused here on an extreme version of strategic naivete, in addition allowing for other misperceptions. The robustness of routines covers also different notions of partial strategic naivete: it extends to partial cursedness as well as all versions of analogy-based equilibrium that partition the type space of other players more finely than the trivial partition.[18] This is easiest

[17] Observe that this notion is tailored to the specific problem we are considering and, in contrast to cursedness, not defined for games more generally.

[18] Jehiel and Koessler [2008], for example, introduce finer partitions of the type space in the Crawford and Sobel [1982] model of communication.

to see when introducing partial sophistication through finer partition in the above analogy-based equilibrium concept: because all types of other players follow the same switching pattern in a routine, and this switching pattern in itself determines a player's best response, it does not matter how a player groups the types of other players. Similarly, since all types of other players have the same switching behavior, a partially cursed player correctly predicts the switching behavior of others in a routine, and hence following the behavior prescribed by the routine remains a best response.

5.3 Suboptimality of Routines

Having characterized routines and shown that their robustness generalizes from the example of Section 3 to our entire class of games, we now turn to lessons on optimal equilibria that generalize from the example. We demonstrate that: (1) it is impossible to implement *ex post* optimal search; (2) optimal equilibria exist and have an intuitive form—they partition the signal space into convex sets; and (3) typically optimal equilibria are cardinal, i.e. players condition on their signal strength in addition to the ranking of signals. Thus, the robustness of (optimal) routines comes typically at the cost of being a suboptimal equilibrium.

The impossibility of *ex post*-optimal search is a simple consequence of the fact that the knowledge required to implement it is distributed across players. *Ex post*-optimal search would require that players calculate the success probability of each action profile conditional on their joint information and then try action profiles in declining order of these probabilities. To see that this is not an equilibrium strategy with the available information, note that for almost any signal vector $\hat{\omega}_i$ of Player i there exists a positive probability set of signal vectors of others players such that the full-information optimal strategy has Player i change her action from period one to period two. At the same time, for the same signal $\hat{\omega}_i$ of player i, there is a positive probability set of signal vectors of other players for which the full-information optimal strategy prescribes that Player i does not change her action between periods one and period two. This behavior cannot be achieved in equilibrium since Player i's behavior can only depend on her own information

Given that *ex post*-optimal search is infeasible, the next question is how well one can do while respecting the players' informational constraints. In order to address this question, we first note that the sets of optimal and of Nash equilibrium profiles can be analyzed in terms of mappings from signals to distributions over action sequences. Since Player i has m^i actions, she can follow one of $\left(m^i\right)^T$ possible action sequences in the T-period game. We denote a typical action sequence of this kind for Player i by λ_i and the set of such action sequences for Player

i by Λ_i. We show in the appendix that in the present environment the sets of optimal and of Nash equilibrium profiles can be fully characterized in terms of the action-sequence mappings $\chi_i : \Omega_i \to \Delta(\Lambda_i)$. This is a consequence of our assumption that players cannot observe and therefore cannot condition their behavior on each others' actions.

Every strategy σ_i of Player i induces a mapping $\chi_i|\sigma_i : \Omega_i \to \Delta(\Lambda_i)$ from signals into distributions over action sequences. Strategies are particularly simple if they are pure and the induced action sequence mappings are measurable with respect to a finite partition. This motivates the following definition:

Definition 1 *If there exists a finite partition $\mathcal{P}$ of the signal space of Player i such that the action-equence mapping $\chi_i|\sigma_i : \Omega_i \to \Delta(\Lambda_i)$ is measurable with respect to $\mathcal{P}$, then σ_i is a* ***partition strategy*** *with respect to $\mathcal{P}$.*

In the example of Section 3 optimal strategies are cutoff strategies and thus partition strategies. In addition, the partition elements are intervals. The following definition generalizes this property to multi-dimensional signal spaces.

Definition 2 *A partition strategy with respect to a partition $\mathcal{P}$ is a* ***convex partition strategy*** *if the elements of $\mathcal{P}$ are convex.*

Our next result shows that optimal strategies exists and that it is without loss of generality to consider strategies that have a simple form. The statement also contains a reminder that as we noted earlier, optimal strategies are equilibria.

Proposition 11 *There exists an optimal strategy profile in convex partition strategies and any optimal profile is an equilibrium profile.*

The proof of Proposition 11 is in the appendix. It first establishes the fact that a player's payoff from an action sequence is linear in her signal for any partial profile of strategies of other players. This observation is then used to argue that for any strategy profile there exists a profile of convex partition strategies that yields an at least equally high payoff and can be described in terms of a bounded number of points. The space of such strategy profiles is compact and the common payoff is continuous in this class. Hence an optimal strategy profil exists.

Next we identify signal distributions for which one can improve on the best ordinal equilibrium. Since optimal strategy profiles are equilibrium profiles in our common interest environment, it suffices to show that one can improve on the best ordinal equilibrium in order to show that the best equilibrium strategy profile is cardinal. Lemma 1 proves this result for a class of distributions with mass points. Proposition 12 shows that in the neighborhood of any distribution, there are distributions without mass points for which the optimal equilibrium is cardinal.

Say that a player's signal distributions has a *mass point at certainty* if there is positive probability that she receives a signal that singles out one of her actions as the one that is part of a success profile. If a player receives such a signal, we say that she is certain. Similarly, say that a player's signal distributions has a *mass point at indifference* if there is positive probability that she receives a signal that assigns equal probability to each of her actions as being part of a success profile. In the event that she receives such a signal, we say that the player is indifferent. Denote by E_i^C the event that i is certain and by E_i^I the event that she is indifferent.

Intuitively, we can exploit the fact that a player who is certain has no reason to switch even if that is what the equilibrium prescribes, and a player who is indifferent weakly prefers switching to an unused action in the short run. Note that in any ordinal equilibrium we can find a player who switches regardless of her signal in period two and another player who does not switch in period two regardless of her signal. If instead we make these two players more sensitive to their signals by having the former not switch when she is certain and the latter switch when she is indifferent, then in the positive-probability event where both conditions hold, there is a strict gain in period two and in all other events in period two, there is no loss. Furthermore, one can show that there is a simple way to compensate for these strategy changes in future periods such that there any potential loss in future periods is no greater than the gain in period two, and hence due to discounting this increases the players' expected payoffs. Hence we get the following result.

Lemma 1 *If all players' signal distributions have mass points at certainty and at indifference, any optimal equilibrium is cardinal.*

Next, we show that the ability to improve on the best ordinal equilibrium does not critically depend on the distribution of signals having mass points.

Proposition 12 *For each Player i, let F_i have an everywhere positive density f_i. Then there exist sequences of distributions $F_{n,i}$ with everywhere positive densities $f_{n,i}$ and an $N > 0$ such that each $F_{n,i}$ converges weakly to F_i and for all $n > N$, any optimal equilibrium is cardinal.*

The proof proceeds by in a first step approximating each player's signal distribution through a sequence of distribution functions that have mass points at indifference and certainty. For these approximating distributions we know that an improvement is possible from Lemma 1.

In a second step, we then approximate the approximations by distribution functions that have everywhere positive density. For close enough approximations, the property that the optimal equilibrium is cardinal carries over. This shows that if it ever were the case that a routine is optimal for a given distribution of signals, there must exist a sequence of distributions converging to this distribution with the property that for each distribution in the sequence, the optimal equilibrium is not a routine. As is obvious from the example in Section 3, the converse does not hold. In this sense, routines are often suboptimal.

6 Possible Extensions

In this paper we interpret organizational routines as ordinal equilibria in a setting where a problem-solving team has private signals regarding the most promising action profile and repeated opportunities to solve a given problem. We emphasize a variety of properties of these routines, among them their functioning as solving coordination problems via simple patterns of behavior, their resilience to changing circumstances, their suboptimality in specific circumstances, and their robustness to various behavioral biases. While we believe our model naturally captures many aspects of organizational routines discussed in the literature, there are others we do not investigate. So far the literature on organizational behavior has not converged on a single definition of organizational routines but it—often verbally—discusses various properties and benefits thereof. For example, Becker [2004] notes that routines facilitate coordination, avoid deliberation costs, improve measurability and monitoring, reduce uncertainty, act as repositories of organizational knowledge and competence, and can serve as reference points for organizational change. We leave for future research some of the benefits routines may have in these regards but we briefly conjecture here that some of these aspects can be fruitfully analyzed in natural variants of the framework we provide. To do so, we focus completely on the two-player, two-action, and uniform-distribution example of Section 3.

Consider, for example, the claim that routines help reduce deliberation costs. Regarding individual deliberation, suppose each player needs to think hard about how to interpret her decentralized knowledge regarding the optimal problem-solving approach—which in the current setting is summarized in her private signal. In an ordinal equilibrium, players need to only think about the rank-order of the various signals while in the optimal cardinal equilibrium we characterize, players in addition have to consider their signal strength in order to contemplate whether they should switch their chosen action following a first-period failure. If such individual deliberation is costly, then routines become more desirable.

Routines may also help in collective deliberation. Consider players that discuss beforehand how to approach upcoming coordination problems, without having seen their signal realization yet. To solve for the optimal strategy profile, players have to exchange detailed knowledge about their signal distributions and then find the optimal strategy. Furthermore, whenever the signal distributions change, player have to reconsider their optimal plan of action. In contrast, far less detailed information needs to be exchanged for players to agree on a routine, and these routines remain valid—although they may become suboptimal—even when circumstances change, thereby potentially significantly reducing the need for future collective deliberation. Analyzing this question is left for future research.

Next, consider the question whether routines improve measurability and monitoring. Suppose different agents differ in their problem-solving ability in the sense that they have different signal distributions over which of their action profiles is likely to be part of a success. Then in a strategy similar to the optimal cardinal equilibrium of our example, it is hard for an outside observe to attribute failure even over time.[19] On the other hand, when using problem-solving routines, it is potentially easy for an outside observer to learn the probability with which each agent can identify her more likely action. Whether and under what circumstance routines help in monitoring is thus an exciting question for future research.

Finally, problem-solving teams with different routines generate different values in our setting. Since these routines are robust, we can naturally interpret them as part of what defines an organization and thus as part of its intangible assets. How and when routines (optimally) adapt in a changing environment is also an interesting question for future research. For example, organizations that start out with identical optimal routines will experience different success realizations, and as a result may be prompted at different times to reevaluate their routines in a changing environment. As a result organizations that start out being identical will at different times operate with different routines and during those times experience performance differences.

[19] Also, of course, play in such type of equilibrium would not be static as a player would update her belief as to what signal distribution characterizes the type of her rival.

A Appendix

A.1 Signal Strength Strategy Representation

In the main body of the text, we argued that formally a behavioral strategy for Player i in the $2 \times 2 \times 2$ example maps signals ω_{i1} into three probabilities: (1) $p_1^i(\omega_{i1})$, (2) $q_1^i(\omega_{i1})$, and (3) $q_2^i(\omega_{i1})$. We next prove that any given behavioral strategy $(p_1^i(\omega_{i1}), q_1^i(\omega_{i1}), q_2^i(\omega_{i1}))$ of Player i induces a payoff-equivalent strategy $(p^i(\alpha), q_h^i(\alpha), q_l^i(\alpha))$ that conditions only on the signal strength, where the payoff equivalence holds for any given strategy of Player j and any given signal realization ω_j.

To see this, consider two different signals ω'_{i1} and ω''_{i1} that give rise to the same α. Without loss of generality, suppose that ω'_{i1} identifies action 1 as the high-probability action so that $\alpha = \omega'_{i1} = 1 - \omega''_{i1}$. Given signal and action independence, the success probabilities of action profiles for the signal realizations (ω'_i, ω_j) and (ω''_i, ω_j) equal:

Signal Realization (ω'_i, ω_j)

	a_{j1}	a_{j2}
a_{i1}	$\alpha\omega_{j1}$	$\alpha(1-\omega_{j1})$
a_{i2}	$(1-\alpha)\omega_{j1}$	$(1-\alpha)(1-\omega_{j1})$

Signal Realization (ω''_i, ω_j)

	a_{j1}	a_{j2}
a_{i1}	$\alpha\omega_{j1}$	$\alpha(1-\omega_{j1})$
a_{i2}	$(1-\alpha)\omega_{j1}$	$(1-\alpha)(1-\omega_{j1})$

Success Probabilities of Action Profiles for the case $\alpha = \omega'_{i1} = 1 - \omega''_{i1}$

Since conditional on having signal strength α both ω'_i and ω''_i are equally likely, Player i's expected payoff when following strategy $(p_1^i(\omega_{i1}), q_1^i(\omega_{i1}), q_2^i(\omega_{i1}))$ conditional on having signal strength α is equal to:

$$\frac{p_1^i(\omega'_{i1})+(1-p_1^i(\omega''_{i1}))}{2}p_1^j(\omega_{j1})\alpha\omega_{j1}+\frac{p_1^i(\omega'_{i1})+(1-p_1^i(\omega''_{i1}))}{2}(1-p_1^j(\omega_{j1}))\alpha(1-\omega_{j1})$$
$$+\frac{(1-p_1^i(\omega'_{i1}))+p_1^i(\omega''_{i1})}{2}p_1^j(\omega_{j1})(1-\alpha)\omega_{j1}+\frac{(1-p_1^i(\omega'_{i1}))+p_1^i(\omega''_{i1})}{2}(1-p_1^j(\omega_{j1}))(1-\alpha)(1-\omega_{j1})$$
$$+\delta\frac{p_1^i(\omega'_{i1})+(1-p_1^i(\omega''_{i1}))}{2}p_1^j(\omega_{j1})$$
$$\Big\{\left[q_1^i(\omega'_{i1})+(1-q_2^i(\omega''_{i1}))\right]q_1^j(\omega_{j1})\ 0+\left[q_1^i(\omega'_{i1})+(1-q_2^i(\omega''_{i1}))\right](1-q_1^j(\omega_{j1}))\alpha(1-\omega_{j1})$$
$$+\left[(1-q_1^i(\omega'_{i1}))+q_2^i(\omega''_{i1}))\right]q_1^j(\omega_{j1})(1-\alpha)\omega_{j1}+\left[(1-q_1^i(\omega'_{i1}))+q_2^i(\omega''_{i1}))\right](1-q_1^j(\omega_{j1}))(1-\alpha)(1-\omega_{j1})\Big\}$$
$$+\delta\frac{p_1^i(\omega'_{i1})+(1-p_1^i(\omega''_{i1}))}{2}(1-p_1^j(\omega_{j1}))$$
$$\Big\{\left[q_1^i(\omega'_{i1})+(1-q_2^i(\omega''_{i1}))\right]q_2^j(\omega_{j1})\alpha\omega_{j1}+\left[q_1^i(\omega'_{i1})+(1-q_2^i(\omega''_{i1}))\right](1-q_2^j(\omega_{j1}))0$$
$$+\left[(1-q_1^i(\omega'_{i1}))+q_2^i(\omega''_{i1}))\right]q_2^j(\omega_{j1})(1-\alpha)\omega_{j1}+\left[(1-q_1^i(\omega'_{i1}))+q_2^i(\omega''_{i1}))\right](1-q_2^j(\omega_{j1}))(1-\alpha)(1-\omega_{j1})\Big\}$$
$$+\delta\frac{(1-p_1^i(\omega'_{i1}))+p_1^i(\omega''_{i1})}{2}p_1^j(\omega_{j1})$$
$$\Big\{\left[q_2^i(\omega'_{i1})+(1-q_1^i(\omega''_{i1}))\right]q_1^j(\omega_{j1})\alpha\omega_{j1}+\left[q_2^i(\omega'_{i1})+(1-q_1^i(\omega''_{i1}))\right](1-q_1^j(\omega_{j1}))\alpha(1-\omega_{j1})$$
$$+\left[(1-q_2^i(\omega'_{i1}))+q_1^i(\omega''_{i1}))\right]q_1^j(\omega_{j1})\ 0+\left[(1-q_2^i(\omega'_{i1}))+q_1^i(\omega''_{i1}))\right](1-q_1^j(\omega_{j1}))(1-\alpha)(1-\omega_{j1})\Big\}$$
$$+\delta\frac{(1-p_1^i(\omega'_{i1}))+p_1^i(\omega''_{i1})}{2}(1-p_1^j(\omega_{j1}))$$
$$\Big\{\left[q_2^i(\omega'_{i1})+(1-q_1^i(\omega''_{i1}))\right]q_2^j(\omega_{j1})\alpha\omega_{j1}+\left[q_2^i(\omega'_{i1})+(1-q_1^i(\omega''_{i1}))\right](1-q_2^j(\omega_{j1}))\alpha(1-\omega_{j1})$$
$$+\left[(1-q_2^i(\omega'_{i1}))+q_1^i(\omega''_{i1}))\right]q_2^j(\omega_{j1})(1-\alpha)\omega_{j1}+\left[(1-q_2^i(\omega'_{i1}))+q_1^i(\omega''_{i1}))\right](1-q_2^j(\omega_{j1}))\ 0\Big\}.$$

Recalling that for signal ω'_i action a_{i1} and for signal ω''_i action a_{i2} is the high probability action, we can define a strategy that conditions only on signal strength α by setting

$$\begin{aligned}p^i(\alpha) &= \frac{p_1^i(\omega'_{i1})+(1-p_1^i(\omega''_{i1}))}{2}\\ q_h^i(\alpha) &= \frac{q_1^i(\omega'_{i1})+(1-q_2^i(\omega''_{i1}))}{2}\\ q_l^i(\alpha) &= \frac{q_2^i(\omega'_{i1})+(1-q_1^i(\omega''_{i1}))}{2}.\end{aligned}$$

Mere inspection of the above expected payoff formulae verifies that both strategies induce the same expected payoff. Similarly, any $(p^i(\alpha),q_h^i(\alpha),q_l^i(\alpha))$ can be converted into a a behavioral strategy $(p_1^i(\omega_{i1}),q_1^i(\omega_{i1}),q_2^i(\omega_{i1}))$ by simply setting $p_1^i(\omega'_{i1})=(1-p_1^i(\omega''_{i1}))=p^i(\alpha)$, $q_1^i(\omega'_{i1})=(1-q_2^i(\omega''_{i1}))=q_h^i(\alpha)$, and $q_2^i(\omega'_{i1})=(1-q_1^i(\omega''_{i1}))=q_l^i(\alpha)$. Of course, whenever one of the probabilities $p^i(\alpha),q_h^i(\alpha)$, and $q_l^i(\alpha)$ lies in the open interval $(0,1)$ there are multiple behavioral strategies $(p_1^i(\omega_{i1}),q_1^i(\omega_{i1}),q_2^i(\omega_{i1}))$ that correspond to the same signal strength strategy and induces the same expected payoff; intuitively, a player can use the fact whether action 1 or action 2 is the high-probability action as a private randomization device in such cases.

A.2 Existence of LH-equilibria

Any candidate LH-equilibrium must satisfy equation (5). Consider equation (5) with $i = 1$. Integrating the left- hand side yields

$$\text{LHS} := \frac{1}{8}c_1 + c_1c_2 - c_1c_2^2 - \frac{1}{8}c_1\delta + \frac{1}{2}c_1c_2^2\delta,$$

and by integrating the right-hand side, we obtain

$$\text{RHS} = \frac{1}{8} + c_2 - c_2^2 - c_1\left(\frac{1}{8} + c_2 - c_2^2 - \frac{3}{8}\delta\right).$$

Now solve the equation LHS = RHS for c_1 as a function of c_2 and δ. This produces

$$c_1 = \frac{1 + 8c_2 - 8c_2^2}{2 - 4\delta + 16c_2 - 16c_2^2 + 4\delta c_2^2}.$$

One obtains the corresponding expression for c_2 by everywhere exchanging the subscripts.

$$c_2 = \frac{1 + 8c_1 - 8c_1^2}{2 - 4\delta + 16c_1 - 16c_1^2 + 4\delta c_1^2}.$$

Multiply both sides of the last equation by the denominator of the expression on the right-hand side to obtain:

$$(2 - 4\delta + 16c_1 - 16c_1^2 + 4\delta c_1^2)c_2 = 1 + 8c_1 - 8c_1^2.$$

Use N to denote the numerator in the expression for c_1 and D to denote the corresponding denominator. Substitute $\frac{N}{D}$ for c_1 in the last equation, multiply both sides by D^2 and subtract the right-hand side from both sides to obtain:

$$(2D^2 + 16DN - 16N^2 - 4\delta D^2 + 4\delta N^2)c_2 - (D^2 + 8DN - 8N^2) = 0.$$

Substituting for N and D results in:

$$\begin{aligned}
\Phi(c_2, \delta) \quad &\equiv \quad -4(3 - 12\delta + 4\delta^2 + c_2^2(48 + 348\delta - 136\delta^2) + 12c_2^4(80 - 56\delta + 11\delta^2) \\
&+ \quad 8c_2^5(-48 + 48\delta - 17\delta^2 + 2\delta^3) \\
&- \quad 8c_2^3(84 - 3\delta - 20\delta^2 + 4\delta^3) + c_2(42 - 69\delta - 24\delta^2 + 16\delta^3)) = 0
\end{aligned}$$

To analyze the polynomial $\Phi(c_2, \delta)$, we will make use of its derivative with respect to c_2, which is given by:

$$\begin{aligned}\Psi(c_2,\delta) &\equiv -4(42-69\delta-24\delta^2+16\delta^3+2c_2(48+348\delta-136\delta^2)+48c_2^3(80-56\delta+11\delta^2)\\ &+ 40c_2^4(-48+48\delta-17\delta^2+2\delta^3)-24c_2^2(84-3\delta-20\delta^2+4\delta^3))\end{aligned}$$

Note the following facts:

1.
$$\begin{aligned}\Phi(c_2=-1,\delta) &= 2700(\delta-3)\\ \Psi(c_2=-1,\delta) &= 60(522-261\delta+32\delta^2);\end{aligned}$$
i.e., Φ is negative and increasing at $c_2=-1$ for all $\delta \in (0,1)$.

2.
$$\Phi(c_2=1,\delta) = 12(1-\delta);$$
i.e., Φ is positive at $c_2=1$ for all $\delta \in (0,1)$.

3.
$$\Phi(c_2=\frac{1}{2},\delta) = -18\delta(3-4\delta+\delta^2)) = -18\delta(3-\delta)(1-\delta);$$
i.e., Φ is negative at $c_2=\frac{1}{2}$ for all $\delta \in (0,1)$.

4. The factor that multiplies the highest power of c_2 in $\Phi(c_2,\delta)$ equals $-4(-48+48\delta-17\delta^2+2\delta^3)$ and therefore is positive for all δ. Hence, $\Phi(c_2,\delta)$ is positive and grows without bound for sufficiently large values of c_2.

5.
$$\Psi(c_2=\frac{1}{5},\delta) = -\frac{12}{125}(342+2277\delta-2336\delta^2+512\delta^3);$$
i.e., the derivative of Φ is negative at $c_2=\frac{1}{5}$ for all $\delta \in (0,1)$.

6.
$$\begin{aligned}\Psi\left(c_2=\frac{75}{100},\delta\right) &= 30+132\delta-\frac{1587\delta^2}{8}+\frac{203\delta^3}{4}\\ &> 30+132\delta-200\delta^2+51\delta^3\\ &= 132\delta(1-\delta)+[30-68\delta^2+51\delta^3]\\ &> 0\end{aligned}$$

7.

$$\Psi(c_2 = 1, \delta) \quad = \quad -4(42 - 69\delta + 32\delta^2).$$

Since $42 - 69\delta + 32\delta^2$ does not have real roots and is positive at $\delta = 0$, the derivative of Φ at $c_2 = 1$ is negative for all $\delta \in (0, 1)$

Facts 1 and 5 imply that Φ has a local extremum in the the interval $(-1, \frac{1}{5})$. Facts 5 and 6 imply that Φ has a local extremum in the the interval $(\frac{1}{5}, \frac{75}{100})$. Facts 6 an 7 imply that Φ has a local extremum in the the interval $(\frac{75}{100}, 1)$. Facts 4 and 7 imply that Φ has a local extremum in the the interval $(1, \infty)$. Since Φ is a 5th-order polynomial, this accounts for all of its local extrema and rules out stationary points that are not local extrema.

Facts 2 and 7 imply that Φ achieves a local maximum, $\overline{\gamma}$, in the the interval $(\frac{75}{100}, 1)$. This and fact 7 imply that Φ is positive in the interval $[\overline{\gamma}, 1]$. Fact 3, the observation that there is a local extremum in $(\frac{1}{5}, \frac{75}{100})$ and the fact that there are exactly two extrema in the interval $(\frac{1}{5}, 1)$ imply that there is a local minimum $\underline{\gamma}$ in $(\frac{1}{5}, \frac{75}{100})$. If $\underline{\gamma} \leq \frac{1}{2}$, then, since Φ has no stationary points that are not extrema, it must be strictly increasing in the interval $[\frac{1}{2}, \overline{\gamma}]$ and therefore has a unique root in this interval and since Φ is positive on $[\overline{\gamma}, 1]$, it has a unique root on $[\frac{1}{2}, 1]$. If $\frac{1}{2} < \underline{\gamma}$, then Φ is decreasing and by fact 3 negative on the interval $[\frac{1}{2}, \underline{\gamma}]$, is strictly increasing on the interval $[\underline{\gamma}, \overline{\gamma}]$ and positive on the interval $[\overline{\gamma}, 1]$ from the argument given above. Hence, it has a unique root in the interval $[\frac{1}{2}, 1]$.

In any candidate LH-equilibrium the equation $\Phi(c_i, \delta) = 0$ has to hold for both $i = 1$ and $i = 2$. Since this equation has a unique solution, it has to be symmetric. Using symmetry, it suffices to solve one of the two equations for equilibrium cutoffs in terms of a common value c, i.e. c must satisfy

$$\int_c^1 c\beta d\beta + \int_{\frac{1}{2}}^c [c\beta\delta + c(1-\beta)]\, d\beta = \int_c^1 [c\beta\delta + (1-c)\beta]\, d\beta + \int_{\frac{1}{2}}^c [c\beta\delta + (1-c)(1-\beta)]\, d\beta,$$

which is equivalent to

$$1 + c(6 + 4\delta) - 24c^2 + c^3(16 - 4\delta) = 0$$

Solve this for δ to obtain δ as a function of c

$$\delta = \frac{1 + 6c - 24c^2 + 16c^3}{4c(-1 + c^2)}$$

The derivative of δ with respect to c equals

$$\frac{1}{8}\left(\frac{1}{(-1+c)^2} + \frac{2}{c^2} + \frac{45}{(1+c)^2}\right)$$

which is positive. Since $\delta(1/2) = 0$ and the δ-function is strictly increasing for all $c < 1$, its invertible. Hence, the candidate solution $c(\delta)$ is increasing.

To verify that that the candidate solution $c(\delta)$ is indeed an equilibrium for a given value of δ, it remains to verify that inequality (6) is satisfied by the symmetric cutoff c. As we showed earlier, this amounts to $\delta > \delta^*$ where $\delta^* \approx 0.861276$.

A.3 Existence of mixed equilibria

We show existence of equilibria in which there is a common cutoff c and common mixing probability ξ such that both player use HH for signals above the cutoff and mix for signals below the cutoff, putting probability ξ on LH and $1-\xi$ on HL. Such equilibria exist if and only if there is an LH equilibrium. The proof proceeds in two steps. First, we show that for any given mixing probability ξ' there exists a cutoff $c(\xi')$ that makes players indifferent at the cutoff between following the sequences HH and LH. Second, we show that if (and only if) there is an LH equilibrium, there is a unique mixing probability ξ such that if players use the cutoff $c(\xi)$, they are indifferent between LH and HL for all signals.

Player 1's payoff from using the action sequence HH when her signal is α and Player 2 plays HH for signals β above c, plays LH with probability ξ and HL with probability $1-\xi$ for signals below c, equals

$$HH(\alpha;\xi,c) = 2\int_c^1 \alpha\beta d\beta + 2\int_{\frac{1}{2}}^{c} \xi\left[\alpha(1-\beta)+\alpha\beta\delta\right] + (1-\xi)\left[\alpha\beta+\alpha(1-\beta)\delta\right] d\beta.$$

If Player 1 uses the action sequence LH instead, her payoff under the same conditions equals

$$\begin{aligned} LH(\alpha;\xi,c) &= 2\int_c^1 [(1-\alpha)\beta + \delta\alpha\beta] d\beta \\ &+ 2\int_{\frac{1}{2}}^{c} \xi\left[(1-\alpha)(1-\beta)+\alpha\beta\delta\right] + (1-\xi)\left[(1-\alpha)\beta+\alpha(1-\beta)\delta\right] d\beta. \end{aligned}$$

In equilibrium these payoffs have to be equal to each other at the equilibrium cutoff. Therefore, let us look at the (scaled) difference between these payoffs when Player 1's signal α equals Player 2's cutoff c:

$$\begin{aligned} \Psi(\xi,c) &\equiv \frac{1}{2}[HH(c;\xi,c) - LH(c;\xi,c)] \\ &= (2c-1-\delta c)\int_c^1 \beta d\beta + \int_{\frac{1}{2}}^{c} \xi\left[(2c-1)(1-\beta)\right] + (1-\xi)\left[(2c-1)\beta\right] d\beta. \end{aligned}$$

It is straightforward to check the following three properties of the function Ψ: $\Psi(\xi,\frac{1}{2}) = -\delta\frac{1}{2}\int_{\frac{1}{2}}^{1}\beta d\beta < 0 \ \forall\xi\in[0,1]$; $\Psi(\xi,1) = \int_{\frac{1}{2}}^{1}\xi(1-\beta)+(1-\xi)\beta d\beta > 0 \ \forall\xi\in[0,1]$; and, $\Psi(\xi,c)$ is continuous in c for all ξ. Therefore, by the intermediate value theorem, for all $\xi \in [0,1]$ there exists a $c(\xi) \in \left(\frac{1}{2}, 1\right)$ such that $\Psi(\xi, c(\xi)) = 0$.

Note that at any solution $c(\xi)$ of the equation $\Psi(\xi, c) = 0$, we must have $(2c(\xi) - 1 - \delta c(\xi)) < 0$. From this fact it follows that $\partial\Psi(\xi, c(\xi))/\partial c > 0$ for all $\xi \in [0,1]$. This and the fact that Ψ is continuously differentiable implies that for any $\xi \in [0,1]$ the solution $c(\xi)$ is unique. (To see this in more detail, suppose first that the set of solutions has an accumulation point $c^*(\xi)$, i.e. all open neighborhoods of $c^*(\xi)$ contain a solution other than $c^*(\xi)$. By continuity of Ψ, $c^*(\xi)$ is itself a solution and therefore $\partial\Psi(\xi, c^*(\xi))/\partial c > 0$. This, however, is inconsistent with $c^*(\xi)$ being an accumulation point of the set of solutions. This implies that for every point there is a sufficiently small open neighborhood that contains no more than finitely many solutions. This implies that every compact interval of $\mathbb{R}$ contains only finitely many solutions. Now suppose that there are at least two solutions, c_1 and $c_2 > c_1$, and consider a compact interval I that contains c_1 and c_2. Since there are only finitely many solutions in I, the set of solutions c that satisfy $c > c_1$ has a smallest element, $\tilde{c}$. Since $\partial\Psi(\xi, c_1)/\partial c > 0$ and $\partial\Psi(\xi, \tilde{c})/\partial c > 0$, there exist c' and c'' with $c_1 < c' < c'' < \tilde{c}$ such that $\Psi(\xi, c') > 0$ and $\Psi(\xi, c'') < 0$. But then continuity of Ψ and the intermediate value theorem imply that there is a solution in the interval $(c_1, \tilde{c})$, which contradicts the definition of $\tilde{c}$.) By the implicit function theorem for any $\xi \in [0,1]$ there exists an $\epsilon(\xi) > 0$ such that $c(\xi)$ is continuously differentiable for all ξ' with $|\xi' - \xi| < \epsilon(\xi)$. Since this is true for all $\xi \in [0,1]$, $c(\xi)$ is continuously differentiable for all $\xi \in [0,1]$.

It remains to show that there is a mixing probability ξ of Player 2 that for all signals α makes Player 1 indifferent between the action sequences LH and HL when Player 2 uses the cutoff $c(\xi)$. We first show this for $\alpha = 1$ and will argue below that this suffices. Player 1's payoff from action sequence LH, given signal $\alpha = 1$ against a Player 2 who mixes with probability ξ and uses cutoff $c(\xi)$ equals

$$LH(1; \xi, c(\xi)) = \int_{c(\xi)}^{1} \delta\beta d\beta + \int_{\frac{1}{2}}^{c(\xi)} \xi\left[\beta\delta\right] + (1-\xi)\left[(1-\beta)\delta\right] d\beta.$$

Given the same signals and strategy of Player 2, Player 1's payoff from the sequence HL equals

$$HL(1; \xi, c(\xi)) = \int_{c(\xi)}^{1} \beta d\beta + \int_{\frac{1}{2}}^{c(\xi)} \xi\left[(1-\beta)\delta\right] + (1-\xi)\left[(\beta\delta\right] d\beta.$$

Note that both functions are continuous in ξ. Next evaluate both functions at $\xi = 0$,

$$LH(1, 0, c(0)) = \int_{c(0)}^{1} \delta\beta d\beta + \int_{\frac{1}{2}}^{c(0)} \left[(1-\beta)\delta\right] d\beta,$$

$$HL(1; 0, c(0)) = \int_{c(0)}^{1} \beta d\beta + \int_{\frac{1}{2}}^{c(0)} \left[(\beta\delta\right] d\beta,$$

and observe that $LH(1,0,c(0)) < HL(1,0,c(0))$.

It remains to examine both functions at $\xi = 1$. Recall that by construction, $HH(c(1);1,c(1)) = LH(c(1),1,c(1))$. If in addition we have $LH(c(1);1,c(1)) \geq HL(c(1),1,c(1))$, then $c(1)$ is the equilibrium cutoff in an LH equilibrium. The condition $LH(c(1);1,c(1)) \geq HL(c(1),1,c(1))$ is, however, equivalent to $LH(1;1,c(1)) \geq HL(1,1,c(1))$ because $LH(c\left(\frac{1}{2}\right);1,c(1)) = HL(c\left(\frac{1}{2}\right),1,c(1))$ and the functions $LH(\alpha;1,c(1))$ and $HL(\alpha,1,c(1))$ are affine in α. Thus, the intermediate value theorem implies that such a mixed equilibrium exists whenever there is an LH equilibrium (and it differs from the LH- equilibrium whenever $LH(c(1);1,c(1)) > HL(c(1),1,c(1)))$.

Conversely, if there is no LH equilibrium, then $LH(1;1,c(1)) < HL(1,1,c(1))$. Note that

$$LH(1;\xi,c(\xi)) - HL(1;\xi,c(\xi)) = \int_{c(\xi)}^{1} (\delta-1)\beta d\beta + \int_{\frac{1}{2}}^{c(\xi)} (2\xi-1)\delta(2\beta-1)d\beta$$

Since $\partial\Psi(\xi,c)/\partial\xi = \int_{\frac{1}{2}}^{c} \left[(2c-1)(1-2\beta)\right] d\beta < 0$ and, as we showed above, $\partial\Psi(\xi,c(\xi))/\partial c > 0$, we have $c'(\xi) > 0$. Therefore $LH(1;\xi,c(\xi)) < HL(1;\xi,c(\xi))$ for all $\xi \in (0,1)$, and thus there is no mixed equilibrium of this type when there is no LH equilibrium.

B Action independence and signal independence

Action independence alone is consistent with perfectly correlated signals about the marginals. As an examples consider the case of m actions per player and n players who both observe a common signal that tells them for each of their actions the (marginal) probability that it is part of a success profile. Unlike in the case we consider in this paper, if the distribution of signals is atomless, there is an equilibrium that attains *ex post* efficient exploration with probability one.

Receiving independent signals about the marginals is consistent with violations of action independence. As an example, consider the case of two players each of whom has three actions and receives two possible signals, $(\frac{1}{2},\frac{1}{2},0)$ or $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$. The four different combinations of signals induce for different distributions, indicated as 3×3-matrices, as follows.

$$\begin{array}{ccc}
 & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} & \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \\
\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} & \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} \frac{1}{3} & 0 & \frac{1}{6} \\ 0 & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 \end{pmatrix} \\
\begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} & \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} & \begin{pmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{pmatrix}
\end{array}$$

Here, unlike in the case we consider in the paper, if the game has maximally three time periods, then there exists an *ex post* efficient equilibrium: First take action 1, then action 2; if both players received the signal $(\frac{1}{2}, \frac{1}{2}, 0)$ the success has been found; otherwise the player who received signal $(\frac{1}{2}, \frac{1}{2}, 0)$ repeats the same two actions and the player who received a signal $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ takes the third action.

C Proofs

Proof of Proposition 4

Proof: Suppose not. Then there exists a Player i, a signal ω_i and an action history h_{it} that has positive probability given σ and ω_i, following which there is positive probability that Player i plays an action $a_{ij'}$ that is not maximal. Hence, there exists an action $a_{ij''}$ with $\text{Prob}\left\{\left\{(a_{ij''}, a_{-i}) \notin A(h_t)\right\} \cap \left\{(a_{ij''}, a_{-i}) | a_{-i} \in A^t_{-i}(\sigma_{-i}, h_{it}, \omega_i)\right\} \middle| \ \omega_i, h_{it}, \sigma_{-i}\right\} > 0$ and $r(a_{ij''}, \omega_i) > r(a_{ij'}, \omega_i)$. Consider the signal $\hat{\omega}_i$ for which $r(a_{ij}, \hat{\omega}_i) = r(a_{ij}, \omega_i)\ \forall j$ and $\hat{\omega}_{ij} = \frac{1}{r(a_{ij''}, \omega_i)}$ for all j with $r(a_{ij}, \hat{\omega}_i) \geq r(a_{ij''}, \omega_i)$ and $\hat{\omega}_{ij} = 0$ otherwise. Since σ_i is an ordinal strategy, it prescribes the same behavior following $(\hat{\omega}_i, h_{it})$ as it does after (ω_i, h_{it}). Now consider the following deviation after history $(\hat{\omega}_i, h_{it})$: play a''_{ij} in period t and then after the resulting history $(\hat{\omega}_i, (h_{it}, a_{ij''}))$ use the continuation play that the original strategy σ_i would have prescribed following $(\hat{\omega}_i, (h_{it}, a_{ij'}))$. Now there are two possibilities: Either playing $a_{ij''}$ does induce a success in period t, or it does not. In the latter case, there is no loss from the deviation since for the signal $\hat{\omega}_i$ there would also not have been a success from using action $a_{ij'}$ in period t and the sequence of realized action profiles following period t is identical to the one induced by the original strategy. In the former case the deviation is profitable because of discounting. □

Proof of Proposition 5

Proof: For any strategy profile σ define $p_t(\sigma)$ as the (unconditional) probability of a success in period t, let $\Theta^+ := \{t \in T | p_t(\sigma) > 0\}$ and $\Theta^0 := \{t \in T | p_t(\sigma) = 0\}$. Recall that for each Player j, pure strategy s_j and signal ω_j the action that is induced in period t is denoted by $a^t_j(s_j, \omega_j)$.

We claim that if a profile σ of ordinal strategies is maximal then $a^t_j(s_j, \omega_j) = a^t_j(s'_j, \omega_j)\ \forall t \in \Theta^+, \forall s_j, s'_j \in \text{supp}(\sigma_j)$, and for almost all $\omega_j \in \Omega_j$. We argue by induction on t. The claim is true in period 1 because with a maximal strategy Player j will take her highest probability action, and the probability of a tie among highest probability actions is zero. Suppose $t \in \Theta^+$ and the claim holds for all $\tau < t$. With σ_{-j} fixed and the claim being true for all $\tau < t$ the maximal action in period t is independent of which $s_j \in \text{supp}(\sigma_j)$ Player j used before time t since by the inductive hypothesis all of these pure strategies have induce the same action sequences before time t, except possibly in the event of a tie in the probabilities that the signal ω_j assigns to actions, which occurs with probability zero.

Since for all $j \neq i$ we have $a_j^t(s_j, \omega_j) = a_j^t(s'_j, \omega_j)$ $\forall t \in \Theta^+, \forall s_j, s'_j \in \text{supp}(\sigma_j)$, for almost all $\omega_j \in \Omega_j$, for periods $t \in \Theta^+$ Player i's maximal action, after having adhered to σ_i before time t, is the same for all $s_{-i} \in \sigma_{-i}$. Now consider $t \in \Theta^0$. Since ordinal strategies are mappings from histories and rankings into actions, there are only finitely many ordinal pure strategies. Therefore for every Player j every pure strategy in the support of σ_j has positive probability. Therefore if $p_t(\sigma) = 0$, then $p_t(\sigma_i, s_{-i}) = 0$ for all s_{-i} in the support of σ_{-i}. This implies that any action that is maximal in period $t \in \Theta^0$ against σ_{-i} remains maximal against every s_{-i} in the support of σ_{-i}. Together these two observations show that if σ_i is maximal against σ_{-i}, then it is also maximal against every s_{-i} in the support of σ_{-i}.

Next, we show that if σ_i is maximal against the partial profile of pure strategies s_{-i}, then it is a best reply against that profile. Take any pure strategy s_i that is in the support of Player i's mixed strategy σ_i. We will show by way of contradiction that s_i is a best reply against s_{-i}. Suppose that this is not the case. Then there exists a signal ω_i and a pure strategy s'_i such that Player i's expected payoff conditional on having observed signal ω_i satisfies $\pi_i(s'_i, s_{-i}; \omega_i) > \pi_i(s; \omega_i)$. Let τ be the first period in which $a^\tau((s'_i, s_{-i}), \omega) \neq a^\tau(s, \omega)$. Note that τ is independent of ω_{-i}. There are two possibilities: Either $a^\tau((s'_i, s_{-i}), \omega) \in A^{\tau-1}(s, \omega)$, or $a^\tau((s'_i, s_{-i}), \omega) \in A \setminus A^{\tau-1}(s, \omega)$ in which case $\omega_i(a_i^\tau((s'_i, s_{-i}), \omega)) \leq \omega_i(a_i^\tau(s, \omega))$ since by assumption s_i assigns a maximal action after every positive probability history. Let $\theta > \tau$ be the first period in which $a_{-i}^\tau(s, \omega) = a_{-i}^\theta(s, \omega)$ in case that such a θ exists. Note that θ is independent of ω.

Consider a strategy s''_i with

$$\begin{aligned}
a_i^t(s''_i, \omega) &= a_i^t(s'_i, \omega) \;\; \forall t \neq \tau, \theta \\
a_i^\tau(s''_i, \omega) &= a_i^\tau(s_i, \omega) \\
a_i^\theta(s''_i, \omega) &= a_i^\tau(s'_i, \omega).
\end{aligned}$$

Evidently, either this raises the probability of finding a success in period τ by the same amount that it lowers it in period θ, or the two probabilities are the same. Because of discounting, in both cases replacing s'_i with s''_i weakly raises the payoff for the signal ω_i.

If there is no $\theta > \tau$ with $a_{-i}^\tau(s, \omega) = a_{-i}^\theta(s, \omega)$, replace s'_i with a strategy s''_i such that

$$\begin{aligned}
a_i^t(s''_i, \omega) &= a_i^t(s'_i, \omega) \;\; \forall t \neq \tau \\
a_i^\tau(s''_i, \omega) &= a_i^\tau(s, \omega).
\end{aligned}$$

Evidently, also in this case, the payoff of type ω_i weakly increases.

Iterating this procedure generates a sequence of action profiles that converges to $a^t(s, \omega)$. Furthermore the payoff of type ω_i is non-decreasing at each step of the iteration, contradicting the assumption that s'_i induces a strictly higher payoff for type ω_i than s_i against s_{-i}. This confirms that the strategy s_i is a best response against s_{-i}.

Finally, observe that this is true for every s_{-i} in the support of σ_{-i} and for every s_i in the support of σ_i. It follows that σ_i is a best reply against σ_{-i}. $\square$

Proof of Proposition 7

Proof: If we use $(n_1, n_2, \ldots, n_I)$, $n_i \in \{n \in \mathbb{N}_1 | n \leq m^i\}$ to label the rank-ordered action profile in which Player i plays her n_ith ranked action, the path in which rank-ordered profiles are played in lexicographic order, $(1, \ldots 1, 1), (1, \ldots, 1, 2) \ldots$, is maximal. Hence, Proposition 5 implies that there is always an ordinal equilibrium that induces a search path without repetitions.

From Proposition 4 we know that in any ordinal equilibrium in any period with a positive success probability each player uses a maximal action. Therefore, every ordinal equilibrium induces a deterministic sequence of times t at which a novel rank-labeled action profile a^t is played; at any other time s players must induce a distribution over rank-labeled action profiles that have been used earlier and for those times we introduce a generic symbol $*$ that represents "repetition". Call any sequence $\{b^t\}_{t=1}^T$ where each b^t is either a novel rank-labeled action profile a^t or a repetition $*$ an ordinal search path. Clearly, among ordinal search paths, those that induce repetitions (before all rank-labeled action profiles have been exhausted) are dominated. Since there are only finitely many ordinal search paths without repetitions, there must be a payoff maximizing one.

Consider any ordinal equilibrium σ that induces a payoff maximizing search path. Proposition 4 implies that under σ players choose maximal actions in any period with a positive success probability by Proposition 4. Hence, the expected payoff from the profile σ is the present discounted value of expected payoffs from profiles of maximal actions. These expected payoffs can be obtained as follows: Given any signal vector ω, and assuming that Player i takes action a_{ij_i} the expected success probability is $\omega_{1j_1} \times \omega_{2j_2} \times \cdots \times \omega_{Ij_I}$. In case the action taken by players i corresponds to the n_ith order statistic of her signal ω_i, the expected success probability equals $\omega_{1(n_1)} \times \omega_{2(n_2)} \times \cdots \times \omega_{I(n_I)}$. If each Player i follows the rule $a_{i(n_i)}$ the expected success probability equals

$$\int x_1 \times x_2 \times \cdots \times x_I dF_{n_1,n_2,\ldots n_I}(x_1, x_2, \ldots, x_I)$$

where $F_{n_1,n_2,\ldots n_I}$ is the joint distribution of the n_ith order statistics of all players i. By independence, if we let F_{n_i} denote the distribution of the n_ith order statistic of Player i's signal ω_i, this equals

$$\begin{aligned} & \int x_1 dF_{n_1}(x_1) \times \int x_2 dF_{n_2}(x_2) \times \cdots \times \int x_I dF_{n_I}(x_I) \\ = \quad & \overline{\omega}_{1(n_1)} \times \overline{\omega}_{2(n_2)} \times \cdots \times \overline{\omega}_{I(n_I)}. \end{aligned}$$

Because of discounting, the strategy profile σ must prescribes to play the rank-labeled action profiles $(a_{1(n_1)}, a_{2(n_2)}, \ldots, a_{I(n_I)})$ in the order of the probabilities $\overline{\omega}_{1(n_1)} \times \overline{\omega}_{2(n_2)} \times \cdots \times \overline{\omega}_{I(n_I)}$. □

Proof of Proposition 8

Proof: We proved that every ordinal equilibrium is an *ex post* equilibrium in the text.

For the converse consider a pure-strategy *ex post* equilibrium $\tilde{\sigma}$. With pure strategies, if signals are public, a player knows in every period the set of profiles she can induce in that period. For example, if it is known that in period t the partial strategy profile $\tilde{\sigma}_{-i}$ induces the partial action profile a_{-i}, then Player i can induce the set of profiles $P(a_{-i}) := \{a' \in A | a'_{-i} = a_{-i}\}$. These sets partition the set of all strategy profiles A. Therefore, whenever a player can induce the action profile a, the set of profiles she can induce is $P(a_{-i})$; call this her option set. With pure strategies and ω public, in any period τ in which her option set, as determined by profile $\tilde{\sigma}$ equals $O(\tau, \tilde{\sigma})$ Player i also knows the subset of profiles $N(\tau, \tilde{\sigma}_{-i}, h_{i\tau}) \subseteq O(\tau, \tilde{\sigma})$ that have not already been chosen. Since actions are not observable, Player i's choice in period t does not affect her opponents' choices in periods $\tau > t$. Also, if her option set in period t is $O(t, \tilde{\sigma}) = P(a_{-i})$, then her choice in period t, does not directly affect $N(\tau, \tilde{\sigma}_{-i}, h_{i\tau})$ in periods $\tau > t$ with $O(\tau, \tilde{\sigma}) \neq P(a_{-i})$. Therefore, her choice in period t only determines the probability of a success in that period and the composition $N(\tau', \tilde{\sigma}_{-i}, h_{i\tau})$ in periods $\tau' > t$ with $O(\tau', \tilde{\sigma}) = O(t, \tilde{\sigma})$. Therefore in period t she effectively faces the problem making an optimal sequence of choices from the set $N(t, \tilde{\sigma})$, where each choice induces a fixed probability of a success. Given discounting, it is optimal to induce profiles in $O(t, \tilde{\sigma})$ in decreasing order of the magnitude of these probabilities, i.e. to take a maximal action in every period. Therefore in an *ex post* equilibrium players must use maximal strategies conditional on signals being public. A pure strategy that is maximal with publicly known signals remains maximal with private signals because only the ranking of one's own signal matters in the determination of whether a action is maximal. Hence, the pure-strategy *ex post* equilibrium $\tilde{\sigma}$ is an ordinal equilibrium. □

Proof of Proposition 9

Proof: Let the strategy profile σ be an *ex post* equilibrium profile. In order to derive a contradiction, suppose there is a period and a player who does not play a maximal action in that period. Let τ be the first period in which this is the case and let player i be the player who does not play a maximal action in period τ. Now consider the case in which each Player $j \neq i$ has received a signal $\hat{\omega}_j$ that puts probability zero on all actions that are ranked lower than their maximal action in period τ. Then all these players must play their maximal action, denoted a_j^*, in period τ. This is the case since all other actions in the current period induce a zero success probability, since a player can always follow the same action sequence in the future as the one prescribed by σ_j and since there is no effect on the future play of other players from Player j's current choice. Note that each Player k must play her maximal action a_k^* in period τ with positive probability. This is the case because it is strictly optimal to do so for the signal that puts zero probability on all lower ranked actions, and for any action sequence λ her payoffs are continuous in her signals. In an *ex post* equilibrium, i's maximal action must remain maximal when all players $j \neq i$ play their maximal action. Suppose not, then a lower ranked action than i's maximal action has positive probability of success when all players $j \neq i$ play a_j^* but the action a_i^* does not have positive probability of success. Furthermore, this remains true when we restrict attention to signal realizations for which Player i plays a_i^* according to σ_i. For such signals i would want to change her behavior *ex post.* For the signal realization where each Player $j \neq i$ observes $\hat{\omega}_j$ and an action a_i that Player i plays with positive probability in period τ according to σ_i, the action profile (a_i, a_{-i}) has a positive probability of a success in any period $t \geq \tau$ only if all players j play a_j^*. To see this, first note that if any Player j uses a lower ranked action than her maximal action a_j^*, the success probability is zero. Second, observe that since $(a_1^*, \ldots, a_I^*)$ is maximal in period τ, for any action $a_{i\tau}$ that i plays with positive probability in period τ any action profile $(a_1', \ldots, a_{i-1}', a_{i,\tau}, a_{i+1}', \ldots, a_I')$ with $a_j' \succ a_j^*$ for some $j \neq i$ and $a_j' \succeq a_j^*$ for all $j \neq i$ has zero probability of inducing a success. This can be shown as follows: Since $(a_1^*, \ldots, a_I^*)$ is maximal, it follows from before that this action profile is played with positive probability. Hence $(a_1', \ldots, a_{i-1}^*, a_{i\tau}, a_{i+1}^*, \ldots, a_I^*)$ has zero probability of inducing a success. Similarly, the profile $(a_1^*, a_2', a_2^*, \ldots, a_{i-1}^*, a_{i\tau}, a_{i+1}^*, \ldots, a_I^*)$ has zero probability of inducing a success. Thus both of these action profiles have been played before and because by assumption in periods prior to τ only maximal actions have been played, the profile $(a_1', a_2', a_2^*, \ldots, a_{i-1}^*, a_{i\tau}, a_{i+1}^*, \ldots, a_I^*)$ must have have been played before. $\square$

Proof of Proposition 10

Proof: Let s be a pure-strategy ordinal equilibrium. For every Player k and any signal ω_k that that player might receive, denote by $\lambda_k(s_k, \omega_k)$ the path of Player k's actions that is induced by her strategy s_k and signal ω_k.

Note that in an ordinal equilibrium for any Player k and any two signals ω_k and $\tilde{\omega}_k$ with $r(\tilde{\omega}_k) = r(\omega_k)$ we have $\lambda_k(s_k, \omega_k) = \lambda_k(s_k, \tilde{\omega}_k)$. Since in addition we are only considering pure-strategy equilibria, it suffices to check that for every Player i, every signal $\omega_i \in \Omega_i$ of that player, every partial profile of action sequences $\lambda_{-i} \in \mathcal{L}\left(\cup_{\tilde{\omega}_{-i}}\{\lambda_{-i}(s_{-i}, \tilde{\omega}_{-i})\}\right)$ of other players, and every own action sequence $\lambda_i(s_i, \omega_i)$, one has

$$\lambda_i(s_i, \omega_i) \in \arg\max_{\lambda_i} u(\lambda, \omega) \;\; \forall \; \omega_{-i}.$$

This, however, is implied by three facts: (1) fixing any ordinal profile of other players, a maximal strategy for Player i against that profile is a best response against that profile; (2) a period is promising for an ordinal strategy profile if and only if it promising for any permutations of any equilibrium action sequences of other players; and, (3) for Player i the property of a strategy being maximal is preserved as long as the behavior pattern of other players does not change, i.e. as long as every other player plays some permutation of one of her equilibrium action sequences.

$\square$

Proof of Proposition 11:

We prove this result by showing that it suffices to limit the search for an optimal profile to a restricted class of strategy profiles, that this class is compact and that the payoff is continuous in this class. For notational convenience, we use ω^i and ω_i interchangeably to denote Player i's signal in this proof.

As a preliminary step, we first note that the sets of optimal and of Nash equilibrium profiles can be analyzed in terms of mappings from signals to distributions over action sequences. Since Player i has m^i actions, she can follow one of $\left(m^i\right)^T$ possible action sequences in the T-period game. We denote a typical action sequence of this kind for Player i by λ_i and the set of such action sequences for Player i by Λ_i. We show below that in the present environment the sets of optimal and of Nash equilibrium profiles can be fully characterized in terms of the action-sequence mappings $\chi_i : \Omega_i \to \Delta(\Lambda_i)$.

This follows from the following four observations: (1) Any strategy σ_i induces an action-sequence mapping $\chi_i|\sigma_i$ that assigns the same probability to action sequences as does σ_i. Conversely, (2) for any action-sequence mapping $\tilde{\chi}_i$ we can find a strategy $\tilde{\sigma}_i$ for Player i such that $\tilde{\chi}_i = \chi_i|\tilde{\sigma}_i$. Then (3) if $\chi_j|\sigma_j = \chi_j|\tau_j$ for all $j \neq i$ and σ_i is a best reply to σ_{-i}, then σ_i is also a best reply to τ_{-i}; and, (4) any strategy τ_i with $\chi_i|\tau_i = \chi_i|\sigma_i$ is also a best reply to σ_{-i} and τ_{-i}. Thus for any optimal strategy profile σ there exists an action-sequence mapping $\chi|\sigma$ that induces the same payoff and conversely for any action sequence mapping χ there exists a strategy profile that induces that mapping. Similarly, for any Nash equilibrium σ the profile of action-sequences $\chi|\sigma$ retains all relevant information about σ in the sense that any other strategy profile τ with $\chi|\tau$ is a Nash equilibrium that induces the same outcome. Conversely, for any profile χ of action-sequence mappings we can check the best-reply property of Nash equilibrium directly, without specifying strategies σ_i beyond the requirement that they induce χ_i.

For observation (1) note that for any behaviorally mixed strategy σ_i the action-sequence mapping χ_i that is defined by

$$\chi_i((a_i(1),\ldots,a_i(T))|\omega_i) := \sigma_i(a_i(1)|\omega_i)\times\sigma_i(a_i(2)|a_i(1),\omega_i)\times\ldots\times\sigma_i(a_i(T)|a_i(1),\ldots,a_i(T-1),\omega_i)$$

for all action sequences $(a_i(1)),\ldots,a_i(T))$ assigns the same probabilities to action sequences. For observation (2) note that for given χ_i any behaviorally mixed strategy that satisfies

$$\sigma_i(a_i(t)|a_i(1),\ldots,a_i(t-1),\omega_i) = \frac{\sum_{\{\lambda_i|\lambda_i(t')=a_i(t'),t'\leq t\}}\chi_i(\lambda_i|\omega_i)}{\sum_{\{\lambda_i|\lambda_i(t')=a_i(t'),t'\leq t-1\}}\chi_i(\lambda_i|\omega_i)}$$

assigns the same probability to action sequences. Observation (3) follows from the fact that an action sequence, despite not being a fully specified strategy, does fully determine behavior after all histories that can be induced by other players. Hence, when Player i deviates her rivals action-sequence mappings fully determine their response to i's deviation. Finally, observation (4) is a simple consequence of the fact that in the present setting any two strategies of Player i that induce the same action-sequence mappings induce the same outcome and therefore the same payoff for Player i.

Definition A1 *If there exists a finite partition of the signal space of Player i such that her strategy σ_i prescribes the same action sequence everywhere on the interior of a given partition element, σ_i is a partition strategy.*

Definition A2 *A convex-partition strategy is a partition strategy based on a partition all of whose elements are convex.*

Since Player i has m^i actions, she can follow one of $\left(m^i\right)^T$ possible action sequences in the T-period game. Let λ^i_t be the action taken by Player i in period t given her action sequence λ^i. Each Player i's strategy can be viewed as a function that maps her signal ω^i into a distribution over action sequences λ^i. Accordingly, we use $\sigma^i_{\lambda^i}\left(\omega^i\right)$ to denote the probability that Player i plays action sequence λ^i after observing the signal ω^i. $\omega^i_{\lambda^i_t}$ denotes the probability that Player i's signal ω^i assigns to the period-t element of her action sequence λ^i.

Since we have a common-interest game, we can focus on Player 1 as representative for all other players. Denote the joint signal distribution of all players other than Player 1 by

$H\left(\omega^2,\ldots,\omega^N\right)$. Then Player 1's payoff in the T-period game as a function of her chosen action sequence λ^1, her signal ω^1 and the strategies of other players σ_{-1} equals

$$\begin{aligned}
&\pi\left(\lambda^1,\omega^1;\sigma_{-1}\right)= \\
&\quad \int \sum_{\lambda^2}\cdots\sum_{\lambda^N}\left(\sigma^2_{\lambda^2}\left(\omega^2\right)\times\ldots\times\sigma^N_{\lambda^N}\left(\omega^N\right)\right) \\
&\quad \times\left\{\omega^1_{\lambda^1_1}\ldots\omega^N_{\lambda^N_1}+\delta\omega^1_{\lambda^1_2}\ldots\omega^N_{\lambda^N_2}\,1_{\left\{\left(\lambda^1_2,\lambda^2_2,\ldots,\lambda^N_2\right)\neq\left(\lambda^1_1,\lambda^2_1,\ldots,\lambda^N_1\right)\right\}}+\ldots\right. \\
&\quad \left.+\delta^T\omega^1_{\lambda^1_T}\ldots\omega^T_{\lambda^N_T}\,1_{\left\{\left(\lambda^1_T,\lambda^2_T,\ldots,\lambda^N_T\right)\notin\left\{\left(\lambda^1_1,\lambda^2_1,\ldots,\lambda^N_1\right),\ldots,\left(\lambda^1_{T-1},\lambda^2_{T-1},\ldots,\lambda^N_{T-1}\right)\right\}\right\}}\right\}dH\left(\omega^2,\ldots,\omega^N\right)
\end{aligned}$$

Inspection of the above payoff function yields the following observation:

Lemma A1 *Player i's payoffs are linear in Player i's signal ω^i, for any given strategies σ_{-i} of other players and any action sequence $\lambda^i=\left(\lambda^i_1,\ldots,\lambda^i_T\right)$ of Player i.*

Next we show that Lemma A1 implies that any strategy profile can be replaced by a convex-partition strategy profile with an at least equally high payoff. Moreover, the latter profile can be described via a bounded number of points.

Lemma A2 *For any strategy profile σ, there exists a profile of convex-partition strategies $\tilde{\sigma}$ such that $\pi\left(\tilde{\sigma}\right)\geq\pi\left(\sigma\right)$ and in which each element of Player i's partition is a convex polytope with at most M^i vertices, where $M^i\equiv\binom{(m^i)^T-1+m^i}{m^i-1}$.*

Proof: Take σ_{-i} as given. Since Player i's payoff from a given action sequence λ^i is linear in her signal ω^i, the set of signals for which a given action profile is optimal satisfies $\left((m^i)^T-1\right)$ linear inequalities that ensure that the payoff from λ^i is higher than from that from any other action sequence $\tilde{\lambda}^i$; an additional m^i inequalities and one equation ensure that the set of signals is a subset of the (m^i-1)-dimensional unit simplex. Note that in the m^i-1 dimensional signal space, a vertex is defined by at least m^i-1 equations. Therefore, in this space an upper bound on the number of vertices of a polytope that is characterized by $k>m^i-1$ inequalities is $\binom{k}{m^i-1}$. Thus, the set of signals for which a given action profile is optimal must be a convex polytope with at most $\binom{(m^i)^T-1+m^i}{m^i-1}=M^i$ vertices. Hence, there exists a best response to σ_{-i} that partitions the signal space of Player i into convex polytopes each of which have at most M^i vertices.

Take any strategy profile σ. Since the set of best responses of Player 1 always includes a convex-partition strategy in which each element of Player 1's partition is a convex polytope with at most M^1 vertices, we can replace σ_1 by such a convex-partition strategy $\tilde{\sigma}_1$ without lowering payoffs. We then have a strategy profile given by $\sigma' = (\tilde{\sigma}_1, \sigma_{-i})$. By the same argument as above, we can replace Player 2's strategy with a convex partition strategy $\tilde{\sigma}_2$ in which each element of Player 2's partition is a convex polytope with at most M^2 vertices, again without lowering payoffs. Iterating, we get a convex-partition strategy profile $\tilde{\sigma}$ such that $\pi(\tilde{\sigma}) \geq \pi(\sigma)$ and in which each element of Player i's partition is a polytope with at most M^i vertices. $\square$

Proof: (of Proposition 11) Let $\overline{\pi} := \sup_{\{\sigma \in \Sigma^T\}} \pi(\sigma)$ and note that $\overline{\pi} < 1$. Consider a sequence of strategy profiles σ_n that satisfies $\lim_{n\to\infty} \pi(\sigma_n) = \overline{\pi}$. By Lemma A2, for each strategy profile σ_n in the sequence, we can find a profile of convex-partition strategies $\tilde{\sigma}_n$ with $\pi(\tilde{\sigma}_n) \geq \pi(\sigma_n)$. Evidently, the sequence $\tilde{\sigma}_n$ satisfies $\lim_{n\to\infty} \pi(\tilde{\sigma}_n) = \overline{\pi}$.

Each convex-partition profile $\tilde{\sigma}_n$ can be represented as a point in a compact Euclidian space: Recall that Player i has $(m^i)^T$ possible action sequences. A convex-partition strategy of Player i assigns each of those action sequences to the interior of a convex polytope with at most $M^i = \binom{(m^i)^T - 1 + m^i}{m^i - 1}$ elements. Therefore, a convex-partition strategy of Player i can be viewed as a point in the set $\Xi_i := \Delta^{(m^i-1)\times M^i \times ((m^i)^T - 1)}$, where the first $m^i - 1$ components describe a point in the signal space, the second $m^i - 1$ components describe a point in the signal space and so on; the first M^i such points are the vertices of the convex polytope on which Player i uses his first action sequence (if the convex polytope assigned to the action sequence has less than M^i vertices, simply repeat one of the vertices; if it has empty interior, the corresponding action sequence is not used with positive probability), likewise the kth M^i-tuple of $m^i - 1$-tuples corresponds to the vertices of the convex polytope on which Player i uses her kth action sequence; it suffices to specify the convex polytopes associated with $((m^i)^T - 1)$ action sequences, because the convex polytope associated with the remaining action sequence is specified by default.

Hence, there exists a convergent subsequence $\tilde{\sigma}_{n_k}$. Denote the limit of this sequence by $\overline{\sigma}$ and note that $\overline{\sigma}$ is a convex-partition strategy profile. For any $\epsilon \in (0, 1)$ and for each Player i, we can find a closed subset $\Phi_i(\epsilon)$ of the signal space such that all elements of $\Phi_i(\epsilon)$ belong to the interior of elements of the partition induced by $\overline{\sigma}_i$ and the probability that i's signal is in $\Phi_i(\epsilon)$ satisfies $\text{Prob}\{\Phi_i(\epsilon)\} > 1 - \epsilon$. Since the boundary of each partition element varies continuously with the vertices defining that element, we also have that for large k everywhere on $\Phi_i(\epsilon)$, the strategy profiles $\overline{\sigma}$ and $\tilde{\sigma}_{n_k}$ induce the same action sequence. Hence, the profiles of action sequences induced by the two strategy profiles differ with as most probability $1 - (1 - \epsilon)^I$. Since the maximum payoff difference from any two strategy profiles is bounded, this implies that the expected payoff from the profile $\overline{\sigma}$ must equal $\overline{\pi}$. $\square$

Proof of Lemma 1

Proof: From Proposition 4 it follows that for any ordinal equilibrium σ, there is a pure-strategy ordinal equilibrium s that is payoff equivalent to σ. For a given ω, label Player i' actions in the sequence in which they are first used by s. Label actions that are not used by s arbitrarily. Since s is an ordinal equilibrium, a_{i1} is the action of Player i with the highest probability of success. In any ordinal equilibrium s, there will be one player, i, who switches in period two, and another player, k, who does not switch in period two. Modify the behavior of these two players as follows: Let i never switch from her first-period action when she is certain. Have k switch in period two to her action a_{k2} when she is indifferent. Have k otherwise not change her behavior, except that in case there exists a first period $\tau > 2$ in which $a^\tau(s,\omega) = (a_{k2}, a^2_{-k}(s,\omega))$, she takes the action a^2_{k1}, instead of a^2_{k2} in period τ. Formally, define s' such that $s'_{-\{i,k\}} = s_{-\{i,k\}}$, i.e. s coincides with s' for all players other than i and k, and

$$\begin{aligned}
a_i^t(s_i',\omega) &= a_i^t(s_i,\omega)\ \forall\omega_i \in \Omega_i \setminus E_i^C,\ \forall t\\
a_i^t(s_i',\omega) &= a_i^1(s_i,\omega)\ \forall\omega_i \in E_i^C,\ \forall t\\
a_k^t(s_k',\omega) &= a_k^t(s_k,\omega)\ \forall\omega_k \in \Omega_k \setminus E_k^I,\ \forall t\\
a_k^2(s_k',\omega) &\neq a_k^1(s_k,\omega)\ \forall\omega_k \in E_k^I\\
a_k^\tau(s_k',\omega) &= a_k^1(s_k,\omega)\ \forall\omega_k \in E_k^I\\
a_k^t(s_k',\omega) &= a_k^t(s_k,\omega)\ \forall\omega_k \in E_k^I,\ \forall t \neq 2,\tau
\end{aligned}$$

There are four possible cases: (1) If Player i is uncertain and Player k is not indifferent, then the sequence in which cells are examined under the modified strategy profile s' is the same as in the original equilibrium s, and therefore payoffs are the same as well. (2) If Player i is certain and Player k is not indifferent, then Player i is using a dominant action and all other players are following the same behavior as under s_{-i}. Consequently, the expected payoff cannot be lower than from all players using strategy s. (3) If Player i is uncertain and Player k is indifferent, the only effect of changing from s to s' is that the order in which two cells are visited is reversed. Furthermore, these cells are only distinguished by Player k's action and Player k is indifferent. Hence payoffs are unchanged in this case. (4) If Player i is certain and Player k is indifferent, the cell examined in period two has a positive success probability under s', whereas that probability is zero under s. Furthermore, since Player i is using a dominant action, and all players other than players i and k do not change their behavior, the overall effect of switching from s to s' is to move the examination of higher probability profiles forward. Therefore, in this case the expected payoff increases. $\square$

Proof of Proposition 12

Proof: Let $e_{ij} \in \Omega_i$ be the signal for Player i that assigns probability one to the jth action of Player i being required for a success profile, and let $z_i \in \Omega$ be the signal that assigns equal probability to each action of Player i being required for a success profile. Define E_i to be the distribution of Player i's signals that assigns probability one to the set of signals $\{e_{i1}, \ldots, e_{iJ(i)}, z_i\}$ and equal probability to all signals in that set.

Let $\zeta_n \in (0,1)$ and $\zeta_n \to 0$. Define $G_{n,i} = \zeta_n E_i + (1-\zeta_n)F_i$ as the distribution that draws ω_i with probability ζ_n from the distribution E_i and with probability $(1-\zeta_n)$ from F_i and let $G_n = \prod_{i=1}^{I} G_{n,i}$. Then $\{G_{n,i}\}_{n=1}^{\infty}$ is a sequence of distributions converging weakly to F_i, denoted $G_{n,i} \overset{w}{\to} F_i$, where each $G_{n,i}$ has mass points at indifference and at certainty. Let $E_{i,k}$ be the distribution of Player i's signals that assigns probability one to the set of signals $\tilde{\Omega}_i \subset \Omega_i$ that are within (Hausdorff) distance $\frac{1}{k}$ from the set $\{e_{i1}, \ldots, e_{im^i}, z_i\}$ and that is uniform on $\tilde{\Omega}_i$. Define $H_{n,i,k} = \zeta_n E_{i,k} + (1-\zeta_n)F_i$ and let $H_{n,k} = \prod_{i=1}^{I} H_{n,i,k}$. Then each $\{H_{n,i,k}\}_{k=1}^{\infty}$ is a sequence of distribution functions with $H_{n,i,k} \overset{w}{\to} G_{n,i}$ where each $H_{n,i,k}$ has an everywhere positive density, and $H_{n,k} \overset{w}{\to} G_n$.

An optimal ordinal strategy examines a new cell in every period in which that is still feasible. Since there are only finitely many paths of play that do so, an optimal ordinal strategy σ_n^k for $H_{n,k}$ exists. Finiteness of the set of such play paths also implies that there is a subsequence of $\{H_{n,k}\}_{k=1}^{\infty}$ for which (after reindexing) each $\{\sigma_n^k\}_{k=1}^{\infty}$ induces the same play path. From now on consider this subsequence, and pick a strategy σ_n that induces this path of play.

Given a signal realization ω, denote Player i's expected payoff from the strategy profile σ by $v_i(\sigma, \omega)$. Then, for any strategy profile σ and signal distribution F, Player i's expected payoff $U_i(\sigma, F)$ is

$$U_i(\sigma, F) = \int v_i(\sigma, \omega) dF(\omega).$$

Let $1_{\{\sigma,a,t\}}$ be the indicator function of the event that profile a is visited for the first time in period t under strategy σ and let $P(a|\omega)$ stand for the probability that the profile a is a success given the signal vector ω. Then, for an ordinal strategy $\tilde{\sigma}$, Player i's payoff for a fixed ω has the form

$$v_i(\tilde{\sigma}, \omega) = \sum_{t=1}^{T} \delta^{t-1} \left(\sum_{a \in A} 1_{\{\tilde{\sigma},a,t\}}(\omega) P(a|\omega) \right).$$

Here $P(a|\omega)$ is a polynomial in the elements of ω and therefore varies continuously with ω. Since $\tilde{\sigma}$ is ordinal, for any time t the quantity $\sum_{a \in A} 1_{\{\tilde{\sigma},a,t\}}(\omega) P(a|\omega)$ varies continuously with ω: This holds because variations in ω that do not change the ranking of actions do not affect the value of the indicator function, and at points $\tilde{\omega}$ where the indicator function switches from assigning the value 1 to a' to assigning it to a'', we have $P(a'|\tilde{\omega}) = P(a''|\tilde{\omega})$. Taken together, these observations imply that $v_i(\tilde{\sigma}, \omega)$ is continuous in ω. Hence, by weak convergence of $H_{n,k}$ to G_n, we have

$$U_i(\tilde{\sigma}, H_{n,k}) \to U_i(\tilde{\sigma}, G_n),$$

for any ordinal strategy $\tilde{\sigma}$. Therefore, σ_n must be an optimal ordinal strategy for G_n.

For a given G_n, denote by σ'_n the improvement strategy for σ_n, that we constructed in the proof of Lemma 1. For any given σ'_n, σ_n and $\epsilon \in (0, \frac{1}{4})$, we define the strategy σ^ϵ_n as follows:

$$\sigma^\epsilon_{i,n}(\omega_i) = \begin{cases} \sigma_{i,n}(\omega_i) \text{ if } |\omega_i - e_{ij}| > \epsilon \text{ and } |\omega_i - z_i| > \epsilon \\ \frac{\epsilon - x}{\epsilon}\sigma'_{i,n}(z_i) + \frac{x}{\epsilon}\sigma_{i,n}(\omega_i) \text{ if } |\omega_i - z_i| = x \leq \epsilon \\ \frac{\epsilon - x}{\epsilon}\sigma'_{i,n}(e_{ij}) + \frac{x}{\epsilon}\sigma_{i,n}(\omega_i) \text{ if } |\omega_i - e_{ij}| = x \leq \epsilon \end{cases}$$

Note that the payoff $v_i(\sigma^\epsilon_n, \omega)$ is a continuous function of the signal vector ω. Hence, weak convergence implies that

$$U_i(\sigma^\epsilon_n, H_{n,k}) \to U_i(\sigma^\epsilon_n, G_n) \text{ as } k \to \infty.$$

By construction, we have

$$U_i(\sigma'_n, G_n) > U_i(\sigma_n, G_n),$$

and

$$U_i(\sigma^\epsilon_n, G_n) \to U_i(\sigma'_n, G_n) \text{ as } \epsilon \to 0.$$

Since σ_n is ordinal, the payoff $v_i(\sigma_n, \omega)$ is a continuous function of the signal vector ω. Hence, weak convergence implies that

$$U_i(\sigma_n, H_{n,k}) \to U_i(\sigma_n, G_n) \text{ as } k \to \infty.$$

Combining these observations, we conclude that for any n, we can find $k(n)$ and $\epsilon(n)$ such that

$$U_i(\sigma^{\epsilon(n)}_n, H_{n,k(n)}) > U_i(\sigma_n, H_{n,k(n)}).$$

To conclude, simply let $F_n = H_{n,k(n)}$. □

References

ALONSO, R., W. DESSEIN AND N. MATOUSCHEK [2008], "When Does Coordination Require Centralization?" *American Economic Review* **98**, 145-179.

ALPERN, S. [1976], "Hide and Seek Games," Seminar, Institut für höhere Studien, Wien, 26 July.

ALPERN, S. [2002], "Rendezvous Search: A Personal Prespective," *Operations Research* **50**, .

ALPERN, A. AND S. GAL [2003], *The Theory of Search Games and Rendezvous,* Dordrecht, The Netherlands: Kluwer Academic Publishers.

ARGOTE L., S. BECKMAN AND D. EPPLE [1990]: "The Persistence and Transfer of Learning in Industrial Settings," *Management Science* **36**, 140-154.

AUSTEN-SMITH, D., AND J. S. BANKS [1996]: "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review* **90**, 34-45.

BANERJEE, A. [1992], "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* **107**, 797-818.

BECKER, MARKUS C. [2004], "Organizational Routines: A Review of the Literature," *Industrial and Corporate Change* **13**, 643-677.

BIKCHANDANI, S., D. HIRSHLEIFER AND I. WELCH [1992], "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy* **100**, 992-1026.

BLUME, A. AND A.M. FRANCO [2007], "Decentralized Learning from Failure," *Journal of Economic Theory* **133**, 504-523.

BLUME, A., J. DUFFY AND A.M. FRANCO [2009], "Decentralized Organizational Learning: An Experimental Investigation," *American Economic Review* **99**, 1178-1205.

BLUME, A. AND P. HEIDHUES [2006], "Private Monitoring in Auctions," *Journal of Economic Theory* **131**, 179211.

CHAMLEY, C.P. [2004], *Rational Herds: Economic Models of Social Learning,* Cambridge, UK: Cambridge University Press.

Chassang, S. [2010], "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts," *American Economic Review* **100**, 448-465.

Chew W.B., T. Bresnahan and K. Clark [1990], "Measurement, Coordination, and Learning in a Multiplant Network," in *Measures for Manufacturing Excellence*, ed. R Kaplan, 129-162, Harvard Business School Press: Boston, MA.

Cohen, M.D. and P. Bacdayan [1994], "Organizational Routines Are Stored As Procedural Memory: Evidence from a Laboratory Study," *Organization Science* **5**, 554-568.

Crawford, V. and H. Haller [1990], "Learning how to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica* **58**, 571-595.

Crawford, V. and J. Sobel [1982], "Strategic Information Transmission," *Econometrica* **50**, 1431-1451.

Crémer, J., L. Garicano and A. Prat [2007], "Language and the Theory of the Firm," *Quarterly Journal of Economics* **122**, 373-407.

Dewatripont, M. and J. Tirole [2005], "Modes of Communication," *Journal of Political Economy* **113**, 1217-1238.

Ellison, G. and R. Holden [2010], "A Theory of Rule Development," MIT working paper.

Eyster, E. and M. Rabin [2005], "Cursed Equilibrium," *Econometrica* **73**, 1623-1672.

Gibbons, R. [2010], "Inside Organizations: Pricing, Politics, and Path Dependence," *Annual Review of Economics* **2**, 337-365.

Gossner, O. and N. Vieille [2003], "Strategic Learning in Games with Symmetric Information," *Games and Economic Behavior* **42**, 25-47.

Grant, R.M. [1996], "Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration," *Organization Science* **7**, 375-387.

Hayek, F.A. [1945], "The Use of Knowledge in Society," *American Economic Review* **35**, 519-530.

Jehiel, P. [2005], "Analogy-based Expectation Equilibrium," *Journal of Economic Theory* **123**, 81-104.

JEHIEL, P. AND F. KOESSLER [2008], "Revisiting Games of Incomplete Information with Analogy-based Expectations" *Games and Economic Behavior* **62**, 533-557.

MCLENNAN, A. [1998], "Consequences of of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents," *American Political Science Review* **92**, 413-418.

MARCH, J. G., AND SIMON. H. A. [1958], *Organizations,* John Wiley: New York, NY.

MILLER, D. [2011], "Robust Collusion with Private Information," *Review of Economic Studies*, forthcoming.

NELSON, R.R. AND S.G. WINTER [1982], *An Evolutionary Theory of Economic Change,* Belnap Press/Harvard University Press: Cambridge, MA.

POLANYI, M. [1966], *The Tacit Dimension,* Doubleday & Company, Inc.: Garden City, NY.

SALTER W. [1960], *Productivity and Technical Change,* Cambridge University Press: Cambridge, UK.

STASSER, G. AND W. TITUS [1985], "Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussions," *Journal of Personality and Social Psychology* **48**, 1467-1478.

SYVERSON, C. [2011], "What Determines Productivity?," *Journal of Economic Literature* **49**, 326-365.

UNITED STATES. JOINT CHIEFS OF STAFF. SPECIAL OPERATIONS REVIEW GROUP. [1980] *Rescue Mission Report.* Washington, DC. Available at: `http://www.history.navy.mil/library/online/hollowayrpt.htm#execsum`

WISEMAN, T. [2005], "A Partial Folk Theorem for Games with Unknown Payoff Distributions," *Econometrica* **73**, 629-645.

Recent ESMT Working Paper

	ESMT No.
Dynamic coordination via organizational routines Andreas Blume, University of Pittsburgh April M. Franco, Rotman School of Management, University of Toronto Paul Heidhues, ESMT	**11-10**
Naïve and capricious: Stumbling into the ring of self-control conflict Kristian Ove R. Myrseth, ESMT Conny Wollbrant, University of Gothenburg	**11-09**
Decomposing first mover advantages in the mobile telecommunications industry J.P. Eggers, NYU Stern School of Business Michał Grajek, ESMT Tobias Kretschmer, LMU Munich	**11-03 (R1)**
Vertical coordination through renegotiation Özlem Bedre-Defolie, ESMT	**11-08**
Merger efficiency and welfare implications of buyer power Özlem Bedre-Defolie, ESMT Stéphane Caprice, Toulouse School of Economics	**11-07**
Patent examination at the State Intellectual Property Office in China? Johannes Liegsalz, BMW AG Stefan Wagner, ESMT	**11-06**
When and how does corporate social responsibility encourage customer orientation? Daniel Korschun, LeBow College of Business, Drexel University C.B. Bhattacharya, ESMT Scott D. Swain, Northeastern University	**11-05**
Is leadership a part of me? Identity centrality, self-role alignment, and the motivation to lead Laura Guillén, ESMT Konstantin Korotov, ESMT	**11-04**

ESMT
European School of Management and Technology
Faculty Publications
Schlossplatz 1
10178 Berlin
Germany

Phone: +49 (0) 30 21231-1279
publications@esmt.org
www.esmt.org